

Tomáš Výrost – Eduard Baumöhl – Štefan Lyócsa

---

# **Kvantitatívne metódy v ekonómii III.**

---

ELFA 2013

Tomáš Výrost  
Eduard Baumöhl  
Štefan Lyócsa

# **Kvantitatívne metódy v ekonómii III.**

Košice, 2013

Recenzenti:

**Dr. h. c. prof. RNDr. Michal Tkáč, CSc.**

Katedra hospodárskej informatiky a matematiky, Podnikovohospodárska fakulta so sídlom  
v Košiciach, Ekonomická univerzita v Bratislave

**Ing. Silvia Megyesiová, PhD.**

Katedra hospodárskej informatiky a matematiky, Podnikovohospodárska fakulta so sídlom  
v Košiciach, Ekonomická univerzita v Bratislave

**Mgr. Svatopluk Svoboda**

Accenture Services s.r.o, Praha, Česká Republika

Publikácia neprešla jazykovou korektúrou. Za odbornú stránku a jazykovú úpravu textu  
zodpovedajú autori.

Umiestnenie: <http://www.econometrics.sk>

Dostupné od: 07 / 2013

Vydanie prvé

Rozsah: 18.1 AH

© Autori:

Ing. Tomáš Výrost, PhD. – Ing. Eduard Baumöhl, PhD. – Ing. Štefan Lyócsa, PhD.

Podnikovohospodárska fakulta so sídlom v Košiciach, Ekonomická univerzita v Bratislave

2013

Všetky práva vyhradené.

**ISBN 978-80-8086-211-4**

# OBSAH

<b>ÚVOD</b>	<b>5</b>
<b>1 VEKTORY A MATICE</b>	<b>7</b>
1.1 Stručná rekapitulácia poznatkov z teórie množín	7
1.2 Definícia pojmov vektor a matica	10
1.3 Operácie s maticami	17
1.3.1 Transponovaná matica	19
1.3.2 Násobenie matíc	21
1.3.3 Vlastnosti násobenia matíc	25
1.3.4 Inverzná matica a jej vlastnosti	29
1.3.5 Stopa matice	34
1.3.6 Maticové derivácie	36
1.3.7 Blokové matice	37
1.3.8 Kroneckerov súčin	39
<b>2 VYBRANÉ POZNATKY LINEÁRNEJ ALGEBRY</b>	<b>42</b>
2.1 Riešenie sústav lineárnych rovníc	42
2.2 Polia a vektorové priestory	76
2.3 Lineárna nezávislosť vektorov a báza vektorového priestoru	82
2.4 Hodnosť matice	87
2.5 Zmena bázy a lineárne zobrazenia	89
2.6 Skalárny súčin, Euklidovský priestor, norma vektora a ortogonalita	99
2.7 Ortogonálna projekcia	103
<b>3 ZÁKLADY KLASICKEJ EKONOMETRIE</b>	<b>106</b>
3.1 Stručná rekapitulácia poznatkov z teórie pravdepodobnosti	106
3.1.1 Pravdepodobnostný priestor	106
3.1.2 Náhodná premenná	109
3.1.3 Stredná hodnota, kovariancia, disperzia a variančno-kovariančná matica	115

<b>3.2</b>	<b>Formulovanie ekonometrického modelu</b>	<b>126</b>
3.2.1	Typy premenných a charakter dát	128
3.2.2	Jednoduché lineárne vzťahy medzi premennými	130
3.2.3	Aplikácie lineárnych modelov	135
<b>3.3</b>	<b>Odhad lineárneho modelu</b>	<b>141</b>
3.3.1	Vzťah ekonometrického modelu a jeho odhadu	144
3.3.2	Odhad modelu metódou najmenších štvorcov	146
3.3.3	Predpoklady lineárneho modelu	155
3.3.4	Odhad lineárneho modelu metódou maximálnej vierohodnosti	160
3.3.5	Geometrická interpretácia metódy najmenších štvorcov	169
<b>3.4</b>	<b>Vlastnosti odhadu lineárneho modelu</b>	<b>172</b>
3.4.1	Základné vzťahy pre odhadované regresné koeficienty	172
3.4.2	Vlastnosti bodového odhadu – neskreslenosť	176
3.4.3	Vlastnosti bodového odhadu – efektívnosť	177
3.4.4	Vlastnosti bodového odhadu – konzistentnosť	181
3.4.5	Odhad rozptylu poruchového člena	183
3.4.6	Koeficient determinácie	185
<b>3.5</b>	<b>Porušenie predpokladov lineárneho modelu</b>	<b>193</b>
3.5.1	Dôsledky porušenia predpokladov o homoskedasticite a autokorelácii	194
3.5.2	Zovšeobecnená metóda najmenších štvorcov	196
3.5.3	Heteroskedasticita	199
3.5.4	Autokorelácia	204
3.5.5	Multikolinearita	210
<b>4</b>	<b>SIMULAČNÉ METÓDY TESTOVANIA PREDPOKLADOV LINEÁRNEHO MODELU</b>	<b>213</b>
<b>4.1</b>	<b>Testovanie normality rezíduí</b>	<b>213</b>
4.1.1	Monte – Carlo Jarque – Berov test	214
<b>4.2</b>	<b>Testovanie heteroskedasticity rezíduí</b>	<b>225</b>
4.2.1	Breusch – Paganov test	225
4.2.2	Whiteov test	226
4.2.3	Whiteov test – Bootstrap	228
<b>4.3</b>	<b>Testovanie autokorelácie rezíduí</b>	<b>231</b>
4.3.1	HR robust LM Breusch - Godfreyov test	232
4.3.2	Bootstrap HR robust LM Breusch - Godfreyov test	234
<b>5</b>	<b>ÚVOD DO ANALÝZY ČASOVÝCH RADOV</b>	<b>235</b>

<b>5.1</b>	<b>Autoregresívne procesy a procesy klzavých priemerov</b>	<b>235</b>
<b>5.2</b>	<b>Modely ARMA, ARCH a GARCH</b>	<b>242</b>
<b>5.3</b>	<b>Stacionarita časových radov</b>	<b>245</b>
5.3.1	ADF-test	254
5.3.2	ADF-GLS test	261
5.3.3	KPSS test	265
5.3.4	Lee – Strazicich (2003, 2004) test	267
5.3.5	Carrion-i-Silvestre – Sansó (2007) test	269
5.3.6	Ostatné testy na stacionaritu	270
<b>5.4</b>	<b>Problém falošnej regresie</b>	<b>272</b>
<b>5.5</b>	<b>Transformácia časových radov</b>	<b>277</b>
<b>5.6</b>	<b>Kauzalita v Grangerovom zmysle</b>	<b>283</b>
<b>6</b>	<b>PRÍKLADY</b>	<b>290</b>
<b>6.1</b>	<b>Lineárne modely a ich diagnostika</b>	<b>290</b>
<b>6.2</b>	<b>Stacionarita a kauzalita v Grangerovom zmysle</b>	<b>331</b>
	<b>PRÍLOHA 1</b>	<b>364</b>
	<b>PRÍLOHA 2</b>	<b>368</b>
	<b>PRÍLOHA 3</b>	<b>370</b>
	<b>PRÍLOHA 4</b>	<b>375</b>
	<b>ZOZNAM LITERATÚRY</b>	<b>379</b>
	<b>ZOZNAM OBRÁZKOV</b>	<b>388</b>
	<b>ZOZNAM TABULIEK</b>	<b>390</b>
	<b>ZOZNAM PROGRAMOVÝCH KNIŽNÍC</b>	<b>391</b>

## Úvod

Predkladané skriptá nadväzujú na našu publikáciu Kvantitatívne metódy v ekonómii I. a Kvantitatívne metódy v ekonómii II. Vznikali v čase, keď sme sa podieľali na výučbe rovnomenných predmetov na Podnikovohospodárskej fakulte so sídlom v Košiciach, Ekonomickej univerzity v Bratislave (EUBA-PHF).

Na pôde fakulty sme stáli pri vzniku spomínaných výberových predmetov. Naším cieľom bolo ponúknuť študentom možnosť rozšíriť si vedomosti v oblasti kvantitatívnych metód. Prvé dva diely tejto série skript sa zaoberali deskriptívnou a indukčnou štatistikou v prostredí štatistického programu R. Ukázalo sa, že zvládnutie relevantnej teórie a paralelne aj získanie zručností v štatistickom programe, akým je R, predstavuje úlohu trvajúcu viac ako jeden semester. R predstavuje veľmi užitočný, komplexný a flexibilný štatistický program, ktorý je aktívne udržiavaný komunitou vývojárov a tvorí voľnú (open-source) alternatívu k drahším komerčným systémom. Táto komplexnosť a flexibilita však kladie určité nároky aj na používateľa – naučiť sa pracovať v tomto prostredí vyžaduje určité úsilie. Po absolvovaní dvoch semestrov majú naši študenti možnosť pokračovať v štúdiu voliteľného predmetu Finančná ekonometria. Obsahom tohto predmetu je okrem výsledkov klasickej ekonometrie aj jednoduchý úvod k problematike časových radov a modelovanie stacionárnych časových radov. Práve oblasť ekonometrie je spracovaná v tomto treťom pokračovaní série skript.

Okrem samotnej ekonometrickej teórie sme tieto skriptá poňali širšie a preto je ich prvá časť venovaná poznatkom z maticovej a lineárnej algebry. Zistili sme, že študenti mávajú s touto časťou matematiky určité problémy, čo komplikuje výklad v rámci ekonometrie. Teóriu ekonometrie je možné prezentovať veľmi efektívne práve prostredníctvom maticových zápisov. Prvá časť týchto skript je zamýšľaná aj ako referenčný text využiteľný študentmi v situáciách, keď im využívané maticové zápisy a operácie spôsobujú ťažkosti.

Ďalším dôvodom pre zahrnutie poznatkov lineárnej algebry bolo ich využitie pri predmete Meranie efektívnosti produkčných jednotiek I. a II., v rámci ktorých sme sa zaoberali hodnotením efektívnosti založeným na metóde analýzy obalu dát (angl. Data Envelopment Analysis, DEA). Tieto metódy súvisia s úlohami lineárneho programovania, ktoré sú taktiež založené na lineárnej algebre: využívajú sa napríklad lineárne kombinácie vektorov, Gauss-Jordanova eliminačná metóda, elementárna zmena bázy. Poznatky

z lineárnej algebry umožňujú prezentovanú ekonometrickú teóriu rozšíriť aj o modernejšie témy, ktoré by sa inak nevyučovali, napríklad geometrickú interpretáciu metódy najmenších štvorcov.

Skriptá sú rozdelené do šiestich kapitol. Prvá kapitola zhŕňa poznatky o vektoroch a maticiach, ako aj základných operáciách s nimi. Druhá kapitola popisuje vybrané poznatky z maticovej algebry. V tejto kapitole sú definované vektorové priestory, ale aj lineárne zobrazenia a ortogonálne projekcie, ktorých využitie neskôr prezentujeme vo vzťahu k metóde najmenších štvorcov.

Tretia kapitola je venovaná klasickej ekonometrii – po krátkej rekapitulácii nutných pojmov z teórie pravdepodobnosti a štatistiky popisujeme spôsob formulácie ekonometrického modelu, jeho odhad pomocou metódy najmenších štvorcov a metódy maximálnej vierohodnosti. Pri odhade regresných koeficientov sa zaoberáme jeho vlastnosťami, ako aj dôsledkami porušenia predpokladov modelu. Štvrtá kapitola nadväzuje na tretiu ako nadstavbová, kde popisujeme niektoré pokročilé techniky testovania predpokladov modelu pomocou simulačných metód.

V piatej kapitole uvádzame jednoduchý úvod do skúmania stacionárnych časových radov. Popísané sú modely ARMA a GARCH. Keďže stacionarita je kľúčovým predpokladom, ktorému s výnimkou špecializovaných publikácií nebýva venovaný veľký priestor, kapitola obsahuje pomerne podrobne rozpracovaný popis rôznych testov, ktoré je možné pre overenie stacionarity využiť, a to vrátane testov, ktoré zohľadňujú prítomnosť štrukturálnych zlomov v časových radoch. Posledná, šiesta kapitola je venovaná príkladom. Aj keď v texte priebežne prezentujeme rôzne ukážky, rozhodli sme sa zaradiť aj niekoľko príkladov, ktoré sú väčšieho rozsahu a komplexnejšie zachycujú ekonometrickú analýzu vo viacerých krokoch. Všetky príklady sú realizované v programe R.

Predkladané skriptá si nekladú za cieľ byť prelomovým učebným textom. Ako sme uviedli vyššie, vznikali ako súhrn poznatkov, s ktorými sme pracovali v rámci nami vyučovaných predmetov. Tieto skriptá sa zaoberajú dvomi oblasťami – lineárnou algebrou a ekonometriou. Na Slovensku existujú vynikajúce učebnice v oboch oblastiach – ako príklad môžeme uviesť publikácie *Lineárna algebra a geometria* (Zlatoš, 2011) a *Ekonometria* (Hatrák, 2007). Cieľom týchto skript v žiadnom prípade nie je tieto (a iné) učebnice nahrádzať – je ním snaha o vytvorenie pomôcky primárne pre študentov EUBA PHF. Sekundárne sú skriptá určené našim diplomantom ako aj iným čitateľom, ktorým môžu poslúžiť ako prvé priblíženie poznatkov z daných oblastí.

# 1 Vektory a matice

V tejto kapitole uvidíme niektoré základné definície týkajúce sa matematických objektov, ktoré nazývame vektory a matice, ako aj operácie, ktoré s nimi môžeme uskutočňovať. Vektory a matice nám dávajú možnosť veľmi úsporným spôsobom popísať viacero modelov, či už ekonometrických, alebo optimalizačných. Zápis pomocou matíc nám umožňuje namiesto modelov, ktorých rovnice by vyžadovali rádovo desiatky znakov často použiť päť – šesť znakov, pričom pôjde o ekvivalentný zápis bez akéhokoľvek zjednodušenia. Keďže s maticami sú výklad, ako aj praktické výpočty oveľa ľahšie, je výhodné sa s týmto aparátom oboznámiť a následne ho využívať v ďalších častiach.

V celej kapitole sa budeme zaoberať výhradne číselnými maticami a vektormi, ich prvkami teda budú reálne čísla. Budeme taktiež dodržiavať konvenciu v označovaní číselných množín, a písmenami  $\mathbb{N}$ ,  $\mathbb{Z}$ ,  $\mathbb{Q}$  a  $\mathbb{R}$  budeme v tomto poradí označovať prirodzené, celé, racionálne a reálne čísla.

## 1.1 Stručná rekapitulácia poznatkov z teórie množín

V ďalšom texte budeme využívať pojmy pochádzajúce z teórie množín. V tejto podkapitole si v krátkosti zopakujeme niektoré z nich tak, aby sme ich mohli neskôr využívať.

**Množina** predstavuje súbor rôznych objektov, nazývaných **prvkami** množiny. Množiny budeme označovať veľkými písmenami a kurzívou, prvky množín malými písmenami. Ak objekt  $a$  patrí do množiny  $A$ , a teda je jej prvkom, označujeme to  $a \in A$ . Výnimkou sú číselné množiny, ktoré pre ich časté využívanie a ľahkú odlišiteľnosť budeme používať zdvojené písmená. Písmenom  $\mathbb{N}$  bude označovať množinu prirodzených čísel,  $\mathbb{Q}$  množinu racionálnych a  $\mathbb{R}$  množinu reálnych čísel. **Prázdnu množinu** (množinu neobsahujúcu žiadne prvky) budeme označovať  $\emptyset$ .

Uvažujme o ľubovoľných množinách  $A$ ,  $B$  a  $C$ . Ak pre všetky prvky  $a \in A$  platí  $a \in B$ , potom hovoríme, že množina  $A$  je **podmnožinou**  $B$  a zapisujeme to  $A \subseteq B$ .

Ak platí  $A \subseteq B$  a súčasne  $B \subseteq A$ , potom hovoríme o **rovnosti množín**  $A$  a  $B$ , čo zapisujeme  $A = B$ .

Ak je  $A$  podmnožinou  $B$ , ale nie je rovná  $B$ , potom  $A$  nazývame **vlastnou podmnožinou**  $B$ , a značíme to  $A \subset B$ .

**Zjednotením** množín  $A$  a  $B$  nazývame množinu, ktorej prvky sú zároveň prvkami aspoň jednej z množín  $A$  a  $B$ . Zjednotenie množín  $A$  a  $B$  označujeme  $A \cup B$ .

$$A \cup B = \{c; c \in A \vee c \in B\} \quad (1.1)$$

**Prienikom** množín  $A$  a  $B$  nazývame množinu, ktorej prvky patria do množiny  $A$  a súčasne patria aj do množiny  $B$ . Prienik množín  $A$  a  $B$  označujeme  $A \cap B$ . Ak  $A \cap B$ :

$$A \cap B = \{c; c \in A \wedge c \in B\} \quad (1.2)$$

**Rozdielom** množín  $A$  a  $B$  (v tomto poradí) nazývame množinu, ktorej prvky patria do množiny  $A$ , ale nie do množiny  $B$ . **Rozdiel** množín  $A$  a  $B$  označujeme  $A \setminus B$ .

$$A \setminus B = \{c; c \in A \wedge c \notin B\} \quad (1.3)$$

Nech  $A \subseteq B$ . Potom **doplňkom** (komplementom) množiny  $A$  v množine  $B$  nazývame množinu prvkov patriacich do  $B$ , ktoré nepatria do  $A$ . Doplnok množiny  $A$  v množine  $B$  označujeme  $A_B^c$ .

$$A_B^c = \{a^c; a^c \in B \wedge a^c \notin A\} \quad (1.4)$$

Ak je z kontextu zrejmé, že doplnok sa realizuje vzhľadom na množinu  $B$ , označujeme ho skráteno aj  $A^c$ .

**Zobrazenie**  $f: A \rightarrow B$  priraduje každému prvku z  $A$  práve jeden prvok z množiny  $B$ . Ak zobrazenie  $f$  priraduje prvku  $a \in A$  prvok  $b \in B$ , túto skutočnosť zapíšeme ako  $f(a) = b$ .

Zobrazenie  $f$  nazývame surjektívnym ( $f$  je **surjekcia**), ak pre každý prvok  $b \in B$  existuje nejaký prvok  $a_b \in A$ , pre ktorý platí  $f(a_b) = b$ . Ak je  $f$  surjektívne zobrazenie, hovoríme o zobrazení „na množinu  $B$ “, v opačnom prípade o zobrazení „do množiny  $B$ “.

Zobrazenie  $f$  nazývame injektívnym ( $f$  je **injekcia**), ak pre ľubovoľné dva prvky  $a_1, a_2 \in A$ , platí  $a_1 \neq a_2 \Rightarrow f(a_1) \neq f(a_2)$ . Inak povedané, injektívne zobrazenie  $f$  zobrazuje každé dva rôzne prvky z  $A$  na dva rôzne prvky v  $B$ . Injektívne zobrazenia niekedy označujeme aj ako **prosté** zobrazenia.

Ak je zobrazenie  $f$  súčasne surjektívne aj injektívne, a teda splňa obidve vlastnosti, nazývame ho bijektívnym zobrazením ( $f$  je **bijekcia**), resp. vzájomne jednoznačným zobrazením.

Ak existuje **bijekcia** medzi množinou  $A$  a množinou  $\{1, 2, \dots, n\}$  pre  $n \in \mathbb{N}$ , tak hovoríme, že množina  $A$  má **mohutnosť**  $n$  a je konečná. Túto skutočnosť označujeme  $|A| = n$ . Pri konečných množinách predstavuje mohutnosť jej „veľkosť“, resp. počet jej prvkov. Trochu komplikovanejšie je to v prípade množín, ktoré nie sú konečné. Ak existuje bijekcia medzi množinou  $A$  a množinou prirodzených čísel  $\mathbb{N}$ , potom hovoríme, že  $A$  má mohutnosť  $\aleph_0$

(čítame „alef nula“). Množiny, ktoré sú konečné, resp. majú rovnakú mohutnosť  $\aleph_0$  ako množina prirodzených čísel  $\mathbb{N}$ , nazývame **spočítateľné**. Množiny, ktoré uvedenú podmienku nespĺňajú, sú nespočítateľné. Ak existuje bijekcia medzi množinou  $A$  a množinou reálnych čísel  $\mathbb{R}$ , hovoríme, že  $A$  má **mohutnosť kontinua**.

**Potenčná množina**  $\mathcal{P}(A)$  množiny  $A$  je množinou, ktorej prvkami sú všetky podmnožiny  $A$ .

### Príklad 1.1

Vymenujte všetky prvky množiny  $\mathcal{P}(\{1,2,3\})$ .

**Riešenie:** Označme pôvodnú množinu  $A = \{1,2,3\}$ . Potenčná množina je množinou všetkých podmnožín  $A$ . Najprv si musíme uvedomiť, že prázdna množina  $\emptyset$  je podmnožinou každej množiny. Okrem tohto špeciálneho prípadu môžeme z množiny  $A$  postupne vytvárať jednoprvkové, dvojprvkové a jednu trojprvkovú podmnožinu.

Ak vymenujeme všetky možnosti, dostávame

$\emptyset$

$\{1\}, \{2\}, \{3\}$

$\{1, 2\}, \{1, 3\}, \{2, 3\}$

$\{1, 2, 3\}$

Z tohto dôvodu platí

$$\mathcal{P}(\{1,2,3\}) = \{ \emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\} \}$$

Pre konečnú množinu  $A$  s mohutnosťou  $|A| = n$  je počet jej  $k$  - prvkových podmnožín ( $0 \leq k \leq n$ ) rovný kombinačnému číslu:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \quad (1.5)$$

V predchádzajúcom zápise predstavuje výraz  $n!$  faktoriál čísla  $n$ , ktorý je pre 0 definovaný ako  $0! = 1$ , a pre  $n \in \mathbb{N}$  nasledovne:

$$n! = \prod_{i=1}^n i = 1 \cdot 2 \cdot 3 \cdot \dots \cdot (n-2) \cdot (n-1) \cdot n \quad (1.6)$$

Pre počet všetkých podmnožín množiny s mohutnosťou  $n$  (prázdnej, jednoprvkových, dvojprvkových,...) platí vzťah:

$$\binom{n}{0} + \binom{n}{1} + \dots + \binom{n}{n} = \sum_{k=0}^n \binom{n}{k} = 2^n \quad (1.7)$$

### Príklad 1.2

Určite mohutnosť množiny  $\mathcal{P}(A)$  a  $\mathcal{P}(B)$  pre  $A = \{1,2,3\}$  a  $B = \{1,2,3,4,5,6,7,8,9,10\}$ .

**Riešenie:** Ako sme si podrobne ukázali v predchádzajúcom príklade,  $|\mathcal{P}(A)| = 2^3 = 8$ , pričom  $|A| = 3$ .

V prípade množiny  $B$  máme  $|B| = 10$ , takže  $|\mathcal{P}(B)| = 2^{10} = 1024$ . Potenčná množina množiny  $B$  má 1024 prvkov, vrátane prázdnej množiny a celej množiny  $B$  (každá množina je svojou podmnožinou, nie je však svojou vlastnou podmnožinou).

Pre prácu s vektormi má sa definuje aj tzv. karteziánsky súčin množín. **Karteziánskym súčinom** množín  $A$  a  $B$  je množina  $A \times B$ , definovaná ako:

$$A \times B = \{(a,b); a \in A \wedge b \in B\} \quad (1.8)$$

Karteziánsky súčin  $A \times B$  je množina, ktorej prvkami sú usporiadané dvojice, v ktorých prvá zložka je vždy prvkom množiny  $A$  a druhá zložka prvkom množiny  $B$ . V prípade, ak sa zaoberáme karteziánskym súčinom tej istej množiny, napr.  $\mathbb{R} \times \mathbb{R}$ , označujeme to spravidla aj horným indexom v tvare  $\mathbb{R}^2$ . Podobne postupujeme aj pri viacerých súčinoch, napr.  $\mathbb{R} \times \mathbb{R} \times \mathbb{R} = \mathbb{R}^3$ .

### Príklad 1.3

Vymenujte všetky prvky karteziánskeho súčinu  $A \times B$  pre množiny  $A = \{1, 2\}$  a  $B = \{6, 7, 8\}$ .

**Riešenie:** Karteziánsky súčin bude obsahovať usporiadané dvojice. Keďže  $|A| = 2$  a  $|B| = 3$ , pre každý prvok z  $A$  máme tri možné prvky z  $B$ . Dokopy bude  $A \times B$  obsahovať šesť usporiadaných dvojíc:

$$A \times B = \{(1,6), (1,7), (1,8), (2,6), (2,7), (2,8)\}$$

## 1.2 Definícia pojmov vektor a matica

Skôr, než prikrčíme k formálnej definícii pojmu vektor je možno užitočné uviesť určitú intuitívnu situáciu, v ktorej sa prirodzene využíva. Predstavme si, že realizujeme dotazníkový prieskum, a získané odpovede respondentov pripravujeme na ďalšie spracovanie. Pre jednoduchosť uvažujme, že odpovedalo päť respondentov a zapisujeme zatiaľ len odpovede na prvú otázku. V prípade tabuľkového procesora by sme mohli postupovať tak, že by sme odpovede zapísali do stĺpca, pričom odpoveď každého respondenta by sa nachádzala v samostatnom riadku. Mali by sme teda jeden stĺpec celkovo piatich hodnôt (Tabuľka 1).

Tabuľka 1: Odpovede respondentov

	Otázka 1
Respondent $R_1$	11
Respondent $R_2$	21
Respondent $R_3$	31
Respondent $R_4$	41
Respondent $R_5$	51

*Zdroj: vlastné spracovanie*

Ak by sme od tabuľkového procesora prešli k matematike, otázkou je, aký objekt by bol vhodný na uchovanie informácie o odpovediach. Prvou možnosťou, ktorá nás môže napadnúť, je množina. Ak by boli odpovede respondentov postupne 11, 21, 31, 41 a 51, tak naším kandidátom by mohla byť množina  $A = \{11, 21, 31, 41, 51\}$ . Na prvý pohľad je množinový zápis vyhovujúci – uchováva všetky odpovede tak ako boli získané. Dôvod, prečo sa na uchovávanie informácií nevyužívajú množiny je, že nedovoľujú identifikovať odpoveď konkrétneho respondenta. Z definície rovnosti množín vieme, že dve množiny sú si rovné práve vtedy, ak obsahujú rovnaké prvky. Teda napríklad aj množina  $B = \{51, 41, 31, 21, 11\}$ , ale aj  $C = \{11, 51, 21, 41, 31\}$  sú totožné – platí, že  $A = B$  a  $B = C$ , teda aj  $A = C$ .

Nedostatok množinového zápisu je pre naše účely zrejmý – na jednej strane síce umožňuje stanoviť, aké boli odpovede respondentov, ale nie je jasné, kto ako odpovedal. Napríklad z toho, že  $51 \in A$  vieme, že niektorý z respondentov zvolil odpoveď 51. Nevieme však, ktorý. Toto je veľmi zásadná nevýhoda, pretože nás veľmi často zaujímajú vzťahy medzi odpoveďami na rôzne otázky, ako napríklad: je pravda, že ak respondent odpovedal na jednu otázku nízkou hodnotou, potom spravidla odpovedal podobne aj na inú otázku? (napr. „je pravda, že ak respondent uviedol, že patrí do nižšej vekovej skupiny, spravidla kupuje náš tovar častejšie“? Ak by sa potvrdila takáto hypotéza, bolo by to možné využiť napríklad v marketingu pri voľbe cieľovej skupiny a trhovej segmentácii).

Nech  $n \in \mathbb{N}$ . **Vektorom** budeme nazývať usporiadanú  $n$ -ticu reálnych čísel. Vektory označujeme malými tučnými písmenami, napr.  $\mathbf{b} = (b_1, b_2, \dots, b_n)$ , kde  $b_i \in \mathbb{R}$  a  $i \in \mathbb{N}$ . Číslo  $b_i$  nazývame  $i$ -tou zložkou vektora  $\mathbf{b}$ .

Všimnime si, že spôsob, akým sme zapísali vektor  $\mathbf{b}$  zdanlivo odporuje našej intuícii, o ktorej sme rozprávali v súvislosti s prepisovaním dotazníkov do tabuľkového procesora. Rozpor spočíva v tom, že kým v prípade odpovedí sme číselné údaje ukladali do stĺpcov, v prípade matematickej definície sme vektor zapisovali v riadku. V skutočnosti vyššie

uvedenej definícii vyhovujú obidva zápisy – v oboch prípadoch ide o usporiadané  $n$ -tice, je možné definovať poradie jednotlivých hodnôt a zápis je jednoznačný. V prípadoch, ak je potrebné formu zápisu rozlíšiť, hovoríme o riadkových a stĺpcových vektorech.

#### Príklad 1.4

Majme dva vektory  $\mathbf{a}$  a  $\mathbf{b}$  dané nasledovne:

$$\mathbf{a} = (1,2,3,4), \quad \mathbf{b} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$$

Vektor  $\mathbf{a}$  je riadkový vektor,  $\mathbf{b}$  je stĺpcový vektor.

Ak by sme sa vrátili k nášmu príkladu o spracovávaní výsledkov dotazníkového prieskumu, mohli by sme pokračovať tak, že by sme postupne pridávali odpovede na ďalšie otázky. Je zrejme prirodzené postupne pridávať k existujúcej tabuľke ďalšie stĺpce (Tabuľka 2).

Tabuľka 2: Odpovede respondentov (pokračovanie)

	Otázka 1	Otázka 2	Otázka 3	Otázka 4
Respondent $R_1$	11	12	13	14
Respondent $R_2$	21	22	23	24
Respondent $R_3$	31	32	33	34
Respondent $R_4$	41	42	43	44
Respondent $R_5$	51	52	53	54

*Zdroj: vlastné spracovanie*

V takto zapísaných odpovediach sa môžeme veľmi ľahko zorientovať – vieme napríklad, že odpovede respondenta  $R_5$  na všetky štyri položené otázky boli 51, 52, 53 a 54, čo vidíme z posledného riadku tabuľky. Jeho odpovede by sme mohli zapísať aj pomocou vektora  $\mathbf{r}_5 = (51, 52, 53, 54)$ .

Podobne odpovede všetkých respondentov na prvú otázku, teda 11, 21, 31, 41 a 51 v tabuľke tvoria samostatný stĺpec, dali by sa teda považovať za päťzložkový vektor. Celú tabuľku, resp. jej číselnú časť budeme nazývať maticou.

Nech  $m, n \in \mathbb{N}$ . **Maticou** typu  $m \times n$  budeme nazývať schému  $mn$  reálnych čísel zostavených do  $m$  riadkov a  $n$  stĺpcov. Matice budeme označovať veľkými tučnými písmenami.

Čísla vytvárajúce maticu nazývame **prvkami matice** a označujeme ich malými písmenami, s indexmi identifikujúcimi ich umiestnenie v matici. Napr. prvok  $a_{i,j}$  je prvkom matice nachádzajúci sa v  $i$ -tom riadku a  $j$ -tom stĺpci.

#### Príklad 1.5

Majme maticu

$$\mathbf{A} = \begin{pmatrix} 1 & 3 & 7 & 5 \\ 7 & 6 & 8 & 6 \\ 2 & 6 & 9 & 5 \end{pmatrix}$$

Matica  $\mathbf{A}$  je typu  $3 \times 4$  a  $a_{2,3}=8$ .

Vektory a matice sú podobné objekty – vektor môžeme považovať za špeciálny typ matice, ktorá má len jeden riadok (hovoríme o *riadkovom vektore*) alebo len jeden stĺpec (hovoríme o *stĺpcovom vektore*). Uvedená spojitosť medzi maticami a vektormi nám bude veľmi užitočná, pretože nám umožňuje formulovať vlastnosti práce s maticami, ktoré budú automaticky platiť aj pre vektory (keďže sú len špeciálnymi maticami).

Pre účely tohto textu zavedieme konvenciu, že ak explicitne neuvedieme niečo iné, pod pojmom vektor budeme vždy rozumieť stĺpcový vektor. Tento dohovor bude kľúčový v prípade pokročilejších metód založených na maticovom zápise, ako tomu je napríklad v ekonometrii, ale aj pri iných metódach, napr. analýze obálky dát (DEA).

Matica v predchádzajúcom príklade bola typu  $3 \times 4$ . Niekedy pracujeme s maticami, ktoré majú rovnaký počet riadkov a stĺpcov. Maticu typu  $n \times n$  pre  $n \in \mathbb{N}$  budeme nazývať **štvorcovou maticou**. Keďže v jej prípade je počet riadkov a stĺpcov rovnaký, označujeme jej rád len jedným číslom,  $n$ .

Špeciálnym typom štvorcovej matice rádu  $n$  je tzv. **jednotková matica**, ktorej prvky opisuje tzv. Kroneckerova delta, daná vzťahom:

$$\delta_{i,j} = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases} \quad (1.9)$$

Jednotkovú maticu označujeme  $\mathbf{I}_n$ .

#### Príklad 1.6

Matice  $\mathbf{I}_1$ ,  $\mathbf{I}_2$ ,  $\mathbf{I}_3$  a  $\mathbf{I}_4$  vyzerajú nasledovne:

$$\mathbf{I}_1 = (1) \quad \mathbf{I}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \mathbf{I}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \mathbf{I}_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Jednotkovú maticu je možné popísať aj inak. Všimnime si, že v jej prípade sú takmer všetky prvky nulové. V každom riadku (aj stĺpci) je práve jeden prvok nenulový, presnejšie rovný jednej. V prvom riadku je to prvý, v druhom riadku druhý, v treťom tretí a vo štvrtom riadku štvrtý prvok. Rovné jednej sú teda tie prvky, ktoré sú v tom istom riadku ako aj stĺpci. Pri formálnom zápise to sú prvky  $a_{1,1}$ ,  $a_{2,2}$ ,  $a_{3,3}$  a  $a_{4,4}$ , čo samozrejme presne zodpovedá vyššie uvedenej definícii. O prvkoch  $a_{i,i}$  pre  $i \in \{1,2,\dots,n\}$  hovoríme, že tvoria **hlavnú diagonálu** štvorcovej matice. S pomocou tohto pojmu môžeme jednotkovú maticu popísať ako štvorcovú maticu, ktorá má na hlavnej diagonále čísla 1 a mimo nej 0. Ak má matica nenulové prvky len na hlavnej diagonále, nazývame ju **diagonálna**.

Na záver ešte môžeme demonštrovať, ako je možné definovať vektory a matice v prostredí programu R. Vektor vytvoríme pomocou funkcie `c()` – z mnohých možných mnemotechnických pomôcok môžeme zvoliť napríklad anglické slovo `create`, teda vytvoriť.

```
> c(11,12,13,14)
[1] 11 12 13 14
```

Ako vidíme, R realizovalo požadovaný príkaz a vrátilo nám požadovaný vektor. Tento vektor môžeme priradiť do samostatnej premennej. Pomerne jednoducho je možné overiť aj veľkosť vytvoreného vektora (zistiť počet jeho zložiek):

```
> R1 <- c(11,12,13,14)
> R1
[1] 11 12 13 14
> length(R1)
[1] 4
```

Z vektorov môžeme vybrať ktorýkoľvek prvok, napríklad tretí v poradí. Tak isto sa dá zvoliť aj viac ako jeden prvok, napríklad druhý a štvrtý:

```
> R1
[1] 11 12 13 14
> R1[3]
[1] 13
> R1[c(2,4)]
[1] 12 14
```

Pri práci s vektormi môžeme zložiť aj viac vektorov do jedného.

```
> R1
[1] 11 12 13 14
> R2
[1] 21 22 23 24
> c(R1,R2,99,98,97)
[1] 11 12 13 14 21 22 23 24 99 98 97
```

Ak sa vrátíme k nášmu príkladu o dotazníkovom prieskume, mohli by sme definovať aj odpovede ostatných respondentov:

```
> R1 <- c(11, 12, 13, 14)
> R2 <- c(21, 22, 23, 24)
> R3 <- c(31, 32, 33, 34)
> R4 <- c(41, 42, 43, 44)
> R5 <- c(51, 52, 53, 54)
```

Z týchto vektorov je v R možné vytvoriť maticu. Spôsobov, akým to urobiť, je niekoľko. Jednou možnosťou je vytvoriť maticu spájaním po riadkoch (angl. „bind by rows”)

```
> rbind(R1,R2,R3,R4,R5)
  [,1] [,2] [,3] [,4]
R1   11   12   13   14
R2   21   22   23   24
R3   31   32   33   34
R4   41   42   43   44
R5   51   52   53   54
```

Maticu by sme mohli vytvoriť aj spájaním po stĺpcoch (angl. „bind by columns”), ale je zrejmé, že by to nevedlo k požadovanému výsledku – tentoraz by boli jednotlivé otázky v riadkoch, a odpovede respondentov by boli v stĺpcoch, teda presne opačne, ako v predchádzajúcom prípade:

```
> cbind(R1,R2,R3,R4,R5)
  R1 R2 R3 R4 R5
[1,] 11 21 31 41 51
[2,] 12 22 32 42 52
[3,] 13 23 33 43 53
[4,] 14 24 34 44 54
```

Úplne nezávislou možnosťou je definovať maticu príkazom `matrix()`. Tomuto príkazu je potrebné zadať číselné údaje, ktoré maticu vytvoria, počet jej riadkov a stĺpcov, ako aj informáciu, či sa matica bude naplňať po riadkoch alebo stĺpcoch.

```
> vsetkyOdpovede <- c(R1,R2,R3,R4,R5)
> vsetkyOdpovede
```

```

[1] 11 12 13 14 21 22 23 24 31 32 33 34 41 42 43 44 51 52 53 54
> matica <- matrix(vsetkyOdpovede, nrow = 5, ncol = 4, byrow =
  TRUE)
> matica
      [,1] [,2] [,3] [,4]
[1,]  11  12  13  14
[2,]  21  22  23  24
[3,]  31  32  33  34
[4,]  41  42  43  44
[5,]  51  52  53  54

```

Keďže vektor sme si definovali aj ako špecifický prípad matice, je možné vektor nadefinovať aj cez funkciu `matrix()`.

```

> matica <- matrix(R1, nrow = 1, ncol = 4)
> matica
      [,1] [,2] [,3] [,4]
[1,]  11  12  13  14
> matica <- matrix(R1, nrow = 4, ncol = 1)
> matica
      [,1]
[1,]  11
[2,]  12
[3,]  13
[4,]  14

```

Pripomeňme, že podobne ako vektory, je možné indexovať aj matice. Ak by sme chceli načítať odpoveď štvrtého respondenta na otázku číslo 2, urobíme to takto:

```

> matica[4,2]
[1] 42

```

Z matice je možné vyberať aj vektory, ktoré ju tvoria. Môžeme vyberať tak riadkové, ako aj stĺpcové vektory. V takomto prípade nezadáme ako indexy dve čísla ako doteraz (označujúce riadok a stĺpec), ale len jedno z nich. Je nutné však stále zadávať aj oddeľujúcu čiarku, aby bolo možné určiť, aký typ vektora požadujeme. Rovnako je možné vybrať súčasne niekoľko riadkov, prípadne stĺpcov.

```

> matica
      [,1] [,2] [,3] [,4]
[1,]  11  12  13  14
[2,]  21  22  23  24
[3,]  31  32  33  34
[4,]  41  42  43  44
[5,]  51  52  53  54
> matica[,2]
[1] 12 22 32 42 52
> matica[3,]

```

```
[1] 31 32 33 34
> matica[c(3,4),]
      [,1] [,2] [,3] [,4]
[1,]    31    32    33    34
[2,]    41    42    43    44
```

### 1.3 Operácie s maticami

V tejto podkapitole si priblížime základné operácie s maticami. Pripomeňme, že v zmysle zavedenej konvencie môžeme chápať vektory ako špeciálne prípady matíc, ktoré majú len jeden riadok, prípadne stĺpec. Všetky maticové operácie je preto možné využiť aj na počítanie s vektormi. Zároveň stále platí, že budeme pracovať so stĺpcovými vektormi, pokiaľ ich v texte explicitne nedefinujeme ako riadkové.

Uvažujme o dvoch maticiach **A** a **B**, ktoré sú rádu  $m \times n$ , kde  $m, n \in \mathbb{N}$ . **Súčtom matíc A a B** dostávame maticu **C** rádu  $m \times n$ , pre ktorej prvky bude platiť:

$$c_{i,j} = a_{i,j} + b_{i,j} \quad (1.10)$$

kde  $i \in \{1, 2, \dots, m\}$  a  $j \in \{1, 2, \dots, n\}$ . Dá sa teda povedať, že výsledkom súčtu matíc je matica rovnakého typu ako sčítavané matice, a jej prvky sú súčtom zodpovedajúcich si prvkov matíc **A** a **B**. Neutrálnym prvkom pre operáciu sčítania je matica, ktorej všetky prvky sú rovné nule.

Ukážme si najprv, ako sa počíta súčet dvoch vektorov:

```
> x <- c(1, 2, 3, 4, 5)
> y <- c(10, 20, 30, 40, 50)
> x + y
[1] 11 22 33 44 55
```

Tento príklad bol pomerne triviálny a zdá sa, že nás pri tejto operácii nemôže nič prekvapiť. Skúsme však nasledujúci príkaz:

```
> x <- c(1, 2, 3, 4, 5, 6)
> y <- c(1, 2)
> x + y
[1] 2 4 4 6 6 8
```

R uskutočnilo požadovanú operáciu napriek tomu, že sčítavané dva vektory nemajú rovnakú veľkosť. Takáto operácia nezodpovedá vyššie uvedenej definícii a vzniká otázka, akým spôsobom je treba výsledok interpretovať. Ak sa na výsledok pozrieme bližšie, vidíme, že dostávame presne to isté, ako keby sme sčítavali nasledovné dva vektory:

```
> x <- c(1, 2, 3, 4, 5, 6)
> y <- c(1, 2, 1, 2, 1, 2)
```

```
> x + y
[1] 2 4 4 6 6 8
```

R teda v prípade, ak je dĺžka jedného vektora menšia ako dĺžka druhého vektora, doplní kratší vektor tak, aby sa dĺžky vyrovnali a až potom realizuje sčítanie. Dopĺňa pritom z už známych hodnôt kratšieho vektora a to postupne, od prvej až po poslednú zložku. Táto vlastnosť, v R nazývaná ako recyklácia vektora (angl. recycling) umožňuje v mnohých situáciách zjednodušiť prebiehajúce výpočty. Ak by sme napríklad chceli namiesto vektora  $x$  vypočítať odchýlky jeho zložiek od priemeru jeho hodnôt, uskutočnili by sme to nasledovne:

```
> x <- c(1, 2, 3, 4, 5, 6)
> mean(x)
[1] 3.5
> odchylky <- x-mean(x)
> odchylky
[1] -2.5 -1.5 -0.5 0.5 1.5 2.5
```

Funkcia `mean()` vypočíta priemer hodnôt vektora, v našom prípade je to 3.5. Výraz `x-mean(x)` využíva recykláciu: od vektora odpočítavame číslo, čo je vlastne vektor dĺžky jedna. R preto kratší vektor recykluje, a vytvorí vektor obsahujúci samé hodnoty 3.5, ktorý je rovnako dlhý ako  $x$ . Výsledkom rozdielu týchto dvoch vektorov je vektor odchýlky definovaný vyššie.

Súčet je v R možné robiť aj s maticami:

```
> A <- matrix(c(1, 2, 3, 3, 2, 1), ncol = 3, byrow = TRUE)
> B <- matrix(c(1, 0, 2, 3, 4, 5), ncol = 3, byrow = TRUE)
> A
  [,1] [,2] [,3]
[1,]  1   2   3
[2,]  3   2   1
> B
  [,1] [,2] [,3]
[1,]  1   0   2
[2,]  3   4   5
> A+B
  [,1] [,2] [,3]
[1,]  2   2   5
[2,]  6   6   6
```

Ďalšou jednoduchou operáciou je násobenie skalárom – ak máme nejaké číslo  $c \in \mathbb{R}$  a maticu  $A$ , potom pod násobením skalárom  $c$  matice  $A$  budeme rozumieť maticu  $B = cA$ , pre prvky ktorej bude platiť:

$$b_{i,j} = ca_{i,j} \tag{1.11}$$

Číslo  $c$  teda vynásobíme každý prvok matice  $\mathbf{A}$ .

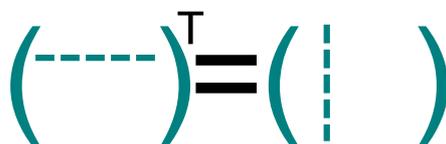
```
> A <- matrix(c(1, 2, 4, 2, 4, 5, 2, 2, 1, 2, 3, 0, 1, 0, 3, 2),
  nrow = 4, byrow = TRUE)
> A
      [,1] [,2] [,3] [,4]
[1,]    1    2    4    2
[2,]    4    5    2    2
[3,]    1    2    3    0
[4,]    1    0    3    2
> 2*A
      [,1] [,2] [,3] [,4]
[1,]    2    4    8    4
[2,]    8   10    4    4
[3,]    2    4    6    0
[4,]    2    0    6    4
```

### 1.3.1 Transponovaná matica

Ak  $\mathbf{A}$  je matica rádu  $m \times n$ , potom k nej **transponovanú maticu** rádu  $n \times m$  označujeme  $\mathbf{A}^T$  a pre jej prvky platí:

$$a_{i,j} = a_{j,i}^T \quad (1.12)$$

kde  $i \in \{1,2,\dots,m\}$  a  $j \in \{1,2,\dots,n\}$ . Zjednodušene povedané, z matice  $\mathbf{A}$  dostaneme k nej transponovanú maticu tak, že vymeníme jej riadky za stĺpce. Túto operáciu si zjednodušene môžeme demonštrovať aj graficky (Obrázok 1).



Obrázok 1: Transponovaná matica

Zdroj: vlastné spracovanie

V programe R slúži na vytvorenie transponovanej matice funkcia  $t()$ .

```
> A <- matrix(c(1, 2, 3, 3, 2, 1), ncol = 3, byrow = TRUE)
> A
      [,1] [,2] [,3]
[1,]    1    2    3
[2,]    3    2    1
> At = t(A)
> At
      [,1] [,2]
[1,]    1    3
[2,]    2    2
[3,]    3    1
> t(At)
```

	[ , 1]	[ , 2]	[ , 3]
[ 1, ]	1	2	3
[ 2, ]	3	2	1

Všimnime si, že kým matica  $\mathbf{A}$  bola rádu  $2 \times 3$ , transponovaná matica  $\mathbf{A}^T$  je rádu  $3 \times 2$ . Taktiež vidíme, že prvý stĺpec matice  $\mathbf{A}^T$  je tvorený prvkami prvého riadku matice  $\mathbf{A}$ , podobne je tomu aj v prípade druhého stĺpca matice  $\mathbf{A}^T$ , ktorý je tvorený prvkami druhého riadku matice  $\mathbf{A}$ . Ako poslednú vlastnosť si môžeme všimnúť, že platí:

$$(\mathbf{A}^T)^T = \mathbf{A} \quad (1.13)$$

teda ak transponujeme transponovanú, dostávame pôvodnú maticu.

Je možné sa stretnúť aj s maticami, pri ktorých transponovaním dostávame tú istú maticu. Matice spĺňajúce túto vlastnosť nazývame **symetrické**. Je hneď zrejmé, že symetrická môže byť len štvorcová matica (pripomeňme, že ak je pôvodná matica rádu  $m \times n$ , transponovaná je  $n \times m$ , pričom štvorcová matica je vždy rádu  $n \times n$ ). Formálne je podmienka symetrickosti matice vyjadrená nasledovne:

$$\mathbf{A} = \mathbf{A}^T \quad (1.14)$$

### Príklad 1.7

Matica

$$\mathbf{A} = \begin{pmatrix} 1 & 3 & 7 \\ 3 & 6 & 8 \\ 7 & 8 & 9 \end{pmatrix} \quad \mathbf{A} = \begin{pmatrix} 1 & 3 & 7 \\ 3 & 6 & 8 \\ 7 & 8 & 9 \end{pmatrix}$$

je symetrická.

Všimnime si, že symetrické matice majú jednu vlastnosť: sú „osovo súmerné“ okolo hlavnej diagonály. Prvky na hlavnej diagonále môžu byť akékoľvek, keďže ide o prvky v tvare  $a_{i,i}$  a tie nezmenia po transponovaní svoju pozíciu. Na druhej strane však, ak máme napríklad prvok  $a_{2,3}$  rovný 8, potom v symetrickej matici nutne musíme mať aj  $a_{3,2} = a_{2,3} = 8$ .

### Príklad 1.8

Doplňte maticu  $\mathbf{A}$  tak, aby bola symetrická.

$$\mathbf{A} = \begin{pmatrix} 1 & 4 & - \\ - & - & 2 \\ 3 & - & 9 \end{pmatrix}$$

**Riešenie:** Zadanie úlohy spĺňa nekonečne veľa možných riešení. Jedným z nich je napríklad:

$$\mathbf{A} = \begin{pmatrix} 1 & 4 & 3 \\ 4 & 6 & 2 \\ 3 & 2 & 9 \end{pmatrix}$$

Napriek tomu, že riešení je nekonečne veľa, pre každé z nich musí platiť, že  $a_{1,2} = a_{2,1} = 4$ ,  $a_{3,1} = a_{1,3} = 3$ ,  $a_{2,3} = a_{3,2} = 2$ . Jediný prvok, ktorý sa môže meniť ľubovoľne, je  $a_{2,2}$ . Ak by sme v úlohe nevynechali žiaden prvok z hlavnej diagonály, riešenie by bolo jediné (alebo žiadne, ak by sme už v zadaní vytvorili spor).

### 1.3.2 Násobenie matíc

Ďalšou základnou operáciou s maticami je ich násobenie. Nech  $m, n, o \in \mathbb{N}$ , matica  $\mathbf{A}$  je typu  $m \times n$ ,  $\mathbf{B}$  je matica typu  $n \times o$ . Potom súčinom matíc  $\mathbf{A}$  a  $\mathbf{B}$  (v tomto poradí) je matica  $\mathbf{C}$  typu  $m \times o$ , pre prvky ktorej platí:

$$c_{i,j} = \sum_{k=1}^n a_{i,k} b_{k,j} \quad (1.15)$$

kde  $i \in \{1, 2, \dots, m\}$  a  $j \in \{1, 2, \dots, n\}$ .

Operácia súčinu matíc je na prvý pohľad pomerne komplikovaná, ozrejmiť nám ju ale môže grafické zobrazenie (Obrázok 2).

$$i \begin{pmatrix} \text{---} \end{pmatrix} \begin{pmatrix} | \end{pmatrix} = \begin{pmatrix} \text{---} \end{pmatrix} \begin{pmatrix} | \end{pmatrix}$$

Obrázok 2: Súčin matíc

Zdroj: vlastné spracovanie

V obrázku je matica  $\mathbf{A}$  označená modrou a matica  $\mathbf{B}$  zelenou farbou. Vidíme, že ak si vezmeme ľubovoľný prvok  $c_{ij}$  v matici  $\mathbf{C}$ , môžeme ho vypočítať tak, že z matice  $\mathbf{A}$  vyberieme  $i$ -tý riadkový vektor, z matice  $\mathbf{B}$   $j$ -tý stĺpcový vektor, ich prislúchajúce si zložky vynásobíme a výsledky sčítame. Teda ak z prvej matice vyberáme  $i$ -tý riadok, aj výsledné číslo v matici  $\mathbf{C}$  bude v  $i$ -tom riadku. Ak z druhej matice vyberáme  $j$ -tý stĺpec, aj výsledné číslo v matici  $\mathbf{C}$  bude v  $j$ -tom stĺpci. To je naznačené aj v poslednej matici v obrázku.

### Príklad 1.9

Vynásobte matice **A** a **B** dané nasledovne:

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

**Riešenie:**

$$\mathbf{AB} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 \cdot 1 + 2 \cdot 3 & 1 \cdot 2 + 2 \cdot 4 \\ 0 \cdot 1 + 1 \cdot 3 & 0 \cdot 2 + 1 \cdot 4 \end{pmatrix} = \begin{pmatrix} 7 & 10 \\ 3 & 4 \end{pmatrix}$$

Pre väčšiu názornosť môžeme zdôrazniť, ako vypočítať napríklad prvok  $c_{1,2}$ :

$$\mathbf{AB} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 \cdot 1 + 2 \cdot 3 & 1 \cdot 2 + 2 \cdot 4 \\ 0 \cdot 1 + 1 \cdot 3 & 0 \cdot 2 + 1 \cdot 4 \end{pmatrix} = \begin{pmatrix} 7 & 10 \\ 3 & 4 \end{pmatrix}$$

Je potrebné si všimnúť, že definícia súčinu neumožňuje násobiť akékoľvek matice. Podobne tomu bolo aj v prípade súčtu – sčítavať sa dali len matice rovnakých rádo. V prípade súčinu je tomu trochu inak. Prepíšme najprv požiadavky kladené na súčin matíc nasledovne:

$$\begin{array}{ccc} \mathbf{A} & \mathbf{B} & = & \mathbf{C} \\ m \times n & n \times o & & m \times o \end{array}$$

Nutnou podmienkou pre násobenie matíc je, aby mali vhodné rozmery, t. j. počet stĺpcov matice **A** (v našom zápise je to  $n$ ) musí byť rovnaký, ako počet riadkov v matici **B** (opäť je to  $n$ ). Dôvod pre túto požiadavku je vidieť aj vyššie (Obrázok 2): pri súčine násobíme prislúchajúce si zložky vektorov z oboch matíc, z tohto dôvodu musia byť riadkové vektory z **A** a stĺpcové vektory z **B** rovnako veľké.

Taktiež si môžeme všimnúť, že výsledok súčinu bude mať počet riadkov rovnaký, ako v matici **A** a počet stĺpcov rovnaký ako v matici **B**.

Všetky uvedené skutočnosti si je možné jednoducho zapamätať pomocou nasledovného pravidla: ak chceme násobiť matice, najprv si napíšme rozmery matíc v tom poradí, ako ich násobíme, napr.  $2 \times 3$  a  $3 \times 4$ . Matice je možné násobiť, ak je druhé a tretie z čísel rovnaké (tu je rovné 3). Súčinom bude matica o rozmeroch, ktoré dostaneme vynechaním druhého a tretieho čísla, teda  $2 \times 4$ .

### Príklad 1.10

Vynásobme matice **A** a **C** dané nasledovne:

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \quad \mathbf{C} = \begin{pmatrix} 1 & 2 & 5 \\ 3 & 4 & 6 \end{pmatrix}$$

#### Riešenie:

Matica **A** je rozmeru  $2 \times 2$  a matica **C**  $2 \times 3$ . Matice je možné násobiť a výsledná matica bude rádu  $2 \times 3$ .

$$\mathbf{AC} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 5 \\ 3 & 4 & 6 \end{pmatrix} = \begin{pmatrix} 1 \cdot 1 + 2 \cdot 3 & 1 \cdot 2 + 2 \cdot 4 & 1 \cdot 5 + 2 \cdot 6 \\ 0 \cdot 1 + 1 \cdot 3 & 0 \cdot 2 + 1 \cdot 4 & 0 \cdot 5 + 1 \cdot 6 \end{pmatrix} = \begin{pmatrix} 7 & 10 & 17 \\ 3 & 4 & 6 \end{pmatrix}$$

V programe R je možné matice násobiť pomocou príkazu `%*%`:

```
> A <- matrix(c(1, 2, 0, 1), nrow = 2, byrow = TRUE)
> B <- matrix(c(1, 2, 5, 3, 4, 6), nrow = 2, byrow = TRUE)
> A
     [,1] [,2]
[1,]    1    2
[2,]    0    1
> B
     [,1] [,2] [,3]
[1,]    1    2    5
[2,]    3    4    6
> A %*% B
     [,1] [,2] [,3]
[1,]    7   10   17
[2,]    3    4    6
```

Ďalej ukážeme jednu veľmi dôležitú skutočnosť, ktorú je potrebné mať na pamäti pri tejto operácii: **násobenie matíc** vo všeobecnosti **nie je komutatívne**. Pri počítaní s reálnymi číslami sme na platnosť komutatívnosti zvyknutí:  $2 \cdot 3 = 3 \cdot 2 = 6$ . Je teda jedno, v akom poradí čísla medzi sebou násobíme. Ako ukážeme v nasledujúcom príklade, na túto vlastnosť sa pri násobení matíc nemôžeme spoľahnúť.

### Príklad 1.11

Nech sú matice **A**, **B** a **C** dané nasledovne:

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad \mathbf{C} = \begin{pmatrix} 1 & 2 & 5 \\ 3 & 4 & 6 \end{pmatrix}$$

Vypočítajte súčiny **AB**, **AC**, **BA** a **CA**.

#### Riešenie:

Prvé dva súčiny sme už vypočítali v príkladoch vyššie. Vieme, že:

$$\mathbf{AB} = \begin{pmatrix} 7 & 10 \\ 3 & 4 \end{pmatrix} \quad \mathbf{AC} = \begin{pmatrix} 7 & 10 & 17 \\ 3 & 4 & 6 \end{pmatrix}$$

Skúsme preto vypočítať posledné dva prípady:

$$\mathbf{BA} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 \cdot 1 + 2 \cdot 0 & 1 \cdot 2 + 2 \cdot 1 \\ 3 \cdot 1 + 4 \cdot 0 & 3 \cdot 2 + 4 \cdot 1 \end{pmatrix} = \begin{pmatrix} 1 & 4 \\ 3 & 10 \end{pmatrix}$$

Je zrejmé, že  $\mathbf{AB} \neq \mathbf{BA}$ , keďže:

$$\mathbf{AB} = \begin{pmatrix} 7 & 10 \\ 3 & 4 \end{pmatrix} \neq \begin{pmatrix} 1 & 4 \\ 3 & 10 \end{pmatrix} = \mathbf{BA}$$

Podobne pri súčine  $\mathbf{CA}$  nastávajú problémy. Keďže rozmery matice  $\mathbf{C}$  sú  $2 \times 3$  a matice  $\mathbf{A}$  sú  $2 \times 2$ , tieto dve matice sa násobiť vôbec nedajú. Vzniká tým situácia, keď súčin  $\mathbf{AC}$  môžeme bez problémov vypočítať, ale súčin  $\mathbf{CA}$  nie je definovaný.

O operácii by sme hovorili, že je komutatívna, len ak by to platilo pre všetky matice, o čom vieme, že to nie je pravda. Príklad na dve matice, pri ktorých zhodou okolností dostávame rovnaký výsledok pri oboch súčinoch si ukážeme v R:

```
> A <- matrix(c(2, 3, 0, 5), nrow = 2, byrow = TRUE)
> B <- matrix(c(4, 2, 0, 6), nrow = 2, byrow = TRUE)
> A
  [,1] [,2]
[1,]  2   3
[2,]  0   5
> B
  [,1] [,2]
[1,]  4   2
[2,]  0   6
> A %*% B
  [,1] [,2]
[1,]  8  22
[2,]  0  30
> B %*% A
  [,1] [,2]
[1,]  8  22
[2,]  0  30
```

Z praktického hľadiska má chýbajúca komutatívnosť následky v dvoch prípadoch. V prvom rade nemôžeme ľubovoľne meniť poradie, v akom matice násobíme. Na druhej strane si musíme uvedomiť, že ak by sme dostali za úlohu vynásobiť maticu  $\mathbf{A}$  maticou  $\mathbf{B}$ , úloha by nebola jednoznačná. Z tohto dôvodu zavádzame pojmy *násobenie zľava* (napríklad

v súčine **AB** násobíme maticu **B** maticou **A** zľava) a násobenie sprava (napríklad v súčine **AB** násobíme maticu **A** maticou **B** sprava).

### 1.3.3 Vlastnosti násobenia matíc

Napriek tomu, že násobenie matíc nie je komutatívne, má aj niekoľko vlastností, ktoré poznáme z násobenia reálnych čísel. Jednou z nich je **asociatívnosť** násobenia – znamená, že aj keď nemôžeme medzi sebou vymieňať matice pri ich vzájomnom súčine, nemusíme pri ňom u troch a viac matíc písať zátvorky – výsledok bude vždy rovnaký. Korektnejšie môžeme asociatívnosť popísať nasledovným vzťahom:

$$(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC}) \quad (1.16)$$

Je teda jedno, či najprv vynásobíme matice **A** a **B**, a výsledok vynásobíme sprava **C**, alebo najprv vynásobíme **B** a **C** a výsledok zľava **A**. V dôsledku asociatívnosti budeme podobné súčiny písať jednoducho ako **ABC**.

Podobne ako aj v reálnych číslach, násobenie matíc je distributívne vzhľadom na sčítanie, a to tak sprava, ako aj zľava:

$$(\mathbf{A}+\mathbf{B})\mathbf{C} = \mathbf{AC}+\mathbf{BC} \quad (1.17)$$

$$\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC} \quad (1.18)$$

Distributívnosť zjednodušené znamená, že pri maticiach môžeme „roznásobiť zátvorky“. Vzhľadom na predchádzajúce poznámky o komutatívnosti súčiny je však veľmi dôležité dať si pozor na to, z ktorej strany roznásobujeme. V prípade vzťahu (1.17) násobíme súčet maticou **C** sprava – aj po roznásobení musí byť **C** v súčinoch vždy na pravej strane. Podobne vo vzťahu (1.18) násobíme maticou **A** zľava, musí preto zostať vľavo aj po roznásobení.

Distributívnosť umožňuje aj opačnú operáciu, t. j. „vybrať výraz pred zátvorku“. Ak by sme mali napríklad výraz **XA + XB + XC**, mohli by sme namiesto neho napísať **X(A + B + C)**. Ak by sme však mali výraz **AX + BX + CX**, v takomto prípade by sme museli „vybrať výraz za zátvorku“, a to takto **(A + B + C)X**.

V predchádzajúcej podkapitole sme si zaviedli pojem transponovanej matice. Je užitočné skúmať aj spôsob, ako by bolo možné **transponovať súčin matíc**. Platí

$$(\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T \quad (1.19)$$

a teda transponovaný súčin matíc je rovný súčinu transponovaných matíc, ale v opačnom poradí. Na prvý pohľad nemusí byť celkom zrejmé, prečo je táto vlastnosť

pravdivá. Aby sme sa o nej presvedčili, museli by sme vykonať matematický dôkaz. Uvedme myšlienku tohto dôkazu v grafickej podobe (Obrázok 3).

$$\left( \begin{pmatrix} \text{---} \end{pmatrix} \begin{pmatrix} \text{ |} \end{pmatrix} \right)^T = \begin{pmatrix} \text{---} \text{+} \end{pmatrix}^T = \begin{pmatrix} \text{ |} \text{---} \end{pmatrix}$$

$$\begin{pmatrix} \text{ |} \end{pmatrix}^T \begin{pmatrix} \text{---} \end{pmatrix}^T = \begin{pmatrix} \text{---} \end{pmatrix} \begin{pmatrix} \text{ |} \end{pmatrix} = \begin{pmatrix} \text{ |} \text{---} \end{pmatrix}$$

Obrázok 3: Transponovanie súčinu matíc

Zdroj: vlastné spracovanie

V obrázku je matica  $\mathbf{A}$  označená modrou a matica  $\mathbf{B}$  zelenou farbou. V prvom riadku je ľavá strana vzťahu (1.19) –  $(\mathbf{AB})^T$ , v druhom riadku pravá –  $\mathbf{B}^T\mathbf{A}^T$ .

V prvom riadku postupujeme tak, že najprv podľa definície vynásobíme obidve matice, a výsledok nakoniec transponujeme, čím sa vymenia riadky a stĺpce matice.

V druhom riadku si všimneme poradie matíc, keďže zelená matica je vyobrazená ako prvá. Obe matice transponujeme (vymeníme v nich riadky a stĺpce) a až potom ich vynásobíme. Vidíme však, že v oboch prípadoch nakoniec prichádzame k tomu istému výsledku, a teda rovnosť vo vzťahu (1.19) naozaj platí.

### Príklad 1.12

Nech  $\mathbf{X}$  je ľubovoľná matica rádu  $m \times n$ , kde  $m, n \in \mathbb{N}$ . Vypočítajte transponovanú maticu súčinu  $\mathbf{X}^T\mathbf{X}$ .

**Riešenie:** Skôr, než sa pustíme do výpočtu samotnej transponovanej matice overme, či vôbec úloha dáva zmysel. Vystáva otázka, či je súčin  $\mathbf{X}^T\mathbf{X}$  definovaný pre každú maticu rádu  $m \times n$ . Keďže  $\mathbf{X}^T$  je k  $\mathbf{X}$  transponovaná matica, bude mať rád  $n \times m$ . V súčine  $\mathbf{X}^T\mathbf{X}$  násobíme matice  $n \times m$  a  $m \times n$ . Je zřejmé, že matice sú vhodných rozmerov, násobenie je možné a výsledná matica bude mať rozmer  $n \times n$ . Uvedený súčin je naozaj definovaný pre ľubovoľnú maticu  $\mathbf{X}$ .

Transponujme teda súčin  $\mathbf{X}^T\mathbf{X}$ . Podľa (1.19) a (1.14) dostávame (vymeníme poradie činiteľov a jednotlivé matice transponujeme):

$$(\mathbf{X}^T\mathbf{X})^T = \mathbf{X}^T(\mathbf{X}^T)^T = \mathbf{X}^T\mathbf{X}$$

Ukázali sme teda nie len to, že je súčin  $X^T X$  definovaný pre všetky matice  $X$ , ale že je súčin  $X^T X$  vždy symetrickou maticou.

Môžeme ukázať ešte tri často využívané operácie.

Prvou bude tzv. **vnútorný súčin**, ktorý sa v našom prípade môže nazývať aj **skalárny súčin** (keďže vzhľadom na charakter vektorov s ktorými pracujeme obidva názvy popisujú rovnakú operáciu, budeme ich voľne zamieňať). Ak máme dva vektory  $x$  a  $y$  rádu  $n$ , potom ich skalárnym súčinom, označovaným  $x \cdot y$  nazývame maticový súčin:

$$x \cdot y = x^T y = y^T x \quad (1.20)$$

Skalárny súčin sa dá vyjadriť ekvivalentne aj nasledovne:

$$x \cdot y = \sum_{i=1}^n x_i y_i \quad (1.21)$$

Význam skalárneho súčinu, jeho vlastnosti a využitie si priblížime v samostatnej kapitole, na tomto mieste nám postačí jeho definícia vo vzťahu k operácii násobenia matíc.

V programe R je možné skalárny súčin vypočítať rôznym spôsobom. Definujme si preto ľubovoľné dva vektory a vyskúšajme najprv operátor súčinu

```
> x <- c(1, 2, 3, 4, 5, 6)
> x
[1] 1 2 3 4 5 6
> y <- c(1, 0, 2, 0, 3, 0)
> y
[1] 1 0 2 0 3 0
> x*y
[1] 1 0 6 0 15 0
```

R zjavne pod  $x*y$  príkazom rozumie vynásobenie prislúchajúcich si zložiek vektorov. Pre výpočet skalárneho súčinu nám už zostáva len spočítať zložky výsledného vektora

```
> sum(x*y)
[1] 22
```

Alternatívou je využitie príkazu `crossprod`

```
> crossprod(x, y)
[,1]
[1,] 22
```

Ak by sme uvažovali o tom, čo by vzniklo násobením, pri ktorom by sme vymenili poradie, v ktorom vektory transponujeme, dostali by sme druhú operáciu – takzvaný **vonkajší súčin vektorov**,  $xy^T$  a  $yx^T$ . Kým skalárny súčin je reálne číslo (pri násobení vektorov  $1 \times n$

a  $n \times 1$  dostávame výsledok rozmeru  $1 \times 1$ ), výsledkom vonkajšieho súčinu je matica: maticu o rozmeroch  $n \times 1$  násobíme ďalšou typu  $1 \times n$  a dostávame maticu  $n \times n$ .

V programe R môžeme na výpočet vonkajšieho súčinu použiť funkciu `outer`.

```
> x <- c(1, 2, 3, 4, 5, 6)
> y <- c(1, 0, 2, 0, 3, 0)
> outer(x, y)
      [,1] [,2] [,3] [,4] [,5] [,6]
[1,]    1    0    2    0    3    0
[2,]    2    0    4    0    6    0
[3,]    3    0    6    0    9    0
[4,]    4    0    8    0   12    0
[5,]    5    0   10    0   15    0
[6,]    6    0   12    0   18    0
> x %o% y
      [,1] [,2] [,3] [,4] [,5] [,6]
[1,]    1    0    2    0    3    0
[2,]    2    0    4    0    6    0
[3,]    3    0    6    0    9    0
[4,]    4    0    8    0   12    0
[5,]    5    0   10    0   15    0
[6,]    6    0   12    0   18    0
```

Identický výsledok dostaneme aj pri použití operátora `%o%`. Všimnime si, že pre maticové násobenie sme využívali operátor `%*%`, keďže ide o príbuzné operácie, R analógiu zachováva.

Zjavne nemusí vždy platiť, že  $\mathbf{xy}^T = \mathbf{yx}^T$ , ale určite platí vzťah:

$$\mathbf{xy}^T = (\mathbf{yx}^T)^T \tag{1.22}$$

```
> x = c(1, 2, 3, 4, 5, 6)
> y = c(1, 0, 2, 0, 3, 0)
> outer(x, y)
      [,1] [,2] [,3] [,4] [,5] [,6]
[1,]    1    0    2    0    3    0
[2,]    2    0    4    0    6    0
[3,]    3    0    6    0    9    0
[4,]    4    0    8    0   12    0
[5,]    5    0   10    0   15    0
[6,]    6    0   12    0   18    0
> outer(y, x)
      [,1] [,2] [,3] [,4] [,5] [,6]
[1,]    1    2    3    4    5    6
[2,]    0    0    0    0    0    0
[3,]    2    4    6    8   10   12
[4,]    0    0    0    0    0    0
[5,]    3    6    9   12   15   18
[6,]    0    0    0    0    0    0
```

Nakoniec ešte zavedieme pojem tzv. **vektora jednotiek**, označovaného gréckym písmenom jota ( $\mathbf{1}$ ). Je to vektor rádu  $n$ , ktorého všetky prvky sú rovné jednej. Pri jeho pomenovaní sme opatrný, pretože pojem jednotkový vektor má iný význam – uprednostníme preto slovné spojenie „vektor jednotiek“. V texte by sme mohli rozlišovať označenie vektora jota podľa jeho dĺžky, napr.  $\mathbf{1}_2$  alebo  $\mathbf{1}_4$ . Vo väčšine prípadov bude jeho rozmer zrejmý z kontextu a preto z hľadiska preferencie jednoduchšieho zápisu ho budeme uvádzať len gréckym písmenom. Vektor  $\mathbf{1}$  je zaujímavý preto, že nám umožňuje sčítať zložky iného vektora jednoducho tým, že spomínané dva vektory medzi sebou maticovo vynásobíme.

### Príklad 1.13

Vynásobte vektor  $\mathbf{x} = (1,3,6,2)^T$  vektorom  $\mathbf{1} = (1,1,1,1)^T$ .

#### Riešenie:

Keďže vektory považujeme za matice, ide o maticové násobenie. Obidva vektory sú rádu 4, priamo by sa ale násobiť nedali. Vektor  $\mathbf{x}$  má maticovo rozmer  $4 \times 1$ , rovnako aj  $\mathbf{1}$ .

Prichádzajú teda do úvahy dve možnosti: buď môžeme násobiť  $\mathbf{x}^T \mathbf{1}$ , alebo naopak  $\mathbf{1}^T \mathbf{x}$ .

V prvom prípade dostávame:

$$\mathbf{x}^T \mathbf{1} = (1,3,6,2) \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = 1 \cdot 1 + 3 \cdot 1 + 6 \cdot 1 + 2 \cdot 1 = 12$$

a v druhom:

$$\mathbf{1}^T \mathbf{x} = (1,1,1,1) \begin{pmatrix} 1 \\ 3 \\ 6 \\ 2 \end{pmatrix} = 1 \cdot 1 + 1 \cdot 3 + 1 \cdot 6 + 1 \cdot 2 = 12$$

### 1.3.4 Inverzná matica a jej vlastnosti

V predchádzajúcich častiach sme ukázali, že chýbajúca komutatívnosť ovplyvňuje prácu s maticami dosť zásadným spôsobom. Veľmi častou využívanou maticou, pri ktorej nás komutatívnosť nemusí veľmi trápiť, je jednotková matica, ktorú sme definovali vyššie. Pripomeňme, že ide o štvorcovú maticu, ktorá má na hlavnej diagonále jednotky a mimo hlavnej diagonály nuly. Výnimočnosť tejto matice je v tom, že predstavuje neutrálny prvok pri operácii násobenia, podobne ako číslo jedna pri násobení reálnych čísel.

Pre ľubovoľnú štvorcovú maticu  $\mathbf{A}$  rádu  $n$  platí:

$$\mathbf{A}\mathbf{I}_n = \mathbf{I}_n\mathbf{A} = \mathbf{A} \quad (1.23)$$

a teda jej súčin so zodpovedajúcou jednotkovou maticou (v akomkoľvek poradí) nám dáva tú istú maticu  $\mathbf{A}$ .

Analógia s reálnymi číslami so sebou prináša otázku, či aj v prípade matíc existuje koncept inverzného prvku k matici, podobne ako je tomu pri recipročných číslach v  $\mathbb{R}$ . Pripomeňme, že ku každému nenulovému číslu  $a \in \mathbb{R}$  existuje číslo  $1/a$  (označuje sa aj  $a^{-1}$ , pre ktoré platí:

$$a \cdot a^{-1} = 1 \quad (1.24)$$

a teda ich súčin je rovný jednej – neutrálnemu prvku pre operáciu násobenia v  $\mathbb{R}$ . Jediným prvkom, ktorý v  $\mathbb{R}$  nemá recipročné číslo, je nula – neutrálny prvok pre operáciu sčítania. Podobné vlastnosti majú aj niektoré matice. Neutrálnym prvkom pre operáciu násobenia je jednotková matica  $\mathbf{I}_n$ . Je tomu však tak len pre štvorcové matice – to je prvé obmedzenie.

Nech  $\mathbf{A}$  je štvorcová matica rádu  $n \in \mathbb{N}$ . Ak k matici  $\mathbf{A}$  existuje matica  $\mathbf{B}$  taká, že platí:

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I}_n \quad (1.25)$$

potom maticu  $\mathbf{B}$  nazývame **inverznou** k matici  $\mathbf{A}$  a označujeme ju  $\mathbf{A}^{-1}$ .

Podobne ako v prípade reálnych čísel platí, že k nulovej matici neexistuje inverzná matica. Okrem neutrálneho prvku pre operáciu sčítania (nulovej matici) však existuje ešte veľké množstvo ďalších štvorcových matíc, ku ktorým neexistuje inverzná matica. Matice, ku ktorým existuje inverzná matica, nazývame **regulárne**, matice ku ktorým neexistuje voláme **singulárne**. Inverzná matica je *jedinečná*: ku každej štvorcovej matici buď neexistuje žiadna, alebo existuje jediná inverzná matica. V skutočnosti je pravdou, že matice  $\mathbf{A}$  a  $\mathbf{A}^{-1}$  sú navzájom inverzné, t. j. okrem toho, že  $\mathbf{A}^{-1}$  je inverznou maticou k  $\mathbf{A}$ , je aj  $\mathbf{A}$  inverznou maticou k  $\mathbf{A}^{-1}$ . Táto skutočnosť sa dá zapísať dvojako:

$$(\mathbf{A}^{-1})^{-1} = \mathbf{A} \quad (1.26)$$

$$\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}_n \quad (1.27)$$

Dôsledkom tohto tvrdenia je fakt, že ak je matica  $\mathbf{A}$  regulárna, tak aj  $\mathbf{A}^{-1}$  je regulárna.

#### Príklad 1.14

Majme matice  $\mathbf{A}$  a  $\mathbf{B}$  dané nasledovne:

$$\mathbf{A} = \begin{pmatrix} 2 & 2 & 1 & 2 \\ 1 & 1 & 1 & 2 \\ 2 & 2 & 1 & 1 \\ 1 & 2 & 1 & 2 \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & -1 & 0 & 1 \\ -3 & 2 & 2 & 0 \\ 1 & 0 & -1 & 0 \end{pmatrix}$$

Overme, či  $\mathbf{B} = \mathbf{A}^{-1}$ .

**Riešenie:** Tvrdenie overíme v programe R.

```
> A <- matrix(c(2, 2, 1, 2, 1, 1, 1, 2, 2, 2, 1, 1, 1, 2, 1, 2),
  nrow = 4, byrow = TRUE)
> A
  [,1] [,2] [,3] [,4]
[1,]   2   2   1   2
[2,]   1   1   1   2
[3,]   2   2   1   1
[4,]   1   2   1   2
> B <- matrix(c(1, 0, 0, -1, 0, -1, 0, 1, -3, 2, 2, 0, 1, 0, -1,
  0), nrow = 4, byrow = T)
> B
  [,1] [,2] [,3] [,4]
[1,]   1   0   0  -1
[2,]   0  -1   0   1
[3,]  -3   2   2   0
[4,]   1   0  -1   0
> A %*% B
  [,1] [,2] [,3] [,4]
[1,]   1   0   0   0
[2,]   0   1   0   0
[3,]   0   0   1   0
[4,]   0   0   0   1
```

Vidíme, že platí  $\mathbf{AB} = \mathbf{I}_n$  a teda je pravda, že  $\mathbf{B} = \mathbf{A}^{-1}$ .

V programe R je možné vypočítať inverznú maticu k regulárnej matici pomocou príkazu `solve()`. Algoritmus, ktorým sa výpočet inverznej matice realizuje, ako aj z neho vyplývajúci pôvod názvu funkcie `solve()` prenecháme na neskôr, teraz sa skôr sústredíme na vlastnosti inverznej matice.

```
> A <- matrix(c(2, 2, 1, 2, 1, 1, 1, 2, 2, 2, 1, 1, 1, 2, 1, 2),
  nrow = 4, byrow = TRUE)
> A
  [,1] [,2] [,3] [,4]
[1,]   2   2   1   2
[2,]   1   1   1   2
[3,]   2   2   1   1
[4,]   1   2   1   2
```

```

> solve(A)
      [,1] [,2] [,3] [,4]
[1,]    1    0    0   -1
[2,]    0   -1    0    1
[3,]   -3    2    2    0
[4,]    1    0   -1    0

```

Nech  $\mathbf{A}$  a  $\mathbf{B}$  sú regulárne štvorcové matice rádu  $n$  a  $\mathbf{A}^{-1}$ ,  $\mathbf{B}^{-1}$  sú k nim inverzné matice. Potom platí:

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1} \quad (1.28)$$

$$(\mathbf{A}^{-1})^T = (\mathbf{A}^T)^{-1} \quad (1.29)$$

Prvý vzťah hovorí, že operácia výpočtu inverznej matice súčinu sa trochu podobá na operáciu transponovania – v oboch prípadoch je možné získať výsledok aplikáciou zvolenej operácie (transponovania, inverzie) na jednotlivé matice v súčine, ale v obrátenom poradí (porovnajzte vzťahy (1.19) a (1.28)).

Zo vzťahu (1.29) vyplývajú hneď niekoľko záverov. Po prvé, ak je matica  $\mathbf{A}$  regulárna, a teda k nej existuje inverzná matica, tak aj matica  $\mathbf{A}^T$  je regulárna (má svoju inverznú maticu).

Po druhé, ak má matica  $\mathbf{A}$  inverznú maticu  $\mathbf{A}^{-1}$ , potom transponovaná matica  $\mathbf{A}^T$  má inverznú maticu  $(\mathbf{A}^T)^{-1}$ , ktorú však nemusíme počítať: dostaneme ju z  $\mathbf{A}^{-1}$  obyčajným transponovaním  $(\mathbf{A}^{-1})^T$ .

Po tretie, vzťah (1.29) sa dá interpretovať tak, že poradie operácií transponovania a maticovej inverzie môžeme voľne zamieňať: je jedno, či vypočítame  $(\mathbf{A}^{-1})^T$  alebo  $(\mathbf{A}^T)^{-1}$ , výsledok bude rovnaký.

Z poslednej menovanej vlastnosti vyplýva napríklad aj to, že **inverzná matica k symetrickej matici je symetrická matica**. Formálne to môžeme formulovať nasledovne. Nech  $\mathbf{C}$  je symetrická regulárna matica. Potom platí:

$$(\mathbf{C}^{-1})^T = \mathbf{C}^{-1} \quad (1.30)$$

Symetrickosť inverznej matice vyplýva zo symetrickosti  $\mathbf{C}$  nasledovne. Začnime s výrazom  $(\mathbf{C}^{-1})^T$ . Vďaka (1.29) vieme, že poradie operácií môžeme vymeniť a máme  $(\mathbf{C}^{-1})^T = (\mathbf{C}^T)^{-1}$ . Keďže je však  $\mathbf{C}$  symetrická matica,  $\mathbf{C}^T = \mathbf{C}$ , a teda v konečnom dôsledku  $(\mathbf{C}^{-1})^T = (\mathbf{C}^T)^{-1} = \mathbf{C}^{-1}$ . Z rovnosti prvého a posledného výrazu podľa (1.14) vyplýva, že  $\mathbf{C}^{-1}$  je symetrická.

### Príklad 1.15

V jednej z prechádzajúcich úloh (Príklad 1.12) sme si ukázali, že matica  $\mathbf{X}^T\mathbf{X}$  je symetrická. Uvažujme o matici:

$$\mathbf{X} = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & -1 & 0 & 1 \\ -3 & 2 & 2 & 0 \\ 1 & 0 & -1 & 0 \\ 2 & -2 & 0 & 1 \end{pmatrix}$$

Ukážme na jej príklade, či súčin  $\mathbf{X}^T\mathbf{X}$ , ako aj k nemu inverzná matica  $(\mathbf{X}^T\mathbf{X})^{-1}$  budú symetrické matice.

**Riešenie:** Úlohu vyriešime pomocou programu R.

```
> X <- matrix(c(1, 0, 0, -1, 0, -1, 0, 1, -3, 2, 2, 0, 1, 0, -1,
0, 2, -2, 0, 1), nrow = 5, byrow = T)
> X
      [,1] [,2] [,3] [,4]
[1,]    1    0    0   -1
[2,]    0   -1    0    1
[3,]   -3    2    2    0
[4,]    1    0   -1    0
[5,]    2   -2    0    1
```

Vidíme, že samotná matica  $\mathbf{X}$  rozmeru  $5 \times 4$  nie je symetrická. Skúsme teda súčin  $\mathbf{X}^T\mathbf{X}$

```
> t(X) %*% X
      [,1] [,2] [,3] [,4]
[1,]   15  -10   -7    1
[2,]  -10    9    4   -3
[3,]   -7    4    5    0
[4,]    1   -3    0    3
```

Vidíme, že  $\mathbf{X}^T\mathbf{X}$  je symetrickou maticou. Zostáva nám ešte vypočítať  $(\mathbf{X}^T\mathbf{X})^{-1}$

```
> solve(t(X) %*% X)
      [,1]      [,2]      [,3]      [,4]
[1,] 1.3548387 1.6451613 0.5806452 1.1935484
[2,] 1.6451613 2.3548387 0.4193548 1.8064516
[3,] 0.5806452 0.4193548 0.6774194 0.2258065
[4,] 1.1935484 1.8064516 0.2258065 1.7419355
```

Matica  $(\mathbf{X}^T\mathbf{X})^{-1}$  je tiež symetrická.

### 1.3.5 Stopa matice

V niektorých špeciálnych prípadoch je užitočné využiť jednu z charakteristík štvorcových matíc, konkrétne tzv. **stopu matice** (angl. trace). Stopa štvorcovej matice  $\mathbf{A}$  rádu  $n$  je definovaná ako súčet prvkov na jej hlavnej diagonále, označujeme ju  $\text{tr}(\mathbf{A})$ . Platí pre ňu:

$$\text{tr}(\mathbf{A}) = \sum_{i=1}^n a_{i,i} \quad (1.31)$$

#### Príklad 1.16

Majme maticu  $\mathbf{B}$  danú nasledovne

$$\mathbf{B} = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & -1 & 0 & 1 \\ -3 & 2 & 2 & 0 \\ 1 & 0 & -1 & 0 \end{pmatrix}$$

Vypočítajme stopu  $\text{tr}(\mathbf{B})$ .

**Riešenie:**

$$\text{tr}(\mathbf{B}) = 1 + (-1) + 2 + 0 = 2$$

V programe R nie je stopa matice priamo k dispozícii. Je možné ju počítať pomocou funkcie `tr()` po doinštalovaní knižnice `psych`. Veľmi ľahko ju však môžeme vypočítať aj s malou pomocou funkcie `diag()`, ktorá vráti vektor tvorený prvkami hlavnej diagonály matice. Stopa je potom už len ich jednoduchým súčtom.

```
> B <- matrix(c(1, 0, 0, -1, 0, -1, 0, 1, -3, 2, 2, 0, 1, 0, -1,
0), nrow = 4, byrow = T)
> B
      [,1] [,2] [,3] [,4]
[1,]    1    0    0   -1
[2,]    0   -1    0    1
[3,]   -3    2    2    0
[4,]    1    0   -1    0
> diag(B)
[1] 1 -1 2 0
> sum(diag(B))
[1] 2
```

Stopa matice je zaujímavá preto, že jej vlastnosti môžeme využiť pri niektorých výpočtoch, ktoré by inak boli veľmi zdĺhavé. Prvou, pomerne zrejmovou vlastnosťou stopy matice je jej aditívnosť.

$$\text{tr}(\mathbf{A}) + \text{tr}(\mathbf{B}) = \text{tr}(\mathbf{A} + \mathbf{B}) \quad (1.32)$$

$$\text{tr}(c\mathbf{A}) = c \cdot \text{tr}(\mathbf{A}) \quad (1.33)$$

$$\text{tr}(\mathbf{A}) = \text{tr}(\mathbf{A}^T) \quad (1.34)$$

kde  $\mathbf{A}$ ,  $\mathbf{B}$  sú matice rádu  $n$  a  $c \in \mathbb{R}$ .

Platnosť vzťahu (1.32) je zrejmá, ak si ho rozpíšeme nasledovne:

$$\begin{aligned} & \text{tr} \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix} + \text{tr} \begin{pmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,n} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m,1} & b_{m,2} & \cdots & b_{m,n} \end{pmatrix} \\ &= \text{tr} \begin{pmatrix} a_{1,1} + b_{1,1} & a_{1,2} + b_{1,2} & \cdots & a_{1,n} + b_{1,n} \\ a_{2,1} + b_{2,1} & a_{2,2} + b_{2,2} & \cdots & a_{2,n} + b_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} + b_{m,1} & a_{m,2} + b_{m,2} & \cdots & a_{m,n} + b_{m,n} \end{pmatrix} \end{aligned} \quad (1.35)$$

```
> A <- matrix(c(1, 2, 4, 2, 4, 5, 2, 2, 1, 2, 3, 0, 1, 0, 3, 2),
  nrow = 4, byrow = TRUE)
> A
  [,1] [,2] [,3] [,4]
[1,]  1   2   4   2
[2,]  4   5   2   2
[3,]  1   2   3   0
[4,]  1   0   3   2
> B <- matrix(c(1, 0, 0, -1, 0, -1, 0, 1, -3, 2, 2, 0, 1, 0, -1,
  0), nrow = 4, byrow = T)
> B
  [,1] [,2] [,3] [,4]
[1,]  1   0   0  -1
[2,]  0  -1   0   1
[3,] -3   2   2   0
[4,]  1   0  -1   0
> A+B
  [,1] [,2] [,3] [,4]
[1,]  2   2   4   1
[2,]  4   4   2   3
[3,] -2   4   5   0
[4,]  2   0   2   2
> sum(diag(A))
[1] 11
> sum(diag(B))
[1] 2
> sum(diag(A+B))
[1] 13
```

Zaujímavejšia je však nasledujúca vlastnosť, ktorá už nie je tak ľahko viditeľná:

$$\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA}) \quad (1.36)$$

Aj napriek tomu, že násobenie matíc nie je komutatívne, stopa súčinu dvoch matíc je rovnaká bez ohľadu na poradie, v ktorom ich násobíme. Podotkneme pritom, že súčin  $\mathbf{AB}$  aj  $\mathbf{BA}$  existuje, pretože  $\mathbf{A}$  a  $\mathbf{B}$  sú štvorcové matice (stopu sme definovali na štvorcových maticiach).

```

> A <- matrix(c(1, 2, 4, 2, 4, 5, 2, 2, 1, 2, 3, 0, 1, 0, 3, 2),
  nrow = 4, byrow = TRUE)
> B <- matrix(c(1, 0, 0, -1, 0, -1, 0, 1, -3, 2, 2, 0, 1, 0, -1,
  0), nrow = 4, byrow = T)
> A %*% B
      [,1] [,2] [,3] [,4]
[1,]   -9    6    6    1
[2,]    0   -1    2    1
[3,]   -8    4    6    1
[4,]   -6    6    4   -1
> B %*% A
      [,1] [,2] [,3] [,4]
[1,]    0    2    1    0
[2,]   -3   -5    1    0
[3,]    7    8   -2   -2
[4,]    0    0    1    2
> sum(diag(A %*% B))
[1] -5
> sum(diag(B %*% A))
[1] -5

```

Na tomto mieste sa ponúka zovšeobecnenie na súčin troch a viacerých matíc. Mohlo by sa zdať, že aj pri viacerých maticiach by sme mohli ľubovoľne meniť poradie násobenia a stopa by sa nezmenila. V skutočnosti sa stopa nemení len pri cyklických permutáciách činiteľov, ktoré vznikajú tak, že vždy prvý činiteľ presunieme na koniec súčinu a proces opakujeme:

$$\text{tr}(\mathbf{ABC}) = \text{tr}(\mathbf{BCA}) = \text{tr}(\mathbf{CAB}) \quad (1.37)$$

Z uvedeného vyplýva, že vo všeobecnosti nie je pravda, že  $\text{tr}(\mathbf{ABC}) = \text{tr}(\mathbf{ACB})$ .

### 1.3.6 Maticové derivácie

Podobne ako v prípade funkcie reálnej premennej, môžeme definovať pojem derivácie a následne ho rozšíriť na funkciu viac premenných. Ďalej je možné definovať operáciu derivácie podľa vektora alebo matice. Presnejšie vymedzenie týchto operácií ďaleko presahuje rozsah, ako aj úroveň výkladu tejto publikácie. Z tohto dôvodu uvedieme len niektoré vybrané vlastnosti týchto derivácií, opierajúc sa o intuíciu čitateľa získanú štúdiom derivácií funkcií.

Ak máme funkciu  $n$  premenných  $f: \mathbb{R}^n \rightarrow \mathbb{R}$ , potom **gradient** definujeme ako vektor prvých parciálnych derivácií podľa jednotlivých premenných, teda:

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = \left( \frac{\partial f(\mathbf{x})}{\partial x_1}, \frac{\partial f(\mathbf{x})}{\partial x_2}, \dots, \frac{\partial f(\mathbf{x})}{\partial x_n} \right)^T \quad (1.38)$$

Ak by sme chceli zapísať druhé derivácie, vektor by už nepostačoval: druhé parciálne derivácie tvoria tzv. **Hessovu maticu**:

$$H(f) = \begin{pmatrix} \frac{\partial^2 f(\mathbf{x})}{\partial x_1^2} & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_2} & \dots & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_1} & \frac{\partial^2 f(\mathbf{x})}{\partial x_2^2} & \dots & \frac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_1} & \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_2} & \dots & \frac{\partial^2 f(\mathbf{x})}{\partial x_n^2} \end{pmatrix} \quad (1.39)$$

Nech  $\mathbf{A}$  je štvorcová matica rádu  $n$  a  $\mathbf{x}$  je  $n$ -rozmerný stĺpcový vektor. Pre derivácie lineárnych a kvadratických foriem podľa vektora  $\mathbf{x}$  platia nasledovné vzťahy:

$$\frac{\partial \mathbf{Ax}}{\partial \mathbf{x}} = \mathbf{A}^T \quad (1.40)$$

$$\frac{\partial \mathbf{x}^T \mathbf{A}}{\partial \mathbf{x}} = \mathbf{A} \quad (1.41)$$

$$\frac{\partial \mathbf{x}^T \mathbf{Ax}}{\partial \mathbf{x}} = (\mathbf{A} + \mathbf{A}^T) \mathbf{x} \quad (1.42)$$

Ak by v poslednom vzťahu bola matica  $\mathbf{A}$  symetrická, potom by platilo aj:

$$\frac{\partial \mathbf{x}^T \mathbf{Ax}}{\partial \mathbf{x}} = (\mathbf{A} + \mathbf{A}^T) \mathbf{x} = (\mathbf{A} + \mathbf{A}) \mathbf{x} = 2\mathbf{Ax} \quad (1.43)$$

### 1.3.7 Blokové matice

V niektorých prípadoch, napríklad ak je matica, s ktorou pracujeme veľmi veľká, môže byť výhodné rozdeliť si ju do blokov a pracovať s jej stručnejšou reprezentáciou (je tomu tak napríklad pri odhade metódou SUR, z angl. *Seemingly Unrelated Regression* v ekonometrii).

**Príklad 1.17**

Maticu

$$\mathbf{A} = \begin{pmatrix} 11 & 12 & 21 & 22 \\ 13 & 14 & 23 & 24 \\ 31 & 32 & 41 & 42 \\ 33 & 34 & 43 & 44 \end{pmatrix}$$

môžeme pomocou štyroch blokov:

$$\mathbf{A}_1 = \begin{pmatrix} 11 & 12 \\ 13 & 14 \end{pmatrix}, \quad \mathbf{A}_2 = \begin{pmatrix} 21 & 22 \\ 23 & 24 \end{pmatrix}, \quad \mathbf{A}_3 = \begin{pmatrix} 31 & 32 \\ 33 & 34 \end{pmatrix}, \quad \mathbf{A}_4 = \begin{pmatrix} 41 & 42 \\ 43 & 44 \end{pmatrix}$$

zapísať ako blokovú nasledovne:

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_2 \\ \mathbf{A}_3 & \mathbf{A}_4 \end{pmatrix}$$

**Blokovú maticu**  $\mathbf{A}$  rádu  $m \times n$  pozostávajúcu z  $k$  blokov v riadku a  $l$  blokov v stĺpci môžeme vo všeobecnosti zapísať takto:

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{1,1} & \mathbf{A}_{1,2} & \cdots & \mathbf{A}_{1,l} \\ \mathbf{A}_{2,1} & \mathbf{A}_{2,2} & \cdots & \mathbf{A}_{2,l} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_{k,1} & \mathbf{A}_{k,2} & \cdots & \mathbf{A}_{k,l} \end{pmatrix} \quad (1.44)$$

Veľmi frekventovaným typom matice je takzvaná **blokovo-diagonálna matica**, ktorej diagonálne bloky pozostávajú z matíc, a prvky mimo nich sú nulové:

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{A}_n \end{pmatrix} \quad (1.45)$$

Operácie s diagonálnymi maticami presne zodpovedajú definíciám bežných maticových operácií, avšak všetky čiastkové operácie musia byť realizovateľné, čo je obzvlášť dôležité pri ich násobení. Napríklad sčítanie medzi dvomi blokovými maticami rovnakého typu je definované nasledovne:

$$\begin{pmatrix} \mathbf{A}_{1,1} & \mathbf{A}_{1,2} & \cdots & \mathbf{A}_{1,l} \\ \mathbf{A}_{2,1} & \mathbf{A}_{2,2} & \cdots & \mathbf{A}_{2,l} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_{k,1} & \mathbf{A}_{k,2} & \cdots & \mathbf{A}_{k,l} \end{pmatrix} + \begin{pmatrix} \mathbf{B}_{1,1} & \mathbf{B}_{1,2} & \cdots & \mathbf{B}_{1,l} \\ \mathbf{B}_{2,1} & \mathbf{B}_{2,2} & \cdots & \mathbf{B}_{2,l} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{B}_{k,1} & \mathbf{B}_{k,2} & \cdots & \mathbf{B}_{k,l} \end{pmatrix} = \begin{pmatrix} \mathbf{A}_{1,1} + \mathbf{B}_{1,1} & \mathbf{A}_{1,2} + \mathbf{B}_{1,2} & \cdots & \mathbf{A}_{1,l} + \mathbf{B}_{1,l} \\ \mathbf{A}_{2,1} + \mathbf{B}_{2,1} & \mathbf{A}_{2,2} + \mathbf{B}_{2,2} & \cdots & \mathbf{A}_{2,l} + \mathbf{B}_{2,l} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_{k,1} + \mathbf{B}_{k,1} & \mathbf{A}_{k,2} + \mathbf{B}_{k,2} & \cdots & \mathbf{A}_{k,l} + \mathbf{B}_{k,l} \end{pmatrix} \quad (1.46)$$

Súčin blokových matic je definovaný podobne ako súčin obyčajných matic, len s tým rozdielom, že operácie násobenia realizujeme na blokoch. Ak je  $\mathbf{A}$  matica rádu  $m \times n$  pozostávajúca z  $k$  blokov v riadku a  $l$  blokov v stĺpci a matica  $\mathbf{B}$  je rádu  $n \times o$  pozostávajúca z  $l$  blokov v riadku a  $s$  blokov v stĺpci, potom súčinom  $\mathbf{AB}$  dostaneme maticu  $\mathbf{C}$ , ktorá je rozmeru  $m \times o$  a má  $k$  blokov v riadku a  $s$  blokov v stĺpci daných vzťahom:

$$\mathbf{C}_{i,j} = \sum_{t=1}^l \mathbf{A}_{i,t} \mathbf{B}_{t,j} \quad (1.47)$$

### 1.3.8 Kroneckerov súčin

Kroneckerov súčin predstavuje operáciu, ktorá sa zďaleka nevyužíva tak často, ako súčin matic, ale pri niektorých situáciách poskytuje veľmi úsporný spôsob, ako zapísať a vytvoriť niektoré matice. Využíva sa napríklad v niektorých častiach ekonometrie.

Ak Kroneckerovým súčinom násobíme dve matice  $\mathbf{A}$  a  $\mathbf{B}$  (označujeme  $\mathbf{A} \otimes \mathbf{B}$  o rozmeroch  $m \times n$  a  $p \times q$ , výsledkom je matica  $mn \times pq$ , ktorá je, využívajúc poznatky o blokových maticiach, daná výrazom:

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{1,1} \mathbf{B} & a_{1,2} \mathbf{B} & \cdots & a_{1,n} \mathbf{B} \\ a_{2,1} \mathbf{B} & a_{2,2} \mathbf{B} & \cdots & a_{2,n} \mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} \mathbf{B} & a_{m,2} \mathbf{B} & \cdots & a_{m,n} \mathbf{B} \end{pmatrix} \quad (1.48)$$

Ak by sme túto maticu rozpísali úplne po jednotlivých prvkoch, vyzeralo by to takto,

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{1,1}b_{1,1} & a_{1,1}b_{1,2} & \cdots & a_{1,1}b_{1,q} & \cdots & \cdots & a_{1,n}b_{1,1} & a_{1,n}b_{1,2} & \cdots & a_{1,n}b_{1,q} \\ a_{1,1}b_{2,1} & a_{1,1}b_{2,2} & \cdots & a_{1,1}b_{2,q} & \cdots & \cdots & a_{1,n}b_{2,1} & a_{1,n}b_{2,2} & \cdots & a_{1,n}b_{2,q} \\ \vdots & \vdots & \ddots & \vdots & \cdots & \cdots & \vdots & \vdots & \ddots & \vdots \\ a_{1,1}b_{p,1} & a_{1,1}b_{p,2} & \cdots & a_{1,1}b_{p,q} & \cdots & \cdots & a_{1,n}b_{p,1} & a_{1,n}b_{p,2} & \cdots & a_{1,n}b_{p,q} \\ \vdots & \vdots & \cdots & \vdots & \ddots & \cdots & \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots & \ddots & \cdots & \vdots & \vdots & \cdots & \vdots \\ a_{m,1}b_{1,1} & a_{m,1}b_{1,2} & \cdots & a_{m,1}b_{1,q} & \cdots & \cdots & a_{m,n}b_{1,1} & a_{m,n}b_{1,2} & \cdots & a_{m,n}b_{1,q} \\ a_{m,1}b_{2,1} & a_{m,1}b_{2,2} & \cdots & a_{m,1}b_{2,q} & \cdots & \cdots & a_{m,n}b_{2,1} & a_{m,n}b_{2,2} & \cdots & a_{m,n}b_{2,q} \\ \vdots & \vdots & \ddots & \vdots & \cdots & \cdots & \vdots & \vdots & \ddots & \vdots \\ a_{m,1}b_{p,1} & a_{m,2}b_{p,2} & \cdots & a_{m,1}b_{p,q} & \cdots & \cdots & a_{m,n}b_{p,1} & a_{m,n}b_{p,2} & \cdots & a_{m,n}b_{p,q} \end{pmatrix} \quad (1.49)$$

Z uvedeného zápisu je zrejmé, prečo je niekedy výhodné využívať zápisy pomocou blokových matic – vedie to k stručnejším, čitateľnejším a ekvivalentným výrazom.

### Príklad 1.18

Kroneckerovým súčinom vynásobme matice:

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix}$$

**Riešenie:**

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} 1 \cdot \begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix} & 2 \cdot \begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix} \\ 0 \cdot \begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix} & 1 \cdot \begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 1 & 3 & 2 & 6 \\ 2 & 1 & 4 & 2 \\ \hline 0 & 0 & 1 & 3 \\ 0 & 0 & 2 & 1 \end{pmatrix}$$

Ako sme už spomínali, Kroneckerov súčin sa používa pomerne zriedka. Nebudeme mu preto poskytovať veľa priestoru, uveďme len niektoré jeho významnejšie vlastnosti.

Nech  $\mathbf{A}$ ,  $\mathbf{B}$  a  $\mathbf{C}$  sú matice takých rádo, pri ktorých sú definované nasledovné operácie (ide hlavne o súčty a súčiny matic, keďže Kroneckerov súčin je definovaný pre akékoľvek matice). Platí:

$$\mathbf{A} \otimes (\mathbf{B} + \mathbf{C}) = \mathbf{A} \otimes \mathbf{B} + \mathbf{A} \otimes \mathbf{C} \quad (1.50)$$

$$(\mathbf{A} + \mathbf{B}) \otimes \mathbf{C} = \mathbf{A} \otimes \mathbf{C} + \mathbf{B} \otimes \mathbf{C} \quad (1.51)$$

$$(k\mathbf{A}) \otimes \mathbf{B} = \mathbf{A} \otimes (k\mathbf{B}) = k(\mathbf{A} \otimes \mathbf{B}) \quad (1.52)$$

$$(\mathbf{A} \otimes \mathbf{B}) \otimes \mathbf{C} = \mathbf{A} \otimes (\mathbf{B} \otimes \mathbf{C}) \quad (1.53)$$

$$(\mathbf{A} \otimes \mathbf{B}) \otimes (\mathbf{C} \otimes \mathbf{D}) = \mathbf{A} \otimes \mathbf{B} \otimes \mathbf{C} \otimes \mathbf{D} \quad (1.54)$$

$$(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1} \quad (1.55)$$

$$(\mathbf{A} \otimes \mathbf{B})^T = \mathbf{A}^T \otimes \mathbf{B}^T \quad (1.56)$$

Prvý vzťah hovorí o distributívnosti Kroneckerovho násobenia zľava vzhľadom na operáciu sčítania, druhý o distributívnosti násobenia sprava vzhľadom na operáciu sčítania. Tretí vzťah nám hovorí, že je jedno, ktorú maticu násobíme skalárom, výsledok je rovnaký, ako keby sme skalárom násobili Kroneckerov súčin. Štvrtý vzťah hovorí o asociatívnosti Kroneckerovho súčinu, nemusíme teda pri ňom písať zátvorky a môžeme výrazy písať aj takto  $\mathbf{A} \otimes \mathbf{B} \otimes \mathbf{C}$ . Predposledná vlastnosť platí len ak k  $\mathbf{A}$  a  $\mathbf{B}$  existujú inverzné matice: inverzná matica ku Kroneckerovmu súčinu je Kroneckerov súčin inverzných matíc. Posledná vlastnosť hovorí, že transponovanie je distributívne vzhľadom na Kroneckerov súčin.

## 2 Vybrané poznatky lineárnej algebry

V tejto časti sa budeme zaoberať niektorými výsledkami lineárnej algebry. Algebra je vo všeobecnosti časťou matematiky, ktorá sa dá zjednodušene opísať ako časť skúmajúca rôzne matematické štruktúry. Pre študentov je algebra menej prítlačivá možno preto, že je veľmi abstraktná a nie je pri nej vždy okamžite jasné, ako je nejaký poznatok prakticky využiteľný. Táto abstraktnosť je ale zároveň aj jej najväčšou silou – skúmané štruktúry sú často definované len veľmi malým množstvom obmedzujúcich podmienok, a preto sa hodia na veľmi veľké množstvo objektov (táto vlastnosť platí aj pre algebru všeobecne – lineárna algebra je len jej vybranou časťou). Vďaka tomu existuje aj veľa aplikácií, ktoré v konečnom dôsledku vychádzajú z lineárnej algebry. V rámci našich textov ich budeme využívať v ekonometrii. V kapitole uvedieme len niektoré vybrané výsledky z pomerne bohatej teórie, ktorou je lineárna algebra. Zvedavého čitateľa odkážeme napr. na špecializované publikácie (napr. Zlatoš, 2011).

### 2.1 Riešenie sústav lineárnych rovníc

So sústavami lineárnych rovníc sa zrejme už každý stretol počas štúdia na strednej alebo vysokej škole. Táto kapitola sa snaží niektoré podstatné výsledky zopakovať, ale jej zmysel je hlbší. Popri štúdiu systémov lineárnych rovníc si zopakujeme niektoré hlavné algoritmy, ktorými k ich riešeniu dochádzame. Tieto algoritmy sú extrémne užitočné neskôr, napríklad v teórii lineárneho programovania, pri výpočte inverzných matic, a používajú sa aj pri niektorých problémoch týkajúcich sa vektorových priestorov, ktoré si predstavíme v nasledujúcich častiach.

Pod **lineárnou rovnicou o  $n$  neznámych**  $x_1, x_2, \dots, x_n$  rozumieme rovnicu v tvare:

$$a_1x_1 + a_2x_2 + \dots + a_nx_n = b \quad (2.1)$$

kde  $a_1, a_2, \dots, a_n, b \in \mathbb{R}$ .

**Sústavu (resp. systém) lineárnych rovníc** dostávame vtedy, ak máme takýchto rovníc niekoľko, pričom čísla  $x_1, x_2, \dots, x_n \in \mathbb{R}$  budeme nazývať **riešením systému**  $m$  rovníc vtedy, ak budú vyhovovať každej z nich (Zlatoš, 2011):

$$\begin{aligned} a_{1,1}x_1 + a_{1,2}x_2 + \dots + a_{1,n}x_n &= b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \dots + a_{2,n}x_n &= b_2 \\ &\vdots \end{aligned} \quad (2.2)$$

$$a_{m,1}x_1 + a_{m,2}x_2 + \dots + a_{m,n}x_n = b_m$$

kde  $a_{i,j} \in \mathbb{R}$  pre  $i = 1, 2, \dots, m$  a  $j = 1, 2, \dots, n$  a  $b_1, b_2, \dots, b_m \in \mathbb{R}$ .

Vzhľadom na maticové operácie, ktoré sme definovali v predchádzajúcich častiach môžeme sústavu (2.2) zapísať úspornejšie takto:

$$\mathbf{Ax} = \mathbf{b} \quad (2.3)$$

kde:

$$\mathbf{A} = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \quad (2.4)$$

To, že je maticový zápis (2.3) ekvivalentný (2.2) môžeme ukázať tak, že zrealizujeme maticový súčin (2.3):

$$\mathbf{Ax} = \mathbf{b} \quad (2.5)$$

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \quad (2.6)$$

$$\begin{pmatrix} a_{1,1}x_1 + a_{1,2}x_2 + \dots + a_{1,n}x_n \\ a_{2,1}x_1 + a_{2,2}x_2 + \dots + a_{2,n}x_n \\ \vdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \dots + a_{m,n}x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \quad (2.7)$$

V poslednej rovnici máme na pravej aj na ľavej strane dva vektory. Dva vektory sú si rovné vtedy, ak majú rovnaký rozmer a všetky ich prislúchajúce si prvky sú totožné.

V prvom rade sa vyjadríme k rozmerom daných vektorov. Matica  $\mathbf{A}$  v (2.5) má rozmer  $m \times n$ , vektor  $\mathbf{x}$   $n \times 1$  a vektor  $\mathbf{b}$   $m \times 1$ . Súčin  $\mathbf{Ax}$  je teda definovaný, a jeho výsledkom je stĺpcový vektor  $m \times 1$ . Pravá a ľavá strana rovnice (2.7) má teda naozaj rovnaký rozmer.

Z druhej podmienky rovnosti vektorov vyplýva, že ak sa majú vektory na pravej a ľavej strane rovnice (2.7) rovnať, musia sa rovnať aj ich zložky, čo vedie presne k systému rovníc (2.2). Maticový zápis (2.3) je teda naozaj ekvivalentný sústave (2.2).

Uvedená skutočnosť je jednou z hlavných príčin, prečo majú aj z praktického hľadiska vedomosti o maticových operáciách zmysel – zápis je s ich pomocou oveľa jednoduchší, čitateľnejší a zároveň ekvivalentný pôvodnému problému.

Maticu  $\mathbf{A}$  budeme nazývať maticou sústavy a vektor  $\mathbf{b}$  vektorom pravých strán. V niektorých prípadoch bude pre nás užitočné uvádzať všetky koeficienty za riešenú sústavu lineárnych rovníc v jednej matici. Ide tu o koeficienty obsiahnuté v matici  $\mathbf{A}$  rozmeru  $m \times n$  a vo vektore  $\mathbf{b}$  rozmeru  $m \times 1$ . Vidíme, že matica  $\mathbf{A}$  aj vektor  $\mathbf{b}$  majú rovnaký počet riadkov – maticu všetkých koeficientov systému rovníc preto dostaneme tak, že k matici  $\mathbf{A}$  pridáme vektor  $\mathbf{b}$  ako dodatočný stĺpec. Túto maticu budeme nazývať **rozšírenou maticou sústavy** (angl. *augmented matrix*), ktorá sa bežne označuje  $(\mathbf{A}|\mathbf{b})$ .

### Príklad 2.1

Zapíšte nasledovnú sústavu rovníc v maticovom tvare:

$$2x_1 + x_2 - x_3 + 5x_4 - x_5 = 4$$

$$x_1 + 4x_2 + 2x_3 + x_4 + 2x_5 = 8$$

$$x_2 + 2x_3 - x_4 + x_5 = 2$$

$$2x_1 + x_2 - 2x_3 + 3x_4 - 2x_5 = 8$$

$$x_1 + x_2 + x_3 + x_4 + x_5 = 7$$

### Riešenie:

Sústavu rovníc je možné maticovo zapísať nasledovne:

$$\begin{pmatrix} 2 & 1 & -1 & 5 & -1 \\ 1 & 4 & 2 & 1 & 2 \\ 0 & 1 & 2 & -1 & 1 \\ 2 & 1 & -2 & 3 & -2 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 4 \\ 8 \\ 2 \\ 8 \\ 7 \end{pmatrix}$$

Matica sústavy a vektor pravých strán sú preto dané:

$$\mathbf{A} = \begin{pmatrix} 2 & 1 & -1 & 5 & -1 \\ 1 & 4 & 2 & 1 & 2 \\ 0 & 1 & 2 & -1 & 1 \\ 2 & 1 & -2 & 3 & -2 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 4 \\ 8 \\ 2 \\ 8 \\ 7 \end{pmatrix}$$

Rozšírená matica sústavy má tvar:

$$(\mathbf{A} | \mathbf{b}) = \left( \begin{array}{ccccc|c} 2 & 1 & -1 & 5 & -1 & 4 \\ 1 & 4 & 2 & 1 & 2 & 8 \\ 0 & 1 & 2 & -1 & 1 & 2 \\ 2 & 1 & -2 & 3 & -2 & 8 \\ 1 & 1 & 1 & 1 & 1 & 7 \end{array} \right)$$

Otázkou zostáva, ako je možné sústavu (2.3) vyriešiť. Pozorný čitateľ predchádzajúcej kapitoly by mohol prísť na jednu možnosť, založenú na využívaní inverznej matice:

$$\mathbf{Ax} = \mathbf{b} \quad (2.8)$$

$$\mathbf{A}^{-1}\mathbf{Ax} = \mathbf{A}^{-1}\mathbf{b} \quad (2.9)$$

$$\mathbf{I}_n\mathbf{x} = \mathbf{A}^{-1}\mathbf{b} \quad (2.10)$$

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b} \quad (2.11)$$

Na prvý pohľad jednoduché riešenie má však niekoľko nezodpovedaných otázok. V prvom rade si pripomenieme, že nie ku každej matici  $\mathbf{A}$  existuje inverzná matica  $\mathbf{A}^{-1}$ . V predchádzajúcej kapitole sme definovali inverznú maticu len pre regulárne štvorcové matice. Matica sústavy  $\mathbf{A}$  má rozmer  $m \times n$ , teda počet riadkov je rovnaký ako počet rovníc a počet stĺpcov je rovný počtu premenných (neznámych). Ak matica sústavy nie je štvorcová – ak nie je počet rovníc rovný počtu neznámych, potom týmto spôsobom nemôžeme sústavu riešiť, inverzná matica  $\mathbf{A}^{-1}$  by neexistovala.

Povedzme ale, že zhodou okolností je táto podmienka splnená, počet rovníc sa rovná počtu neznámych a matica  $\mathbf{A}$  je štvorcová. Ani v takomto prípade však ešte nemáme riešenie sústavy rovníc zaručené. Pripomeňme totiž, že nie ku každej štvorcovej matici  $\mathbf{A}$  existuje inverzná matica. Len po splnení aj podmienky regularity matice  $\mathbf{A}$  by bolo možné riešiť sústavu lineárnych rovníc spôsobom naznačeným v (2.11).

Podotknime ešte, že pri vhodnej matici  $\mathbf{A}$  by sme sústavu vedeli vyriešiť v programe R. Spôsob, akým program R počíta (hľadá) inverznú maticu sme si zatiaľ neukázali. V skutočnosti existuje priamy vzťah medzi riešením sústavy  $\mathbf{Ax} = \mathbf{b}$  so štvorcovou maticou sústavy a inverznou maticou. Jej výpočet si preto necháme na neskôr, zatiaľ si len ukážeme navrhovaný postup v programe R.

### Príklad 2.2

Riešme v programe R nasledovnú sústavu lineárnych rovníc:

$$x_1 + x_2 - 2x_3 + x_4 + 3x_5 - x_6 = 4$$

$$2x_1 - x_2 + x_3 + 2x_4 + x_5 - 3x_6 = 20$$

$$x_1 + 3x_2 - 3x_3 - x_4 + 2x_5 + x_6 = -15$$

$$5x_1 + 2x_2 - x_3 - x_4 + 2x_5 + x_6 = -3$$

$$-3x_1 - x_2 + 2x_3 + 3x_4 + x_5 + 3x_6 = 16$$

$$4x_1 + 3x_2 + x_3 - 6x_4 - 3x_5 - 2x_6 = -27$$

**Riešenie:**

Najprv do R vložíme maticu sústavy a vektor pravých strán:

```
> A <- matrix(c(1, 1, -2, 1, 3, -1, 2, -1, 1, 2, 1, -3, 1, 3, -
  3, -1, 2, 1, 5, 2, -1, -1, 2, 1, -3, -1, 2, 3, 1, 3, 4, 3, 1,
  -6, -3, -2), nrow = 6, byrow = TRUE)
> A
      [,1] [,2] [,3] [,4] [,5] [,6]
[1,]    1    1   -2    1    3   -1
[2,]    2   -1    1    2    1   -3
[3,]    1    3   -3   -1    2    1
[4,]    5    2   -1   -1    2    1
[5,]   -3   -1    2    3    1    3
[6,]    4    3    1   -6   -3   -2
> b <- c(4, 20, -15, -3, 16, -27)
> b
[1]  4 20 -15 -3 16 -27
```

Podľa (2.11) by sme mali ďalej vypočítať inverznú maticu  $\mathbf{A}^{-1}$ :

```
> solve(A)
      [,1]      [,2]      [,3]      [,4]      [,5]      [,6]
[1,] -0.3028169  0.15492958  0.08450704  0.197183099 -0.105633803 -0.09859155
[2,] -0.5492958  0.54460094  0.81220657 -0.215962441  0.234741784  0.10798122
[3,]  0.3309859 -0.03755869 -0.26291080 -0.002347418  0.328638498  0.25117371
[4,] -0.9154930  0.68544601  0.79812207 -0.082159624  0.002347418 -0.20892019
[5,]  1.0915493 -0.53521127 -0.74647887  0.091549296  0.183098592  0.20422535
[6,] -0.1549296 -0.14553991 -0.01877934  0.178403756  0.023474178 -0.08920188
```

Posledným krokom je prenášobenie vektora pravých strán  $\mathbf{b}$  inverznou maticou k matici sústavy zľava ( $\mathbf{A}^{-1}\mathbf{b}$ ):

```
> solve(A) %*% b
      [,1]
[1,]    1
[2,]   -2
[3,]    3
[4,]    4
[5,]    2
[6,]   -1
```

Dostávame tak stĺpcový vektor riešení sústavy rovníc. V skutočnosti R umožňuje vypočítať sústavu aj priamo, pomocou príkazu `solve`:

```
> solve(A, b)
[1]  1 -2  3  4  2 -1
```

Všimnime si, že R tiež poukazuje úzky súvis medzi inverznou maticou a riešením sústavy lineárnych rovníc – oboje je možné vypočítať príkazom `solve`.

V tejto časti budeme pokračovať tým, že si vysvetlíme, ako je možné jednak vyriešiť sústavu lineárnych rovníc, ale aj zistiť, kedy to nie je možné, resp. koľko riešení rovnica má (žiadne, jedno, alebo nekonečne veľa). Keďže je to efektívnejšie, budeme v ďalšom texte pracovať hlavne v zápise pomocou rozšírenej matice sústavy.

Aby sme pochopili základnú myšlienku riešenia sústavy rovníc, je potrebné si všimnúť dve vlastnosti.

V prvom rade existujú sústavy, pri ktorých je veľmi ľahké určiť ich riešenie. Takouto rozšírenou maticou sústavy je napríklad:

$$\left( \begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 8 \\ 0 & 0 & 1 & 11 \end{array} \right) \quad (2.12)$$

Keďže táto rozšírená matica sústavy zodpovedá rovniciam:

$$\begin{aligned} x_1 &= 2 \\ x_2 &= 8 \\ x_3 &= 11 \end{aligned} \quad (2.13)$$

tak je zrejmé, že riešením je vektor (2, 8, 11).

Existujú rozšírené matice sústavy (a teda aj sústavy lineárnych rovníc), ktoré majú rovnaké množiny riešení. Napríklad pri nasledovných dvoch sústavách:

$$\left( \begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 8 \\ 0 & 0 & 1 & 11 \end{array} \right) \quad \left( \begin{array}{ccc|c} 2 & 0 & 0 & 4 \\ 0 & 3 & 0 & 24 \\ 0 & 0 & 4 & 44 \end{array} \right) \quad (2.14)$$

vieme, že pri prvej je riešením vektor (2, 8, 11). Druhej rozšírenej matici sústavy zodpovedá systém rovníc:

$$\begin{aligned} 2x_1 &= 4 \\ 3x_2 &= 24 \\ 4x_3 &= 44 \end{aligned} \quad (2.15)$$

ktorých riešením je tiež vektor (2,8,11).

Je evidentné, že je možné definovať nekonečne veľa rôznych sústav rovníc, ktoré majú tú istú množinu riešení (existujú aj komplikovanejšie alternatívy, ako sme prezentovali v texte). Všetky sústavy lineárnych rovníc, ktoré majú tú istú množinu riešení, nazývame

**ekvivalentnými sústavami.** Dve matice **A** a **B** budeme označovať ekvivalentnými pomocou symbolu „~“ takto:  $A \sim B$ .

Teraz máme všetko potrebné k tomu, aby sme navrhli „stratégiu“ riešenia sústav lineárnych rovníc. Ak dostaneme úlohu riešiť sústavu lineárnych rovníc, tak vieme, že existuje nekonečne veľa sústav rovníc, ktoré sú s ňou ekvivalentné. Keďže existujú aj sústavy, z ktorých dokážeme veľmi jednoducho určiť riešenie, zdá sa byť vhodné nahradzovať pôvodnú sústavu rovníc novou sústavou, ktorá je jednoduchšia, s pôvodnou ekvivalentná a umožňuje ľahšie zistiť jej riešenie a pokračovať v tom dotedy, kým nenájdeme riešenie.

Úpravy, ktorými sa dostávame od jednej sústavy lineárnych rovníc k inej, s ňou ekvivalentnej, nazývame **ekvivalentnými úpravami**. Ekvivalentné úpravy tvoria nasledovné operácie

- výmena dvoch riadkov rozšírenej matice sústavy,
- vynásobenie ktoréhokoľvek riadku rozšírenej matice sústavy nenulovým číslom  $a \in \mathbb{R} \setminus \{0\}$ ,
- pre  $a \in \mathbb{R}$ , pripočítanie  $a$ -násobku jedného riadku matice k inému riadku.

Ukážme si, ako je možné ekvivalentné úpravy realizovať v programe R.

```
> A <- matrix(c(0, 1, 0, 0, 5, 0, 0, 0, -3, 0, 1, 0, 3, 0, -1,
1), nrow = 4, byrow = T)
> A
      [,1] [,2] [,3] [,4]
[1,]    0    1    0    0
[2,]    5    0    0    0
[3,]   -3    0    1    0
[4,]    3    0   -1    1
```

Pripomeňme, že konkrétnu hodnotu z matice **A** môžeme vybrať pomocou indexovania, napríklad prvok  $a_{4,3}$  dostaneme nasledovne:

```
> A[4, 3]
[1] -1
```

Druhý riadok, prípadne tretí stĺpec sa dá vybrať tak, že pri indexovaní zadáme len jeden rozmer:

```
> A[2, ]
[1] 5 0 0 0
> A[, 3]
[1] 0 0 1 -1
```

Podobne sa dá vybrať aj niekoľko riadkov, resp. stĺpcov naraz (napríklad prvé dva):

```

> A[c(1, 2),]
      [,1] [,2] [,3] [,4]
[1,]    0    1    0    0
[2,]    5    0    0    0
> A[,c(1, 2)]
      [,1] [,2]
[1,]    0    1
[2,]    5    0
[3,]   -3    0
[4,]    3    0

```

Zároveň je možné využiť predchádzajúce príkazy aj na modifikáciu prvkov, riadkov, alebo stĺpcov matice jednoduchým priradením:

```

> A
      [,1] [,2] [,3] [,4]
[1,]    0    1    0    0
[2,]    5    0    0    0
[3,]   -3    0    1    0
[4,]    3    0   -1    1
> A[2, 3] <- 80
> A
      [,1] [,2] [,3] [,4]
[1,]    0    1    0    0
[2,]    5    0   80    0
[3,]   -3    0    1    0
[4,]    3    0   -1    1
> A[1,] <- c(11, 12, 13, 14)
> A
      [,1] [,2] [,3] [,4]
[1,]   11   12   13   14
[2,]    5    0   80    0
[3,]   -3    0    1    0
[4,]    3    0   -1    1
> A[,4] <- c(-1, -2, -3, -4)
> A
      [,1] [,2] [,3] [,4]
[1,]   11   12   13   -1
[2,]    5    0   80   -2
[3,]   -3    0    1   -3
[4,]    3    0   -1   -4

```

S pomocou týchto poznatkov si môžeme ukázať, ako je možné realizovať vyššie spomínané ekvivalentné úpravy.

Najprv si ukážme **výmenu dvoch riadkov matice**:

```

> A
      [,1] [,2] [,3] [,4]
[1,]    0    1    0    0
[2,]    5    0    0    0
[3,]   -3    0    1    0

```

```

[4,] 3 0 -1 1
> A[c(1, 4),] <- A[c(4, 1),]
> A
      [,1] [,2] [,3] [,4]
[1,] 3 0 -1 1
[2,] 5 0 0 0
[3,] -3 0 1 0
[4,] 0 1 0 0

```

Kľúčový je výraz  $A[c(1, 4),] \leftarrow A[c(4, 1),]$ . Jeho úvodná časť  $A[c(1, 4),]$  hovorí, že výsledok operácie priradíme do matice **A**, konkrétne do prvého a štvrtého riadku (v tomto poradí). To, čo sa bude do týchto riadkov priraďovať, je dané pravou stranou výrazu:  $A[c(4, 1),]$  čo nám z matice **A** vyberie štvrtý a prvý riadok (v tomto poradí). Výsledkom je, že do prvého riadku priradíme obsah štvrtého, a do štvrtého riadku obsah prvého riadku, čo je vlastne to isté, ako ich výmena.

**Vynásobenie riadku** nenulovým číslom má ešte ľahší zápis, vychádzajúci z rovnakého princípu:

```

> A
      [,1] [,2] [,3] [,4]
[1,] 0 1 0 0
[2,] 5 0 0 0
[3,] -3 0 1 0
[4,] 3 0 -1 1
> A[4,] <- 4*A[4,]
> A
      [,1] [,2] [,3] [,4]
[1,] 0 1 0 0
[2,] 5 0 0 0
[3,] -3 0 1 0
[4,] 12 0 -4 4

```

Do štvrtého riadku priraďujeme štvornásobok bývalého štvrtého riadku matice. **Pripočítanie násobku jedného riadku** matice k inému riadku je už po predchádzajúcich dvoch operáciách veľmi jednoduché:

```

> A
      [,1] [,2] [,3] [,4]
[1,] 0 1 0 0
[2,] 5 0 0 0
[3,] -3 0 1 0
[4,] 3 0 -1 1
> A[2,] <- A[2,] + 2*A[1,]
> A
      [,1] [,2] [,3] [,4]
[1,] 0 1 0 0
[2,] 5 2 0 0

```

[3,]	-3	0	1	0
[4,]	3	0	-1	1

V predchádzajúcom príklade sme k druhému riadku matice pripočítali dvojnásobok prvého. Presnejšie, z matice **A** sme vybrali prvý riadok, vynásobili ho dvomi, pripočítali k druhému riadku matice **A** a výsledok vložili do druhého riadku matice **A**.

Je možné si všimnúť skutočnosť, že až kým R nevloží výsledok do matice, resp. jej časti, je možné pracovať s jej pôvodnými hodnotami. Nehrozí preto problém, že by sa nové a pôvodné hodnoty pomiešali.

### Príklad 2.3

V nasledujúcej matici (**A|b**):

- vymeňme prvý a druhý riadok,
- vynásobme prvý riadok jednou päťtinou,
- pripočítajme tretí riadok k štvrtému,
- pripočítajme trojnásobok prvého riadku k tretiemu.

$$(\mathbf{A} | \mathbf{b}) = \left( \begin{array}{cccc|c} 0 & 1 & 0 & 0 & 2 \\ 5 & 0 & 0 & 0 & 5 \\ -3 & 0 & 1 & 0 & 0 \\ 3 & 0 & -1 & 1 & 4 \end{array} \right)$$

**Riešenie:** Postup opäť vykonáme v programe R.

```
> A <- matrix(c(0, 1, 0, 0, 2, 5, 0, 0, 0, 5, -3, 0, 1, 0, 0, 3,
0, -1, 1, 4), nrow = 4, byrow = TRUE)
> A
      [,1] [,2] [,3] [,4] [,5]
[1,]    0    1    0    0    2
[2,]    5    0    0    0    5
[3,]   -3    0    1    0    0
[4,]    3    0   -1    1    4
```

Teraz vymeňme prvý a druhý riadok:

```
> A[c(2, 1),] <- A[c(1, 2),]
> A
      [,1] [,2] [,3] [,4] [,5]
[1,]    5    0    0    0    5
[2,]    0    1    0    0    2
[3,]   -3    0    1    0    0
[4,]    3    0   -1    1    4
```

Vynásobme prvý riadok jednou pätinou:

```
> A[1,] <- A[1,]/5
> A
      [,1] [,2] [,3] [,4] [,5]
[1,]    1    0    0    0    1
[2,]    0    1    0    0    2
[3,]   -3    0    1    0    0
[4,]    3    0   -1    1    4
```

Pripočítajme tretí riadok k štvrtému:

```
> A[4,] <- A[4,]+A[3,]
> A
      [,1] [,2] [,3] [,4] [,5]
[1,]    1    0    0    0    1
[2,]    0    1    0    0    2
[3,]   -3    0    1    0    0
[4,]    0    0    0    1    4
```

Pripočítajme trojnásobok prvého riadku k tretiemu:

```
> A[3,] <- A[3,] + 3*A[1,]
> A
      [,1] [,2] [,3] [,4] [,5]
[1,]    1    0    0    0    1
[2,]    0    1    0    0    2
[3,]    0    0    1    0    3
[4,]    0    0    0    1    4
```

Po všetkých úpravách sme sa dostali ekvivalentnými úpravami k rozšírenej matici systému

$$\left( \begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & 4 \end{array} \right)$$

Táto rozšírená matica systému je ekvivalentná s  $(\mathbf{A}|\mathbf{b})$ , ich množina riešení je preto rovnaká. Keďže z poslednej matice ľahko vidíme, že  $\mathbf{x} = (1, 2, 3, 4)^T$ , práve sme vyriešili našu prvú sústavu lineárnych rovníc.

V predchádzajúcom príklade sme videli, že ak má sústava lineárnych rovníc riešenie, je možné ho nájsť pomocou riadkových ekvivalentných úprav. Tento postup sa nazýva Gaussova eliminačná metóda. Metóda samotná spočíva v sérii ekvivalentných riadkových

úprav, ktorými maticu prevedieme na tzv. stupňovitý tvar. Pre jeho definíciu je dôležitý pojem vedúci prvok riadku  $i$  matice  $\mathbf{A}$ .

**Vedúcim prvok riadku  $i$  matice  $\mathbf{A}$**  nazývame prvý nenulový prvok v  $i$ -tom riadkovom vektore matice  $\mathbf{A}$ . Ak je riadok  $i$  nulový ( $i$ -tý riadkový vektor je nulovým vektorom), nemá vedúci prvok.

#### Príklad 2.4

Nájdime vedúci prvok prvého a tretieho riadku matice  $\mathbf{A}$ :

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 3 & 4 & 1 \\ 0 & 8 & 1 & 2 & 2 \\ 0 & 0 & 0 & 4 & 3 \\ 0 & 0 & 0 & 3 & 4 \end{pmatrix}$$

**Riešenie:** Prvý riadkový vektor matice  $\mathbf{A}$  je  $(1, 0, 3, 4, 1)$  a jeho prvým nenulovým prvkom je 1. Tretí riadkový vektor matice  $\mathbf{A}$  je  $(0, 0, 0, 4, 3)$  a jeho prvým nenulovým prvkom je 4. Vedúcim prvkom prvého riadku matice  $\mathbf{A}$  je preto prvok  $a_{1,1} = 1$  a vedúcim prvkom tretieho riadku je  $a_{3,4} = 4$ .

O matici  $\mathbf{A}$  hovoríme, že je v **stupňovitom tvare**, ak sa v nej všetky nulové riadky nachádzajú pod nenulovými, a súčasne sa pod každým vedúcim prvkom v stĺpci nachádzajú len nulové prvky.

#### Príklad 2.5

Posúďte, či sú matice  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$  a  $\mathbf{D}$  v stupňovitom tvare:

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 3 & 4 & 1 \\ 0 & 8 & 1 & 2 & 2 \\ 0 & 0 & 0 & 4 & 3 \\ 0 & 0 & 0 & 0 & 4 \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} 1 & 2 & 5 & 4 & 1 \\ 0 & 4 & 1 & 2 & 2 \\ 0 & 0 & 0 & 4 & 3 \\ 0 & 0 & 2 & 3 & 4 \end{pmatrix} \quad \mathbf{C} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 2 & 4 & 1 & 2 & 2 \\ 0 & 1 & 2 & 4 & 3 \\ 0 & 0 & 2 & 3 & 4 \end{pmatrix} \quad \mathbf{D} = \begin{pmatrix} 1 & 1 & 3 \\ 2 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

**Riešenie:** Matica  $\mathbf{A}$  je v stupňovitom tvare, pretože jednak neobsahuje nulové riadky, a pod vedúcimi prvkami sú len nulové prvky. Vedúcimi prvkami riadkov sú  $a_{1,1} = 1$ ,  $a_{2,2} = 8$ ,  $a_{3,4} = 4$ ,  $a_{4,5} = 4$ . V stĺpcoch sa pod každým vedúcim prvkom nachádzajú len nulové prvky.

Matica  $\mathbf{B}$  nie je v stupňovitom tvare, pretože pod vedúcim prvkom tretieho riadku  $a_{3,4} = 4$  sa nachádza nenulový prvok  $a_{4,4} = 3$ .

Matica **C** nie je v stupňovitom tvare, prvý riadok je nulový, a predchádza nasledujúcim nenulovým riadkom.

Matica **D** nie je v stupňovitom tvare, pretože pod vedúcim prvkom prvého riadku  $a_{1,1} = 1$  sa nachádza nenulový prvok  $a_{2,1} = 2$  a  $a_{3,1} = 1$ .

Gaussova eliminačná metóda spočíva v tom, že najprv maticu ekvivalentnými úpravami prevedieme na stupňovitý tvar, a následne spätnou substitúciou vyjadríme riešenia rovnice.

### Príklad 2.6

V programe R ekvivalentnými úpravami preved'ite maticu **A|b** na stupňovitý tvar.

$$\mathbf{A | b} = \left( \begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 0 & 2 \\ -2 & 2 & -2 & 1 & 1 & 13 \\ 3 & -2 & 3 & -2 & -1 & -15 \\ -1 & 1 & -1 & 1 & 0 & 6 \\ -2 & 2 & -2 & 2 & 1 & 14 \end{array} \right)$$

**Riešenie:** Maticu najprv vložíme do R:

```
> A <- matrix(c(1, 0, 0, 0, 0, 2, -2, 2, -2, 1, 1, 13, 3, -2, 3,
-2, -1, -15, -1, 1, -1, 1, 0, 6, -2, 2, -2, 2, 1, 14), nrow =
5, byrow = T)
> A
      [,1] [,2] [,3] [,4] [,5] [,6]
[1,]    1    0    0    0    0    2
[2,]   -2    2   -2    1    1   13
[3,]    3   -2    3   -2   -1  -15
[4,]   -1    1   -1    1    0    6
[5,]   -2    2   -2    2    1   14
```

Pri prevode na stupňovitý tvar budeme využívať definované ekvivalentné úpravy. Vidíme, že v prvom riadku máme vedúci prvok v prvom stĺpci. Pripočítajme preto vždy vhodný násobok prvého riadku k ostatným tak, aby sme pod týmto vedúcim prvkom dostávali stále nulu. K druhému riadku preto pripočítame dvojnásobok prvého, k tretiemu  $-3$  násobok prvého, k štvrtému pripočítame prvý a k piatemu pripočítame dvojnásobok prvého riadku.

```
> A[2,] = A[2,] + 2*A[1,]
> A[3,] = A[3,] - 3*A[1,]
> A[4,] = A[4,] + A[1,]
> A[5,] = A[5,] + 2*A[1,]
> A
```

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]
[1,]	1	0	0	0	0	2
[2,]	0	2	-2	1	1	17
[3,]	0	-2	3	-2	-1	-21
[4,]	0	1	-1	1	0	8
[5,]	0	2	-2	2	1	18

Vidíme, že dostávame novú maticu, ktorá má pod vedúcim prvkom v prvom riadku samé nuly. Keďže sú všetky vykonané úpravy ekvivalentné, je aj výsledná matica ekvivalentná s pôvodnou. Vynásobme druhý riadok číslom  $2^{-1}$ , aby sme dostali vedúci prvok v druhom riadku rovný 1.

```
> A[2,] <- A[2,]/2
> A
      [,1] [,2] [,3] [,4] [,5] [,6]
[1,]    1    0    0 0.0  0.0  2.0
[2,]    0    1   -1 0.5  0.5  8.5
[3,]    0   -2    3 -2.0 -1.0 -21.0
[4,]    0    1   -1  1.0  0.0  8.0
[5,]    0    2   -2  2.0  1.0 18.0
```

Upravme teraz maticu tak, aby mala pod vedúcim prvkom druhého riadku samé nuly. K tretiemu riadku pripočítame dvojnásobok druhého riadku, od štvrtého odpočítame druhý a k piatemu pripočítame mínus dvojnásobok druhého riadku:

```
> A[3,] <- A[3,] + 2*A[2,]
> A[4,] <- A[4,] - A[2,]
> A[5,] <- A[5,] - 2*A[2,]
> A
      [,1] [,2] [,3] [,4] [,5] [,6]
[1,]    1    0    0 0.0  0.0  2.0
[2,]    0    1   -1 0.5  0.5  8.5
[3,]    0    0    1 -1.0  0.0 -4.0
[4,]    0    0    0  0.5 -0.5 -0.5
[5,]    0    0    0  1.0  0.0  1.0
```

Naše úpravy spôsobili, že pod vedúcim prvkom druhého riadku dostávame nuly. Zhodou okolností je táto požiadavka splnená aj pre vedúci prvok v treťom riadku. Zostáva teda ešte postup zopakovať pre vedúci prvok vo štvrtom riadku – najprv ho vynásobíme dvomi, aby sa vedúci prvok rovnal jednej.

```
> A[4,] <- 2*A[4,]
> A
      [,1] [,2] [,3] [,4] [,5] [,6]
```

[1,]	1	0	0	0.0	0.0	2.0
[2,]	0	1	-1	0.5	0.5	8.5
[3,]	0	0	1	-1.0	0.0	-4.0
[4,]	0	0	0	1.0	-1.0	-1.0
[5,]	0	0	0	1.0	0.0	1.0

Vidíme, že na to, aby bola matica v stupňovitom tvare, stačí k piatemu riadku pripočítať (-1) násobok štvrtého.

```
> A[5,] <- A[5,] - A[4,]
> A
      [,1] [,2] [,3] [,4] [,5] [,6]
[1,]    1    0    0 0.0 0.0 2.0
[2,]    0    1   -1 0.5 0.5 8.5
[3,]    0    0    1 -1.0 0.0 -4.0
[4,]    0    0    0 1.0 -1.0 -1.0
[5,]    0    0    0 0.0 1.0 2.0
```

Táto matica je v stupňovitom tvare, a teda:

$$\mathbf{A} = \left( \begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 0 & 2 \\ -2 & 2 & -2 & 1 & 1 & 13 \\ 3 & -2 & 3 & -2 & -1 & -15 \\ -1 & 1 & -1 & 1 & 0 & 6 \\ -2 & 2 & -2 & 2 & 1 & 14 \end{array} \right) \sim \left( \begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 0 & 2 \\ 0 & 1 & -1 & 0.5 & 0.5 & 8.5 \\ 0 & 0 & 1 & -1 & 0 & -4 \\ 0 & 0 & 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{array} \right)$$

Podotknime ešte, že bez ohľadu na počet a jedinečnosť riešenia sústavy lineárnych rovníc, stupňovitý tvar nie je daný jednoznačne. Napríklad je zrejmé, že ak by sme ktorýkoľvek riadok vynásobili nenulovým číslom, nová matica bude ekvivalentná s maticou  $\mathbf{A}$  (ide o ekvivalentnú úpravu), matica zostane v stupňovitom tvare, ale nebude rovná našej výslednej matici. Podobných ekvivalentných úprav by sa samozrejme dalo urobiť oveľa viac.

Spätné dosadzovanie, ako druhú časť Gaussovej eliminačnej metódy realizujeme nasledovne. Vychádzajme z výslednej matice v prechádzajúcom príklade, teda:

$$\left( \begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 0 & 2 \\ 0 & 1 & -1 & 0.5 & 0.5 & 8.5 \\ 0 & 0 & 1 & -1 & 0 & -4 \\ 0 & 0 & 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{array} \right)$$

Táto rozšírená matica sústavy zodpovedá rovniciam:

$$\begin{aligned}
x_1 &= 2 \\
x_2 - x_3 + 0.5 \cdot x_4 + 0.5 \cdot x_5 &= 8.5 \\
x_3 - x_4 &= -4 \\
x_4 - x_5 &= -1 \\
x_5 &= 2
\end{aligned}
\tag{2.16}$$

Okamžite vidíme, že  $x_1 = 2$  a  $x_5 = 2$ . Riešenie pre  $x_4$  dostaneme tak, že sčítame posledné dve rovnice, resp. k štvrtej rovnici pripočítame poslednú piatu:

$$\begin{aligned}
(x_4 - x_5) + x_5 &= -1 + 2 \\
x_4 &= 1
\end{aligned}
\tag{2.17}$$

Podobne môžeme pokračovať aj pre zvyšné premenné. Uvedomme si však, že tieto operácie sú podobné, ako tie, ktorými sme vypočítavali stupňovitý tvar matice, len v obrátenom poradí – začneme poslednou rovnicou a snažíme sa „vynulovať“ prvky nachádzajúce sa v matici nad vedúcimi prvkami riadkov (maticu sa teda snažíme previesť na diagonálnu).

### Príklad 2.7

Preved'me nasledujúcu maticu na diagonálnu spätným dosadzovaním:

$$\left( \begin{array}{ccccc|c}
1 & 0 & 0 & 0 & 0 & 2 \\
0 & 1 & -1 & 0.5 & 0.5 & 8.5 \\
0 & 0 & 1 & -1 & 0 & -4 \\
0 & 0 & 0 & 1 & -1 & -1 \\
0 & 0 & 0 & 0 & 1 & 2
\end{array} \right)$$

**Riešenie:** Najprv pripočítame piaty riadok k štvrtému, a potom  $2^{-1}$ -násobok piateho riadku odpočítame od druhého:

```

> A[4,] <- A[4,]+A[5,]
> A[2,] <- A[2,]-A[5,]/2
> A
      [,1] [,2] [,3] [,4] [,5] [,6]
[1,]    1    0    0 0.0    0  2.0
[2,]    0    1   -1 0.5    0  7.5
[3,]    0    0    1 -1.0    0 -4.0
[4,]    0    0    0  1.0    0  1.0
[5,]    0    0    0  0.0    1  2.0

```

Nakoniec pripočítame štvrtý riadok k tretiemu a  $2^{-1}$ -násobok štvrtého riadku odpočítame od druhého:

```

> A[3,] <- A[3,] + A[4,]
> A[2,] <- A[2,] - A[4,]/2
> A
      [,1] [,2] [,3] [,4] [,5] [,6]
[1,]    1    0    0    0    0    2
[2,]    0    1   -1    0    0    7
[3,]    0    0    1    0    0   -3
[4,]    0    0    0    1    0    1
[5,]    0    0    0    0    1    2

```

Zostáva už len pripočítať tretí riadok k druhému:

```

> A[2,] <- A[2,] + A[3,]
> A
      [,1] [,2] [,3] [,4] [,5] [,6]
[1,]    1    0    0    0    0    2
[2,]    0    1    0    0    0    4
[3,]    0    0    1    0    0   -3
[4,]    0    0    0    1    0    1
[5,]    0    0    0    0    1    2

```

Dostávame maticu v tzv. kanonickom tvare, v ktorom na hlavnej diagonále sú jednotky a mimo hlavnej diagonály sú prvky matice rovné nule. Riešením rovnice je vektor  $\mathbf{x} = (2, 4, -3, 1, 2)$ .

Gaussova eliminačná metóda spočíva v prevode matice ekvivalentnými úpravami na stupňovitý tvar. Takzvaná Gaussova – Jordanova eliminačná metóda sa od Gaussovej líši v tom, že sa priamo vypočítava diagonálna matica a to tak, že pri každom vedúcom prvku „vynulujeme“ tak všetky prvky pod, ako aj nad daným vedúcim.

### Príklad 2.8

Vyriešte sústavu rovníc danú nasledovnou rozšírenou maticou sústavy pomocou Gaussovej – Jordanovej eliminačnej metódy.

$$\left( \begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ -2 & 1 & -2 & 0 & -2 \\ 1 & 0 & 1 & 0 & 2 \\ -1 & 1 & -1 & 1 & 2 \end{array} \right)$$

**Riešenie:** Najprv vložíme do R požadovanú maticu:

```

> A <- matrix(c(1, 0, 0, 0, 1, -2, 1, -2, 0, -2, 1, 0, 1, 0, 2,
-1, 1, -1, 1, 2), nrow = 4, byrow = T)
> A

```

	[,1]	[,2]	[,3]	[,4]	[,5]
[1,]	1	0	0	0	1
[2,]	-2	1	-2	0	-2
[3,]	1	0	1	0	2
[4,]	-1	1	-1	1	2

V prvok kroku „vynulujeme“ prvky pod vedúcim prvkom prvého riadku:

```
> A[2,] <- A[2,] + 2*A[1,]
> A[3,] <- A[3,] - A[1,]
> A[4,] <- A[4,] + A[1,]
> A
```

	[,1]	[,2]	[,3]	[,4]	[,5]
[1,]	1	0	0	0	1
[2,]	0	1	-2	0	0
[3,]	0	0	1	0	1
[4,]	0	1	-1	1	3

V druhom vynulujeme prvky matice pod vedúcim prvkom druhého riadku (nad nemusíme, lebo zhodou okolností je prvok  $a_{1,2} = 0$ ).

```
> A[4,] <- A[4,]-A[2,]
> A
```

	[,1]	[,2]	[,3]	[,4]	[,5]
[1,]	1	0	0	0	1
[2,]	0	1	-2	0	0
[3,]	0	0	1	0	1
[4,]	0	0	1	1	3

Teraz postup zopakujeme aj pre vedúci prvok tretieho riadku. Všimnime si, že „vynulujeme“ aj prvok v druhom riadku – toto je rozdiel oproti obyčajnej Gaussovej eliminačnej metóde.

```
> A[2,] <- A[2,] + 2*A[3,]
> A[4,] <- A[4,] - A[3,]
> A
```

	[,1]	[,2]	[,3]	[,4]	[,5]
[1,]	1	0	0	0	1
[2,]	0	1	0	0	2
[3,]	0	0	1	0	1
[4,]	0	0	0	1	2

Vidíme, že po tejto úprave je rozšírená matica sústavy v kanonickom tvare, a riešením je vektor  $\mathbf{x} = (1, 2, 1, 2)$ .

Gaussova a Gaussova – Jordanova metóda nám poskytujú algoritmy, ako vypočítať riešenie sústavy lineárnych rovníc. Zostáva sa nám ešte vyjadriť k otázke, či všetky sústavy lineárnych rovníc majú riešenie, prípadne kedy tomu tak nie je.

### Príklad 2.9

Vyriešte sústavu rovníc danú nasledovnou rozšírenou maticou sústavy:

$$(\mathbf{A} | \mathbf{b}) = \left( \begin{array}{ccc|c} 1 & 2 & 3 & 8 \\ 2 & 0 & 1 & 4 \\ 5 & 2 & 5 & 32 \end{array} \right)$$

**Riešenie:** Postupujeme Gaussovou – Jordanovou eliminačnou metódou.

```
> A <- matrix(c(1, 2, 3, 8, 2, 0, 1, 4, 5, 2, 5, 32), nrow = 3,
  byrow = T)
> A
      [,1] [,2] [,3] [,4]
[1,]    1    2    3    8
[2,]    2    0    1    4
[3,]    5    2    5   32
```

Začnime vedúcim prvkom prvého riadku:

```
> A[2,] <- A[2,] - 2*A[1,]
> A[3,] <- A[3,] - 5*A[1,]
> A
      [,1] [,2] [,3] [,4]
[1,]    1    2    3    8
[2,]    0   -4   -5  -12
[3,]    0   -8  -10   -8
```

Vynásobme druhý riadok číslom  $(-4)^{-1}$  a postupujme vedúcim prvkom druhého riadku:

```
> A[2,] <- A[2,]/(-4)
> A
      [,1] [,2] [,3] [,4]
[1,]    1    2  3.00    8
[2,]    0    1  1.25    3
[3,]    0   -8 -10.00   -8
> A[3,] <- A[3,] + 8*A[2,]
> A
      [,1] [,2] [,3] [,4]
[1,]    1    2  3.00    8
[2,]    0    1  1.25    3
[3,]    0    0  0.00   16
```

V treťom riadku je vedúcim prvkom prvok v poslednom stĺpci, čo nezodpovedá žiadnej premennej z vektora  $\mathbf{x}$ . Touto metódou nemôžeme pokračovať – zdá sa, že neexistuje spôsob, ako maticu previesť na kanonický tvar.

Ak si všimneme lepšie posledný riadok matice v predchádzajúcom príklade, zistíme, že sa dostávame k sporu:

$$0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 = 16 \quad (2.18)$$

Keďže výraz na ľavej strane je rovný nule, nie je možné, aby posledná rovnosť platila – rovnica nemá riešenie. Skúsme preto, čo by sa stalo, ak by sme neriešili systém v R sami, ale využili vstavanú funkciu `solve()`.

```
> A <- matrix(c(1, 2, 3, 2, 0, 1, 5, 2, 5), nrow = 3, byrow=T)
> b <- c(8, 4, 32)
> A
      [,1] [,2] [,3]
[1,]    1    2    3
[2,]    2    0    1
[3,]    5    2    5
> b
[1]  8  4 32
> solve(A, b)
Error in solve.default(A, b) :
  Lapack routine dgesv: system is exactly singular
```

Pri pokuse o riešenie tejto sústavy rovníc R nedokázal vypočítať riešenie. Informuje nás o tom však pomerne zložitým chybovým hlásením, ktoré nie je úplne zrozumiteľné. Kľúčová je v ňom informácia o singularite systému – to nám môže pripomenúť diskusiu na začiatku tejto podkapitoly, kde sme uvažovali o riešení systému v tvare:

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b} \quad (2.19)$$

Ak si spomenieme, že tento zápis závisel na existencii inverznej matice  $\mathbf{A}^{-1}$  a vieme, že štvorcové matice, ku ktorým neexistuje inverzná matica sa nazývajú singularne, potom nám dané chybové hlásenie vie byť užitočné – hovorí nám presne, kde je chyba. Hovorí to ale spôsobom, ktorý je nečitateľný, ak nemáme potrebné vedomosti. Je preto výhodné poznať algoritmy riešenia sústav lineárnych rovníc, aby bolo možné lepšie pochopiť, čo nám nami používaný štatistický softvér poskytuje a ako k odpovediam prichádza.

Teraz už vieme, že ak riešenie sústavy lineárnych rovníc existuje, môžeme ho vypočítať Gaussovou eliminačnou metódou. Dokážeme aj rozpoznať situáciu, v ktorej systém obsahuje sporné vzťahy, takže nemá riešenie. Existuje ešte tretia možná situácia, ktorú nám naznačí nasledovný príklad.

### Príklad 2.10

Vyriešte sústavu rovníc danú nasledovnou rozšírenou maticou sústavy:

$$(\mathbf{A} | \mathbf{b}) = \left( \begin{array}{cccc|c} 1 & 2 & 3 & 2 & 8 \\ 2 & 0 & 1 & 2 & 4 \\ 5 & 2 & 5 & 6 & 16 \end{array} \right)$$

**Riešenie:** Postupujeme Gaussovou – Jordanovou eliminačnou metódou.

```
> A <- matrix(c(1, 2, 3, 2, 1, 8, 2, 0, 1, 2, 1, 4, 5, 2, 5, 6,
 3, 16), nrow=3, byrow=T)
> A
      [,1] [,2] [,3] [,4] [,5] [,6]
[1,]    1    2    3    2    1    8
[2,]    2    0    1    2    1    4
[3,]    5    2    5    6    3   16
```

Keďže matica je veľmi podobná tej v predchádzajúcom príklade, aj pri riešení postupujeme analogicky:

```
> A <- matrix(c(1, 2, 3, 2, 1, 8, 2, 0, 1, 2, 1, 4, 5, 2, 5, 6,
 3, 16), nrow=3, byrow=T)
> A[2,] <- A[2,]-2*A[1,]
> A[3,] <- A[3,]-5*A[1,]
> A
      [,1] [,2] [,3] [,4] [,5] [,6]
[1,]    1    2    3    2    1    8
[2,]    0   -4   -5   -2   -1  -12
[3,]    0   -8  -10   -4   -2  -24
> A[2,] <- A[2,]/(-4)
> A
      [,1] [,2]  [,3] [,4]  [,5] [,6]
[1,]    1    2   3.00  2.0   1.00   8
[2,]    0    1   1.25  0.5   0.25   3
[3,]    0   -8 -10.00 -4.0  -2.00  -24
> A[3,] <- A[3,]+8*A[2,]
> A[1,] <- A[1,]-2*A[2,]
> A
      [,1] [,2] [,3] [,4] [,5] [,6]
[1,]    1    0  0.50  1.0  0.50   2
[2,]    0    1  1.25  0.5  0.25   3
[3,]    0    0  0.00  0.0  0.00   0
```

Opäť sa dostávame do stavu, keď maticu nie je možné previesť na kanonický tvar, sústavu preto nevieme doteraz prezentovaným spôsobom vyriešiť.

Prepíšme si však rozšírenú maticu sústavy z predchádzajúceho príkladu do rovníc:

$$\begin{aligned}
x_1 + 0.5x_3 + x_4 + 0.5x_5 &= 2 \\
x_2 + 1.25x_3 + 0.5x_4 + 0.25x_5 &= 3 \\
0 &= 0
\end{aligned}
\tag{2.20}$$

Tento krát posledná rovnica nepredstavuje spor, je to výraz, ktorý je vždy pravdivý. Ako taký by sme ho mohli zo sústavy aj vylúčiť – nijakým spôsobom riešenie neobmedzuje. Dostávame tak dve rovnice o piatich neznámych. Vzhľadom na voľnosť, ktorú tým dostávame pri voľbe parametrov je zrejmé, že sa vždy bude dať zvoliť takých päť čísel, aby boli splnené požadované dve rovnice. V skutočnosti ak zvolíme  $x_3, x_4$  a  $x_5 \in \mathbb{R}$  ľubovoľne, riešením pre  $x_1$  a  $x_2$  bude vzhľadom na (2.20) vzťah:

$$\begin{aligned}
x_1 &= 2 - 0.5x_3 - x_4 - 0.5x_5 \\
x_2 &= 3 - 1.25x_3 - 0.5x_4 - 0.25x_5
\end{aligned}
\tag{2.21}$$

Odtiaľ dostávame, že riešením sústavy lineárnych rovníc je akýkoľvek vektor:

$$\mathbf{x} = (2 - 0.5x_3 - x_4 - 0.5x_5, 3 - 1.25x_3 - 0.5x_4 - 0.25x_5, x_3, x_4, x_5)
\tag{2.22}$$

Ak by sme to zhrnuli, sústava lineárnych rovníc môže a nemusí mať riešenie. V prípade, že môžeme rozšírenú maticu systému pomocou Gaussovej alebo Jordanovej eliminačnej metódy previesť na kanonický tvar, sústava má jediné a jednoznačne určené riešenie. Tieto metódy nám to riešenie umožňujú nájsť. V prípade, ak systém rovníc v sebe obsahuje spor, systém nemá riešenie. Ak ale pri riešení zistíme, že systém neobsahuje spor, no po vylúčení nulových riadkov z rozšírenej matice systému zostáva menej rovníc, ako je neznámych, systém bude mať nekonečne veľa riešení. Možnosti sú teda tri: nula, jedno, alebo nekonečne veľa riešení v závislosti od konkrétneho riešeného systému, iná možnosť nemôže nastať.

V súvislosti s riešením sústav lineárnych rovníc nám zostáva ešte vysvetliť poslednú vlastnosť, ktorú sme spomínali už na začiatku kapitoly – súvis s inverznou maticou. Ukázali sme si, že ak je matica  $\mathbf{A}$  regulárna, môžeme sústavu lineárnych rovníc riešiť pomocou inverznej matice  $\mathbf{A}^{-1}$ . V skutočnosti vždy, keď má sústava lineárnych rovníc práve jedno riešenie je matica  $\mathbf{A}$  regulárna. Platí to aj naopak – ak je matica  $\mathbf{A}$  regulárna, sústava má jedinečné riešenie. Tvrdenia, že matica má jedno riešenie, a že je regulárna, sú ekvivalentné. Hlbší súvis tohto tvrdenia nám pomôže vysvetliť nasledovný príklad.

### Príklad 2.11

Vyriešte sústavu lineárnych rovníc dané nasledovnými rozšírenými maticami systému:

$$\left( \begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{array} \right) \quad \left( \begin{array}{cccc|c} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{array} \right) \quad \left( \begin{array}{cccc|c} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{array} \right) \quad \left( \begin{array}{cccc|c} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{array} \right)$$

### Riešenie:

Riešme v programe R prvú sústavu rovníc:

```
A <- matrix(c(1, 0, 0, 0, 1, 1, 1, 0, 0, 0, 1, 1, 1, 0, 0, 0, 0, 0, 0, 1, 0), nrow=4, byrow=T)
> A
      [,1] [,2] [,3] [,4] [,5]
[1,]    1    0    0    0    1
[2,]    1    1    0    0    0
[3,]    1    1    1    0    0
[4,]    0    0    0    1    0
> A[2,] <- A[2,]-A[1,]
> A[3,] <- A[3,]-A[1,]
> A
      [,1] [,2] [,3] [,4] [,5]
[1,]    1    0    0    0    1
[2,]    0    1    0    0   -1
[3,]    0    1    1    0   -1
[4,]    0    0    0    1    0
> A[3,] <- A[3,]-A[2,]
> A
      [,1] [,2] [,3] [,4] [,5]
[1,]    1    0    0    0    1
[2,]    0    1    0    0   -1
[3,]    0    0    1    0    0
[4,]    0    0    0    1    0
```

Výsledná matica je v kanonickom tvare, riešením je vektor  $\mathbf{x}_1 = (1, -1, 0, 0)^T$ . Ak si pozorne všimneme zostávajúce rozšírené matice, vidíme, že v každej z nich je matica sústavy rovnaká, líši sa vždy len pravá strana. Keďže prevod na kanonický tvar sa týka práve matice sústavy, znamená to, že vo všetkých štyroch prípadoch musíme realizovať rovnaké operácie.

```
> A <- matrix(c(1, 0, 0, 0, 0, 1, 1, 0, 0, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 1, 0), nrow=4, byrow=T)
> A[2,] <- A[2,]-A[1,]
> A[3,] <- A[3,]-A[1,]
> A[3,] <- A[3,]-A[2,]
```

```
> A
      [,1] [,2] [,3] [,4] [,5]
[1,]    1    0    0    0    0
[2,]    0    1    0    0    1
[3,]    0    0    1    0   -1
[4,]    0    0    0    1    0
```

V prípade tretej sústavy:

```
> A <- matrix(c(1, 0, 0, 0, 0, 1, 1, 0, 0, 0, 1, 1, 1, 0, 1, 0,
0, 0, 1, 0), nrow=4, byrow=T)
> A[2,] <- A[2,]-A[1,]
> A[3,] <- A[3,]-A[1,]
> A[3,] <- A[3,]-A[2,]
> A
      [,1] [,2] [,3] [,4] [,5]
[1,]    1    0    0    0    0
[2,]    0    1    0    0    0
[3,]    0    0    1    0    1
[4,]    0    0    0    1    0
```

V prípade štvrtej sústavy:

```
> A <- matrix(c(1, 0, 0, 0, 0, 1, 1, 0, 0, 0, 1, 1, 1, 0, 0, 0,
0, 0, 1, 1), nrow=4, byrow=T)
> A[2,] <- A[2,]-A[1,]
> A[3,] <- A[3,]-A[1,]
> A[3,] <- A[3,]-A[2,]
> A
      [,1] [,2] [,3] [,4] [,5]
[1,]    1    0    0    0    0
[2,]    0    1    0    0    0
[3,]    0    0    1    0    0
[4,]    0    0    0    1    1
```

Riešením sústav lineárnych rovníc sú vektory  $\mathbf{x}_1 = (1, -1, 0, 0)^T$ ,  $\mathbf{x}_2 = (0, 1, -1, 0)^T$ ,  $\mathbf{x}_3 = (0, 0, 1, 0)^T$  a  $\mathbf{x}_4 = (0, 0, 0, 1)^T$ .

Z príkladu je vidieť, že v prípade, ak by sme riešili niekoľko sústav lineárnych rovníc, ktoré majú rovnakú maticu sústavy, a líšia sa len pravými stranami, ich riešenie je len opakovaním toho istého postupu – prevodu matice na kanonický tvar. V predchádzajúcom príklade sa vo všetkých štyroch prípadoch opakovali nasledovné kroky: odpočítanie prvého riadku od druhého, odpočítanie prvého riadku od tretieho a odpočítanie druhého riadku od tretieho. Je zjavne neefektívne opakovať ten istý výpočet pre každú pravú stranu samostatne.

V prípade, ak by sme potrebovali riešiť  $s \in \mathbb{N}$  sústav rovníc líšiacich sa len vektorom pravých strán:

$$\mathbf{Ax}_1 = \mathbf{b}_1 \quad \mathbf{Ax}_2 = \mathbf{b}_2 \quad \dots \quad \mathbf{Ax}_s = \mathbf{b}_s \quad (2.23)$$

potom modifikujeme rozšírenú maticu sústavy, a napíšeme ju v tvare:

$$(\mathbf{A} | \mathbf{b}_1 \mathbf{b}_2 \dots \mathbf{b}_s) = \left( \begin{array}{cccc|cccc} a_{1,1} & a_{1,2} & \dots & a_{1,n} & b_{1,1} & b_{1,2} & \dots & b_{1,s} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} & b_{2,1} & b_{2,2} & \dots & b_{2,s} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} & b_{m,1} & b_{m,2} & \dots & b_{m,s} \end{array} \right) \quad (2.24)$$

Namiesto toho, aby sme za maticu sústavy  $\mathbf{A}$  napísali stĺpcový vektor pravej strany  $\mathbf{b}_1$  tam pridáme všetkých  $s$  stĺpcových vektorov pravých strán. Ak túto maticu prevedieme na kanonický tvar, dostaneme na pravej strane v stĺpoch riešenia všetkých rovníc. Je tomu tak preto, lebo úpravy, ktorými maticu prevádzame na kanonický tvar vôbec nezávisia od vektorov  $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_{1s}$ , ale len od matice sústavy  $\mathbf{A}$ .

### Príklad 2.12

Postupom popísaným v predchádzajúcom odseku vyriešme sústavy lineárnych rovníc dané nasledovnými rozšírenými maticami systému:

$$\left( \begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{array} \right) \quad \left( \begin{array}{cccc|c} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{array} \right) \quad \left( \begin{array}{cccc|c} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{array} \right) \quad \left( \begin{array}{cccc|c} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{array} \right)$$

### Riešenie:

Najprv vytvoríme navrhovanú modifikovanú rozšírenú maticu systému:

$$\left( \begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array} \right)$$

Jej riešením v  $\mathbb{R}$  dostávame:

```
>A <- matrix(c(1, 0, 0, 0, 1, 0, 0, 0, 1, 1, 0, 0, 0, 1, 0, 0,
  1, 1, 1, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 1), nrow=4,
  byrow=T)
> A
  [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8]
[1,]  1    0    0    0    1    0    0    0
[2,]  1    1    0    0    0    1    0    0
[3,]  1    1    1    0    0    0    1    0
```

```

[4,] 0 0 0 1 0 0 0 1
> A[2,] <- A[2,] - A[1,]
> A[3,] <- A[3,] - A[1,]
> A
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8]
[1,] 1 0 0 0 1 0 0 0
[2,] 0 1 0 0 -1 1 0 0
[3,] 0 1 1 0 -1 0 1 0
[4,] 0 0 0 1 0 0 0 1
> A[3,] <- A[3,] - A[2,]
> A
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8]
[1,] 1 0 0 0 1 0 0 0
[2,] 0 1 0 0 -1 1 0 0
[3,] 0 0 1 0 0 -1 1 0
[4,] 0 0 0 1 0 0 0 1

```

Riešením sústav lineárnych rovníc sú vektory  $\mathbf{x}_1 = (1, -1, 0, 0)^T$ ,  $\mathbf{x}_2 = (0, 1, -1, 0)^T$ ,  $\mathbf{x}_3 = (0, 0, 1, 0)^T$  a  $\mathbf{x}_4 = (0, 0, 0, 1)^T$ . Dostávame rovnaký výsledok ako v predchádzajúcom príklade, tentoraz sme však výsledky získali pre všetky štyri sústavy naraz.

V príklade sme z pôvodného systému rovníc:

$$\mathbf{Ax}_1 = \mathbf{b}_1 \quad \mathbf{Ax}_2 = \mathbf{b}_2 \quad \dots \quad \mathbf{Ax}_s = \mathbf{b}_s \quad (2.25)$$

modifikovali rozšírenú maticu sústavy a previedli ju na ekvivalentný tvar:

$$(\mathbf{A} | \mathbf{b}_1 \mathbf{b}_2 \dots \mathbf{b}_s) = \left( \begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array} \right) \sim \left( \begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array} \right) \quad (2.26)$$

Všimnime si, že vektory pravých strán boli v pôvodnom zadaní úlohy volené tak, že vytvárali jednotkovú maticu. Rozšírená matica má zasa po vyriešení rovníc jednotkovú maticu na ľavej strane (keďže je v kanonickom tvare) a na pravej strane má riešenia systému.

Ak si spojíme predchádzajúce dva vzťahy, tak vidíme, že pre  $\mathbf{x}_1 = (1, -1, 0, 0)^T$ ,  $\mathbf{x}_2 = (0, 1, -1, 0)^T$ ,  $\mathbf{x}_3 = (0, 0, 1, 0)^T$  a  $\mathbf{x}_4 = (0, 0, 0, 1)^T$  musí platiť:

$$\mathbf{Ax}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \mathbf{Ax}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad \mathbf{Ax}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad \mathbf{Ax}_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \quad (2.27)$$

Ak teda vytvoríme maticu z vektorov pre  $\mathbf{x}_1$ ,  $\mathbf{x}_2$ ,  $\mathbf{x}_3$  a  $\mathbf{x}_4$  tak, že budú tvoriť stĺpce matice  $\mathbf{B}$ :

$$\mathbf{B} = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (2.28)$$

potom bude platiť:

$$\mathbf{AB} = \mathbf{A}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4) = \mathbf{A} \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \mathbf{I}_4 \quad (2.29)$$

Keďže v súčine  $\mathbf{AB}$  dostávame jednotkovú maticu, musí platiť, že:

$$\mathbf{B} = \mathbf{A}^{-1} \quad (2.30)$$

Našli sme teda inverznú maticu k matici  $\mathbf{A}$ . Vo všeobecnosti je tento postup použiteľný na výpočet inverznej matice – k štvorcovej matici  $\mathbf{A}$ , ku ktorej hľadáme inverznú maticu, doplníme maticou pravých strán, ktorá má podobu jednotkovej matice rovnakého rádu, ako má  $\mathbf{A}$ . Ekvivalentnými úpravami prevedieme rozšírenú maticu systému na kanonický tvar, a ak získame jediné riešenie, na pravej strane rozšírenej matice získame maticu, ktorá je inverzná k  $\mathbf{A}$ .

### Príklad 2.13

Nájdime inverznú maticu k matici:

$$\mathbf{A} = \begin{pmatrix} 1 & 3 & 1 \\ 1 & 1 & 2 \\ 2 & 3 & 4 \end{pmatrix}$$

#### Riešenie:

Najprv vytvoríme rozšírenú maticu sústavy. V skutočnosti by sme mali riešiť tri sústavy, kde by pravé strany tvorili jednotlivé stĺpcové vektory jednotkovej matice  $\mathbf{I}_3$ . Urobíme to ale rýchlejšie, ak namiesto vektora pravých strán dosadíme rovno maticu  $\mathbf{I}_3$ .

```
> A <- matrix(c(1, 3, 1, 1, 0, 0, 1, 1, 2, 0, 1, 0, 2, 3, 4, 0,
0, 1), nrow=3, byrow=TRUE)
> A
      [,1] [,2] [,3] [,4] [,5] [,6]
[1,]    1    3    1    1    0    0
[2,]    1    1    2    0    1    0
[3,]    2    3    4    0    0    1
```

Ďalej pripočítame mínus jeden násobok prvého riadku k druhému, ako aj mínus dvojnásobok prvého k tretiemu riadku.

```
> A[2,] <- A[2,]-A[1,]
> A[3,] <- A[3,]-2*A[1,]
> A
      [,1] [,2] [,3] [,4] [,5] [,6]
[1,]    1    3    1    1    0    0
[2,]    0   -2    1   -1    1    0
[3,]    0   -3    2   -2    0    1
```

Druhý riadok vynásobíme číslom  $(-2)^{-1}$ , aby bol vedúci riadok rovný jednej.

```
> A[2,] <- -A[2,]/2
> A[2,] <- -A[2,]/2
> A
      [,1] [,2] [,3] [,4] [,5] [,6]
[1,]    1    3  1.0  1.0  0.0    0
[2,]    0    1 -0.5  0.5 -0.5    0
[3,]    0   -3  2.0 -2.0  0.0    1
```

V ďalšom kroku odpočítame trojnásobok druhého riadku od prvého, a pripočítame trojnásobok druhého riadku k tretiemu (postupujeme Gaussovou – Jordanovou metódou a nulujeme rovno aj riadky nad vedúcim prvkom).

```
> A[1,] <- A[1,]-3*A[2,]
> A[3,] <- A[3,]+3*A[2,]
> A
      [,1] [,2] [,3] [,4] [,5] [,6]
[1,]    1    0  2.5 -0.5  1.5    0
[2,]    0    1 -0.5  0.5 -0.5    0
[3,]    0    0  0.5 -0.5 -1.5    1
```

Tretí riadok vynásobíme dvomi:

```
> A[3,] <- 2*A[3,]
> A
      [,1] [,2] [,3] [,4] [,5] [,6]
[1,]    1    0  2.5 -0.5  1.5    0
[2,]    0    1 -0.5  0.5 -0.5    0
[3,]    0    0  1.0 -1.0 -3.0    2
```

K druhému riadku pripočítame pol násobok tretieho a od prvému riadku odpočítame dva a pol násobok tretieho riadku.

```

> A[2,] <- A[2,]+0.5*A[3,]
> A[1,] <- A[1,]-2.5*A[3,]
> A
      [,1] [,2] [,3] [,4] [,5] [,6]
[1,]    1    0    0    2    9   -5
[2,]    0    1    0    0   -2    1
[3,]    0    0    1   -1   -3    2

```

Vidíme, že rozšírená matica je v kanonickom tvare, inverznou maticou k matici **A** je matica:

$$\begin{pmatrix} 2 & 9 & -5 \\ 0 & -2 & 1 \\ -1 & -3 & 2 \end{pmatrix}$$

Uvedenú skutočnosť môžeme aj overiť:

```

> A <- matrix(c(1, 3, 1, 1, 1, 2, 2, 3, 4), nrow=3, byrow=T)
> A
      [,1] [,2] [,3]
[1,]    1    3    1
[2,]    1    1    2
[3,]    2    3    4
> D <- matrix(c(2, 9, -5, 0, -2, 1, -1, -3, 2), nrow=3, byrow=T)
> D
      [,1] [,2] [,3]
[1,]    2    9   -5
[2,]    0   -2    1
[3,]   -1   -3    2
> A %*% D
      [,1] [,2] [,3]
[1,]    1    0    0
[2,]    0    1    0
[3,]    0    0    1

```

Keďže súčinom daných matíc dostávame jednotkovú maticu  $\mathbf{I}_3$ , skutočne ide o navzájom inverzné matice.

Na základe týchto poznatkov je možné definitívne pochopiť súvislosť medzi inverznou maticou a riešením sústavy lineárnych rovníc. Na začiatku podkapitoly sme si ukázali, že prirodzeným spôsobom, ako by bolo možné uvažovať o riešení systému lineárnych rovníc je výraz  $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ . Problémom bolo jednak to, že nie vždy existuje inverzná matica  $\mathbf{A}^{-1}$ , no súčasne sme doposiaľ nepoznali spôsob, ako inverznú maticu vypočítať (poznali sme príkaz v R, ale nepovedali sme si ako funguje). Na základe poznatkov z tejto kapitoly vieme

- 1) kedy existuje riešenie sústavy lineárnych rovníc
- 2) koľko riešení sústava môže mať
- 3) ako

ju vyriešiť, resp. ukázať, že je neriešiteľná. Pokiaľ by existovalo jediné riešenie sústavy, potom je možné využiť vzťah  $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ , no na výpočet inverznej matice  $\mathbf{A}^{-1}$  musíme vykonať tie isté operácie, ako na riešenie sústavy  $\mathbf{Ax} = \mathbf{b}$  Gaussovou, resp. Gaussovou – Jordanovou metódou. Aj z tohto dôvodu sa v  $\mathbb{R}$  výpočet sústavy lineárnych rovníc a výpočet inverznej matice realizuje tým istým príkazom `solve()` – výpočet je v oboch prípadoch vlastne rovnaký.

Z tohto pohľadu neprináša výpočet inverznej matice pri výpočte riešenia sústavy lineárnych rovníc veľkú výhodu – skôr naopak, ak sa zahrnú do úvahy aj určité charakteristiky numerickej stability výsledkov pri počítačovom spracovaní. Väčšina programov preto rieši sústavy rovníc inak ako cez inverznú maticu.

Na záver tejto podkapitoly si ukážeme ešte jednu možnosť, ako nazerať na inverznú maticu. Tento pohľad sa hodí v ďalšom výklade, ak by sme riešili určité lineárne transformácie.

V predchádzajúcej časti sme si definovali elementárne (riadkovo) ekvivalentné úpravy, ktorými boli:

- výmena dvoch riadkov rozšírenej matice sústavy,
- vynásobenie ktoréhokoľvek riadku rozšírenej matice sústavy nenulovým číslom  $a \in \mathbb{R} \setminus \{0\}$ ,
- pre  $a \in \mathbb{R}$ , pripočítanie  $a$ -násobku jedného riadku matice k inému riadku.

Pre každú z týchto ekvivalentných úprav môžeme definovať tzv. transformačnú maticu. Namiesto toho, aby sme realizovali ekvivalentnú úpravu, môžeme vhodným spôsobom prenásobiť skúmanú maticu transformačnou, a výsledok bude rovnaký.

Transformačnú maticu  $\mathbf{T}_{i,j}$ , pomocou ktorej dosiahneme výmenu riadkov  $i$  a  $j$  ( $i, j \in \mathbb{N}$ ), dostaneme z jednotkovej matice  $\mathbf{I}_n$  výmenou jej  $i$ -tého a  $j$ -tého riadku.

#### Príklad 2.14

Zostrojme maticu transformácie pre výmenu prvého a tretieho riadku v matici:

$$\mathbf{A} = \left( \begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{array} \right)$$

**Riešenie:** Vychádzať budeme z jednotkovej matice  $\mathbf{I}_4$ . Jednotková a transformačná matica vyzerajú nasledovne:

$$\mathbf{I}_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \mathbf{T}_{1,3} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Prenásobme maticou  $\mathbf{T}_{1,3}$  maticu  $\mathbf{A}$  zľava:

```
>A <- matrix(c(1, 0, 0, 0, 1, 1, 1, 0, 0, 0, 1, 1, 1, 0, 0, 0,
  0, 0, 1, 0), nrow=4, byrow=T)
> A
  [,1] [,2] [,3] [,4] [,5]
[1,]   1   0   0   0   1
[2,]   1   1   0   0   0
[3,]   1   1   1   0   0
[4,]   0   0   0   1   0
> T13 <- matrix(c(0, 0, 1, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0,
  1), nrow=4, byrow=T)
> T13
  [,1] [,2] [,3] [,4]
[1,]   0   0   1   0
[2,]   0   1   0   0
[3,]   1   0   0   0
[4,]   0   0   0   1
> T13 %*% A
  [,1] [,2] [,3] [,4] [,5]
[1,]   1   1   1   0   0
[2,]   1   1   0   0   0
[3,]   1   0   0   0   1
[4,]   0   0   0   1   0
```

Vidíme, že naozaj došlo k výmene prvého a tretieho riadku.

Transformačnú maticu  $\mathbf{T}_i^m$ , pomocou ktorej vynásobíme  $i$ -tý riadok ( $i \in \mathbb{N}$ ) číslom  $m \in \mathbb{R}$ , dostaneme z jednotkovej matice  $\mathbf{I}_n$  nahradením  $i$ -tého diagonálneho prvku číslom  $m$ .

### Príklad 2.15

Zostrojme maticu transformácie, ak chceme vynásobiť číslom 8 tretí riadok matice:

$$\mathbf{A} = \left( \begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{array} \right)$$

**Riešenie:** Vychádzať budeme z jednotkovej matice  $\mathbf{I}_4$ . Jednotková a transformačná matica vyzerajú nasledovne:

$$\mathbf{I}_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \mathbf{T}_3^8 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Prenásobme maticou  $\mathbf{T}_3^8$  maticu  $\mathbf{A}$  zľava:

```
>A <- matrix(c(1, 0, 0, 0, 1, 1, 1, 0, 0, 0, 1, 1, 1, 0, 0, 0,
  0, 0, 1, 0), nrow=4, byrow=T)
> A
  [,1] [,2] [,3] [,4] [,5]
[1,]   1   0   0   0   1
[2,]   1   1   0   0   0
[3,]   1   1   1   0   0
[4,]   0   0   0   1   0
> T3_8 <- matrix(c(1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 8, 0, 0, 0, 0,
  1), nrow=4, byrow=T)
> T3_8
  [,1] [,2] [,3] [,4]
[1,]   1   0   0   0
[2,]   0   1   0   0
[3,]   0   0   8   0
[4,]   0   0   0   1
> T3_8 %*% A
  [,1] [,2] [,3] [,4] [,5]
[1,]   1   0   0   0   1
[2,]   1   1   0   0   0
[3,]   8   8   8   0   0
[4,]   0   0   0   1   0
```

Transformačnú maticu  $\mathbf{T}_{ij}^m$ , pomocou ktorej pripočítame k  $i$ -tému riadku  $m$ -násobok  $j$ -tého riadku ( $i, j \in \mathbb{N}$ ,  $m \in \mathbb{R}$ ), dostaneme z jednotkovej matice  $\mathbf{I}_n$  nahradením prvku v  $i$ -tom riadku a  $j$ -tom stĺpci číslom  $m$ .

### Príklad 2.16

Zostrojme maticu transformácie, ak chceme k prvému riadku pripočítať štvornásobok tretieho riadku matice:

$$\mathbf{A} = \left( \begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{array} \right)$$

**Riešenie:** Vychádzať budeme z jednotkovej matice  $\mathbf{I}_4$ . Jednotková a transformačná matica vyzerajú nasledovne:

$$\mathbf{I}_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \mathbf{T}_{1,3}^4 = \begin{pmatrix} 1 & 0 & 4 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Prenásobme maticou  $\mathbf{T}_{1,3}^4$  maticu  $\mathbf{A}$  zľava:

```
>A <- matrix(c(1, 0, 0, 0, 1, 1, 1, 0, 0, 0, 1, 1, 1, 0, 0, 0,
  0, 0, 1, 0), nrow=4, byrow=T)
> A
  [,1] [,2] [,3] [,4] [,5]
[1,]   1   0   0   0   1
[2,]   1   1   0   0   0
[3,]   1   1   1   0   0
[4,]   0   0   0   1   0
> T13_4 <- matrix(c(1, 0, 4, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0,
  1), nrow=4, byrow=T)
> T13_4
  [,1] [,2] [,3] [,4]
[1,]   1   0   4   0
[2,]   0   1   0   0
[3,]   0   0   1   0
[4,]   0   0   0   1
> T13_4 %*% A
  [,1] [,2] [,3] [,4] [,5]
[1,]   5   4   4   0   1
[2,]   1   1   0   0   0
[3,]   1   1   1   0   0
[4,]   0   0   0   1   0
```

Všetky uvedené transformácie je možné aj skladať, napríklad je možné prenásobiť maticu zľava maticou  $\mathbf{T}_{2,1}^{-5}$  a  $\mathbf{T}_{3,1}^{-2}$ , v dôsledku čoho v matici od druhého riadku odpočítame päťnásobok prvého, a od tretieho riadku dvojnásobok prvého.

### Príklad 2.17

Vynásobte zľava v tomto poradí maticami  $\mathbf{T}_{2,1}^{-5}$  a  $\mathbf{T}_{3,1}^{-2}$  maticu  $\mathbf{A}$  danú:

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 5 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

**Riešenie:**

```
> A <- matrix(c(1, 0, 0, 0, 5, 1, 0, 0, 2, 0, 1, 0, 0, 0, 0,
  1), nrow=4, byrow=T)
```

```

> A
      [,1] [,2] [,3] [,4]
[1,]    1    0    0    0
[2,]    5    1    0    0
[3,]    2    0    1    0
[4,]    0    0    0    1
>T21__5 <- matrix(c(1, 0, 0, 0, -5, 1, 0, 0, 0, 0, 1, 0, 0, 0,
0, 1), nrow=4, byrow=T)
> T21__5
      [,1] [,2] [,3] [,4]
[1,]    1    0    0    0
[2,]   -5    1    0    0
[3,]    0    0    1    0
[4,]    0    0    0    1
>T31__2 <- matrix(c(1, 0, 0, 0, 0, 1, 0, 0, -2, 0, 1, 0, 0, 0,
0, 1), nrow=4, byrow=T)
> T31__2
      [,1] [,2] [,3] [,4]
[1,]    1    0    0    0
[2,]    0    1    0    0
[3,]   -2    0    1    0
[4,]    0    0    0    1
> T31__2 %*% (T21__5 %*% A)
      [,1] [,2] [,3] [,4]
[1,]    1    0    0    0
[2,]    0    1    0    0
[3,]    0    0    1    0
[4,]    0    0    0    1

```

Z predchádzajúceho príkladu vyplýva, že jednak môžeme transformácie kombinovať, no aj to, že je treba dať pozor na poradie, v ktorom ich realizujeme.

Zároveň je však vidieť, že požadovanou sériou operácií sme dostali jednotkovú maticu. Úpravy, ktoré sme realizovali, prevádzajú maticu  $\mathbf{A}$  na jednotkovú. Dostali sme teda:

$$\mathbf{T}_{2,1}^{-5} (\mathbf{T}_{3,1}^{-2} \mathbf{A}) = \mathbf{I}_n \quad (2.31)$$

Vďaka asociatívnosti násobenia matic (*1.16*) je ale táto rovnica identická s rovnicou:

$$(\mathbf{T}_{2,1}^{-5} \mathbf{T}_{3,1}^{-2}) \mathbf{A} = \mathbf{I}_n \quad (2.32)$$

Ak označíme:

$$\mathbf{B} = \mathbf{T}_{2,1}^{-5} \mathbf{T}_{3,1}^{-2} \quad (2.33)$$

Tak máme:

$$\mathbf{BA} = \mathbf{I}_n \quad (2.34)$$

To ale znamená, že  $\mathbf{B}$ , a teda aj súčin  $\mathbf{T}_{2,1}^{-5} \mathbf{T}_{3,1}^{-2}$  predstavujú inverznú maticu k  $\mathbf{A}$ :

$$\mathbf{B} = \mathbf{T}_{2,1}^{-5} \mathbf{T}_{3,1}^{-2} = \mathbf{A}^{-1} \quad (2.35)$$

Inverznú maticu k matici  $\mathbf{A}$  teda môžeme chápať aj tak, že predstavuje súčin všetkých transformačných matic, ktoré pomocou ekvivalentných úprav transformujú maticu  $\mathbf{A}$  na jednotkovú maticu.

### Príklad 2.18

Overme platnosť tvrdenia o interpretácii inverznej matice na maticiach z predchádzajúceho príkladu.

#### Riešenie:

Z predchádzajúceho príkladu máme:

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 5 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \mathbf{T}_{2,1}^{-5} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -5 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \mathbf{T}_{3,1}^{-2} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Odtiaľ:

```
> A <- matrix(c(1, 0, 0, 0, 5, 1, 0, 0, 2, 0, 1, 0, 0, 0, 0, 1),
  nrow=4, byrow=T)
> T21__5 <- matrix(c(1, 0, 0, 0, -5, 1, 0, 0, 0, 0, 1, 0, 0, 0,
  0, 1), nrow=4, byrow=T)
> T31__2 <- matrix(c(1, 0, 0, 0, 0, 1, 0, 0, -2, 0, 1, 0, 0, 0,
  0, 1), nrow=4, byrow=T)
> B <- T21__5 %*% T31__2
> B
      [,1] [,2] [,3] [,4]
[1,]    1    0    0    0
[2,]   -5    1    0    0
[3,]   -2    0    1    0
[4,]    0    0    0    1
> B %*% A
      [,1] [,2] [,3] [,4]
[1,]    1    0    0    0
[2,]    0    1    0    0
[3,]    0    0    1    0
[4,]    0    0    0    1
```

## 2.2 Polia a vektorové priestory

Už na základnej škole sa každý z nás stretol s pojmom reálnych čísel. Spravidla sa vyučuje súvislosť bodov na reálnej osi s číslami, a pokračuje sa v rámci aritmetiky so základnými operáciami počítania. S reálnymi číslami pracujeme tak dlho, že sú nám mnohé ich vlastnosti samozrejmé a pripadajú nám zjavné. Ako sme však videli v predchádzajúcej časti, kde sme zaviedli operáciu násobenia matic, nie na všetkých objektoch môžeme

odôvodnene predpokladať platnosť tvrdení podobných reálnym číslam (narážame tu hlavne na to, že násobenie matíc nie je komutatívne). Začneme preto tým, že zopakujeme niektoré vlastnosti štruktúry nazývanej pole, ktorým sú aj reálne čísla. Následne uvedieme pojem vektorového priestoru, teda určitej štruktúry, ktorú je možné vytvoriť za pomoci vektorov, definovaných v predchádzajúcej kapitole.

Ak by sme chceli definovať **pole**, potrebujeme k tomu niekoľko objektov. Pôjde o množinu, na ktorej prvkoch budú definované dve operácie, nazývané sčítanie a násobenie s dvoma špeciálnymi prvkami – nulovým a jednotkovým prvkom. Pre jednoduchosť môžeme pracovať rovno s množinou  $\mathbb{R}$ , aj keď nasledujúce definície umožňujú definovať veľmi veľa rôznych, v niektorých prípadoch veľmi neobvyklých polí.

Operácie sčítania a násobenia, označované  $+$  a  $\cdot$  musia spĺňať nasledovné podmienky: pre každé  $a, b, c \in \mathbb{R}$  platí:

$$a + b = b + a \quad (2.36)$$

$$a + (b + c) = (a + b) + c \quad (2.37)$$

$$a + 0 = a \quad (2.38)$$

$$(\forall a \in \mathbb{R})(\exists d \in \mathbb{R}): a + d = 0 \quad (2.39)$$

$$a \cdot b = b \cdot a \quad (2.40)$$

$$a \cdot (b \cdot c) = (a \cdot b) \cdot c \quad (2.41)$$

$$1 \cdot a = a \quad (2.42)$$

$$(\forall a \in \mathbb{R} \setminus \{0\})(\exists d \in \mathbb{R}): a \cdot d = 1 \quad (2.43)$$

$$a \cdot (b + c) = (a \cdot b) + (a \cdot c) \quad (2.44)$$

Podmienky (2.36) a (2.40) hovoria o komutatívnosti sčítania a násobenia. Vďaka nim je pri reálnych číslach jedno, ako vymeníme sčítance a činitele – výsledok to nezmení. Ako sme videli v predchádzajúcich častiach, pri maticiach je podmienka komutatívnosti splnená pri sčítaní, nie však pri násobení.

Vzťahy (2.37) a (2.41) hovoria o asociatívnosti sčítania a násobenia. Ak máme sčítať alebo vynásobiť niekoľko reálnych čísel, je jedno, v akom poradí to urobíme, súčet resp. súčin zostanú rovnaké. Ako sme poznamenali už pri maticiach, vďaka asociatívnosti si môžeme zjednodušiť zápis tým, že nemusíme uvádzať zátvorky, výrazy možno ponechať v tvare  $a + b + c$ , resp.  $a \cdot b \cdot c$ . Pri operácii súčinu sa navyše často samotný operátor vynecháva, a píšeme len  $abc$ .

Vzťahy (2.38) a (2.42) sú o neutrálnych prvkoch pre operácie sčítania a násobenia. Neutrálnym prvok pre sčítanie je 0 – ak ho pripočítame k akémukoľvek číslu  $a$ , dostaneme opäť  $a$ . Podobne je tomu tak v prípade čísla 1 a násobenia:  $1 \cdot a$  je stále  $a$ .

Vlastnosť (2.39) zaručuje existenciu opačného prvku k ľubovoľnému číslu  $a \in \mathbb{R}$ . Číslo  $d$  vo vzťahu sa často označuje aj  $-a$ .

Podobne vzťah (2.43) definuje existenciu recipročného čísla ku každému nenulovému číslu  $a$ . V tomto prípade sa recipročné číslo  $d$  označuje tiež  $1/a$ , alebo  $a^{-1}$ . Všimnime si tiež, že recipročné čísla existujú ku každému číslu, okrem neutrálneho prvku pre operáciu sčítania (0).

Posledný vzťah (2.44) vyjadruje distributívnosť násobenia vzhľadom na sčítanie, umožňuje nám „roznásobiť“ zátvorky. Ak zavedieme konvenciu, že v prípade našich zápisov má operácia násobenia prednosť pred operáciou sčítania, môžeme ho vyjadriť aj nasledovne

$$a \cdot (b + c) = a \cdot b + a \cdot c \quad (2.45)$$

Aby sme mohli nejakú množinu spolu so zvolenými dvomi operáciami a dvoma neutrálnymi prvkami prehlásiť za pole, museli by pre všetky prvky množiny platiť všetky vyššie uvedené vzťahy. Ak by sme sa zamýšľali nad tým, či môže existovať aj pole, ktorého prvkami by boli matice s operáciami, ktoré sme definovali v predchádzajúcej časti, tak už teraz vieme, že kvôli operácii maticového súčinu to nebude možné. Keďže tento súčin nie je komutatívny, nemohol by platiť vzťah (2.40).

Napriek tomu, že nedokážeme vytvoriť pole matíc, je pole reálnych čísel a jeho vlastnosti kľúčové pre definíciu inej, nesmierne dôležitej štruktúry: vektorového priestoru.

Podobne ako vyššie, **vektorovým priestorom** nad poľom  $\mathbb{R}$  nazývame podľa Zlatoša (2011) množinu  $V$ , operácie sčítania  $+$  a skalárneho násobenia  $\cdot$  a vektor  $0 \in V$ , ktorý spĺňa nižšie uvedené podmienky. Uvedme ešte, že operácia sčítania  $+$  nie je operáciou medzi reálnymi číslami (teda operáciou poľa  $\mathbb{R}$ ), ale súčtom dvoch vektorov z  $V$ , ktorej výsledkom je opäť vektor z  $V$ . Podobne súčin  $\cdot$  nie je súčinom reálnych čísel (v  $\mathbb{R}$ ), ale skalárnym násobením vektora: súčin je definovaný medzi reálnym číslom z  $\mathbb{R}$  a vektorom z  $V$ , pričom výsledok je opäť vektor z  $V$ . Pozor je treba dať tiež na to, že rozlišujeme pojmy skalárny súčin a násobenie skalárom – ide o rôzne operácie. Formálne,  $+$  je zobrazením  $V \times V \rightarrow V$  a  $\cdot$  je zobrazením  $\mathbb{R} \times V \rightarrow V$ .

Vo vektorovom priestore platia pre všetky  $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$  a  $a, b \in \mathbb{R}$  nasledovné vzťahy (Zlatoš, 2011):

$$\mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z} \quad (2.46)$$

$$\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x} \quad (2.47)$$

$$\mathbf{x} + \mathbf{0} = \mathbf{x} \quad (2.48)$$

$$(\forall \mathbf{x} \in V)(\exists \mathbf{d} \in V): \mathbf{x} + \mathbf{d} = \mathbf{0} \quad (2.49)$$

$$a \cdot (b \cdot \mathbf{x}) = (a \cdot b) \cdot \mathbf{x} \quad (2.50)$$

$$1 \cdot \mathbf{x} = \mathbf{x} \quad (2.51)$$

$$a \cdot (\mathbf{x} + \mathbf{y}) = (a \cdot \mathbf{x}) + (a \cdot \mathbf{y}) \quad (2.52)$$

$$(a + b) \cdot \mathbf{x} = (a \cdot \mathbf{x}) + (b \cdot \mathbf{x}) \quad (2.53)$$

Vzťah (2.46) je opäť o asociatívnosti, podobne ako u reálnych čísel, vzťah (2.47) požaduje komutatívnosť sčítania vektorov, (2.48) hovorí, že  $\mathbf{0}$  je neutrálny prvok pre operáciu sčítania, (2.49) zaručuje existenciu opačného vektora, (2.51) hovorí o neutrálnosti 1 vzhľadom na operáciu násobenia skalárom, a napokon (2.52) a (2.53) nám hovoria o distributívnosti. V texte podobne ako pri reálnych číslach niekedy znak násobenia skalárom vynechávame.

Pozastavíme sa pri podmienke (2.50): je veľmi dôležité si uvedomiť, že aj keď symboly pre operácie v rámci vektorového priestoru sú rovnaké ako v prípade poľa  $\mathbb{R}$ , ide o úplne rôzne veci. Vo vzťahu (2.50) vystupujú na ľavej strane výrazu dve násobenia skalárom (operácie  $V$ ): jednak násobenie  $b \cdot \mathbf{x}$ , jeho výsledkom je však opäť vektor, takže  $a \cdot (b \cdot \mathbf{x})$  je taktiež násobením vektora  $b \cdot \mathbf{x}$  skalárom  $a$ . Naopak na pravej strane máme najprv obyčajné násobenie v  $\mathbb{R}$  ( $a \cdot b$ ), ktorého výsledkom potom prenášobujeme v rámci násobenia skalárom vektor  $\mathbf{x}$ . Na pravej strane teda máme jedno násobenie čísel a jedno násobenie vektora skalárom. Uvedený príklad zdôrazňuje opatrnosť, s akou musíme pristupovať k čítaniu podobných zápisov.

Predchádzajúca pasáž môže pôsobiť trochu abstraktne (to sme napokon aj sľúbili v úvode k tejto kapitole). Ukážme si preto príklad vektorového priestoru, s ktorým budeme pracovať ďalej. Ak sa vrátíme k predchádzajúcej kapitole, pracovali sme v nej s vektormi, ktoré sme definovali ako usporiadané  $n$ -tice reálnych čísel. Množinu všetkých takých  $n$ -rozmerných vektorov budeme označovať symbolom  $\mathbb{R}^n$ . Vektorový priestor budeme vytvárať nad poľom reálnych čísel  $\mathbb{R}$ , operácia sčítania bude rovnaká ako v predchádzajúcich kapitolách a násobenie skalárom definujeme pre  $\mathbf{x} \in \mathbb{R}^n$  a  $a \in \mathbb{R}$  nasledovne:

$$a \cdot (x_1, x_2, \dots, x_n) = (ax_1, ax_2, \dots, ax_n) \quad (2.54)$$

Ak dodáme, že za nulový vektor budeme považovať vektor, ktorého všetky prvky sú rovné nule v  $\mathbb{R}$ , teda  $\mathbf{0} = (0, 0, \dots, 0) \in \mathbb{R}^n$ , potom máme všetko potrebné, aby sme mohli definovať vektorový priestor. Vlastnosti (2.46) až (2.53) nebudeme dokazovať, ale ukážeme si ich na príklade (postup pri dôkazoch by bol analogický).

### **Príklad 2.19**

Zvoľme tri vektory  $\mathbb{R}^5$ , konkrétne  $\mathbf{x} = (1, 2, 3, 4, 5)$ ,  $\mathbf{y} = (1, 2, 1, 2, 1)$  a  $\mathbf{z} = (1, 0, -1, 2, 1)$ , ako aj čísla  $a = 2$  a  $b = 3$ . Ukážme vždy aspoň na jednom príklade platnosť vzťahov (2.46) až (2.53).

Vzťah (2.46):

$$\begin{aligned} \mathbf{x} + (\mathbf{y} + \mathbf{z}) &= (1, 2, 3, 4, 5) + ((1, 2, 1, 2, 1) + (1, 0, -1, 2, 1)) = (1, 2, 3, 4, 5) + (2, 2, 0, 4, 2) = (3, \\ &4, 3, 8, 7) \\ &= (2, 4, 4, 6, 6) + (1, 0, -1, 2, 1) = ((1, 2, 3, 4, 5) + (1, 2, 1, 2, 1)) + (1, 0, -1, 2, 1) = (\mathbf{x} \\ &+ \mathbf{y}) + \mathbf{z} \end{aligned}$$

Vzťah (2.47):

$$\mathbf{x} + \mathbf{y} = (1, 2, 3, 4, 5) + (1, 2, 1, 2, 1) = (2, 4, 4, 6, 6) = (1, 2, 1, 2, 1) + (1, 2, 3, 4, 5) = \mathbf{y} + \mathbf{x}$$

Vzťah (2.48):

$$\mathbf{x} + \mathbf{0} = (1, 2, 3, 4, 5) + (0, 0, 0, 0, 0) = (1, 2, 3, 4, 5) = \mathbf{x}$$

Vzťah (2.49):

Pre  $\mathbf{z} = (1, 0, -1, 2, 1)$  existuje  $\mathbf{d} = (-1, 0, 1, -2, -1) \in \mathbb{R}^5$  také, že  $\mathbf{z} + \mathbf{d} = (0, 0, 0, 0, 0) = \mathbf{0}$ .

Vzťah (2.50):

$$a \cdot (b \cdot \mathbf{x}) = 2 \cdot (3 \cdot \mathbf{x}) = 2 \cdot (3, 6, 9, 12, 15) = (6, 12, 18, 24, 30) = (2 \cdot 3) \cdot (1, 2, 3, 4, 5) = (a \cdot b) \cdot \mathbf{x}$$

Vzťah (2.51):

$$1 \cdot \mathbf{x} = 1 \cdot (1, 2, 3, 4, 5) = (1, 2, 3, 4, 5) = \mathbf{x}$$

Vzťah (2.52):

$$\begin{aligned} a \cdot (\mathbf{x} + \mathbf{y}) &= 2 \cdot ((1, 2, 3, 4, 5) + (1, 2, 1, 2, 1)) = 2 \cdot (2, 4, 4, 6, 6) = (4, 8, 8, 12, 12) = \\ &= (2, 4, 6, 8, 10) + (2, 4, 2, 4, 2) = 2 \cdot (1, 2, 3, 4, 5) + 2 \cdot (1, 2, 1, 2, 1) = (a \cdot \mathbf{x}) + (a \cdot \mathbf{y}) \end{aligned}$$

Vzťah (2.53):

$$\begin{aligned} (a+b) \cdot \mathbf{x} &= (2+3) \cdot (1, 2, 3, 4, 5) = 5 \cdot (1, 2, 3, 4, 5) = (5, 10, 15, 20, 35) = (2, 4, 6, 8, 10) + (3, \\ &6, 9, 12, 15) \\ &= 2 \cdot (1, 2, 3, 4, 5) + 3 \cdot (1, 2, 3, 4, 5) = (a \cdot \mathbf{x}) + (b \cdot \mathbf{x}) \end{aligned}$$

Podstatnou vlastnosťou vektorového priestoru je, že je uzavretý vzhľadom na operácie násobenia vektorov skalárom, ako aj súčtu vektorov. Znamená to, že ak sčítame ľubovoľné dva vektory z toho istého vektorového priestoru  $V$ , nikdy sa nemôžeme dostať mimo tento vektorový priestor – výsledkom týchto operácií musí byť vždy vektor z  $V$ .

Uvažujme teraz o tom, že by sme mali  $n \in \mathbb{N}$  vektorov  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$  vybraných z vektorového priestoru  $V$ , tvoriacich usporiadanú  $n$ -tícu  $\mathbf{X}=(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$ . **Lineárnym obalom**  $\mathbf{X}$ , ktorý označíme  $[\mathbf{X}]$ , nazveme množinu:

$$[\mathbf{X}] = \{a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \dots + a_n\mathbf{x}_n; a_1, a_2, \dots, a_n \in \mathbb{R}\} \quad (2.55)$$

Ak výraz  $a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \dots + a_n\mathbf{x}_n$  nazveme **lineárnou kombináciou** vektorov  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ , potom predstavuje lineárny obal  $\mathbf{X}$  množinu všetkých lineárnych kombinácií vektorov, ktoré  $\mathbf{X}$  tvoria.

Vektory  $\mathbf{X}$  sme zvolili tak, že patrili do vektorového priestoru  $V$ . Zaujímavou skutočnosťou je, že lineárny obal  $[\mathbf{X}]$  bude sám osebe vektorovým priestorom. Najprv si uvedomme, že nulový vektor  $\mathbf{0}$  bude určite patriť do  $[\mathbf{X}]$ . Aby sme tomu uverili, stačí si uvedomiť, že lineárna kombinácia  $0\mathbf{x}_1 + 0\mathbf{x}_2 + \dots + 0\mathbf{x}_n = \mathbf{0}$  určite patrí do  $[\mathbf{X}]$ .

Keďže  $[\mathbf{X}]$  obsahuje všetky možné lineárne kombinácie vektorov, ktoré ho tvoria, určite bude uzavretý vzhľadom na operácie súčtu a násobenia skalárom – množina  $[\mathbf{X}]$  bola presne takto vytvorená. Čo sa týka vlastností požadovaných od vektorového priestoru, teda vzťahov (2.46) až (2.53), keďže boli všetky splnené vo vektorom priestore  $V$ , ktorého prvkami sú aj vektory z  $\mathbf{X}$ , ako aj  $[\mathbf{X}]$ , určite platia aj vo  $[\mathbf{X}]$ . Týmto je ale možné jednoznačne uzavrieť, že  $[\mathbf{X}]$  je naozaj vektorovým priestorom. Neznamená to však nutne, že  $[\mathbf{X}] = V$ , teda že lineárny obal je totožný s pôvodným vektorovým priestorom – môže ísť o tzv. **podpriestor vektorového priestoru**. O vektoroch tvoriacich  $\mathbf{X}$  hovoríme, že generujú vektorový priestor  $[\mathbf{X}]$ , a o priestore  $[\mathbf{X}]$  hovoríme, že je generovaný vektormi  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ .

Ak existujú vektory  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$  kde  $m \in \mathbb{N}$  generujúce celý vektorový priestor  $V$ , potom ho nazývame **konečnorozmerným**.

### Príklad 2.20

V predchádzajúcom príklade sme si naznačili, že  $\mathbb{R}^5$  je vektorovým priestorom. Uvažujme množinu  $\mathbf{X} = ((2, 0, 0, 0, 0), (0, 8, 0, 0, 0))$ . Úlohou je charakterizovať lineárny obal  $[\mathbf{X}]$ .

#### Riešenie:

Prvkami lineárneho obalu budú všetky vektory v tvare:

$$a_1(2, 0, 0, 0, 0) + a_2(0, 8, 0, 0, 0)$$

Skúsme si tieto vektory bližšie charakterizovať. Vektory, ktoré môžu vzniknúť z prvého výrazu, teda  $a_1(2, 0, 0, 0, 0)$  predstavujú všetko vektory, ktoré majú nenulový prvý člen a ostatné prvky nulové. Ide teda o všetky vektory v tvare  $(a, 0, 0, 0, 0)$  kde  $a \in \mathbb{R}$ . Podobne z výrazu  $a_2(0, 8, 0, 0, 0)$  môžu vzniknúť len vektory v tvare  $(0, b, 0, 0, 0)$  kde  $b \in \mathbb{R}$ . V dôsledku toho všetky lineárne kombinácie  $a_1(2, 0, 0, 0, 0) + a_2(0, 8, 0, 0, 0)$  budú zahŕňať všetky možné vektory v tvare  $(a, b, 0, 0, 0)$ , kde  $a, b \in \mathbb{R}$ .

Je zrejmé, že nulový vektor bude patriť do lineárneho obalu  $[X]$  – je to preto, že pre  $a = b = 0$  dostávame vektor  $\mathbf{0}$ . Ďalej je zrejmé, že ak sčítame dva vektory z  $[X]$ , t. j. dva vektory, ktoré môže mať len prvé dva prvky nenulové, ich súčet môže mať tiež len prvé dva prvky nenulové, a teda bude patriť do  $[X]$ . To isté bude platiť aj o skalárnych násobkoch vektorov z  $[X]$ . Pri troche snahy je možné ukázať, že pre takéto vektory budú platiť všetky vzťahy (2.46) až (2.53). Lineárny obal  $[X]$  je teda vektorovým priestorom. Zároveň platí, že  $[X]$  je vlastnou podmnožinou  $\mathbb{R}^5$ ,  $[X] \subset \mathbb{R}^5$ , keďže  $\mathbb{R}^5$  je vektorovým priestorom obsahujúcim aj vektory, ktoré nie sú prvkom  $[X]$ , napríklad vektor  $(0, 0, 1, 1, 1)$ . Lineárny obal  $[X]$  je preto podpriestorom vektorového priestoru  $\mathbb{R}^5$ .

### 2.3 Lineárna nezávislosť vektorov a báza vektorového priestoru

V predchádzajúcej podkapitole sme si definovali lineárnu kombináciu vektorov, ako výraz  $a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \dots + a_n\mathbf{x}_n$ . V tejto podkapitole s pomocou nej definujeme kľúčovú vlastnosť množiny vektorov, ktorá sa nazýva lineárna nezávislosť. Jej pochopenie je dôležité pri riešení sústav rovníc, lineárnom programovaní, ako aj v ekonometrii, kde sa o ňu opiera jeden z predpokladov, bez ktorých nie je možné zvolený model vôbec vypočítať.

Podobne ako v predchádzajúcej časti vychádzajme z  $n$ -tice vektorov  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ . O tejto  $n$ -tici budeme hovoriť, že je **lineárne závislá**, ak existujú také čísla  $c_1, c_2, \dots, c_n \in \mathbb{R}$ , kde  $(c_1, c_2, \dots, c_n) \neq \mathbf{0}$ , že platí:

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \dots + c_n\mathbf{x}_n = \mathbf{0} \quad (2.56)$$

Je jasné, že ak by sme nekládli podmienku  $(c_1, c_2, \dots, c_n) \neq \mathbf{0}$ , čísla  $(0, 0, \dots, 0)$  by vždy vyhovovali predchádzajúcemu vzťahu (išlo by o tzv. triviálnu lineárnu kombináciu). Nás však zaujíma, či existuje okrem tohto triviálneho riešenia ešte nejaké iné. O  $n$ -tici vektorov budeme hovoriť, že je **lineárne nezávislá**, ak nie je lineárne závislá (a teda  $(c_1, c_2, \dots, c_n) = \mathbf{0}$  by bolo jediným spôsobom, ako by platilo  $c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \dots + c_n\mathbf{x}_n = \mathbf{0}$ ).

**Príklad 2.21**

O vektoroch  $\mathbf{a} = (1, 2, 6, 8)$ ,  $\mathbf{b} = (2, 0, 1, 3)$  a  $\mathbf{c} = (8, 4, 15, 25)$  rozhodnite, či sú lineárne závislé.

**Riešenie:**

Táto úloha je trochu nefér, pretože úplný algoritmus umožňujúci posúdiť lineárnu nezávislosť množiny vektorom sme si ešte neukázali. Napriek tomu sa však môžeme presvedčiť, že:

$$\begin{aligned} -2\mathbf{a} - 3\mathbf{b} + \mathbf{c} &= -2(1, 2, 6, 8) - 3(2, 0, 1, 3) + (8, 4, 15, 25) \\ &= (-2, -4, -12, -16) + (-6, 0, -3, -9) + (8, 4, 15, 25) \\ &= (-8, -4, -15, -25) + (8, 4, 15, 25) = \mathbf{0} \end{aligned}$$

Z toho vyplýva, že pre  $(c_1, c_2, c_3) = (-2, -3, 1)$  je podmienka splnená, a teda vektory  $\mathbf{a}$ ,  $\mathbf{b}$  a  $\mathbf{c}$  sú lineárne závislé.

Ekvivalentným, a možno trochu názornejším spôsobom, ako je možné hovoriť o lineárnej závislosti vektorov je tento: vektory sú lineárne závislé, ak je možné jeden z nich vyjadriť ako lineárnu kombináciu ostatných.

Platnosť predchádzajúceho tvrdenia si môžeme ukázať pomerne jednoducho. Ak by vektory  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$  boli lineárne závislé, potom existujú také čísla  $c_1, c_2, \dots, c_n \in \mathbb{R}$  spĺňajúce podmienku (2.56), že nie všetky z nich sú nulové. Bez ujmy na všeobecnosti predpokladajme, že nenulové je  $c_n$ . Potom ale zo vzťahu (2.56) dostávame:

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \dots + c_n\mathbf{x}_n = \mathbf{0} \quad (2.57)$$

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \dots + c_{n-1}\mathbf{x}_{n-1} = -c_n\mathbf{x}_n \quad (2.58)$$

$$(-c_1c_n^{-1})\mathbf{x}_1 + (-c_2c_n^{-1})\mathbf{x}_2 + \dots + (-c_{n-1}c_n^{-1})\mathbf{x}_{n-1} = \mathbf{x}_n \quad (2.59)$$

Z tejto úvahy je zrejmé, že ak sú vektory lineárne závislé (2.57), potom dokážeme jeden z vektorov vyjadriť ako lineárnu kombináciu ostatných, čo je vidieť na ľavej strane (2.59). Keďže obidva vzťahy sú ekvivalentné, platí to aj naopak – ak je jeden z vektorov možné vyjadriť ako lineárnu kombináciu ostatných, potom sú lineárne závislé.

**Príklad 2.22**

Pri vektoroch  $\mathbf{a} = (1, 2, 6, 8)$ ,  $\mathbf{b} = (2, 0, 1, 3)$  a  $\mathbf{c} = (8, 4, 15, 25)$  sme sa presvedčili, že sú lineárne závislé. Vyjadriť teraz vektor  $\mathbf{c}$  a ako lineárnu kombináciu ostatných

**Riešenie:**

V minulom príklade sme sa presvedčili, že platí:

$$-2\mathbf{a} - 3\mathbf{b} + \mathbf{c} = \mathbf{0}$$

Odtiaľ teda dostávame:

$$\mathbf{c} = 2\mathbf{a} + 3\mathbf{b}$$

$$\mathbf{a} = 2^{-1}(\mathbf{c} - 3\mathbf{b})$$

Platnosť vzťahov overme:

$$2\mathbf{a} + 3\mathbf{b} = 2(1, 2, 6, 8) + 3(2, 0, 1, 3) = (2, 4, 12, 16) + (6, 0, 3, 9) = (8, 4, 15, 25) = \mathbf{c}$$

$$2^{-1}(\mathbf{c} - 3\mathbf{b}) = 2^{-1}((8, 4, 15, 25) - 3(2, 0, 1, 3)) = 2^{-1}((8, 4, 15, 25) - (6, 0, 3, 9)) = 2^{-1}(2, 4, 12, 16) = (1, 2, 6, 8) = \mathbf{a}$$

Lineárne nezávislé vektory majú v lineárnej algebre veľmi veľký význam, pretože umožňujú definovať tzv. bázu vektorového priestoru.

Ak je  $V$  konečnorozmerný vektorový priestor a  $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$  je usporiadaná  $n$ -tíca lineárne nezávislých vektorov, ktoré ho generujú, potom hovoríme, že tieto vektory tvoria **bázu vektorového priestoru**  $V$  a  $V$  má rozmer (**dimenziu**) rovnú  $n$ . Rozmer vektorového priestoru označujeme  $\dim(V)$ .

V predchádzajúcom tvrdení je dôležité, že vektory musia byť lineárne nezávislé. Navyše musia generovať celý vektorový priestor  $V$ , takže ľubovoľný vektor  $\mathbf{y} \in V$  sa musí dať vyjadriť ako lineárna kombinácia vektorov  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ . Ak sú všetky tieto podmienky splnené, potom počet vektorov  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$  určujeme dimenziu vektorového priestoru ( $n$ ).

Platí, že každý vektorový priestor má aspoň jednu bázu a každá báza toho istého vektorového priestoru má rovnaký počet prvkov. Nemôže sa nám preto stať, že by sme mali problém so stanovením dimenzie vektorového priestoru – každá jeho báza nám dá tú istú odpoveď.

### Príklad 2.23

Overte, či vektory  $\mathbf{e}_1 = (1, 0, 0)$ ,  $\mathbf{e}_2 = (0, 1, 0)$  a  $\mathbf{e}_3 = (0, 0, 1)$  tvoria bázu vektorového priestoru  $\mathbb{R}^3$ .

#### Riešenie:

Je zrejmé, že sú lineárne nezávislé, keďže žiaden z nich nie je možné vyjadriť ako lineárnu kombináciu ostatných. Overme si to napríklad tak, že sa pokúsime vyjadriť si vektor  $\mathbf{e}_1$  ako lineárnu kombináciu ostatných. Ľubovoľná lineárna kombinácia vektorov  $\mathbf{e}_2$  a  $\mathbf{e}_3$  je daná výrazom:

$$a_2\mathbf{e}_2 + a_3\mathbf{e}_3$$

pre  $a_2, a_3 \in \mathbb{R}$ . Všetky tieto lineárne kombinácie budú vyzerat' nasledovne:

$$a_2\mathbf{e}_2 + a_3\mathbf{e}_3 = a_2(0, 1, 0) + a_3(0, 0, 1) = (0, a_2, a_3)$$

V žiadnom prípade teda nemôžeme vyjadriť vektor  $\mathbf{e}_1 = (1, 0, 0)$  ako vektor  $(0, a_2, a_3)$ , keďže prvá zložka vektora je vždy rôzna. Podobne sa dá argumentovať aj v prípade ostatných vektorov, takže žiaden z nich sa nedá vyjadriť ako lineárna kombinácia ostatných, čo ale znamená, že sú naozaj lineárne nezávislé.

Vektory  $\mathbf{e}_1, \mathbf{e}_2$  a  $\mathbf{e}_3$  zároveň generujú celý vektorový priestor  $\mathbb{R}^3$ . Ak by tomu tak bolo, potom ľubovoľný vektor  $\mathbf{x} = (x_1, x_2, x_3)$  by sa mal dať napísať ako lineárna kombinácia vektorov  $\mathbf{e}_1, \mathbf{e}_2$  a  $\mathbf{e}_3$ . Ich ľubovoľná lineárna kombinácia sa dá napísať v tvare:

$$a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + a_3\mathbf{e}_3 = a_1(1, 0, 0) + a_2(0, 1, 0) + a_3(0, 0, 1) = (a_1, a_2, a_3)$$

Odtiaľ vidíme, že ak zvolíme  $a_1 = x_1, a_2 = x_2$  a  $a_3 = x_3$ , potom naozaj:

$$\mathbf{x} = a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + a_3\mathbf{e}_3$$

Každý vektor  $\mathbf{x}$  je možné vyjadriť ako lineárnu kombináciu vektorov  $\mathbf{e}_1, \mathbf{e}_2$  a  $\mathbf{e}_3$ . Tieto vektory preto generujú vektorový priestor  $V$ , čo môžeme zapísať ako  $[(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)] = V$ .

V dôsledku splnenia všetkých podmienok môžeme konštatovať, že vektory  $\mathbf{e}_1, \mathbf{e}_2$  a  $\mathbf{e}_3$  tvoria bázu vektorového priestoru  $\mathbb{R}^3$ . Priestor  $\mathbb{R}^3$  má rozmer, resp. dimenziu rovnú 3.

Báza vektorového priestoru  $\mathbb{R}^n$  zložená z vektorov  $\mathbf{e}_1 = (1, 0, \dots, 0), \mathbf{e}_2 = (0, 1, \dots, 0), \dots, \mathbf{e}_n = (0, 0, \dots, 1)$  sa nazýva **kanonická báza** vektorového priestoru  $\mathbb{R}^n$ . V skutočnosti môže mať vektorový priestor nekonečne veľa rôznych báz. Báza vektorového priestoru nie je nutne jediná. Ak má vektorový priestor  $V$  dimenziu  $n$ , potom ľubovoľných  $n$  lineárne nezávislých vektorov patriacich do  $V$  tvorí bázu  $V$ . Báza **nie je určená jednoznačne**.

#### Príklad 2.24

Overte, či vektory  $\mathbf{a}_1 = (1, 1, 0), \mathbf{a}_2 = (0, 1, 1), \mathbf{a}_3 = (0, 0, 1)$  a  $\mathbf{a}_4 = (0, 1, 0)$  tvoria bázu vektorového priestoru  $\mathbb{R}^3$ .

#### Riešenie:

Ako prvé si uvedomme, že sme v predchádzajúcom príklade už videli vektory, ktoré tvorili bázu vektorového priestoru  $\mathbb{R}^3$ : boli to vektory tvoriace jeho kanonickú bázu. Odtiaľ ale vieme, že tento vektorový priestor má rozmer 3, každá jeho báza bude pozostávať z troch vektorov. Keďže máme overiť, či vektory  $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$  a  $\mathbf{a}_4$  tvoria bázu  $\mathbb{R}^3$ , odpoveď je aj bez bližšieho skúmania jednoznačná: nie. Keďže ide o štyri vektory v trojrozmernom priestore,

bázu tvoriť nemôžu. Podobne by sme mohli zamietnuť tvrdenie, že ktorékoľvek dva vektory tvoria bázu  $\mathbb{R}^3$ : v tomto vektorovom priestore musia bázu tvoriť práve tri vektory.

### Príklad 2.25

Overte, či vektory  $\mathbf{x}_1 = (1, 1, 0)$ ,  $\mathbf{x}_2 = (0, 1, 1)$  a  $\mathbf{x}_3 = (0, 0, 1)$  tvoria bázu vektorového priestoru  $\mathbb{R}^3$ .

**Riešenie:** Vieme, že každá báza vektorového priestoru  $\mathbb{R}^3$  je zložená z práve troch vektorov. Keďže skúmame tri vektory, táto podmienka je splnená. To však ešte nepostačuje na to, aby sme vedeli vektory  $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$  prehlásiť za bázu.

Postupovať by sme mohli dvojako. Ukážeme, že tieto tri vektory sú lineárne nezávislé, a ukážeme, že ľubovoľné  $\mathbf{y} \in \mathbb{R}^3$  je možné vyjadriť ako lineárnu kombináciu  $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ .

V tomto konkrétnom prípade je ľahšie ukázať prvú vlastnosť, a to sporom. Predpokladajme na chvíľu, že vektory  $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$  sú lineárne závislé. Potom je možné nájsť čísla  $a_2, a_3 \in \mathbb{R}$  tak, že jeden z vektorov je možné vyjadriť ako lineárnu kombináciu ostatných. Pokúsme sa preto vyjadriť prvý vektor ako lineárnu kombináciu druhého a tretieho vektora. Každá ich lineárna kombinácia bude v tvare:

$$a_2\mathbf{x}_2 + a_3\mathbf{x}_3 = a_2(0,1,1) + a_3(0,0,1) = (0, a_2, a_2 + a_3)$$

Keďže má existovať lineárna kombinácia, ktorá sa rovná vektoru  $\mathbf{x}_1$ , má platiť:

$$a_2\mathbf{x}_2 + a_3\mathbf{x}_3 = \mathbf{x}_1$$

teda:

$$(0, a_2, a_2 + a_3) = (1, 1, 0)$$

čo nemôže byť splnené pre žiadne  $a_2, a_3 \in \mathbb{R}$  (všimnime si prvé zložky vektorov, ktoré sa nemôžu rovnať). Dostávame teda spor, a pôvodné tvrdenie, že vektory  $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$  sú lineárne závislé je nepravdivé. Ukážme ešte, že ľubovoľné  $\mathbf{y} = (y_1, y_2, y_3) \in \mathbb{R}^3$  je možné vyjadriť ako lineárnu kombináciu  $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ .

Preskúmajme lineárnu kombináciu:

$$y_1\mathbf{x}_1 + (y_2 - y_1)\mathbf{x}_2 + (y_3 - y_2 + y_1)\mathbf{x}_3$$

Dostávame:

$$\begin{aligned} & y_1\mathbf{x}_1 + (y_2 - y_1)\mathbf{x}_2 + (y_3 - y_2 + y_1)\mathbf{x}_3 \\ &= y_1(1, 1, 0) + (y_2 - y_1)(0, 1, 1) + (y_3 - y_2 + y_1)(0, 0, 1) \\ &= (y_1, y_1, 0) + (0, y_2 - y_1, y_2 - y_1) + (0, 0, y_3 - y_2 + y_1) \end{aligned}$$

$$\begin{aligned}
&= (y_1, y_1 + y_2 - y_1, y_2 - y_1) + (0, 0, y_3 - y_2 + y_1) \\
&= (y_1, y_2, y_2 - y_1) + (0, 0, y_3 - y_2 + y_1) \\
&= (y_1, y_2, y_2 - y_1 + y_3 - y_2 + y_1) \\
&= (y_1, y_2, y_3)
\end{aligned}$$

Z toho ale vyplýva, že ľubovoľný vektor  $\mathbf{y} = (y_1, y_2, y_3) \in \mathbb{R}^3$  môžeme vyjadriť ako lineárnu kombináciu  $y_1\mathbf{x}_1 + (y_2 - y_1)\mathbf{x}_2 + (y_3 - y_2 + y_1)\mathbf{x}_3$  vektorov  $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ . Lineárny obal týchto vektorov bude zahŕňať celý vektorový priestor  $\mathbb{R}^3$ . Keďže sú splnené všetky podmienky definície, vektory  $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$  tvoria bázu vektorového priestoru  $\mathbb{R}^3$ .

V skutočnosti je možné ukázať, že obidve požadované vlastnosti sú ekvivalentné: ak máme v  $n$ -rozmernom vektorovom priestore  $V$   $n$  lineárne nezávislých vektorov, budú tvoriť aj jeho bázu, a teda akýkoľvek vektor z  $V$  je možné vyjadriť ako ich kombináciu. Naopak, ak platí  $[\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] = V$  a  $\dim(V) = n$ , potom sú vektory  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$  lineárne nezávislé a tvoria bázu  $V$ .

## 2.4 Hodnosť matice

V časti 1.3.4 sme sa zaoberali pojmom inverznej matice, v časti 2.1 sme zase poukázali na vzťah, ktorý existuje medzi riešením sústavy lineárnych rovníc a inverznými maticami. Vysvetlili sme, že inverzná matica existuje len k štvorcovým maticiam, a aj to len k niektorým – takzvaným regulárnym maticiam. V časti o riešení lineárnych rovníc sme si ukázali aj spôsob, ako je možné nájsť inverznú maticu k regulárnej štvorcovej matici.

V skutočnosti regulárnosť matice súvisí aj s koncepciou, ktorú sme si ukázali v predchádzajúcej časti – s lineárnymi kombináciami vektorov.

Uvažujme napríklad o matici:

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 2 & 1 & 1 & 2 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

Označme jej riadkové vektory  $\mathbf{a}_1 = (1, 0, 1, 1)$ ,  $\mathbf{a}_2 = (2, 1, 1, 2)$ ,  $\mathbf{a}_3 = (1, 0, 0, 1)$ ,  $\mathbf{a}_4 = (0, 1, 1, 0)$  a stĺpcové vektory  $\mathbf{a}_1 = (1, 2, 1, 0)^T$ ,  $\mathbf{a}_2 = (0, 1, 0, 1)^T$ ,  $\mathbf{a}_3 = (1, 1, 0, 1)^T$ ,  $\mathbf{a}_4 = (1, 2, 1, 0)^T$ .

Ak by sme v programe R skúsili vypočítať inverznú maticu k  $\mathbf{A}$ , dostali by sme tento výsledok:

```
> A <- matrix(c(1, 0, 1, 1, 2, 1, 1, 2, 1, 0, 0, 1, 0, 1, 1, 0),
  nrow=4, byrow=T)
> A
      [,1] [,2] [,3] [,4]
[1,]    1    0    1    1
[2,]    2    1    1    2
[3,]    1    0    0    1
[4,]    0    1    1    0
> solve(A)
Error in solve.default(A) :
  Lapack routine dgesv: system is exactly singular
```

Zisťujeme, že matica  $\mathbf{A}$  je singularárna, a teda k nej neexistuje inverzná matica. Skúsme sa ale pozrieť na riadkové vektory  $\mathbf{a}_1$ ,  $\mathbf{a}_2$ ,  $\mathbf{a}_3$  a  $\mathbf{a}_4$  z hľadiska lineárnej nezávislosti, definovanej v predchádzajúcej podkapitole.

Zrejme platí:

$$\mathbf{a}_2 = 2 \cdot \mathbf{a}_3 + \mathbf{a}_4 \quad (2.60)$$

$$(2, 1, 1, 2) = 2 \cdot (1, 0, 0, 1) + (0, 1, 1, 0) \quad (2.61)$$

Riadkové vektory matice  $\mathbf{A}$  sú lineárne závislé. Vynechajme teda vektor  $\mathbf{a}_2$ , a preskúmajme ďalej zostávajúce riadkové vektory. Je zjavné, že z vektorov  $\mathbf{a}_1$  a  $\mathbf{a}_3$  nemôžeme vo forme lineárnej kombinácie dostať vektor  $\mathbf{a}_4$ . Ak by tomu tak malo byť, potom by malo pre nejaké  $b_1, b_3 \in \mathbb{R}$  platiť:

$$\mathbf{a}_4 = (0, 1, 1, 0) = b_1(1, 0, 1, 1) + b_3(1, 0, 0, 1) = b_1\mathbf{a}_1 + b_3\mathbf{a}_3 \quad (2.62)$$

Ak porovnáme druhé zložky vektorov, vidíme, že bez ohľadu na voľbu  $b_1, b_3 \in \mathbb{R}$  nikdy nemôžeme nájsť takú lineárnu kombináciu vektorov  $\mathbf{a}_1$  a  $\mathbf{a}_3$ , ktorá by bola rovná  $\mathbf{a}_4$ . Zároveň vidíme, že vektor  $\mathbf{a}_1$  nie je násobkom vektora  $\mathbf{a}_3$ . Vektory  $\mathbf{a}_1$ ,  $\mathbf{a}_3$  a  $\mathbf{a}_4$  sú lineárne nezávislé.

Maximálny počet lineárne nezávislých riadkov v matici sa nazýva **riadková hodnosť matice**. Matica  $\mathbf{A}$  z nášho príkladu má riadkovú hodnosť 3, pretože obsahuje tri riadkové vektory, ktoré sú lineárne nezávislé.

Skúsme analogickú úvahu realizovať aj na stĺpcových vektoroch matice  $\mathbf{A}$ . Na prvý pohľad vidíme, že prvý a štvrtý stĺpcový vektor matice  $\mathbf{A}$  sú identické ( $\mathbf{a}_1$  a  $\mathbf{a}_4$ ), v ďalšej úvahe pracujme preto len s prvými tromi stĺpcovými vektormi. Podobne ako pri riadkových vektoroch, kvôli tretím zložkám vektorov nie je možné vyjadriť vektor  $\mathbf{a}_1$  ako lineárnu

kombináciu vektorov  $\mathbf{a}_2$  a  $\mathbf{a}_3$ . Vektor  $\mathbf{a}_2$  nie je násobkom vektora  $\mathbf{a}_3$ , takže v matici  $\mathbf{A}$  vieme nájsť maximálne tri lineárne nezávislé stĺpce.

Maximálny počet lineárne nezávislých stĺpcov v matici sa nazýva **stĺpcová hodnosť matice**. Matica  $\mathbf{A}$  z nášho príkladu má stĺpcovú hodnosť 3.

Dá sa dokázať, že výsledok, ktorý sme získali, konkrétne, že riadková a stĺpcová hodnosť matice je rovnaká, nie je náhodný, ale platí pre všetky matice. Z tohto dôvodu maximálny počet lineárne nezávislých stĺpcov, ako aj riadkov v matici  $\mathbf{A}$  nazývame **hodnosťou matice**  $\mathbf{A}$  a označujeme ju  $h(\mathbf{A})$ .

Jedným z dôsledkov predchádzajúceho tvrdenia je, že ak máme vo všeobecnosti maticu rozmeru  $m \times n$ , potom  $h(\mathbf{A}) \leq \min(m, n)$ .

Pomocou hodnosti matice je možné určiť, kedy k štvorcovej matici rozmeru  $n \times n$  existuje inverzná matica. Je tomu tak vtedy, ak platí:

$$h(\mathbf{A}) = n \quad (2.63)$$

Nasledovné štyri tvrdenia týkajúce sa štvorcovej matice  $\mathbf{A}$  sú ekvivalentné:

- Matica  $\mathbf{A}$  je regulárna.
- Hodnosť matice  $\mathbf{A}$  je  $n$ .
- K matici  $\mathbf{A}$  existuje inverzná matica  $\mathbf{A}^{-1}$ .
- Ak je  $\mathbf{A}$  maticou sústavy  $n$  lineárnych rovníc, táto sústava má jediné riešenie.

## 2.5 Zmena bázy a lineárne zobrazenia

V predchádzajúcich častiach sme si ukázali tzv. kanonickú bázu vektorového priestoru  $\mathbb{R}^n$ , ktorú tvoria vektory  $\mathbf{e}_1 = (1, 0, \dots, 0)$ ,  $\mathbf{e}_2 = (0, 1, \dots, 0)$ , ...,  $\mathbf{e}_n = (0, 0, \dots, 1)$ . Pre jednoduchosť pracujme na chvíľu s bázou vektorového priestoru  $\mathbb{R}^4$ ,  $\mathbf{e}_1 = (1, 0, 0, 0)$ ,  $\mathbf{e}_2 = (0, 1, 0, 0)$ ,  $\mathbf{e}_3 = (0, 0, 1, 0)$  a  $\mathbf{e}_4 = (0, 0, 0, 1)$ .

Keďže tieto vektory tvoria bázu  $\mathbb{R}^4$ , každý vektor v  $\mathbb{R}^4$  sa dá vyjadriť ako lineárna kombinácia vektorov tejto bázy. V skutočnosti je toto vyjadrenie veľmi jednoduché: ak máme vektor  $\mathbf{x} = (x_1, x_2, x_3, x_4) \in \mathbb{R}^4$ , potom platí:

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3 + x_4\mathbf{e}_4 \quad (2.64)$$

$$\mathbf{x} = x_1(1, 0, 0, 0) + x_2(0, 1, 0, 0) + x_3(0, 0, 1, 0) + x_4(0, 0, 0, 1) \quad (2.65)$$

### Príklad 2.26

Vyjadrite vektor  $\mathbf{x} = (2, 4, 8, 16)$  ako lineárnu kombináciu vektorov kanonickej bázy v  $\mathbb{R}^4$ .

**Riešenie:** Riešenie je veľmi jednoduché, platí:

$$\begin{aligned}\mathbf{x} &= 2\mathbf{e}_1 + 4\mathbf{e}_2 + 8\mathbf{e}_3 + 16\mathbf{e}_4 \\ &= 2(1, 0, 0, 0) + 4(0, 1, 0, 0) + 8(0, 0, 1, 0) + 16(0, 0, 0, 1)\end{aligned}$$

Čísla  $x_1, x_2, x_3, x_4$  je možné chápať ako **súradnice vektora  $\mathbf{x}$**  v kanonickej báze (štandardne vektory zapisujeme práve v kanonickej báze). Tento výrok však so sebou prináša logickú otázku: ak sme si práve definovali súradnice vektora v kanonickej báze, a vieme, že báz toho istého vektorového priestoru môže byť nekonečne veľa, znamená to, že je možné definovať súradnice vektora aj vzhľadom na inú, ľubovoľne zvolenú bázu? Odpoveďou na túto otázku je áno.

### Príklad 2.27

Overte, či súradnice vektora  $\mathbf{x} = (2, 4, 8, 16)$  v báze  $\mathbf{a}_1 = (1, 0, 1, 0)$ ,  $\mathbf{a}_2 = (2, 2, 0, 1)$ ,  $\mathbf{a}_3 = (1, 1, 1, 0)$ ,  $\mathbf{a}_4 = (1, 1, 0, 0)$  sú v kanonickej báze  $(-2, 16, 10, -38)$ .

**Riešenie:** Súradnice vektora v báze vyjadrujú, akou lineárnou kombináciou vektorov bázy dostaneme požadovaný vektor. Malo by preto platiť, že:

$$-2\mathbf{a}_1 + 16\mathbf{a}_2 + 10\mathbf{a}_3 - 38\mathbf{a}_4 = \mathbf{x}$$

Túto skutočnosť overme:

$$\begin{aligned}-2\mathbf{a}_1 + 16\mathbf{a}_2 + 10\mathbf{a}_3 - 38\mathbf{a}_4 &= -2(1, 0, 1, 0) + 16(2, 2, 0, 1) + 10\mathbf{a}_3 - 38\mathbf{a}_4 \\ &= (-2, 0, -2, 0) + (32, 32, 0, 16) + 10\mathbf{a}_3 - 38\mathbf{a}_4 \\ &= (30, 32, -2, 16) + 10\mathbf{a}_3 - 38\mathbf{a}_4 \\ &= (30, 32, -2, 16) + 10(1, 1, 1, 0) - 38\mathbf{a}_4 \\ &= (30, 32, -2, 16) + (10, 10, 10, 0) - 38\mathbf{a}_4 \\ &= (40, 42, 8, 16) - 38\mathbf{a}_4 \\ &= (40, 42, 8, 16) - 38(1, 1, 0, 0) \\ &= (40, 42, 8, 16) - (38, 38, 0, 0) \\ &= (2, 4, 8, 16)\end{aligned}$$

Skutočnosť, že je vždy možné vektor v jednej báze  $(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \dots, \mathbf{a}_n)$  vyjadriť v inej báze, napr.  $(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \dots, \mathbf{b}_n)$  nie je prekvapivá. Keďže obidve  $n$ -tice vektorov tvoria bázu nejakého vektorového priestoru  $V$ , potom musí platiť, že obidve bázy generujú  $V$

$$[(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \dots, \mathbf{a}_n)] = V = [(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \dots, \mathbf{b}_n)] \quad (2.66)$$

Vzhľadom na to, že vektory oboch báz generujú celý vektorový priestor  $V$ , každý jeho vektor sa dá vyjadriť ako lineárna kombinácia jednej, ale aj druhej bázy. Koeficienty týchto lineárnych kombinácií predstavujú súradnice v príslušnej báze.

Označme bázu tvorenú vektormi  $(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \dots, \mathbf{a}_n)$  znakom  $\mathbf{a}$ . Súradnice vektora  $\mathbf{x}$  vyjadrené v báze  $\mathbf{a}$  budeme označovať  $\mathbf{x}_a = (x_{1a}, x_{2a}, \dots, x_{na})_a$ . Pre súradnice vektora v kanonickej báze budeme používať zápis ako doposiaľ, teda  $\mathbf{x} = (x_1, x_2, \dots, x_n)$ .

V predchádzajúcom príklade sme realizovali prevod medzi dvomi bázami, alebo takzvanú **zmenu bázy**. Súradnice vektora  $\mathbf{x}_a = (-2, 16, 10, -38)_a^T$  sme v kanonickej báze vypočítali tak, že sme dosadili ich koeficienty do lineárnej kombinácie vektorov tvoriacich  $\mathbf{a}$ . Tento postup sme mohli zapísať aj inak: vektory tvoriace bázu  $\mathbf{a}$ , teda  $\mathbf{a}_1 = (1, 0, 1, 0)^T$ ,  $\mathbf{a}_2 = (2, 2, 0, 1)^T$ ,  $\mathbf{a}_3 = (1, 1, 1, 0)^T$ ,  $\mathbf{a}_4 = (1, 1, 0, 0)^T$  sme mohli zapísať ako stĺpcové vektory matice:

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 2 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} = (\mathbf{a}_1 \mathbf{a}_2 \mathbf{a}_3 \mathbf{a}_4) \quad (2.67)$$

Všimnime si, že posledný výraz, t. j.  $(\mathbf{a}_1 \mathbf{a}_2 \mathbf{a}_3 \mathbf{a}_4)$  nepredstavuje vektor, ale maticu v blokovej forme (pozri podkapitulu 1.3.7). Vo vektoroch oddeľujeme jeho zložky vždy čiarkou.

Ak by súradnice vektora  $\mathbf{x}$  v báze  $\mathbf{a}$  boli  $\mathbf{x}_a = (-2, 16, 10, -38)_a^T$ , potom by jeho súradnice v kanonickej báze boli dané lineárnou kombináciou:

$$\mathbf{x} = -2\mathbf{a}_1 + 16\mathbf{a}_2 + 10\mathbf{a}_3 - 38\mathbf{a}_4 \quad (2.68)$$

Z príkladu vieme, že  $\mathbf{x} = (2, 4, 8, 16)^T$ . Predchádzajúci výraz sa však dá zapísať aj inak:

$$\mathbf{x} = -2\mathbf{a}_1 + 16\mathbf{a}_2 + 10\mathbf{a}_3 - 38\mathbf{a}_4 \quad (2.69)$$

$$= (\mathbf{a}_1 \mathbf{a}_2 \mathbf{a}_3 \mathbf{a}_4) \begin{pmatrix} -2 \\ 16 \\ 10 \\ -38 \end{pmatrix} \quad (2.70)$$

$$= (\mathbf{a}_1 \mathbf{a}_2 \mathbf{a}_3 \mathbf{a}_4) \mathbf{x}_a \quad (2.71)$$

$$= \mathbf{A} \mathbf{x}_a \quad (2.72)$$

Súradnice vektora  $\mathbf{x}$  v kanonickej báze teda vieme získať ako súčin matice vektorov bázy  $\mathbf{A}$  a vektora  $\mathbf{x}_a$  súradníc v báze  $\mathbf{a}$ . V určitom zmysle teda môžeme maticu  $\mathbf{A}$  chápať ako reprezentáciu **transformácie medzi súradnicami** vyjadrenými v rôznych bázach.

V predchádzajúcom texte sme uvažovali o transformácii súradníc vektora  $\mathbf{x}$  z ľubovoľnej bázy  $\mathbf{a}$  do kanonickej bázy. Je možné sa ale opýtať, ako je možné vykonať opačnú transformáciu, teda ako získať súradnice vektora  $\mathbf{x}_a$ , ak poznáme jeho súradnice  $\mathbf{x}$  v kanonickej báze. Už vieme, že platí:

$$\mathbf{x} = \mathbf{A}\mathbf{x}_a \quad (2.73)$$

Aby sme si z predchádzajúceho vzťahu vyjadrili hľadané  $\mathbf{x}_a$ , prenásobíme celú rovnicu maticou  $\mathbf{A}^{-1}$  zľava. Podotknime, že matica  $\mathbf{A}$  sa skladá z vektorov, ktoré tvoria bázu  $\mathbb{R}^4$ , sú preto lineárne nezávislé a hodnosť matice  $\mathbf{A}$  je rovná počtu jej riadkov a stĺpcov. Matica  $\mathbf{A}$  je preto regulárna, a preto k nej určite existuje inverzná matica  $\mathbf{A}^{-1}$ . Dostávame preto:

$$\mathbf{A}^{-1}\mathbf{x} = \mathbf{A}^{-1}\mathbf{A}\mathbf{x}_a \quad (2.74)$$

$$\mathbf{A}^{-1}\mathbf{x} = \mathbf{I}_n\mathbf{x}_a \quad (2.75)$$

$$\mathbf{A}^{-1}\mathbf{x} = \mathbf{x}_a \quad (2.76)$$

Súradnice vektora vyjadreného v kanonickej báze ako  $\mathbf{x}$  môžeme vyjadriť aj v báze  $\mathbf{a}$  ako  $\mathbf{A}^{-1}\mathbf{x}$ .

### Príklad 2.28

Zistite súradnice vektora  $\mathbf{x} = (1, 0, 4, 6)^T$  v báze  $\mathbf{a}$  pozostávajúcej z vektorov  $\mathbf{a}_1 = (1, 0, 0, -1)^T$ ,  $\mathbf{a}_2 = (0, 1, 0, 1)^T$ ,  $\mathbf{a}_3 = (1, 1, 1, 0)^T$ ,  $\mathbf{a}_4 = (-1, 1, 1, 1)^T$ .

**Riešenie:** Matica, ktorou by sme prevádzali súradnice vektorov z bázy  $\mathbf{a}$  do kanonickej bázy pozostáva zo stĺpcových vektorov rovných vektorom bázy. V programe R má nasledovný tvar:

```
> A <- matrix(c(1, 0, 1, -1, 0, 1, 1, 1, 0, 0, 1, 1, -1, 1, 0,
  1), nrow=4, byrow=TRUE)
> A
      [,1] [,2] [,3] [,4]
[1,]    1    0    1   -1
[2,]    0    1    1    1
[3,]    0    0    1    1
[4,]   -1    1    0    1
```

Keďže je našou úlohou uskutočniť opačnú transformáciu, nájdeme inverznú maticu k matici  $\mathbf{A}$ :

```

> Ainv <- solve(A)
> Ainv
      [,1] [,2] [,3] [,4]
[1,]  -1    2  -1   -2
[2,]   0    1  -1    0
[3,]   1   -1   1    1
[4,]  -1    1   0   -1

```

Požadované súradnice  $\mathbf{x}_a$  dostávame ako súčin inverznej matice k matici  $\mathbf{A}$  a súradníc vektora  $\mathbf{x}$  v kanonickej báze.

```

> x <- c(1, 0, 4, 6)
> xa <- Ainv %*% x
> xa
      [,1]
[1,] -17
[2,]  -4
[3,]  11
[4,]  -7

```

Overme ešte, či naozaj ide o súradnice vektora  $\mathbf{x}$  tým, že uskutočníme transformáciu na súradnice kanonickej bázy:

```

> A %*% xa
      [,1]
[1,]    1
[2,]    0
[3,]    4
[4,]    6

```

Súradnice vektora  $\mathbf{x}$  v báze  $\mathbf{a}$  sú  $\mathbf{x}_a = (-17, -4, 11, -7)_a$ .

Teraz vieme, ako je možné vektor v kanonickej báze vyjadriť v ľubovoľnej inej báze, a naopak – z vektora v ľubovoľnej báze vieme získať súradnice v kanonickej báze.

Poslednou zmenou bázy, ktorá prichádza do úvahy, je zmena medzi dvomi ľubovoľnými bázami, z ktorých žiadna nie je kanonická. Ide teda o úlohu vyjadriť vektor so súradnicami  $\mathbf{x}_a$  pomocou súradníc  $\mathbf{x}_b$ , kde  $\mathbf{a}$  a  $\mathbf{b}$  sú iné ako kanonické bázy toho istého vektorového priestoru.

Začnime preto so súradnicami  $\mathbf{x}_a$  v báze  $\mathbf{a}$ . Vieme, že súradnice vektora v kanonickej báze dostaneme nasledovne:

$$\mathbf{x} = \mathbf{A}\mathbf{x}_a \quad (2.77)$$

Ďalej vieme, že súradnice v kanonickej báze vieme transformovať na súradnice v báze  $\mathbf{b}$  takto:

$$\mathbf{x}_b = \mathbf{B}^{-1}\mathbf{x} \quad (2.78)$$

kde  $\mathbf{B}$  predstavuje maticu, ktorej stĺpce zodpovedajú vektorom bázy  $\mathbf{b}$  vyjadrených v kanonickej báze. Celkovo teda dostávame:

$$\mathbf{x}_b = \mathbf{B}^{-1}\mathbf{x} = \mathbf{B}^{-1}\mathbf{A}\mathbf{x}_a \quad (2.79)$$

Súradnice v báze  $\mathbf{b}$  dostávame pomocou transformačnej matice  $\mathbf{B}^{-1}\mathbf{A}$ , ktorá je zložená z dvoch transformácií, zapísaných pomocou  $\mathbf{B}^{-1}$  a  $\mathbf{A}$ .

### Príklad 2.29

Zistite súradnice vektora  $\mathbf{x}_a = (1, 0, 4, 6)_a$  v báze  $\mathbf{b}$ , ak je báza  $\mathbf{a}$  tvorená vektormi  $\mathbf{a}_1 = (2, 2, 1, 4)$ ,  $\mathbf{a}_2 = (1, -2, 2, 4)$ ,  $\mathbf{a}_3 = (4, -3, 1, 1)$ ,  $\mathbf{a}_4 = (1, 1, 2, 4)$  a báza  $\mathbf{b}$  vektormi  $\mathbf{b}_1 = (0, 2, 0, 1)$ ,  $\mathbf{b}_2 = (1, -2, 0, 1)$ ,  $\mathbf{b}_3 = (2, 0, 1, 2)$ ,  $\mathbf{b}_4 = (0, 1, 1, -2)$ .

**Riešenie:** Podľa (2.79) dostaneme požadovaný vektor  $\mathbf{x}_b$  ako súčin  $\mathbf{x}_b = \mathbf{B}^{-1}\mathbf{A}\mathbf{x}_a$ . Postupovať preto budeme tak, že najprv definujeme potrebné matice v programe R.

```
> A <- matrix(c(2, 1, 4, 1, 2, -2, -3, 1, 1, 2, 1, 2, 4, 4, 1,
  4), nrow=4, byrow=T)
> A
  [,1] [,2] [,3] [,4]
[1,]   2   1   4   1
[2,]   2  -2  -3   1
[3,]   1   2   1   2
[4,]   4   4   1   4
> B <- matrix(c(0, 1, 2, 0, 2, -2, 0, 1, 0, 0, 1, 1, 1, 1, 2, -
  2), nrow=4, byrow=T)
> B
  [,1] [,2] [,3] [,4]
[1,]   0   1   2   0
[2,]   2  -2   0   1
[3,]   0   0   1   1
[4,]   1   1   2  -2
```

Ďalej spočítame inverznú maticu k matici  $\mathbf{B}$ :

```
> Binv <- solve(B)
> Binv
  [,1] [,2] [,3] [,4]
[1,]   7   2  -8  -3
[2,]   9   2 -10  -4
[3,]  -4  -1   5   2
[4,]   4   1  -4  -2
```

Matica transformácie medzi súradnicami v báze  $\mathbf{a}$  a  $\mathbf{b}$  je potom rovná:

```
> BinvA <- Binv %*% A
> BinvA
      [,1] [,2] [,3] [,4]
[1,]  -2  -25  11  -19
[2,]  -4  -31  16  -25
[3,]   3   16  -6   13
[4,]  -2  -14   7  -11
```

Požadované súradnice  $\mathbf{x}_b$  dostaneme nasledovne:

```
> xa <- c(1, 0, 4, 6)
> xb <- BinvA %*% xa
> xb
      [,1]
[1,]  -72
[2,]  -90
[3,]   57
[4,]  -40
```

Súradnice vektora v báze  $\mathbf{b}$  sú teda  $\mathbf{x}_b = (-72, -90, 57, -40)$ . Overme ešte, či si obidva vektory súradníc naozaj zodpovedajú. Vyjadrime preto vektory  $\mathbf{x}_a$  a  $\mathbf{x}_b$  pomocou kanonickej bázy:

```
> A %*% xa
      [,1]
[1,]   24
[2,]   -4
[3,]   17
[4,]   32
> B %*% xb
      [,1]
[1,]   24
[2,]   -4
[3,]   17
[4,]   32
```

V predchádzajúcej časti sme videli, že vyjadrenie vektora vo vektorovom priestore  $V$  nie je jedinečné v tom zmysle, že ten istý vektor môžeme vyjadriť pomocou rôznych súradníc v závislosti od bázy, ktorú si zvolíme. Ak máme vektor vyjadrený ako lineárnu kombináciu (súradnice) v jednej báze, existuje matica transformácie, ktorou môžeme ľubovoľný vektor v danej báze vyjadriť ako lineárnu kombináciu (súradnice) v inej báze toho istého vektorového priestoru. Formálne by sme túto transformáciu zmeny bázy mohli nazvať zobrazením z vektorového priestoru  $V$  do toho istého vektorového priestoru  $V$ . Už sme si

ukázali, že k takejto transformácii existuje aj inverzná transformácia – zobrazenie je teda injektívne. Navyše, keďže v oboch prípadoch ide o vyjadrenie pomocou vektorov báz, tie z definície generujú celý vektorový priestor  $V$  a zobrazenie je preto aj surjektívne. Keďže sú splnené obidve podmienky, môžeme prehlásiť, že zmena bázy je bijektívnym zobrazením.

Predstavu zobrazenia medzi vektorovými priestormi je možné zovšeobecniť. Zaujímaj nás budú tzv. lineárne zobrazenia. Uvažujme o dvoch vektorových priestoroch  $U$  a  $V$ . **Lineárnym zobrazením** nazveme každé zobrazenie  $\varphi: V \rightarrow U$ , pre ktoré pre všetky  $\mathbf{x}, \mathbf{y} \in V$  a  $a, b \in \mathbb{R}$  platí:

$$\varphi(a\mathbf{x} + b\mathbf{y}) = a\varphi(\mathbf{x}) + b\varphi(\mathbf{y}) \quad (2.80)$$

Je možné ukázať, že zmena bázy tak, ako sme si ju popísali v predchádzajúcej časti pomocou definovanej matice transformácie je lineárnym zobrazením. Označme tak ako doposiaľ  $\mathbf{x}_a$  súradnice vektora v báze  $\mathbf{a}$  a  $\mathbf{x}$  jeho súradnice v kanonickej báze. Vieme, že ak z vektorov bázy  $\mathbf{a}$  zostavíme maticu  $\mathbf{A}$ , ktorej stĺpce zodpovedajú vektorom bázy, potom platí:

$$\mathbf{x} = \mathbf{A}\mathbf{x}_a \quad (2.81)$$

Týmto predpisom ale môžeme definovať zobrazenie  $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^n$ , ktoré každému  $\mathbf{x}_a$  patriacemu do  $\mathbb{R}^n$  priradí práve jedno  $\mathbf{x}$  patriace do  $\mathbb{R}^n$ .

Potom platí:

$$\varphi(a\mathbf{x} + b\mathbf{y}) = \mathbf{A}(a\mathbf{x} + b\mathbf{y}) \quad (2.82)$$

$$= \mathbf{A}a\mathbf{x} + \mathbf{A}b\mathbf{y} \quad (2.83)$$

$$= a\mathbf{A}\mathbf{x} + b\mathbf{A}\mathbf{y} \quad (2.84)$$

$$= a\varphi(\mathbf{x}) + b\varphi(\mathbf{y}) \quad (2.85)$$

Nami definované zobrazenie  $\varphi$ , zodpovedajúce zmene bázy, je preto lineárnym zobrazením. Každé lineárne zobrazenie, ktoré je zároveň bijektívne, nazývame aj **lineárnym izomorfizmom**. Platí, že každý lineárny izomorfizmus je lineárnym zobrazením, nie každé lineárne zobrazenie je však lineárnym izomorfizmom.

### Príklad 2.30

Báza  $\mathbf{a}$  vektorového priestoru  $\mathbb{R}^5$  je tvorená vektormi  $\mathbf{a}_1 = (1, 1, 2, 1, 2)$ ,  $\mathbf{a}_2 = (0, 0, 1, 0, 2)$ ,  $\mathbf{a}_3 = (2, 1, 1, 1, 0)$ ,  $\mathbf{a}_4 = (1, 0, 0, 1, 0)$  a  $\mathbf{a}_5 = (2, 1, 1, 0, 1)$ . Definujme lineárny izomorfizmus  $\varphi: \mathbb{R}^5 \rightarrow \mathbb{R}^5$  ako zobrazenie  $\mathbf{x} = \mathbf{A}\mathbf{x}_a$  priradujúce každému vektoru  $\mathbf{x}_a$  súradníc v báze  $\mathbf{a}$  jeho súradnice v kanonickej báze  $\mathbf{x}$ .

Na príklade vektorov  $\mathbf{y}_a = (1, 2, 3, 4, 5)$ ,  $\mathbf{z}_a = (2, 0, -1, 1, -1)$ ,  $\mathbf{s} = (1, 0, 1, 0, 1)$  a  $\mathbf{t} = (2, 2, 2, 2, 2)$ . Ukážme vlastnosti lineárneho izomorfizmu  $\varphi$ .

**Riešenie:**

Najprv si definujeme maticu **A** ako maticu vektorov bázy **a**, ktoré sú usporiadané do stĺpcov:

```
> A <- matrix(c(1, 0, 2, 1, 2, 1, 0, 1, 0, 1, 2, 1, 1, 0, 1, 1,
  0, 1, 1, 0, 2, 2, 0, 0, 1), nrow=5, byrow=T)
> A
      [,1] [,2] [,3] [,4] [,5]
[1,]    1    0    2    1    2
[2,]    1    0    1    0    1
[3,]    2    1    1    0    1
[4,]    1    0    1    1    0
[5,]    2    2    0    0    1
```

Matica **A** je maticou transformácie, ktorou prevádzame vektory súradníc v báze **a** do súradníc kanonickej bázy. Definujme ďalej vektory **y<sub>a</sub>** a **z<sub>a</sub>**.

```
> ya <- c(1, 2, 3, 4, 5)
> ya
[1] 1 2 3 4 5
> za <- c(2, 0, -1, 1, -1)
> za
[1] 2 0 -1 1 -1
```

Ďalej si vypočítajme vektory  $\varphi(\mathbf{y}_a)$  a  $\varphi(\mathbf{z}_a)$ :

```
> phi_ya <- A %**% ya
> t(phi_ya)
      [,1] [,2] [,3] [,4] [,5]
[1,]   21    9   12    8   11
> phi_za <- A %**% za
> t(phi_za)
      [,1] [,2] [,3] [,4] [,5]
[1,]   -1    0    2    2    3
```

Pri výpise, resp. zobrazení výsledkov sme využili funkciu `t()`, ktorá umožňuje transponovanie výsledkov. Urobili sme to len z dôvodov úspory miesta, keďže výsledkom maticových súčinov sú stĺpcové vektory (matice rozmeru  $5 \times 1$ ). Skúsme teraz previesť zo súradníc v báze **a** do kanonickej bázy vektor  $2\mathbf{y}_a$ , spočítajme teda  $\varphi(2\mathbf{y}_a)$ .

```
> t(A %**% (2*ya))
      [,1] [,2] [,3] [,4] [,5]
[1,]   42   18   24   16   22
> t(2*phi_ya)
```

```

      [,1] [,2] [,3] [,4] [,5]
[1,]   42   18   24   16   22

```

Vidíme, že platí  $\varphi(2\mathbf{y}_a) = 2\varphi(\mathbf{y}_a)$ . Spočítajme ďalej  $\varphi(\mathbf{y}_a + \mathbf{z}_a)$ .

```

> t(A %*% (ya+za))
      [,1] [,2] [,3] [,4] [,5]
[1,]   20    9   14   10   14
> t(phi_ya+phi_za)
      [,1] [,2] [,3] [,4] [,5]
[1,]   20    9   14   10   14

```

Opäť vidíme, že platí  $\varphi(\mathbf{y}_a + \mathbf{z}_a) = \varphi(\mathbf{y}_a) + \varphi(\mathbf{z}_a)$ . Predchádzajúce operácie môžeme aj spojiť, vyhodnoťme teda výraz  $\varphi(2\mathbf{y}_a + 3\mathbf{z}_a)$ .

```

> t(A %*% (2*ya + 3*za))
      [,1] [,2] [,3] [,4] [,5]
[1,]   39   18   30   22   31
> t(2*phi_ya + 3*phi_za)
      [,1] [,2] [,3] [,4] [,5]
[1,]   39   18   30   22   31

```

Vidíme, že platí  $\varphi(2\mathbf{y}_a + 3\mathbf{z}_a) = 2\varphi(\mathbf{y}_a) + 3\varphi(\mathbf{z}_a)$ . Povedali sme si, že  $\varphi$  je lineárnym izomorfizmom. Existuje preto aj zobrazenie  $\varphi^{-1}$ , ktoré je tiež lineárnym izomorfizmom. Pre ukážku využime dva vektory v kanonickej báze, dané  $\mathbf{s} = (1, 0, 1, 0, 1)$  a  $\mathbf{t} = (2, 2, 2, 2, 2)$ . Budeme potrebovať inverznú maticu  $\mathbf{A}^{-1}$  k matici  $\mathbf{A}$ , ktorou prevedieme súradnice v kanonickej báze na súradnice v báze  $\mathbf{a}$ .

```

> Ainv <- solve(A)
> Ainv
      [,1] [,2] [,3] [,4] [,5]
[1,]   -1    3   -2    1    1
[2,]    1   -4    3   -1   -1
[3,]    1   -4    4   -1   -2
[4,]    0    1   -2    1    1
[5,]    0    2   -2    0    1
> s <- c(1, 0, 1, 0, 1)
> t <- c(2, 2, 2, 2, 2)
> s
[1] 1 0 1 0 1
> t
[1] 2 2 2 2 2

```

Určíme  $\varphi^{-1}(\mathbf{s})$ ,  $\varphi^{-1}(\mathbf{t})$  a  $\varphi^{-1}(3\mathbf{s} - \mathbf{t})$ .

```

> t(Ainv %*% s)
      [,1] [,2] [,3] [,4] [,5]
[1,]   -2    3    3   -1   -1
> t(Ainv %*% t)
      [,1] [,2] [,3] [,4] [,5]
[1,]     4   -4   -4    2    2
> t( 3*(Ainv %*% s) - (Ainv %*% t) )
      [,1] [,2] [,3] [,4] [,5]
[1,]  -10   13   13   -5   -5
> t( Ainv %*% (3*s-t) )
      [,1] [,2] [,3] [,4] [,5]
[1,]  -10   13   13   -5   -5

```

V tomto príklade je vidieť, že  $\varphi$  a  $\varphi^{-1}$  naozaj vykazujú vlastnosti požadované na lineárne izomorfizmy.

Vo všeobecnosti je možné, podobne ako v prípade zmeny bázy, nájsť pre každé lineárne zobrazenie tzv. **maticu lineárnej transformácie**  $\mathbf{A}_\varphi$  tak, že pre každé  $\mathbf{x} \in V$  platí:

$$\varphi(\mathbf{x}) = \mathbf{A}_\varphi \mathbf{x} \quad (2.86)$$

Celé zobrazenie môžeme popísať maticou lineárnej transformácie.

Pomocou lineárnych zobrazení je možné popísať veľa geometrických transformácií: otočenie, zmenu mierky, preklopenie (súmernosť) a skosenie. Jednu transformáciu však pomocou lineárnych zobrazení urobiť nedokážeme – posunutie. Zobrazenie  $f: V \rightarrow U$  v tvare

$$f(\mathbf{x}) = \varphi(\mathbf{x}) + \mathbf{b} = \mathbf{A}_\varphi \mathbf{x} + \mathbf{b} \quad (2.87)$$

priradujúce každému  $\mathbf{x} \in V$  prvok  $f(\mathbf{x}) \in U$ , kde  $\varphi$  je lineárne zobrazenie a  $\mathbf{b} \in V$  nazývame **afinné zobrazenie**. Vidíme, že afinné zobrazenie sa od lineárneho líši len pripočítaním vektora  $\mathbf{b}$ . Práve táto zmena však umožňuje definovať posunutie (transláciu). Inak povedané, lineárne zobrazenie je špeciálnym prípadom afinného zobrazenia pre  $\mathbf{b} = \mathbf{0}$ .

## 2.6 Skalárny súčin, Euklidovský priestor, norma vektora a ortogonalita

V kapitole 1.3.3 sme definovali operáciu skalárneho súčinu vektorov. Ak by sme túto operáciu pridali ku konečnorozmerným vektorovým priestorom, ktoré sme skúmali v predchádzajúcich častiach, dostávame tzv. **Euklidovský priestor**. Pripomeňme, že skalárny súčin vektorov  $\mathbf{x}$  a  $\mathbf{y} \in \mathbb{R}^n$  sme definovali nasledovne:

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^T \mathbf{y} = \mathbf{y}^T \mathbf{x} \quad (2.88)$$

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^n x_i y_i \quad (2.89)$$

Z definície je zrejmé, že skalárny súčin vektorov predstavuje operáciu, ktorej výsledkom je reálne číslo. Zo vzťahu (2.89) vyplýva, že skalárny súčin je komutatívny, keďže  $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$ . Uvedme jeho ďalšie vlastnosti:

$$\mathbf{x} \cdot (\mathbf{y} + \mathbf{z}) = \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \cdot \mathbf{z} \quad (2.90)$$

$$(a\mathbf{x}) \cdot (b\mathbf{y}) = (ab) (\mathbf{x} \cdot \mathbf{y}) \quad (2.91)$$

$$\mathbf{x} \cdot \mathbf{x} \geq 0 \quad (2.92)$$

Vlastnosť (2.90) vyjadruje distributívnosť skalárneho súčinu vzhľadom na operáciu sčítania. Ak si ju spojíme s komutatívnosťou skalárneho súčinu, tak je zrejmé, že musí byť distributívny tak zľava, ako aj sprava, čiže:

$$(\mathbf{y} + \mathbf{z}) \cdot \mathbf{x} = \mathbf{y} \cdot \mathbf{x} + \mathbf{z} \cdot \mathbf{x} \quad (2.93)$$

Vlastnosť (2.91) zasa hovorí, že násobenie vektorov skalárom môžeme nahradiť vynásobením skalárneho súčinu príslušnými číslami, t. j. „vybrať konštanty pred skalárny súčin“.

Nakoniec máme vlastnosť (2.92), ktorá hovorí o nezápornosti skalárneho súčinu vektora so sebou samým. Táto vlastnosť je zrejmá, ak si nerovnicu zapíšeme nasledovne:

$$\mathbf{x} \cdot \mathbf{x} = \mathbf{x}^T \mathbf{x} = \sum_{i=1}^n x_i x_i = \sum_{i=1}^n x_i^2 \geq 0 \quad (2.94)$$

Keďže v tomto prípade je skalárny súčin rovný súčtu štvorcov (druhých mocnín) zložiek vektora  $\mathbf{x}$ , tieto sú vždy nezáporné, a nezáporný bude aj ich súčet. Zároveň v dôsledku tejto úvahy vieme povedať ešte jednu vec: skalárny súčin  $\mathbf{x} \cdot \mathbf{x}$  bude nulový práve vtedy, ak je aj vektor  $\mathbf{x}$  nulový. Ak by mal vektor  $\mathbf{x}$  aspoň jednu nenulovú zložku, výraz  $\mathbf{x}^T \mathbf{x}$  by bol kladný. Platí teda výrok:

$$\mathbf{x} \cdot \mathbf{x} = 0 \Leftrightarrow \mathbf{x} = \mathbf{0} \quad (2.95)$$

Vďaka tejto vlastnosti je možné pomocou skalárneho súčinu definovať **dĺžku vektora**, nazývanú aj normou vektora, označovanú  $\|\mathbf{x}\|$ :

$$\|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}} \quad (2.96)$$

Z vlastností skalárneho súčinu dostávame nasledovné vlastnosti dĺžky vektora:

$$\|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0} \quad (2.97)$$

$$\|c\mathbf{x}\| = |c| \|\mathbf{x}\|$$

$$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$$

Posledná vlastnosť je nazývaná aj **trojuholníkovou nerovnosťou**.

V geometrii má skalárny súčin veľký význam aj preto, že je pomocou neho možné definovať uhol medzi dvoma vektormi  $\theta \in \langle 0, \pi \rangle$ . Platí totiž:

$$\mathbf{x} \cdot \mathbf{y} = \|\mathbf{x}\| \|\mathbf{y}\| \cos \theta \quad (2.98)$$

Práve tento fakt môžeme využiť pri definícii ortogonalít vektorov. Z trigonometrie vieme, že kosínus uhla  $\pi/2$ , resp.  $90^\circ$  je rovný nule. Pri tomto uhle by však boli vektory na seba kolmé. Z tohto dôvodu o dvoch vektoroch  $\mathbf{x}$  a  $\mathbf{y} \in \mathbb{R}^n$  hovoríme, že sú na seba kolmé, resp. že sú vzájomne **ortogonálne** práve vtedy ak:

$$\mathbf{x} \cdot \mathbf{y} = 0 \quad (2.99)$$

Skutočnosť, že vektory  $\mathbf{x}$  a  $\mathbf{y}$  sú vzájomne ortogonálne označujeme nasledovne  $\mathbf{x} \perp \mathbf{y}$ . Nulový vektor  $\mathbf{0}$  je kolmý na všetky vektory. Platí, že ak je vektor  $\mathbf{x}$  kolmý na vektory  $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n$  kde  $n \in \mathbb{N}$ , potom je kolmý aj na ľubovoľnú ich lineárnu kombináciu s koeficientmi  $a_1, a_2, \dots, a_n \in \mathbb{R}$  t. j.:

$$\mathbf{x} \perp (a_1\mathbf{y}_1 + a_2\mathbf{y}_2 + \dots + a_n\mathbf{y}_n) \quad (2.100)$$

Usporiadanú  $m$ -ticu vektorov  $\mathbf{a} = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m)$  nazveme **ortogonálnou**, ak sú každé dva rôzne vektory  $\mathbf{a}_i$  a  $\mathbf{a}_j$  navzájom ortogonálne:

$$\mathbf{a}_i \cdot \mathbf{a}_j = 0, \quad i \neq j; \quad i, j \in \{1, 2, \dots, m\} \quad (2.101)$$

Ak vektory  $\mathbf{a}$  spĺňajú navyše podmienku, že dĺžka každého z vektorov je rovná jednej, teda:

$$\|\mathbf{a}_i\| = 1 \quad (2.102)$$

pre  $i \in \{1, 2, \dots, m\}$ , potom túto  $m$ -ticu nazveme **ortonormálnou**. Tieto pojmy môžeme používať aj vo vzťahu k bázam vektorových priestorov z predchádzajúcej časti. Môžeme tak hovoriť o **ortogonálnych**, resp. **ortonormálnych bázach** vektorového priestoru.

Ortonormálne bázy majú oproti bežným bázam vektorových priestorov niekoľko zaujímavých vlastností. Ak  $\mathbf{a} = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$  je ortonormálna báza vektorového priestoru  $V$ , potom súradnice ľubovoľného vektora  $\mathbf{x}$  môžeme vyjadriť v báze  $\mathbf{a}$  nasledovne:

$$\mathbf{x}_a = (\mathbf{x} \cdot \mathbf{a}_1, \mathbf{x} \cdot \mathbf{a}_2, \dots, \mathbf{x} \cdot \mathbf{a}_n) \quad (2.103)$$

Je potrebné si uvedomiť, že  $\mathbf{x} \cdot \mathbf{a}_i$  pre  $i = 1, 2, \dots, n$  v predchádzajúcom vzťahu predstavuje skalárny súčin, ktorého výsledkom je reálne číslo. Ak by sme definovali  $c_i = \mathbf{x} \cdot \mathbf{a}_i$ , potom predchádzajúci vzťah je ekvivalentný s nasledovným tvrdením:

$$\mathbf{x} = c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \dots + c_n\mathbf{a}_n \quad (2.104)$$

Vektor  $\mathbf{x}$  je teda vyjadrený ako lineárna kombinácia vektorov bázy  $\mathbf{a}$ , tak ako doposiaľ. To, čo je v tomto prípade zaujímavé je, že príslušné koeficienty lineárnej kombinácie môžeme vypočítať pomocou skalárnych súčinov  $\mathbf{x} \cdot \mathbf{a}_i$ .

Keďže táto vlastnosť nie je na prvý pohľad zrejmá, budeme sa jej vysvetleniu venovať trochu obširnejšie. V prvom rade si uvedomme, že keďže vektory tvoriace  $\mathbf{a}$  predstavujú bázu vektorového priestoru  $V$ , vektor  $\mathbf{x}$  môžeme určite vyjadriť v tvare:

$$\mathbf{x} = c_1 \mathbf{a}_1 + c_2 \mathbf{a}_2 + \dots + c_n \mathbf{a}_n = \sum_{i=1}^n c_i \mathbf{a}_i \quad (2.105)$$

Uvedené vyplýva z toho, že  $\mathbf{a}$  je báza, a ako taká generuje  $V$ . Zamyslime sa ďalej nad skalárnymi súčini  $\mathbf{x} \cdot \mathbf{a}_i$ . Tie môžeme zapísať v tvare:

$$\mathbf{x} \cdot \mathbf{a}_i = \left( \sum_{j=1}^n c_j \mathbf{a}_j \right) \cdot \mathbf{a}_i \quad (2.106)$$

$$= \sum_{j=1}^n [(c_j \mathbf{a}_j) \cdot \mathbf{a}_i] \quad (2.107)$$

$$= \sum_{j=1}^n [c_j (\mathbf{a}_j \cdot \mathbf{a}_i)] \quad (2.108)$$

$$= c_i (\mathbf{a}_i \cdot \mathbf{a}_i) \quad (2.109)$$

$$= c_i \quad (2.110)$$

Vo vzťahu (2.106) sme využili to, že vektor  $\mathbf{x}$  môžeme zapísať pomocou vektorov bázy  $\mathbf{a}$ . Vzťah (2.107) vyplýva z distributívnosti skalárneho súčinu (skalárny súčin lineárnej kombinácie vektorov je lineárna kombinácia skalárnych súčinov). V rovnici (2.108) sme využili skutočnosť, že násobenie skalárom môžeme vybrať pred skalárny súčin. Najdôležitejšie kroky však predstavujú posledné dve rovnice. V rovnici (2.109) sme využili skutočnosť, že  $\mathbf{a}$  je ortonormálna báza. V dôsledku toho sú vektory bázy navzájom ortogonálne, a ich skalárny súčin je nulový. Z tohto dôvodu sú všetky súčiny  $\mathbf{a}_j \cdot \mathbf{a}_i$  nulové, s jednou výnimkou v prípade, ak  $i = j$ . Vtedy máme  $\mathbf{a}_j \cdot \mathbf{a}_i = \mathbf{a}_i \cdot \mathbf{a}_i$ . Vieme, že  $\mathbf{a}_i \cdot \mathbf{a}_i = \|\mathbf{a}_i\|^2$ . Avšak o  $\mathbf{a}$  sme predpokladali, že je nielen ortogonálnou, ale aj ortonormálnou bázou – všetky vektory, ktoré ju tvoria, majú preto jednotkovú dĺžku. Potom  $\mathbf{a}_i \cdot \mathbf{a}_i = \|\mathbf{a}_i\|^2 = 1$ . Túto skutočnosť využijeme vo vzťahu (2.110).

Predchádzajúca úvaha dokazuje, že:

$$\mathbf{x}_a = (\mathbf{x} \cdot \mathbf{a}_1, \mathbf{x} \cdot \mathbf{a}_2, \dots, \mathbf{x} \cdot \mathbf{a}_n) \quad (2.111)$$

Keďže ale vieme, že:

$$\mathbf{x} = c_1 \mathbf{a}_1 + c_2 \mathbf{a}_2 + \cdots + c_n \mathbf{a}_n = \sum_{i=1}^n c_i \mathbf{a}_i \quad (2.112)$$

a  $c_i = \mathbf{x} \cdot \mathbf{a}_i$ , potom:

$$\mathbf{x} = \sum_{i=1}^n c_i \mathbf{a}_i = \sum_{i=1}^n (\mathbf{x} \cdot \mathbf{a}_i) \mathbf{a}_i \quad (2.113)$$

## 2.7 Ortogonálna projekcia

V predchádzajúcom texte sme sa zaoberali konečnorozmernými vektorovými priestormi  $V$ , v rámci ktorých sme najčastejšie pracovali s vektorovými priestormi  $\mathbb{R}^n$ . Uvažujme teraz o nejakej jeho podmnožine  $U \subset V$ . V niektorých aplikáciách je zaujímavé skúmať množinu všetkých vektorov, ktoré sú ortogonálne vzhľadom na všetky vektory tvoriace  $U$ . Formálne ide o množinu, ktorú označíme  $U^\perp$ , a definujeme ju nasledovne:

$$U^\perp = \{\mathbf{x} \in V; \forall \mathbf{y} \in U : \mathbf{x} \perp \mathbf{y}\} \quad (2.114)$$

Je možné dokázať, že množina  $U^\perp$  je vektorovým podpriestorom vektorového priestoru  $V$  (a to bez ohľadu na to, či  $U$  bolo vektorovým priestorom). Nie je vždy pravdou, že doplnok ortogonálneho doplnku je pôvodná množina, teda nie je vždy pravda, že:

$$(U^\perp)^\perp = U \quad (2.115)$$

Najjednoduchším argumentom pre toto tvrdenie je aj fakt, že aj keď sme od  $U$  nepožadovali, aby jeho prvky tvorili vektorový priestor,  $(U^\perp)^\perp$  vektorovým priestorom určite je. Na druhej strane však platí nasledujúce tvrdenie:

$$\left( (U^\perp)^\perp \right)^\perp = U^\perp \quad (2.116)$$

Ďalšou zaujímavou vlastnosťou je, že ortogonálny komplement k množine  $U$  je ten istý vektorový priestor, ako ortogonálny komplement k vektorovému priestoru  $[U]$ , a je tým istým vektorovým priestorom, ktorý generujú vektory  $U^\perp$ . Formálne platí:

$$U^\perp = [U]^\perp = [U^\perp] \quad (2.117)$$

Pre ľubovoľný vektor  $\mathbf{u} \in U$  a ľubovoľný vektor  $\mathbf{z} \in U^\perp$  platí, že:

$$\mathbf{u} \cdot \mathbf{z} = 0 \quad (2.118)$$

Uvažujme teraz nejakú lineárnu kombináciu vektorov z  $U$  v tvare:

$$\mathbf{v} = \sum_{i=1}^m c_i \mathbf{u}_i \quad (2.119)$$

Zrejme  $\mathbf{v} \in [U]$ , keďže je vyjadrený ako lineárna kombinácia vektorov patriacich do  $U$ . Pre ľubovoľný vektor  $\mathbf{z} \in U^\perp$  určite platí:

$$\mathbf{v} \cdot \mathbf{z} = \left( \sum_{i=1}^m c_i \mathbf{u}_i \right) \cdot \mathbf{z} \quad (2.120)$$

$$= \sum_{i=1}^m [(c_i \mathbf{u}_i) \cdot \mathbf{z}] \quad (2.121)$$

$$= \sum_{i=1}^m c_i (\mathbf{u}_i \cdot \mathbf{z}) \quad (2.122)$$

$$= \sum_{i=1}^m c_i \cdot 0 \quad (2.123)$$

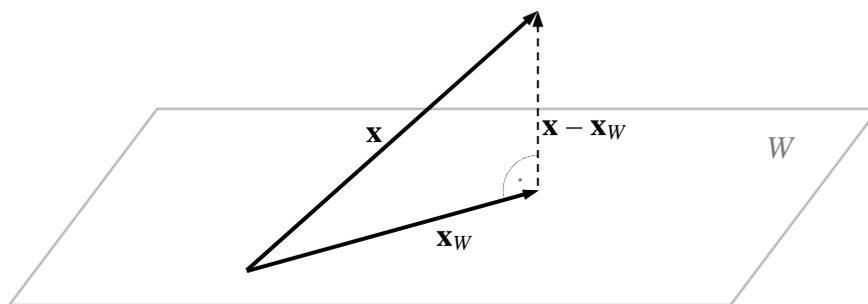
$$= 0 \quad (2.124)$$

Vzťah (2.121) a (2.122) dostávame z vlastností skalárneho súčinu, ktoré sme charakterizovali vyššie pri jeho definícii. Vzťah (2.123) dostávame z ortogonalít  $\mathbf{u}_i$  a  $\mathbf{z}$ , keďže  $\mathbf{u}_i \in U$  a  $\mathbf{z} \in U^\perp$ . Z rovnosti  $\mathbf{v} \cdot \mathbf{z} = 0$  vidíme, že vektory  $\mathbf{z} \in U^\perp$ , ktoré sú kolmé na prvky  $U$  sú kolmé aj na všetky ich lineárne kombinácie patriace do  $[U]$ . Teda  $U^\perp \subseteq [U]^\perp$ . Analogicky je možné ukázať aj opačné tvrdenie – ak by sme mali vektor  $\mathbf{z} \in [U]^\perp$ , teda vektor, ktorý je kolmý na všetky lineárne kombinácie vektorov z  $U$ , potom je určite kolmý aj na samotné vektory tvoriace  $U$ , čo sa dá zapísať ako  $[U]^\perp \subseteq U^\perp$ . Keďže  $U^\perp \subseteq [U]^\perp$  a súčasne  $[U]^\perp \subseteq U^\perp$ , platí  $[U]^\perp = U^\perp$ . Ortogonálny doplnok k množine  $U$  a k vektorovému priestoru  $[U]$  je ten istý vektorový priestor  $U^\perp$ .

Definícia ortogonálnych doplnkov nám umožňuje definovať tzv. **ortogonálnu projekciu** vektora  $\mathbf{x} \in V$  do jeho podpriestoru  $W \in V$ . Ortogonálnu projekciu nazývame aj kolmým priemetom. Hovoríme, že vektor  $\mathbf{x}_W$  je ortogonálnou projekciou vektora  $\mathbf{x}$  do podpriestoru  $W$  práve vtedy, ak:

$$\mathbf{x} - \mathbf{x}_W \in W^\perp \quad (2.125)$$

V prípade, ak by vektor  $\mathbf{x}$  patril do vektorového podpriestoru  $W$ , potom by bol totožný so svojou ortogonálnou projekciou do  $W$  a platilo by  $\mathbf{x} = \mathbf{x}_W$ . Ak  $\mathbf{x}$  nepatrí do  $W$ , potom je vektor  $\mathbf{x}_W$  vektorom ležiacim vo  $W$ , ktorý je najbližší k vektoru  $\mathbf{x}$ . Inak povedané,  $\mathbf{x}_W$  je vektor patriaci do  $W$ , pre ktorý je vzdialenosť  $\|\mathbf{x} - \mathbf{x}_W\|$  najmenšia. Dá sa ukázať, že vektor  $\mathbf{x}_W$  je určený jednoznačne – existuje práve jeden vektor  $\mathbf{x}_W \in W$  spĺňajúci požadovanú podmienku. V tomto zmysle môžeme vzdialenosť  $\|\mathbf{x} - \mathbf{x}_W\|$  chápať ako vzdialenosť vektora  $\mathbf{x}$  od podpriestoru  $W$ .



Obrázok 4: Ortogonálna projekcia

*Zdroj: vlastné spracovanie*

## 3 Základy klasickej ekonometrie

Ekonometria predstavuje časť ekonómie, ktorá sa zaoberá meraním a modelovaním ekonomických veličín. V skutočnosti predstavuje prienik troch vedných oblastí: ekonómie, matematiky a štatistiky. Každá z týchto oblastí má svoje vlastné opodstatnenie a úlohu. Ekonomická teória nám dáva motiváciu skúmať reálne procesy vyskytujúce sa v praxi a často deduktívne predpovedá charakter závislostí a premenné, ktorých vzťahy sa snažíme modelovať. Matematika umožňuje formálny zápis modelu a poskytuje metódy (spravidla optimalizačné), ktorými realizujeme odhad. Výsledky z teórie pravdepodobnosti a štatistiky využívame na zovšeobecnenie výsledkov v prípade, ak model kvantifikujeme na vzorke. Štatistika tiež umožňuje testovanie hypotéz o parametroch modelu a tým aj o vzťahoch medzi premennými.

### 3.1 Stručná rekapitulácia poznatkov z teórie pravdepodobnosti

Ako sme spomínali v úvode tejto kapitoly, ekonometria je založená do značnej miery aj na výsledkoch matematickej štatistiky. Kľúčovú úlohu v nej zohráva teória pravdepodobnosti: na jej základoch je vybudovaná celá inferenčná štatistika, umožňujúca usudzovať o vlastnostiach populácie na základe analýzy vzorky. V tejto kapitole uvedieme niektoré kľúčové poznatky, ktoré sú využívané v kapitolách venovaných ekonometrii.

#### 3.1.1 Pravdepodobnostný priestor

Formálne sa často vo vzťahu k pravdepodobnostiam hovorí o tzv. **pravdepodobnostnom priestore**  $(\Omega, \mathcal{F}, P)$ , ktorý sa skladá z množiny elementárnych javov  $\Omega$ , množiny javov  $\mathcal{F}$  a pravdepodobnostnej miery  $P$ .

**Množina elementárnych javov**  $\Omega$  predstavuje akúkoľvek neprázdnu množinu. Prvky tejto množiny predstavujú elementárne javy. Sú to výsledky náhodného pokusu, ktorý skúmame a sú už ďalej nedeliteľné – predstavujú teda najjednoduchšie výsledky, ktoré pozorujeme.

**Množina javov**  $\mathcal{F}$  predstavuje množinu podmnožín  $\Omega$ , splňajúcu určité vlastnosti. Ak by sme označili  $\mathcal{P}(\Omega)$  potenčnú množinu množiny  $\Omega$ , teda množinu jej všetkých podmnožín, potom  $\mathcal{F} \subseteq \mathcal{P}(\Omega)$ .

Ako klasický príklad si môžeme uviesť situáciu, v ktorej hádzeme hracou kockou. Predstavme si situáciu, ak hodíme hracou kockou a padlo číslo tri. V tomto prípade ak označíme výsledok náhodného pokusu premennou  $x$ , potom  $x = 3$  a  $x \in \mathbb{R}$ .

Čo však v prípade, ak by sme postupovali inak – označme  $X$  výsledok, ktorý dostaneme ak kocku hodíme – kocku sme ale ešte nehodili, keďže ju držíme v ruke. V takomto prípade  $X$  nie je reálne číslo. Je zrejmé, že  $X$  by mohlo byť 1, 2, 3, 4, 5 alebo 6, ale ktoré z nich to naozaj bude, to nevieme, kým kocku nehodíme. Skúsme túto predstavu, ako aj charakteristiku doposiaľ len veľmi všeobecne definovaného objektu  $X$  formalizovať konštrukciou zodpovedajúceho pravdepodobnostného priestoru.

Ak chceme popísať pravdepodobnostný priestor  $(\Omega, \mathcal{F}, P)$ , musíme v prvom rade definovať množinu elementárnych javov  $\Omega$ . Pomerne prirodzenou voľbou by bolo definovať  $\Omega = \{1, 2, 3, 4, 5, 6\}$ . V skutočnosti je teória pravdepodobnosti vybudovaná veľmi elegantne, a pri voľbe prvkov  $\Omega$  nezáleží na tom, o aké objekty ide – vôbec to nemusia byť čísla. Definujme si preto množinu elementárnych javov takto  $\Omega = \{\square, \square, \square, \square, \square, \square\}$ .

Ďalej si potrebujeme definovať množinu javov  $\mathcal{F}$ . V prípade konečných množín elementárnych javov je často postačujúce pracovať s ich celou potenčnou množinou, potom  $\mathcal{F} = \mathcal{P}(\Omega)$ . Pripomeňme, že vzhľadom na mohutnosť množiny  $\Omega$  ( $|\Omega| = 6$ ) bude potenčná množina obsahovať  $2^6 = 64$  prvkov. Budú v nej zastúpené okrem prázdnej aj všetky jednoprvkové, dvojprvkové, ako aj ďalšie podmnožiny  $\Omega$ . Podmnožinám  $\Omega$  budeme hovoriť javy. Takýmto javom je aj napríklad  $\{\square, \square, \square\}$ . Pri interpretácii javov je treba zohľadniť to, že v skutočnosti sme kocku ešte nehodili – jav  $\{\square, \square, \square\}$  nie je možné interpretovať tak, že padla dvojka, štvorka a šestka (teda ako výsledok troch hodov kocky), aj keď nás k tomu môže zápis zvädzať. Jav  $\{\square, \square, \square\}$  interpretujeme tak, že by padla dvojka, štvorka *alebo* šestka, inak povedané, že by padol párný počet bodov. Podobne jav  $\{\square, \square, \square\}$  zodpovedá tomu, že by padol počet bodov väčší ako tri.

Poslednou zložkou pravdepodobnostného priestoru  $(\Omega, \mathcal{F}, P)$  je **pravdepodobnostná miera**  $P$ . Ide o zobrazenie  $P: \mathcal{F} \rightarrow \langle 0,1 \rangle$ , teda o zobrazenie z množiny javov do intervalu reálnych čísel medzi 0 a 1 (vrátane). Nie každé takéto zobrazenie je pravdepodobnostnou mierou. Aby tomu tak bolo,  $P$  musí spĺňať tri vlastnosti:

- $P(\emptyset) = 0$
- $P(\Omega) = 1$

- Pre každú spočítateľnú a po dvojiciach dizjunktnú množinu prvkov  $\mathcal{F}$ , označenú  $F_i$  pre  $i \in I, I \subseteq \mathbb{N}$ , platí:

$$P\left(\bigcup_{i \in I} F_i\right) = \sum_{i \in I} P(F_i) \quad (3.1)$$

Pravdepodobná miera nám umožňuje skúmať pravdepodobnosti, s akými môžu nastať rôzne výsledky skúmaného náhodného pokusu. Prvá z vlastností hovorí, že pravdepodobnosť nemožného javu ( $\emptyset$ ) je nulová. V našom prípade si pod nemožným javom predstavíme to, že na kocke, ktorú sme hodili a na ktorej padla nejaká strana, nič nepadne. Keďže takýto výsledok je nemožný, hovoríme príslušnému javu nemožný jav.

Druhá vlastnosť hovorí, že  $P(\Omega) = 1$ . Pripomeňme, že  $P(\Omega) = P(\{\square, \square, \square, \square, \square, \square\})$ . Takýto jav ale znamená, že padne jednotka, dvojka, trojka, štvorka, päťka alebo šesťka. Keďže iné možnosti (ako napr. študentmi veľmi obľúbená odpoveď, že kocka padne na hranu) nastať nemôžu (inak by tiež boli obsiahnuté v  $\Omega$ ), ide o istý jav, ktorý po hode kockou s istotou nastane. Istý jav má pravdepodobnosť 1.

Tretia vlastnosť, nazývaná aj aditívnosť požaduje, aby sa pravdepodobnosť javu zloženého z nanajvyš spočítateľne veľa navzájom vylučujúcich sa javov mohli vypočítať ako súčet ich jednotlivých pravdepodobností. Podľa tretej vlastnosti napríklad platí:

$$P(\{\square, \square, \square\}) = P(\{\square\}) + P(\{\square\}) + P(\{\square\}) \quad (3.2)$$

čo by sme slovne popísali ako „pravdepodobnosť toho, že na kocke padne párný počet bodov je rovná súčtu pravdepodobností, že na nej padne dvojka, pravdepodobnosti, že na nej padne štvorka a pravdepodobnosti, že na nej padne šesťka“.

Všimnime si, že pravdepodobnostná miera  $P$  je definovaná na množine javov  $\mathcal{F}$ , a nie na množine elementárnych javov  $\Omega$ . Týmto požadujeme, aby sme všetkým javom, ktoré nás zaujímajú (tvoria  $\mathcal{F}$ ) vedeli priradiť pravdepodobnosť.

Na prvý pohľad nie je jasné, prečo by sme mali v prípade pravdepodobnostnej miery rozlišovať medzi  $\Omega$  a  $\mathcal{F}$ , resp. prečo by sme  $\mathcal{F}$  nedefinovali ako všetky možné javy, ktoré vieme vytvoriť z prvkov  $\Omega$ . Ak by sme napríklad pri našej kocke definovali pravdepodobnosť pre každý prvok  $\Omega$ , teda  $P(\{\square\}), P(\{\square\}), \dots, P(\{\square\})$ , potom by sme pomocou týchto pravdepodobností spolu s aditívnosťou pravdepodobnostnej miery vedeli vypočítať pravdepodobnosť akéhokoľvek javu patriaceho do  $\mathcal{F} = \mathcal{P}(\Omega)$ .

V prípade pravdepodobnostných priestorov, v ktorých je  $\Omega$  konečnou množinou žiadne ťažkosti nevznikajú a často volíme  $\mathcal{F} = \mathcal{P}(\Omega)$  tak, ako v nami skúmanom prípade. Pri takejto voľbe obsahuje  $\mathcal{F}$  všetky mysliteľné javy, ktoré vieme získať z elementárnych.

Určité problémy vznikajú v prípade, ak  $\Omega$  nie je konečná. V takomto prípade môže byť  $\mathcal{P}(\Omega)$  veľmi zložitou množinou. Preto sa nepožaduje, aby  $\mathcal{F}$  bolo potenčnou množinou  $\Omega$ , stačí aby tvorilo tzv.  $\sigma$ -algebru. Tou bude  $\mathcal{F}$  vtedy, ak platí:

- $\emptyset \in \mathcal{F}$
- $A \in \mathcal{F} \Rightarrow A_{\Omega}^c \in \mathcal{F}$
- Pre každú spočítateľnú množinu prvkov  $\mathcal{F}$ , označenú  $F_i$  pre  $i \in I, I \subseteq \mathbb{N}$ , platí:

$$\bigcup_{i \in I} F_i \in \mathcal{F} \quad (3.3)$$

Prvá vlastnosť  $\sigma$ -algebry  $\mathcal{F}$  znamená, že  $\mathcal{F}$  by vždy mala obsahovať nemožný jav  $\emptyset$ .

Druhá vlastnosť hovorí, že  $\mathcal{F}$  je uzavretá na doplnky, teda že ak  $A$  patrí do  $\mathcal{F}$ , potom aj doplnok  $A$  v  $\Omega$  ( $A_{\Omega}^c$ ) patrí do  $\mathcal{F}$ . Jav  $A_{\Omega}^c$  nazývame aj javom opačným k javu  $A$ . Uvedomme si, že ak v našom prípade s kockou  $A = \{\square, \square\}$ , potom  $A_{\Omega}^c = \{\square, \square, \square, \square\}$ . Druhá vlastnosť teda požaduje, aby sme ku každému javu mali v  $\mathcal{F}$  aj k nemu opačný jav.

Posledná vlastnosť  $\sigma$ -algebry  $\mathcal{F}$  hovorí, že ak  $F_i$  patria do  $\mathcal{F}$ , potom aj ich zjednotenie patrí do  $\mathcal{F}$ . Posledná vlastnosť vedie k dôsledku, že ak pri našej kocke  $\{\square\} \in \mathcal{F}, \{\square\} \in \mathcal{F}, \dots, \{\square\} \in \mathcal{F}$ , potom nutne  $\mathcal{F} = \mathcal{P}(\Omega)$ .

Všetky vlastnosti  $\sigma$ -algebry  $\mathcal{F}$  sú pomerne jednoduché a zdajú sa byť prirodzené. Práca s nimi, ako aj definícia pravdepodobnostnej miery sa komplikuje, ak  $\Omega$  nie je konečná (alebo spočítateľná) množina. Napríklad pre  $\Omega = \mathbb{R}$  sa za  $\mathcal{F}$  často volí tzv. Borelovská  $\sigma$ -algebra založená na otvorených intervaloch v  $\mathbb{R}$ , ktorej vlastnosti však presahujú rozsah tejto publikácie. Podrobnejšiu diskusiu je možné nájsť v iných publikáciách (Riečan et al., 1984, Neubrunn – Riečan, 1992). Skonštatujme len, že množina javov  $\mathcal{F}$  nie je vždy rovná  $\mathcal{P}(\Omega)$ , ale určite je jej podmnožinou.

### 3.1.2 Náhodná premenná

V predchádzajúcom texte sme popísali definíciu pravdepodobnostného priestoru  $(\Omega, \mathcal{F}, P)$ . Spôsob, akým je definovaný, je pomerne intuitívny a zároveň veľmi elegantný. Pôsobivé je napríklad to, ako jeho definícia nezávisí od skutočného charakteru prvkov  $\Omega$ ,

ktorými nemusia byť číselné objekty. Je však pravdou, že v matematike sa s číslami pracuje lepšie – existuje rozsiahla teória zasahujúca do mnohých oblastí matematiky, ktorá vychádza z práce s číslami (tu nemáme na mysli len aritmetiku, ale napr. aj matematickú analýzu). Z tohto dôvodu môže byť výhodné okrem prvkov množiny  $\Omega$  definovať aj nejaké jej kódovanie pomocou čísel, čím si uľahčíme ďalšiu prácu. Toto kódovanie by malo každému elementárnemu javu, a teda každému prvku  $\Omega$  priradiť nejaké reálne číslo. Táto predchádzajúca veta už naznačuje, že by malo byť možné zostrojiť zobrazenie  $X: \Omega \rightarrow \mathbb{R}$ . Takéto zobrazenie nazývame **náhodnou premennou**.

Ak máme definovaný priestor, tak na základe pravdepodobnosti, ktorá tvorí jeho súčasť, môžeme definovať pravdepodobnosť toho, že náhodná premenná nadobudne nejaké zvolené hodnoty. Povedzme napríklad, že by sme chceli stanoviť pravdepodobnosť toho, že náhodná premenná  $X$  nadobudne hodnotu menšiu alebo rovnú 8. Keďže  $X$  je zobrazenie do množiny reálnych čísel, je možné porovnávať jeho hodnoty s číslom 8 a na základe toho definovať množinu elementárných javov, pre ktoré nadobúda náhodná premenná  $X$  hodnoty menšie ako 8. Označme túto množinu  $X_8 = \{\omega \in \Omega; X(\omega) \leq 8\}$ . Pokiaľ  $X_8 \in \mathcal{F}$ , potom je určite definovaná aj pravdepodobnosť  $P(X_8) = P(\{\omega \in \Omega; X(\omega) \leq 8\})$ . Z dôvodu ľahšej čitateľnosti budeme v ďalšom texte namiesto podobných zápisov používať menej presné, ale jednoduchšie označenie  $P(X \leq 8)$ .

### Príklad 3.1

Definujme pravdepodobnostný priestor a náhodnú premennú  $X$  zodpovedajúcu jednému hodu férovou kockou, kde  $X$  bude nadobúdať hodnoty podľa počtu bodov na kocke.

**Riešenie:** Najprv definujme pravdepodobnostný priestor  $(\Omega, \mathcal{F}, P)$ . V rámci tohto pravdepodobnostného priestoru je množina elementárných javov  $\Omega = \{\square, \square, \square, \square, \square, \square\}$ . Množinu javov  $\mathcal{F}$  položíme rovnú potenčnej množine  $\Omega$ , takže  $\mathcal{F} = \mathcal{P}(\Omega)$ . Nakoniec si definujme pravdepodobnostnú mieru  $P$ , ktorá je zobrazením  $\mathcal{F} \rightarrow \langle 0,1 \rangle$ . Keďže v príklade ide o férovú kocku, každá jej strana by mala mať rovnakú pravdepodobnosť. Preto položíme:

$$P(\{\square\}) = P(\{\square\}) = P(\{\square\}) = P(\{\square\}) = P(\{\square\}) = P(\{\square\}) = 1/6$$

Z vlastnosti aditívnosti je pravdepodobnostná miera spĺňajúce vyššie definované rovnosti určená jednoznačne a je definovaná aj pre ostatné prvky  $\mathcal{F}$ .

Náhodná premenná je zobrazenie  $X: \Omega \rightarrow \mathbb{R}$ . Definujme preto náhodnú premennú  $X$  nasledovne:

$$X(\square) = 1, \quad X(\square) = 2, \quad X(\square) = 3, \quad X(\square) = 4, \quad X(\square) = 5, \quad X(\square) = 6$$

Je zrejmé, že definičný obor  $X$  je rovný  $\{\square, \square, \square, \square, \square, \square\} = \Omega$ . Obor hodnôt  $X$  je množina  $\{1,2,3,4,5,6\}$ .

### Príklad 3.2

Definujme pravdepodobnostný priestor a náhodnú premennú  $X$  zodpovedajúcu jednému hodu férovou kockou, kde  $X$  bude nadobúdať hodnoty 0, 1 a 2 podľa toho, aký zvyšok dostaneme po delení počtu bodov na kocke číslom tri. Následne vypočítajme  $P(X = 1)$  a  $P(X \leq 1)$ .

**Riešenie:** Pravdepodobnostný priestor môžeme definovať identicky ako v predchádzajúcom prípade, teda  $\Omega = \{\square, \square, \square, \square, \square, \square\}$ ,  $\mathcal{F} = \mathcal{P}(\Omega)$  a za pravdepodobnostnú mieru zvolíme mieru spĺňajúcu:

$$P(\{\square\}) = P(\{\square\}) = P(\{\square\}) = P(\{\square\}) = P(\{\square\}) = P(\{\square\}) = 1/6$$

Zmenu však urobíme pri definícii náhodnej premennej  $X: \Omega \rightarrow \mathbb{R}$ .

Na kocke môžu padnúť steny s 1, 2, 3, 4, 5 a 6 bodmi. Platí:

$$\begin{array}{lll} 1 = 0 \cdot 3 + 1 & 2 = 0 \cdot 3 + 2 & 3 = 1 \cdot 3 + 0 \\ 4 = 1 \cdot 3 + 1 & 5 = 1 \cdot 3 + 2 & 6 = 2 \cdot 3 + 0 \end{array}$$

Ak padnú počty bodov 1 alebo 4, dostávame zvyšok po delení tromi rovný jednej. Pre počet bodov 2 alebo 5 máme zvyšok dva. Nakoniec, ak padne 3 alebo 6 bodov, dostávame zvyšok rovný nule.

Náhodnú premennú  $X$  na základe tejto skutočnosti definujeme nasledovne:

$$\begin{array}{ll} X(\square) = 1, & X(\square) = 1, \\ X(\square) = 2, & X(\square) = 2, \\ X(\square) = 0, & X(\square) = 0 \end{array}$$

Tento príklad ukazuje, že náhodná premenná môže byť pre ten istý pravdepodobnostný priestor definovaná rôzne (porovnajme si to s predchádzajúcim príkladom). Definičný obor náhodnej premennej  $X$  je stále  $\Omega = \{\square, \square, \square, \square, \square, \square\}$ . Oborom hodnôt je však množina  $\{0, 1, 2\}$ .

Na tomto príklade si ešte ukážeme, ako je možné spočítať pravdepodobnosť súvisiacu s hodnotou náhodnej premennej. Pre  $P(X = 1)$  z definície dostávame:

$$\begin{aligned}
P(X = 1) &= P(\{\omega \in \Omega; X(\omega) = 1\}) \\
&= P(\{\square, \blacksquare\}) \\
&= P(\{\square\}) + P(\{\blacksquare\}) \\
&= \frac{1}{6} + \frac{1}{6} = \frac{1}{3}
\end{aligned}$$

V predposlednom riadku sme využili aditívnosť pravdepodobnostnej miery  $P$ . Podobne môžeme postupovať aj pri výpočte druhej požadovanej pravdepodobnosti,  $P(X \leq 1)$ .

$$\begin{aligned}
P(X \leq 1) &= P(\{\omega \in \Omega; X(\omega) \leq 1\}) \\
&= P(\{\square, \blacksquare, \square, \blacksquare\}) \\
&= P(\{\square\}) + P(\{\blacksquare\}) + P(\{\square\}) + P(\{\blacksquare\}) \\
&= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{2}{3}
\end{aligned}$$

Pravdepodobnosť definovaná v predchádzajúcom odseku má zásadný význam. Funkciu pravdepodobnosti vyjadrujúcej, že náhodná premenná nadobudne hodnotu menšiu alebo rovnú  $x_0 \in \mathbb{R}$  nazývame **distribučnou funkciou** a označujeme ju  $F_X(x_0)$ , prípadne  $F(x_0)$ , ak je zrejmé, o ktorú náhodnú premennú ide.

Distribučná funkcia je zobrazením  $F: \mathbb{R} \rightarrow \langle 0,1 \rangle$  pre ktorú platí:

$$F_X(x_0) = P(X \leq x_0) \quad (3.4)$$

Význam distribučnej funkcie je v tom, že je definovaná pre všetky náhodné premenné bez ohľadu na ich charakter. V skutočnosti by sme s jej pomocou mohli obísť jednu z komplikácií, s ktorou sa stretávame v teórii pravdepodobnosti – konkrétne, mohli by sme sa vyhnúť rozprave o rôznych typoch náhodných premenných.

V tejto publikácii sa budeme zaoberať dvoma druhmi náhodných premenných. Tým prvým sú tzv. **diskrétné náhodné premenné**. Ak náhodná premenná  $X$  nadobúda nanajvýš spočítateľne veľa rôznych hodnôt (inak povedané, obor hodnôt  $H(X)$  náhodnej premennej  $X$  je konečná, alebo spočítateľná množina), potom o náhodnej premennej  $X$  hovoríme, že je diskrétna. Z vlastností pravdepodobnosti potom pre istý jav platí:

$$P(X \in H(X)) = \sum_{x_0 \in H(X)} P(X = x_0) = 1 \quad (3.5)$$

V tomto prípade je možné definovať funkciu, ktorá každej hodnote náhodnej premennej  $x_0$  vráti jej pravdepodobnosť  $P(X = x_0)$  (v angličtine je táto funkcia nazývaná *probability mass function*, PMF).

Z predchádzajúcej definície je evidentné, že nebude použiteľná pre prípad, ak obor hodnôt náhodnej premennej  $X$  nebude spočítateľnou množinou – v takomto prípade súčet v rovnici (3.5) nemá zmysel. Na základe poznatkov z matematickej analýzy vieme, že v podobnej situácii – ak chceme „spočítať“ hodnoty nejakej funkcie na určitom intervale, ktorý je nespočítateľnou množinou – môžeme využiť integrálny súčet, predstavujúci analógiu sumy vo vzorci (3.5).

Ak pre náhodnú premennú  $X$  existuje nezáporná (Lebesguovská) integrovateľná funkcia  $f$ , nazývaná **hustota pravdepodobnosti** náhodnej premennej  $X$ , pre ktorú pre každé  $a, b \in \mathbb{R}$ ,  $a \leq b$  platí:

$$P(a \leq X \leq b) = \int_a^b f(x) dx \quad (3.6)$$

Potom náhodnú premennú  $X$  nazývame **spojitou náhodnou premennou**. Uvedme ešte, že aj keď hustota pravdepodobnosti  $f$  musí byť nezápornou funkciou, na rozdiel od pravdepodobnosti môže byť jej funkčná hodnota väčšia ako 1. Nanajvýš rovný jednej však musí byť integrál v rovnici (3.6). Samotnú hustotu  $f$  nie je možné interpretovať ako pravdepodobnosť a ani pravdepodobnosťou nie je. Pravdepodobnosťou je až integrál na nejakom intervale.

V skutočnosti je problém súvisiaci s popisáním rôznych náhodných premenných trochu zložitejší. Integrál v predchádzajúcej rovnici, ako aj pravdepodobnostnú mieru  $P$ , ktorú používame v prípade nespočítateľných množín, môžeme veľmi elegantne popísať pomocou teórie miery. S jej pomocou je dokonca možné pracovať s diskretnými a spojitými veličinami jednotne pomocou Lebesguovej, resp. Lebesguovej – Stjeltjesovej miery a integrálu. Táto (aj keď krásna) časť matematiky však leží ďaleko za možnosťami a cieľom tejto publikácie. V ďalšom texte preto budeme s rovnicou (3.6) pracovať ako s obyčajným Riemannovým integrálom (ak existuje, tak je hodnota Riemannovho integrálu aj tak rovná Lebesguovmu). Taktiež sa v ďalšom texte obmedzíme okrem diskretných náhodných premenných len na spojité náhodné premenné, ktoré majú hustotu pravdepodobnosti.

Z rovnice (3.6) je možné ukázať, že medzi distribučnou funkciou náhodnej premennej a hustotou pravdepodobnosti existuje vzťah, konkrétne:

$$F_X(x_0) = P(X \leq x_0) = \int_{-\infty}^{x_0} f(x) dx \quad (3.7)$$

Pre diskretné náhodné premenné platí:

$$F_X(x_0) = \sum_{\substack{x \leq x_0 \\ x \in H(X)}} P(X = x_0) \quad (3.8)$$

Vidíme, že distribučná funkcia existuje skutočne pre všetky náhodné premenné – diskkrétne, aj spojité. V našom výklade sa však líši spôsob, ako ju vypočítame, v závislosti od charakteru náhodnej premennej.

V niektorých prípadoch je potrebné hovoriť aj o tzv. podmienenej pravdepodobnosti. Uvažujme o pravdepodobnostnom priestore  $(\Omega, \mathcal{F}, P)$  a dvoch javoch  $A, B \in \mathcal{F}$ . **Podmienenou pravdepodobnosťou**  $P(A|B)$  nazývame pravdepodobnosť nastania javu  $A$  ak vieme, že nastal jav  $B$ . Pre podmienenú pravdepodobnosť platí:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad (3.9)$$

Pokiaľ sa pravdepodobnosť javu  $A$  nemení v závislosti od nastania javu  $B$ , potom hovoríme, že javy  $A$  a  $B$  sú **nezávislé**. Pre nezávislé javy platia nasledovné vzťahy:

$$P(A|B) = P(A) \quad (3.10)$$

$$P(A \cap B) = P(A) P(B) \quad (3.11)$$

### Príklad 3.3

Definujme pravdepodobnostný priestor a náhodnú premennú  $X$  zodpovedajúcu jednému hodu férovou kockou, kde  $X$  bude nadobúdať hodnoty podľa počtu bodov na kocke. Označme  $A$  jav, že na kocke padne párny počet bodov,  $B$  jav, že padne počet bodov menší ako päť a  $C$  jav, že padne počet bodov menší ako štyri. Vypočítajme  $P(A|B)$  a  $P(A|C)$ .

**Riešenie:** Pravdepodobnostný priestor  $(\Omega, \mathcal{F}, P)$  zodpovedajúci hodu kockou sme definovali už v predchádzajúcich príkladoch. V súlade s ním predstavujú javy  $A, B$  a  $C$  množiny:

$$A = \{ \square, \boxplus, \boxtimes \}$$

$$B = \{ \square, \boxplus, \boxtimes, \boxminus \}$$

$$C = \{ \square, \boxplus, \boxtimes \}$$

Pre výpočet  $P(A|B)$  potrebujeme podľa vzťahu (3.9) stanoviť  $P(A \cap B)$  a  $P(B)$ , pre  $P(A|C)$  zasa  $P(A \cap C)$  a  $P(C)$ . Keďže ide o férovú kocku, dostávame:

$$P(A) = P(\{ \square, \boxplus, \boxtimes \}) = P(\{ \square \}) + P(\{ \boxplus \}) + P(\{ \boxtimes \}) = \frac{1}{2}$$

$$P(B) = P(\{ \square, \boxplus, \boxtimes, \boxminus \}) = P(\{ \square \}) + P(\{ \boxplus \}) + P(\{ \boxtimes \}) + P(\{ \boxminus \}) = \frac{2}{3}$$

$$P(C) = P(\{ \square, \boxplus, \boxtimes \}) = P(\{ \square \}) + P(\{ \boxplus \}) + P(\{ \boxtimes \}) = \frac{1}{2}$$

$$P(A \cap B) = P(\{\square, \boxtimes\}) = P(\{\square\}) + P(\{\boxtimes\}) = \frac{1}{3}$$

$$P(A \cap C) = P(\{\square\}) = \frac{1}{6}$$

Odtiaľ podľa (3.9) dostávame:

$$P(A|B) = P(A \cap B) / P(B) = \frac{1}{2}$$

$$P(A|C) = P(A \cap C) / P(C) = \frac{1}{3}$$

Vidíme, že  $P(A|B) = P(A) = \frac{1}{2}$ . Javy  $A$  a  $B$  sú nezávislé. Keďže  $P(A|C) \neq P(A)$ , javy  $A$  a  $C$  nie sú nezávislé.

Z definície štatistickej nezávislosti vychádza tzv. Bayesova veta. Ak definíciu:

$$P(A | B) = \frac{P(A \cap B)}{P(B)} \quad (3.12)$$

napíšeme v tvare:

$$P(A | B)P(B) = P(A \cap B) = P(B | A)P(A) \quad (3.13)$$

tak z oboch rovníc dostávame tvrdenie **Bayesovej vety**:

$$P(A | B) = \frac{P(B | A)P(A)}{P(B)} \quad (3.14)$$

### 3.1.3 Stredná hodnota, kovariancia, disperzia a variančno-kovariančná matica

Náhodné premenné môžeme popisovať rôznym spôsobom. Veľmi užitočnou charakteristikou je takzvaná **stredná hodnota** náhodnej premennej, označovaná  $E(X)$  (z anglického *expected value*, t. j. očakávaná hodnota). Pre diskrétnu a spojitú náhodnú premennú je definovaná nasledovne:

$$E(X) = \sum_{\omega \in \Omega} X(\omega)P(\{\omega\}) \quad (3.15)$$

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx \quad (3.16)$$

Z rovnice (3.15) je vidieť, že stredná hodnota diskretnej náhodnej premennej predstavuje vážený priemer hodnôt náhodnej premennej, kde váhami je pravdepodobnosť nastania jednotlivých elementárnych javov. Pripomeňme, že vážený aritmetický priemer počítame podľa vzorca:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i n_i \quad (3.17)$$

kde  $x_i$  predstavuje hodnoty, ktorých priemer počítame,  $n_i \geq 0$ ,  $n_i \in \mathbb{R}$  a  $n = \sum_{i=1}^n n_i$ . Ak

je súčet váh rovný jednej, ako v našom prípade, keďže:

$$\sum_{\omega \in \Omega} P(\{\omega\}) = 1 \quad (3.18)$$

tak vidíme, že vzťah (3.15) naozaj predstavuje vážený priemer.

### Príklad 3.4

Definujme pravdepodobnostný priestor a náhodnú premennú  $X$  zodpovedajúcu jednému hodu férovou kockou, kde  $X$  bude nadobúdať hodnoty podľa počtu bodov na kocke. Aká je stredná hodnota  $X$  pri hode kockou?

**Riešenie:** Pravdepodobnostný priestor môžeme definovať identicky ako v predchádzajúcom prípade, teda  $\Omega = \{\square, \square, \square, \square, \square, \square\}$ ,  $\mathcal{F} = \mathcal{P}(\Omega)$  a za pravdepodobnostnú mieru zvolíme mieru spĺňajúcu  $P(\{\omega\}) = 1/6$  pre  $\omega \in \Omega$ .

Strednú hodnotu pri hode kockou je možné vypočítať z definície ako:

$$\begin{aligned} E(X) &= X(\square)P(\{\square\}) + X(\square)P(\{\square\}) + X(\square)P(\{\square\}) \\ &\quad + X(\square)P(\{\square\}) + X(\square)P(\{\square\}) + X(\square)P(\{\square\}) \\ &= 1 \cdot 1/6 + 2 \cdot 1/6 + 3 \cdot 1/6 + 4 \cdot 1/6 + 5 \cdot 1/6 + 6 \cdot 1/6 \\ &= 3.5 \end{aligned}$$

Dostávame tak na prvý pohľad prekvapivý výsledok – ak budeme hádzať kockou, tak „v priemere“ by malo padať číslo tri a pol. Paradoxne to znie preto, že také číslo pri hode kockou nastať nemôže.

V skutočnosti nám stredná hodnota dáva iný význam. Pokúsme sa na chvíľu uvažovať o nasledovnej hre. Povedzme, že by sme dostali úlohu napísať na kus papiera číslo ešte predtým, ako kockou hodíme. Označme toto číslo  $a \in \mathbb{R}$ . Mali by sme za úlohu zvoliť  $a$  tak, aby po jeho zapísaní a následnom hodení kockou bol absolútny rozdiel toho, čo padne na kocke a čísla  $a$  čo najmenší. Máme sa teda snažiť uhádnuť, čo na kocke padne.

Ak by sme zvolili  $a = 1$ , potom najväčšia chyba, ktorej by sme sa mohli dopustiť je 5. Ak by padlo číslo 6, naša chyba by bola  $|6 - a| = |6 - 1| = 5$ . Podobne by tomu bolo, aj keby sme zvolili číslo 6. Naša najhoršia chyba by mohla nastať ak by padlo číslo 1, lebo vtedy by sme znova dostali  $|1 - a| = |1 - 6| = 5$ . V prípade, ak by sme zvolili číslo  $a = 2$  alebo  $a = 5$ , naša chyba by nemohla byť väčšia ako 4. Ak by sme zvolili  $a = 3$  alebo  $a = 4$ ,

pomýlili by sme sa prinajhoršom o tri. Ak však dovolíme, aby  $a$  bolo neceločíselné, tak pri  $a = 3.5$  naša chyba nebude horšia ako 2.5.

Ak by sme teda pred vykonaním pokusu mali odhadnúť jeho výsledok, náš najlepší tip by bol z tohto pohľadu naozaj 3.5.

Stredná hodnota má niekoľko kľúčových vlastností, ktoré si vysvetlíme na príklade diskretnej náhodnej premennej. Tie isté vlastnosti je možné dokázať aj pre spojité náhodné premenné, ak nahradíme všetky súčty integrálmi a hodnoty pravdepodobnosti funkciou hustoty. Keďže pre integrálny súčet platia podobné pravidlá ako pre bežný súčet (napr. linearita integrálu), výsledné vlastnosti budú rovnaké.

Nech  $a \in \mathbb{R}$  a  $X, Y$  sú náhodné premenné. Pre strednú hodnotu platia nasledovné tri vlastnosti, ktoré označíme  $E_1$  až  $E_3$ :

$$E1) E(a) = a$$

$$E2) E(aX) = aE(X)$$

$$E3) E(X+Y) = E(X) + E(Y)$$

Preukážme teraz platnosť týchto vlastností pre diskretne náhodné premenné.

V prvom rade máme vlastnosť  $E1$ , ktorá hovorí, že stredná hodnota konštanty je tá istá konštanta. Z definície strednej hodnoty dostávame:

$$E(a) = \sum_{\omega \in \Omega} aP(\{\omega\}) = a \sum_{\omega \in \Omega} P(\{\omega\}) = a \cdot 1 = a \quad (3.19)$$

Vlastnosť  $E2$  hovorí, že ak náhodnú premennú  $X$  prenásobíme konštantou  $a$ , vznikne nová náhodná premenná  $aX$ , ktorej strednú hodnotu  $E(aX)$  nemusíme počítať z definície. Ak poznáme  $E(X)$ , stačí, ak túto hodnotu prenásobíme číslom  $a$ .

### Príklad 3.5

Definujme pravdepodobnostný priestor a náhodnú premennú  $X$  zodpovedajúcu jednému hodu férovou kockou, kde  $X$  bude nadobúdať hodnoty podľa počtu bodov na kocke. Povedzme, že budeme hrať hru, v ktorej výhercovi vyplatíme 1000-násobok počtu bodov, koľko padne na kocke. Aká je stredná hodnota pri hode kockou a stredná hodnota výhry?

**Riešenie:** Pravdepodobnostný priestor môžeme definovať identicky ako v predchádzajúcom prípade, teda  $\Omega = \{\square, \square, \square, \square, \square, \square\}$ ,  $\mathcal{F} = \mathcal{P}(\Omega)$  a za pravdepodobnostnú mieru zvolíme mieru spĺňajúcu  $P(\{\omega\}) = 1/6$  pre  $\omega \in \Omega$ .

Z predchádzajúceho príkladu vieme, že:

$$E(X) = 3.5$$

Ak by sme chceli vypočítať strednú hodnotu výhry z definície strednej hodnoty, počítali by sme to nasledovne:

$$\begin{aligned} E(1000X) &= 1000X(\square)P(\{\square\}) + 1000X(\square)P(\{\square\}) + 1000X(\square)P(\{\square\}) \\ &\quad + 1000X(\square)P(\{\square\}) + 1000X(\square)P(\{\square\}) + 1000X(\square)P(\{\square\}) \\ &= 1000 \cdot \frac{1}{6} + 2000 \cdot \frac{1}{6} + 3000 \cdot \frac{1}{6} + 4000 \cdot \frac{1}{6} + 5000 \cdot \frac{1}{6} + 6000 \cdot \frac{1}{6} \\ &= 3500 \end{aligned}$$

Podľa vlastnosti E2 vieme ten istý výsledok dostať aj rýchlejšie, keďže:

$$E(1000X) = 1000E(X) = 1000 \cdot 3.5 = 3500$$

Dokážme teraz vlastnosť E2. Platí:

$$E(aX) = \sum_{\omega \in \Omega} aX(\omega)P(\{\omega\}) = a \sum_{\omega \in \Omega} X(\omega)P(\{\omega\}) = aE(X) \quad (3.20)$$

Vlastnosť E3 hovorí o tom, ako je možné vyjadriť strednú hodnotu súčtu náhodných premenných – môžeme ju vypočítať ako súčet ich stredných hodnôt.

### Príklad 3.6

Definujme pravdepodobnostný priestor a náhodné premenné  $X, Y$  zodpovedajúce dvom hodom férovou kockou, kde  $X$  a  $Y$  bude nadobúdať hodnoty podľa počtu bodov na kocke. Aká je stredná hodnota počtu bodov, ktorý padne spolu pri dvoch hodoch?

**Riešenie:** Pravdepodobnostný priestor bude pre obidva hody rovnaký, keďže ide o tú istú kocku. Tak ako doteraz preto máme  $\Omega = \{\square, \square, \square, \square, \square, \square\}$ ,  $\mathcal{F} = \mathcal{P}(\Omega)$  a za pravdepodobnostnú mieru zvolíme mieru spĺňajúcu  $P(\{\omega\}) = \frac{1}{6}$  pre  $\omega \in \Omega$ .

Už vieme, že pre náhodné premenné  $X$  a  $Y$ :

$$E(X) = E(Y) = 3.5$$

Nás ale zaujíma stredná hodnota súčtu, čiže  $E(X + Y)$ . Ak by sme chceli túto strednú hodnotu vypočítať priamo, museli by sme vychádzať z iného pravdepodobnostného priestoru. V prvom rade by sa zmenila množina elementárnych javov. Keďže v tomto prípade je náhodným pokusom to, že dvakrát hodíme kockou, musí elementárny jav obsahovať výsledok oboch hodov. Počet elementárnych javov, a teda aj  $|\Omega| = 36$ .

$$\Omega_2 = \{\square\square, \square\square, \square\square, \square\square, \square\square, \square\square,$$

$\{\{1,1\}, \{1,2\}, \{1,3\}, \{1,4\}, \{1,5\}, \{1,6\}, \{2,1\}, \{2,2\}, \{2,3\}, \{2,4\}, \{2,5\}, \{2,6\}, \{3,1\}, \{3,2\}, \{3,3\}, \{3,4\}, \{3,5\}, \{3,6\}, \{4,1\}, \{4,2\}, \{4,3\}, \{4,4\}, \{4,5\}, \{4,6\}, \{5,1\}, \{5,2\}, \{5,3\}, \{5,4\}, \{5,5\}, \{5,6\}, \{6,1\}, \{6,2\}, \{6,3\}, \{6,4\}, \{6,5\}, \{6,6\}\}$

Aby sme dokončili definíciu pravdepodobnostného priestoru pre dva hody, zvolme  $\mathcal{F}_2 = \mathcal{P}(\Omega_2)$  a za pravdepodobnostnú mieru zvolme mieru spĺňajúcu pre  $\omega \in \Omega_2$ :

$$P(\{\omega\}) = \frac{1}{36}$$

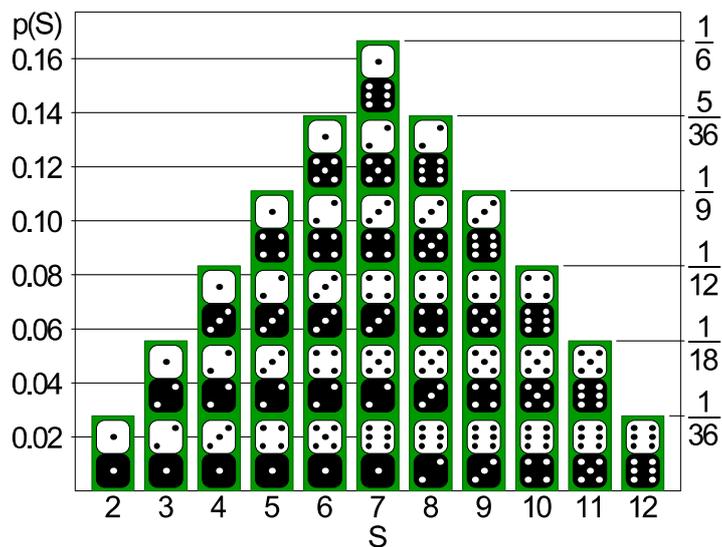
Náhodná premenná  $S = X + Y$  môže nadobúdať hodnoty medzi 2 a 12. Všimnime si, že  $S$  môže nadobudnúť hodnoty viac ako jednou kombináciou hodov, napríklad  $S = 8$  vieme získať piatimi rôznymi spôsobmi (

Obrázok 5):  $\{2,6\}, \{3,5\}, \{4,4\}, \{5,3\}$  a  $\{6,2\}$ . Pre  $E(S) = E(X + Y)$  preto dostávame:

$$E(S) = 2 \frac{1}{36} + 3 \frac{2}{36} + 4 \frac{3}{36} + 5 \frac{4}{36} + 6 \frac{5}{36} + 7 \frac{6}{36} + 8 \frac{5}{36} + 9 \frac{4}{36} + 10 \frac{3}{36} + 11 \frac{2}{36} + 12 \frac{1}{36} = 7$$

Ak by sme namiesto tohto relatívne náročného výpočtu využili vlastnosť E3, dostali by sme okamžite:

$$E(S) = E(X + Y) = E(X) + E(Y) = 3.5 + 3.5 = 7$$



Obrázok 5: Možné výsledky dvoch hodov kockou a ich pravdepodobnosti

Zdroj: Stellmach (2013)

Ďalšou často využívanou charakteristikou náhodných premenných je tzv. kovariancia. **Kovarianciu** medzi náhodnými premennými  $X$  a  $Y$ , označovanú  $\text{cov}(X,Y)$  definujeme pomocou strednej hodnoty nasledovne:

$$\text{cov}(X,Y) = E[(X - E(X))(Y - E(Y))] \quad (3.21)$$

Táto definícia, ktorá na prvý pohľad nepôsobí veľmi intuitívne, má v skutočnosti pomerne ľahké vysvetlenie (Tabuľka 3).

Tabuľka 3: Vzťah medzi hodnotami náhodných premenných a ich kovarianciou

$X - E(X)$	$Y - E(Y)$	$(X - E(X))(Y - E(Y))$
+	+	+
+	-	-
-	+	-
-	-	+

*Zdroj: vlastné spracovanie*

V tabuľke sú zobrazené rôzne možnosti hodnôt jednotlivých výrazov tvoriacich kovarianciu. Povedzme, že náhodná premenná  $X$  v nejakom konkrétnom pozorovaní dosahuje „nadpriemernú“ hodnotu. To by znamenalo, že  $X > E(X)$ , inak povedané  $X - E(X) > 0$ . Táto skutočnosť je v tabuľke vyznačená znakom „+“ v zelenom poli. Naopak podpriemerná hodnota je vyznačená znakom „-“ v červenom poli. Podobne sme vyznačili aj rôzne možnosti pre náhodnú premennú  $Y$  (vynechali sme prípad, ak náhodné premenné nadobúdajú hodnoty presne rovné svojim stredným hodnotám – tieto prípady našu ďalšiu úvahu neovplyvnia).

V poslednom stĺpci máme vyznačené znamienka pre výraz  $(X - E(X))(Y - E(Y))$ , ktorý tvorí jadro kovariance. Všimnime si, že tento výraz je kladný, len ak sa náhodné premenné  $X$  a  $Y$  „správajú rovnako“ v tom zmysle, že buď obidve nadobúdajú zhodne nadpriemerné, alebo zhodne podpriemerné hodnoty.

Naopak sledovaný výraz je záporný vtedy, ak sa  $X$  a  $Y$  „správajú opačne“, teda nadpriemerná hodnota jednej náhodnej premennej je sprevádzaná podpriemernou hodnotou druhej.

Kovarianciu medzi  $X$  a  $Y$  dostávame tak, že z uvedeného výrazu v poslednom stĺpci tabuľky vypočítame strednú hodnotu. Ak táto stredná hodnota, ktorá predstavuje určitý vážený priemer bude kladná, dá sa to s určitou mierou trivializácie chápať aj tak, že „plusové“ hodnoty v tabuľke prevážili „mínusové“, a teda je v tomto zmysle celkovo skôr pravdou, že

nadpriemerné hodnoty jednej náhodnej premennej sú sprevádzané nadpriemernými hodnotami druhej (a podpriemerné hodnoty jednej náhodnej premennej sú sprevádzané podpriemernými hodnotami druhej).

Kovariancia nám preto hovorí o vzťahu medzi náhodnými veličinami. Ide o odlišnú vlastnosť, ako je štatistická nezávislosť o ktorej sme hovorili v kapitole 3.1.2. V ďalšej časti budeme o premenných, pri ktorých môžeme hovoriť o existencii vzájomného vzťahu založenom na kovariancii hovoriť, že sú **korelované**. Ak by kovariancia bola nulová, hovoríme, že sú nekorelované. Ak sú náhodné premenné korelované, ide len o jednu špeciálnu formu vzájomnej závislosti. Existujú aj prípady, ak sú náhodné premenné nekorelované, no napriek tomu nie sú nezávislé – existuje pri nich závislosť, ktorá nie je identifikovateľná kovarianciou. Nezávislosť teda implikuje nekorelovanosť, ale nejde o ekvivalenciu.

Podobne ako pre strednú hodnotu, tak aj pre kovarianciu je užitočné demonštrovať niektoré jej vlastnosti. Označíme ich C1 až C4. Nech  $a \in \mathbb{R}$  a  $X, Y$  a  $Z$  sú náhodné premenné.

$$C1) \text{cov}(a, X) = 0$$

$$C2) \text{cov}(X, Y) = \text{cov}(Y, X)$$

$$C3) \text{cov}(aX, Y) = a \text{cov}(X, Y)$$

$$C4) \text{cov}(X + Y, Z) = \text{cov}(X, Z) + \text{cov}(Y, Z)$$

Prvá vlastnosť, C1, dáva intuitívne dobrý zmysel. Keďže kovariancia charakterizuje určitý vzťah medzi meniacimi sa hodnotami náhodných premenných, kovariancia medzi konštantou a náhodnou premennou by mala byť rovná nule – mali by byť nekorelované.

Preto nie je prekvapivé, že z definície kovariancie dostávame:

$$\text{cov}(a, X) = E[(a - E(a))(X - E(X))] \quad (3.22)$$

$$= E[(a - a)(X - E(X))] \quad (3.23)$$

$$= E[0(X - E(X))] \quad (3.24)$$

$$= 0E[(X - E(X))] \quad (3.25)$$

$$= 0 \quad (3.26)$$

V rovniciach (3.23) a (3.25) sme využili vlastnosť strednej hodnoty E2.

Vlastnosť C2 hovorí o symetrickosti kovariancie. Ak kovariancia charakterizuje vzťah medzi náhodnou premennou  $X$  a  $Y$ , tento vzťah by mal byť rovnaký ako vzťah premennej  $Y$  a  $X$  (v tomto poradí).

$$\text{cov}(X, Y) = E[(X - E(X))(Y - E(Y))] = E[(Y - E(Y))(X - E(X))] = \text{cov}(Y, X) \quad (3.27)$$

Podľa vlastnosti C3 je možné vypočítať kovarianciu medzi náhodnými premennými, z ktorých jednu vynásobíme konštantou  $a$  ako  $a$ -násobok kovariancie pôvodných náhodných premenných.

$$\text{cov}(aX, Y) = E[(aX - E(aX))(Y - E(Y))] \quad (3.28)$$

$$= E[(aX - aE(X))(Y - E(Y))] \quad (3.29)$$

$$= E[a(X - E(X))(Y - E(Y))] \quad (3.30)$$

$$= a E[(X - E(X))(Y - E(Y))] \quad (3.31)$$

$$= a \text{cov}(X, Y)$$

V rovniciach (3.29) a (3.31) sme opäť použili vlastnosť strednej hodnoty E2.

Posledná vlastnosť C4 hovorí, že kovarianciu súčtu náhodných premenných s treťou premennou je možné vypočítať ako súčet kovariancií. Táto vlastnosť má v ekonómii obzvlášť veľký význam v teórii portfólia.

$$\text{cov}(X + Y, Z) = E[(X + Y - E(X + Y))(Z - E(Z))] \quad (3.32)$$

$$= E[(X + Y - E(X) - E(Y))(Z - E(Z))] \quad (3.33)$$

$$= E[(X - E(X) + Y - E(Y))(Z - E(Z))] \quad (3.34)$$

$$= E[(X - E(X))(Z - E(Z)) + (Y - E(Y))(Z - E(Z))] \quad (3.35)$$

$$= E[(X - E(X))(Z - E(Z))] + E[(Y - E(Y))(Z - E(Z))] \quad (3.36)$$

$$= \text{cov}(X, Z) + \text{cov}(Y, Z) \quad (3.37)$$

V rovniciach (3.33) a (3.36) sme využili vlastnosť E3, v rovnici (3.35) sme len roznásobili ľavú vnútornú zátvorku výrazom  $(Z - E(Z))$ .

Poslednou charakteristikou, ktorú si popíšeme v tejto časti, je **disperzia** označovaná  $D(X)$ , nazývaná aj rozptyl náhodnej premennej. Na tomto mieste je potrebné zdôrazniť, že popisujeme charakteristiky náhodných premenných. Hlavne v prípade disperzie – rozptylu vzniká nebezpečenstvo zámieny s výberovými charakteristikami pri skúmaní štatistických vzoriek. Disperzia náhodnej premennej a rozptyl výberového súboru sú úplne rôzne kategórie a netreba ich zamieňať.

Rozptyl náhodnej premennej je daný vzťahom:

$$D(X) = E[(X - E(X))^2] \quad (3.38)$$

Tento vzťah je možné prečítať aj tak, že predstavuje „strednú hodnotu štvorcov odchýlok hodnôt náhodnej premennej od jej priemeru“. Výraz  $X - E(X)$  predstavuje odchýlku od priemeru. To naznačuje, že disperzia je mierou variability – hovorí nám niečo o tom, ako veľmi sa hodnoty  $X$  menia vzhľadom na  $E(X)$ .

Častou otázkou je, prečo sa v definícii používajú štvorce odchýlok, a nie len samotné odchýlky. V skutočnosti by sme dostali:

$$E[X - E(X)] = E(X) - E(X) = 0 \quad (3.39)$$

Tu vidno, prečo  $E(X)$  nazývame strednou hodnotou. Nachádza sa „v strede“ hodnôt  $X$  v tom zmysle, že kladné a záporné odchýlky od  $E(X)$  sa navzájom presne kompenzujú a takto definovaná priemerná odchýlka je vždy nula. Ak chceme charakterizovať variabilitu náhodnej premennej  $X$ , mali by sme takejto vzájomnej kompenzácie zabrániť. Jednou z možností je definovať priemernú absolútnu odchýlku nasledovne:

$$E[|X - E(X)|] \quad (3.40)$$

Vzhľadom na to, že s absolútnou hodnotou sa pracuje viac komplikovane a zároveň existuje určitý súvis s normálnym rozdelením pravdepodobnosti sa častejšie využíva ako miera variabilita práve  $D(X)$ .

Ak si pozorne pozrieme definíciu  $D(X)$ , tak si môžeme všimnúť, že:

$$D(X) = E[(X - E(X))^2] = E[(X - E(X))(X - E(X))] = \text{cov}(X, X) \quad (3.41)$$

Disperzia je teda to isté, ako kovariancia náhodnej premennej samej so sebou. Vďaka tejto skutočnosti je veľmi jednoduché odvodiť vlastnosti disperzie, keďže nám stačí využiť už dokázané vzťahy pre kovarianciu. Pre disperziu a  $a \in \mathbb{R}$  a náhodné premenné  $X$  a  $Y$  sa používajú nasledovné vzťahy:

$$D1) D(a) = 0$$

$$D2) D(aX) = a^2 D(X)$$

$$D3) D(X + Y) = D(X) + 2\text{cov}(X, Y) + D(Y)$$

Vlastnosť D1 hovorí, že konštanta má nulovú variabilitu, čo je intuitívne zřejmé.

$$D(a) = E[(a - E(a))^2] = E[(a - a)^2] = E[0] = 0 \quad (3.42)$$

Podľa vlastnosti D2 ak vynásobíme náhodnú premennú konštantou  $a$ , zmení sa jej rozptyl  $a^2$  násobne. Vyplýva to z nasledujúcich rovníc:

$$D(aX) = E[(aX - E(aX))^2] \quad (3.43)$$

$$= E[(aX - aE(X))^2] \quad (3.44)$$

$$= E[a^2(X - E(X))^2] \quad (3.45)$$

$$= a^2 E[(X - E(X))^2] \quad (3.46)$$

$$= a^2 D(X) \quad (3.47)$$

Posledná vlastnosť D3 hovorí, že rozptyl súčtu dvoch náhodných premenných nie je daný len súčtom ich rozptylov, ale aj kovarianciou. Je teda ovplyvnený aj vzťahom medzi náhodnými premennými. Prečo by tomu malo byť tak, je možné pochopiť na jednoduchšej úvahe. Povedzme, že by sme mali dve náhodné premenné, pričom medzi nimi bude presne opačný vzťah: ak prvá premenná zvýši svoju hodnotu o 8, druhá o 8 poklesne. Ak sa prvá zmení o -5, druhá o 5. Keďže sa obidve premenné menia, ich rozptyly  $D(X)$  a  $D(Y)$  budú nenulové. Ich súčtom je ale v našom prípade vždy konštanta – zmena jednej premennej presne kompenzuje zmenu druhej a súčet je stále rovnaký. Už z vlastnosti D1 vieme, že rozptyl konštanty je nula. Dostávame sa tak do zvláštnej situácie, kedy aj napriek tomu, že prvá a druhá premenná majú nenulovú variabilitu, ich súčet ju má nulovú. Táto koncepcia je základom tzv. Markowitzovej teórie portfólia. Aj keď ju nebudeme bližšie charakterizovať, je zrejmé, že pre disperziu súčtu náhodných premenných nutne musíme brať do úvahy aj vzťah, ktorý medzi nimi je.

Formálne si vlastnosť D3 môžeme ukázať nasledovne:

$$D(X+Y) = \text{cov}(X+Y, X+Y) \quad (3.48)$$

$$= \text{cov}(X, X+Y) + \text{cov}(Y, X+Y) \quad (3.49)$$

$$= \text{cov}(X+Y, X) + \text{cov}(X+Y, Y) \quad (3.50)$$

$$= \text{cov}(X, X) + \text{cov}(Y, X) + \text{cov}(X, Y) + \text{cov}(Y, Y) \quad (3.51)$$

$$= \text{cov}(X, X) + \text{cov}(X, Y) + \text{cov}(X, Y) + \text{cov}(Y, Y) \quad (3.52)$$

$$= D(X) + 2\text{cov}(X, Y) + D(Y) \quad (3.53)$$

V rovnici (3.48), ako aj dvakrát v rovnici (3.51) sme využili vlastnosť C4. V rovnici (3.49), ako aj v (3.52) sme zasa využili vlastnosť C2. Posledná rovnica vyplýva z definície disperzie.

Stredná hodnota náhodnej premennej, jej disperzia, ako aj kovariancia náhodných premenných predstavujú kľúčové charakteristiky, bez ktorých nie je možné vysvetliť teóriu ekonometrie – ekonometrických modelov a ich vlastností. Hore uvedených desať vlastností (E1 – E3, C1 – C4 a D1 – D3) sú pre pochopenie ďalších kapitol absolútne nevyhnutné.

V niektorých prípadoch je potrebné pracovať namiesto jednoduchých náhodných premenných s náhodnými vektormi. Ak sú  $X_1, X_2, \dots, X_n$  pre  $n \in \mathbb{N}$  náhodné premenné, potom náhodný vektor  $\mathbf{X}$  definujeme ako:

$$\mathbf{X} = (X_1, X_2, \dots, X_n)^T \quad (3.54)$$

Pod strednou hodnotou náhodného vektora  $\mathbf{X}$  budeme rozumieť vektor:

$$E(\mathbf{X}) = (E(X_1), E(X_2), \dots, E(X_n))^T \quad (3.55)$$

Môže sa zdať lákavé podobným spôsobom definovať aj vektor rozptylov. Namiesto neho však využívame tzv. **variančno – kovariančnú maticu** v tvare:

$$\text{var}(\mathbf{X}) = \begin{pmatrix} \text{cov}(X_1, X_1) & \text{cov}(X_1, X_2) & \cdots & \text{cov}(X_1, X_n) \\ \text{cov}(X_2, X_1) & \text{cov}(X_2, X_2) & \cdots & \text{cov}(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ \text{cov}(X_n, X_1) & \text{cov}(X_n, X_2) & \cdots & \text{cov}(X_n, X_n) \end{pmatrix} \quad (3.56)$$

Dôvod, prečo maticu  $\text{var}(\mathbf{X})$  nazývame variančno-kovariančná, a nie len kovariančná je ten, že chceme zvýrazniť význam prvkov na jej hlavnej diagonále. Keďže vieme, že platí  $\text{cov}(X, X) = D(X)$ , prvky na hlavnej diagonále matice  $\text{var}(\mathbf{X})$  sú v skutočnosti rozptyly:

$$\text{var}(\mathbf{X}) = \begin{pmatrix} D(X_1) & \text{cov}(X_1, X_2) & \cdots & \text{cov}(X_1, X_n) \\ \text{cov}(X_2, X_1) & D(X_2) & \cdots & \text{cov}(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ \text{cov}(X_n, X_1) & \text{cov}(X_n, X_2) & \cdots & D(X_n) \end{pmatrix} \quad (3.57)$$

Ešte iný spôsob, akým môžeme napísať variančno-kovariančnú maticu (3.56) je tento:

$$\text{var}(\mathbf{X}) = \begin{pmatrix} \text{cov}(X_1, X_1) & \text{cov}(X_1, X_2) & \cdots & \text{cov}(X_1, X_n) \\ \text{cov}(X_2, X_1) & \text{cov}(X_2, X_2) & \cdots & \text{cov}(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ \text{cov}(X_n, X_1) & \text{cov}(X_n, X_2) & \cdots & \text{cov}(X_n, X_n) \end{pmatrix} \quad (3.58)$$

$$= \begin{pmatrix} E((X_1 - E(X_1))(X_1 - E(X_1))) & \cdots & E((X_1 - E(X_1))(X_n - E(X_n))) \\ \vdots & \ddots & \vdots \\ E((X_n - E(X_n))(X_1 - E(X_1))) & \cdots & E((X_n - E(X_n))(X_n - E(X_n))) \end{pmatrix} \quad (3.59)$$

$$= E \begin{pmatrix} (X_1 - E(X_1))(X_1 - E(X_1)) & \cdots & (X_1 - E(X_1))(X_n - E(X_n)) \\ \vdots & \ddots & \vdots \\ (X_n - E(X_n))(X_1 - E(X_1)) & \cdots & (X_n - E(X_n))(X_n - E(X_n)) \end{pmatrix} \quad (3.60)$$

$$= E((\mathbf{X} - E(\mathbf{X}))(\mathbf{X} - E(\mathbf{X}))^T) \quad (3.61)$$

$$= E(\mathbf{X}\mathbf{X}^T) - E(\mathbf{X})E(\mathbf{X})^T \quad (3.62)$$

Vzťahy dostávame aplikáciou z definície kovariancie ako strednej hodnoty. Posledný vzťah dostávame z predposledného nasledovne:

$$\text{var}(\mathbf{X}) = E\left((\mathbf{X} - E(\mathbf{X}))(\mathbf{X} - E(\mathbf{X}))^T\right) \quad (3.63)$$

$$= E\left((\mathbf{X} - E(\mathbf{X}))(\mathbf{X}^T - E(\mathbf{X})^T)\right) \quad (3.64)$$

$$= E\left(\mathbf{X}\mathbf{X}^T - \mathbf{X}E(\mathbf{X})^T - E(\mathbf{X})\mathbf{X}^T + E(\mathbf{X})E(\mathbf{X})^T\right) \quad (3.65)$$

$$= E(\mathbf{X}\mathbf{X}^T) - E(\mathbf{X})E(\mathbf{X})^T - E(\mathbf{X})E(\mathbf{X})^T + E(\mathbf{X})E(\mathbf{X})^T \quad (3.66)$$

$$= E(\mathbf{X}\mathbf{X}^T) - E(\mathbf{X})E(\mathbf{X})^T \quad (3.67)$$

Využívame pritom, že  $E(\mathbf{X}^T) = E(\mathbf{X})^T$ .

Veľmi často sa vyskytuje situácia, ak pracujeme s vektormi alebo maticami, pre ktoré platí  $E(\mathbf{X}) = \mathbf{0}$ . V takomto prípade bude možné variančno-kovariančnú maticu jednoducho zapísať aj ako:

$$\text{var}(\mathbf{X}) = E(\mathbf{X}\mathbf{X}^T) \quad (3.68)$$

## 3.2 Formulovanie ekonometrického modelu

Ekonometria je časťou ekonómie, ktorá sa zaoberá meraním a modelovaním ekonomických veličín. Predstavuje prienik matematiky, štatistiky a ekonómie. Z matematiky ekonometria využíva najmä matematickú logiku a matematickú analýzu.

Hovorí sa, že matematika je jazykom vedy – to platí aj v prípade ekonometrie. Modely, ktoré sú v ekonometrii hlavným objektom záujmu, sú spravidla formulované pomocou matematických zápisov ako rovnice, resp. sústavy rovníc, ktoré charakterizujú vzťahy medzi skúmanými ekonomickými veličinami.

Pri skúmaní určitého ekonomického problému, napríklad inflácie označujeme modelovanú veličinu, ktorej hodnoty sa snažíme popísať ako **závislú (vysvetľovanú)** premennú. Veličiny, ktoré by mali súvisieť so závislou premennou a pomocou ktorých sa snažíme závislú premennú modelovať nazývame **nezávislé (vysvetľujúce)** premenné.

Na prvý pohľad sa zdá, že modelovanie ekonomických veličín bude náročnou úlohou, keďže v skutočnosti s akoukoľvek veličinou, napr. spomínanou infláciou, potenciálne súvisí veľmi veľká množina iných faktorov. Sú zrejmé dva problémy – prvým je úloha identifikovať všetky relevantné premenné, ktoré by mali byť v modeli využité, aby nedošlo k vynechaniu nejakej relevantnej veličiny. Druhým problémom je samotná kvantifikácia modelu.

V skutočnosti málokedy postupujeme tak, že by sme do modelov dosadzovali všetky premenné, ktoré s vysvetľovanou veličinou môžu súvisieť. Spravidla sa obmedzíme na veličiny, ktoré považujeme z hľadiska predmetu analýzy za dôležité. Vplyv všetkých faktorov, ktoré v modeli nevystupujú v explicitnom tvare ako vysvetľujúce premenné, agregujeme do tzv. **náhodného (poruchového)** člena. Náhodný člen v modeli predstavuje súhrn všetkých faktorov, ktoré v modeli explicitne neuvádzame. Predpokladáme pritom, že vzájomné interakcie týchto faktorov v konečnom dôsledku vedú k tomu, že ich súhrnný efekt sa zdá byť z hľadiska vplyvu na modelovanú veličinu náhodný. Pokiaľ tomu tak nie je, a náhodný člen v sebe obsahuje relevantné informácie z hľadiska modelovanej veličiny, zrejme je navrhovaný model **nesprávne špecifikovaný**.

Existencia náhodného člena je priamym dôsledkom abstrakcie pri ekonometrickom modelovaní. Keďže nie je prakticky možné zohľadňovať všetky mysliteľné faktory, práca s poruchovým členom ako náhodnou veličinou umožňuje zohľadniť premenné nezaraďené do modelu vo forme jeho charakteristík, akými sú napr. jeho rozptyl. Zahnutie stochastických (náhodných) veličín do modelu však so sebou prináša nové otázky. Ktorá časť modelu dominuje – vypovedacia schopnosť vysvetľujúcich premenných, alebo nepredvídateľný náhodný člen? Bolo by možné, že kvantifikovaný model a jeho výstupy sú nespoľahlivé, keďže v sebe obsahujú náhodný komponent? A taktiež je tu otázka zo všetkých najdôležitejšia: je možné výsledkom získaných pomocou ekonometrického modelovania danej veličiny veriť, sú spoľahlivé, alebo sú naopak len výsledkom náhody? Práve z dôvodu týchto otázok sa ekonometria obracia na štatistiku a matematickú štatistiku. V indukčnej štatistike sa totiž riešia problémy podobné tým, ktoré sme práve popísali. Snažíme sa v nej na základe skúmania vzorky popísať vlastnosti populácie. V tomto prípade rovnako potrebujeme rozhodnúť, či výsledky získané na vzorke nie sú výsledkom náhody (čo ak na inej vzorke dostaneme diametrálne odlišný výsledok? Náhodou je v tomto prípade to, ktoré prvky populácie sa dostali do vzorky). Prístupy využívané v ekonometrii preto majú silný súvis so štatistickými metódami.

Matematika teda slúži na zápis ekonometrických modelov, prácu s nimi, dokazovanie kľúčových vlastností a umožňuje odvodiť metódy odhadu špecifikovaného modelu. Štatistika slúži pri popise poruchového člena, na formulovanie indukčných záverov a testovaní signifikancie dosiahnutých výsledkov. Ani matematika, ani štatistika však nedefinuje problém, ktorého sa modelovanie bude týkať. Toto je úloha ekonómie. Ekonomické teórie sa dajú naformulovať spôsobom, ktorý umožňuje ich empirickú verifikáciu prostredníctvom

ekonometrických modelov. Matematika a štatistika v tomto duchu nepredstavujú účel, ale nástroje, ktoré sú využívané pre formulovanie, overovanie a rozširovanie ekonomickej teórie.

### **3.2.1 Typy premenných a charakter dát**

Z publikácií zaoberajúcich sa štatistikou je známe, že premenné, ktoré v analýze využívame, môžu mať rôzny charakter. Veľmi známou je klasifikácia podľa Stevensa (1946), ktorá delí premenné na nominálne, ordinálne, intervalové a podielové (angl. ratio scale). Pripomeňme, že nominálne (kategorické) údaje nie je možné navzájom porovnávať, príkladom je farba očí.

Ordinálne premenné umožňujú stanoviť poradie, napr. poradie bežcov v cieľi. Je zmysluplné hovoriť, že prvý bežec dobehol skôr ako tretí, ale na základe tohto faktu sa nedá povedať, že bol trikrát rýchlejší.

Intervalové premenné už umožňujú zmysluplne počítat' aj rozdiely. V ich prípade má zmysel počítat' priemery, rozptyly a pod. Klasicky uvádzaným príkladom pre intervalové premenné je teplota meraná v stupňoch Celzia. Pri tomto type premenných nie je zmysluplné počítat' podiely – nedá sa teda napríklad povedať, že niektorá hodnota je dvakrát väčšia ako iné. Pri teplote meranej stupňami Celzia je to napríklad preto, že nula stupňov neznamená absenciu teploty – ide len o jednu z možných definícií nulovej hodnoty, založenej na určitej fyzikálnej vlastnosti vody pri špecifických podmienkach (trojný bod vody VSMOW).

Premenné, ktoré majú definovanú aj absolútnu nulu, a tým umožňujú zmysluplne definovať aj podiely svojich hodnôt majú tzv. podielovú škálu. Príkladom je napríklad Kelvinova teplotná stupnica (definuje nulovú teplotu ako stav, v ktorom majú častice nulovú kinetickú energiu, odhliadnuc od kvantových efektov).

Charakter premenných, ktoré využívame v ekonometrickom modelovaní do značnej miery vymedzuje aj metódy, ktoré na ich analýzu môžeme použiť. Je zrejmé, že metódy, ktoré predpokladajú minimálne intervalovú škálu merania nie je vhodné používať na ordinálnych údajoch. Toto tvrdenie je spravidla jasné a väčšina čitateľov s ním bude zrejme súhlasiť. V skutočnosti je ale veľmi ľahko možné urobiť chybu, a používať nesprávne metódy pre údaje na nevhodnej škále merania. Hlavným dôvodom, prečo v tomto smere dochádza k chybám je rozsah, v akom sa v súčasnosti využíva výpočtová technika. Pri využívaní rôzneho software, umožňujúceho štatistickú a ekonometrickú analýzu dochádza ku kódovaniu sledovaných premenných pomocou číselných hodnôt. Takto môžeme v počítači zaznamenať

napr. farbu očí spôsobom, kde by sme hodnotu 1 priradili hnedým, hodnotu 2 modrým, 3 zeleným a 4 ostatným farbám. Počítačovému programu nič nebráni spočítať aj na takýchto údajoch aritmetický priemer, čo nás môže priviesť k výsledku, že priemerná farba očí je 2,65. Identifikovanie toho, že ide o nezmyselnú operáciu, je na tom, kto takúto analýzu vykonáva – software takúto chybu nedokáže rozpoznať. V ekonometrii existujú špeciálne metódy, ktoré sú určené pre konkrétne škály merania, napr. pre ordinálne premenné existujú rôzne varianty modelov LOGIT a PROBIT.

Okrem škály merania nás pri ekonometrickom modelovaní nutne zaujíma aj charakter dát. Delíme ich na prierezové, časové rady a panelové.

Prierezové údaje predstavujú súbory údajov, ktoré sú získavané v tom istom čase na skupine rôznych subjektov. Príkladom môžu byť aktuálne stavy zamestnancov vo zvolenej skupine podnikov. Jednoduchý model založený na prierezových údajoch môžeme zapísať v tvare:

$$y_i = \beta_0 + \beta_1 x_i + u_i \quad (3.69)$$

kde  $y_i$  sú hodnoty závislej premennej,  $x_i$  hodnoty nezávislej premennej,  $u_i$  náhodný poruchový člen,  $i = 1, 2, \dots, N$  identifikuje pozorovanie (subjekt),  $N \in \mathbb{N}$  je počet pozorovaní a  $\beta_0, \beta_1 \in \mathbb{R}$  sú koeficienty modelu.

Časové rady dostávame, ak sledujeme hodnoty zvoleného ukazovateľa v rôznych obdobiach, ale len u toho istého subjektu. Príkladom môže byť mesačný vývoj tržieb v konkrétnom podniku. Jeden z možných modelov by mohol mať tvar:

$$y_t = \beta_0 + \beta_1 y_{t-1} + u_t \quad (3.70)$$

kde  $y_t$  sú hodnoty závislej premennej v čase  $t$ ,  $y_{t-1}$  hodnoty závislej premennej v čase  $t-1$ , teda v predchádzajúcom období,  $u_t$  náhodný poruchový člen,  $t = 2, 3, \dots, T$  predstavuje čas,  $T \in \mathbb{N}$  je počet pozorovaní a  $\beta_0, \beta_1 \in \mathbb{R}$  sú koeficienty modelu. Tento model sa síce funkčným tvarom podobá na predchádzajúci, založený na prierezových údajoch, ale je zjavné, že pri ňom modelujeme iný typ dát.

Panelové údaje zahŕňajú oba predchádzajúce aspekty – pri nich máme k dispozícii časové rady za tie isté obdobia za určitú skupinu subjektov. Príkladom by mohli byť údaje o HDP, získané za členské krajiny Európskej únie za roky 1996 – 2003. Typický panelový model je v tvare:

$$y_{it} = \beta_0 + \beta_1 x_{it} + u_{it} \quad (3.71)$$

kde  $y_{it}$  sú hodnoty závislej premennej za prierezovú jednotku  $i$  v čase  $t$ ,  $x_{it}$  hodnoty nezávislej premennej za prierezovú jednotku  $i$  v čase  $t$ ,  $u_{it}$  je náhodný poruchový člen,

$i = 1, 2, \dots, N$  a  $t = 1, 2, \dots, T$ , kde  $N \in \mathbb{N}$  predstavuje počet prierezových jednotiek a  $T \in \mathbb{N}$  je počet časových období.  $\beta_0, \beta_1 \in \mathbb{R}$  sú koeficienty modelu.

### 3.2.2 Jednoduché lineárne vzťahy medzi premennými

Najjednoduchšie ekonometrické modely sa zaoberajú vzťahmi, ktoré je možné popísať pomocou lineárnych funkcií. Vysvetlíme si preto na úvod najjednoduchšie prípady, ktoré je možné zobrazit' v rovine.

Povedzme, že sa zaoberáme vzťahom medzi dvomi premennými, ktoré označíme  $x$  a  $y$ . Aj z učiva na stredných školách je známe, že ak by sa dala závislosť medzi  $x$  a  $y$  vyjadriť ako lineárna funkcia, ktorej grafom je priamka, tak by sme ju mohli zapísať v tvare:

$$y = \beta_0 + \beta_1 x \quad (3.72)$$

V tomto prípade sú koeficienty  $\beta_0$  a  $\beta_1$  reálne čísla, teda  $\beta_0, \beta_1 \in \mathbb{R}$ . Význam koeficientov je pomerne jednoduchý – konštanta  $\beta_0$  definuje, kde priamka pretína os  $y$ . Platí totiž, že pre  $x = 0$  dostávame:

$$y = \beta_0 + \beta_1 0 = \beta_0 \quad (3.73)$$

Priamka teda prechádza bodom  $[0, \beta_0]$ . Koeficient  $\beta_1$  predstavuje smernicu, alebo aj tangens uhla, ktorý zvierá priamka s osou  $x$ . Hovorí teda o sklone priamky. V prípade, ak je  $\beta_1 < 0$ , priamka je klesajúca. Pre  $\beta_1 > 0$  je priamka rastúca.

Posledným prípadom je  $\beta_1 = 0$  – v takomto prípade je priamka rovnobežná s osou  $x$  a hodnoty  $y$  sa nemenia – bez ohľadu na to, akú hodnotu nadobudne premenná  $x$ , premenná  $y$  je rovná  $\beta_0$ . V tomto prípade môžeme aj povedať, že hodnota  $y$  s hodnotou  $x$  v istom zmysle nesúvisí – nech sa  $x$  mení ľubovoľne, táto zmena nie je sprevádzaná zmenou premennej  $y$ .

Prvou vlastnosťou, ktorá nás pri koeficiente sklonu v lineárnom modeli zaujíma je jeho znamienko. Okrem znamienka má často značnú vypovedaciu hodnotu aj hodnota koeficientu (v absolútnej hodnote). Ak je koeficient  $\beta_1$  rôzny od nuly, znamená to, že meniace sa  $x$  je sprevádzané zmenou premennej  $y$ . Ak je znamienko koeficientu kladné, s rastom  $x$  rastie aj  $y$ . Zaujímavé však je aj to, o koľko vzrastie  $y$ .

Aby sme zodpovedali túto otázku, môžeme využiť skutočnosť, že vzťah zatiaľ sledujeme v podobe funkcie. Matematická analýza poskytuje veľmi veľa elegantných nástrojov, ktoré umožňujú skúmať funkcie a ich vlastnosti. Priebeh funkcií sa väčšinou vyšetruje pomocou skúmania ich derivácií.

Ak by sme vypočítali prvú deriváciu pre našu jednoduchú lineárnu funkciu dostali by sme:

$$\frac{dy}{dx} = \beta_1 \quad (3.74)$$

Znamená to, že koeficient  $\beta_1$  je aj prvou deriváciou skúmanej funkcie podľa premennej  $x$ . Opäť tu vidíme súvis s monotónnosťou funkcie – znamienko  $\beta_1$  vypovedá o tom, či je funkcia rastúca alebo klesajúca. Keďže ide o priamku, derivácia je konštantná – rovnaká pre všetky hodnoty  $x$ .

Derivácia je pre nás vhodná aj preto, aby sme posúdili zmeny  $x$  a  $y$ . Tu sa využíva súvis derivácie a diferenciálu funkcie – pre malé zmeny premennej  $x$  môžeme odhadnúť zmeny premennej  $y$  ako jej deriváciu podľa  $x$ . Ak by sme zmenili premennú  $x$  o jednu jednotku, tak na základe tohto odhadu by sme mohli očakávať, že sa premenná  $y$  zmení o  $\beta_1$ . Tento koeficient nám preto hovorí aj o rýchlosti, akou sa mení premenná  $y$ : ak je  $\beta_1$  kladné, tak s rastom hodnoty  $\beta_1$  získavame pre zvyšujúce sa  $x$  rýchlejšie rastúcu, teda „strmšiu“ priamku. Pre  $\beta_1$  záporné priamka klesá tým rýchlejšie, čím je  $\beta_1$  menšie (t. j. väčšie v absolútnej hodnote).

Čo v prípade, ak by sme chceli skúmať vzťah premennej  $y$  a niekoľkých premenných  $x$ ? Povedzme, že budeme mať takéto premenné tri: označme ich  $x_1$ ,  $x_2$  a  $x_3$ . Vzťah týchto premenných a premennej  $y$  by sme mohli modelovať v tvare reálnej funkcie troch premenných:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 \quad (3.75)$$

Vidíme, že tvar tejto funkcie pripomína predchádzajúcu funkciu. Jej grafom už nebude priamka – keďže máme dohromady štyri premenné ( $x_1$ ,  $x_2$ ,  $x_3$  a  $y$ ), pre jej zobrazenie by sme potrebovali štvorrozmerný graf. Situáciu by bolo možné vizualizovať v prípade, ak by sme mali premenné tri – napríklad  $x_1$ ,  $x_2$  a  $y$ . V takomto prípade by body spĺňajúce podmienku danú rovnicou vytvárali rovinu. V prípade štyroch a viacerých premenných si už graf funkcie predstaviť nevieme – graf funkcie, ktorú sme si definovali voláme nadrovina.

V predchádzajúcich odsekoch sme si popísali význam koeficientov  $\beta_0$  a  $\beta_1$ . Ich význam zostáva zachovaný aj naďalej. Interpretácia koeficientov  $\beta_2$  a  $\beta_3$  je podobná interpretácii koeficientu  $\beta_1$ . Ak koeficient  $\beta_1$  vyššie predstavoval deriváciu  $y$  podľa  $x$ , potom v našom novom modeli platí:

$$\frac{\partial y}{\partial x_1} = \beta_1, \quad \frac{\partial y}{\partial x_2} = \beta_2, \quad \frac{\partial y}{\partial x_3} = \beta_3 \quad (3.76)$$

Koeficienty sú teda hodnotami parciálnych derivácií podľa jednotlivých premenných. Tu upozorníme na jednu skutočnosť – tým, že ide o parciálne derivácie, nie je celkom správne tvrdiť „ak sa premenná  $x_1$  zmení o jednotku, tak sa  $y$  zmení o  $\beta_1$ “, ako to bolo predtým. Správne tvrdenie by bolo „ak sa premenná  $x_1$  zmení o jednotku, tak sa pri nezmenených hodnotách  $x_2$  a  $x_3$  hodnota  $y$  zmení o  $\beta_1$ “. Koeficient nám teda hovorí, ako sa zmení hodnota  $y$  pri zmene  $x_1$  ceteris paribus, teda za inak nezmenených podmienok.

Pri empirickej práci veľmi často hľadáme vzťah medzi premennými. Bežne sa stáva, že vieme získať hodnoty premennej  $y$ , ako aj premenných  $x_1$ ,  $x_2$ ,  $x_3$  a následne sa snažíme zistiť, akým modelom by bolo možné vhodne opísať ich vzájomný vzťah. Veľmi často za týmto účelom realizujeme výskum, v rámci ktorého zbierame údaje o premenných, ktoré nás zaujímajú. Na základe týchto pozorovaní sa snažíme odhadnúť model, ktorý by získané údaje vhodne popisoval.

Povezme, že sme presvedčení, že medzi skúmanými premennými predpokladáme vzťah v tvare:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 \quad (3.77)$$

Úlohou v takomto prípade je identifikovať koeficienty  $\beta_0$ ,  $\beta_1$ ,  $\beta_2$  a  $\beta_3$  na základe pozorovaných hodnôt  $x_1$ ,  $x_2$ ,  $x_3$  a  $y$ . Ak by sme mali  $n \in \mathbb{N}$  takýchto pozorovaní, mohli by sme ich zapísať ako systém nasledovne:

$$\begin{aligned} y_1 &= \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \beta_3 x_{13} \\ y_2 &= \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \beta_3 x_{23} \\ &\vdots \\ y_n &= \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \beta_3 x_{n3} \end{aligned} \quad (3.78)$$

Uvedený systém si je možné predstaviť nasledovne. Povedzme, že skúmame vzťah medzi veličinami  $x_1$ ,  $x_2$ ,  $x_3$  a  $y$ , ktorými budú napríklad vek, pohlavie, počet rokov dosiahnutého vzdelania a mesačný príjem. Ak by sme poznali koeficienty modelu, mohli by sme dať odpoveď na otázku, či ženy zarábajú viac alebo menej ako muži (podľa znamienka a hodnoty  $\beta_2$ ), či skúsenosti vyjadrené vekom majú súvis s mesačným príjmom ( $\beta_1$ ) a ako súvisí vzdelanie s mesačným príjmom ( $\beta_3$ ).

Aby sme tieto neznáme koeficienty identifikovali, mohli by sme realizovať dotazníkový prieskum a zistiť hodnoty pre všetky premenné u zvolenej vzorky respondentov.

Získané údaje by sme mohli označiť tak, že každej premennej pridáme dolný index identifikujúci respondenta. Potom by napríklad  $y_2$  označovalo mesačný príjem druhého respondenta,  $x_{11}$  vek prvého respondenta a  $x_{23}$  počet rokov dosiahnutého vzdelania druhého

respondenta (pri premenných  $x$  predstavuje prvý dolný index číslo respondenta a druhý identifikuje o ktorú premennú ide).

Pokiaľ je u všetkých respondentov vzťah medzi premennými  $x_1, x_2, x_3$  a  $y$  rovnaký, potom pre každého z nich má platiť rovnica definovaná vyššie. Zapísaním rovnice pre každého z  $n$  respondentov dostávame systém rovníc (3.78).

Všimnime si, že predpoklad o tom, že vzťah medzi skúmanými premennými je rovnaký u všetkých respondentov je vyjadrený tým, že koeficienty  $\beta_0, \beta_1, \beta_2$  a  $\beta_3$  majú len jeden dolný index – sú spoločné pre všetkých respondentov. V takomto prípade má význam vytvárať vzorku a zisťovať údaje od viacerých respondentov, pretože odpoveď každého z nich nám pomáha identifikovať ten istý vzťah, ktorý je pre nich spoločný.

Pokiaľ by pre každého respondenta naozaj platil vzťah daný vyššie spomínanou sústavou rovníc, bolo by možné stanoviť veľkosť vzorky tak, aby sme s určitosťou vedeli identifikovať každý z koeficientov  $\beta_0, \beta_1, \beta_2$  a  $\beta_3$ . Pokiaľ by sme tieto koeficienty považovali za neznáme, systém (3.78) predstavuje sústavu rovníc o štyroch neznámych. Podľa Frobeniovej vety by na jednoznačnú identifikáciu koeficientov stačili štyri pozorovania (ak by neboli lineárne závislé). Je tomu tak preto, lebo predpokladáme, že vzťah medzi premennými platí pre každého respondenta presne.

Tento predpoklad je jednoznačne veľmi silným predpokladom, ktorý je v praxi len málokedy splnený. Hlavným dôvodom je fakt, že vzťahy medzi veličinami v podobe modelov väčšinou predstavujú určitú abstrakciu. V praxi je totiž nemožné zahrnúť vplyv všetkých efektov, ktoré môžu súvisieť so skúmanou premennou  $y$ . Väčšinou si vyberáme menšiu skupinu tých, ktoré považujeme za najdôležitejšie -  $x_1, x_2, x_3$ . Tým, že efekt ostatných vplyvov zanedbávame, dopúšťame sa určitej chyby. V dôsledku toho pre prvého respondenta:

$$y_1 \neq \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \beta_3 x_{13} \quad (3.79)$$

vyjadríme rozdiel medzi skutočnou a vypočítanou hodnotou ako:

$$u_1 = y_1 - (\beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \beta_3 x_{13}) \quad (3.80)$$

Ekvivalentne potom môžeme napísať:

$$y_1 = \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \beta_3 x_{13} + u_1 \quad (3.81)$$

Objekt  $u_1$  nazývame poruchový, alebo tiež náhodný člen. Predstavuje súhrn všetkých veličín, ktoré sme v modeli nezohľadnili, ale majú vzťah k hodnote  $y_1$ . Keďže pri rôznych pozorovaniach bude náhodný člen nadobúdať rôzne hodnoty (mesačný príjem jedného môže byť ovplyvnený tým, že žije v krajine s vyššími príjmami, mesačný príjem iného môže byť ovplyvnený tým, že má prácu len na čiastočný úväzok a podobne), je ťažké, resp. nemožné ho

presne kvantifikovať. Z tohto dôvodu sa k nemu pristupuje ako k náhodnej premennej – aj keď konkrétnu hodnotu u konkrétneho respondenta nedokážeme predikovať, vieme využiť poznatky pravdepodobnosti a štatistiky a popisovať jeho vlastnosti, akými sú stredná hodnota, rozptyl a podobne.

Skúmaný systém by sa po pridaní poruchových členov, ktoré zohľadňujú, že vzťah medzi premennými  $x_1, x_2, x_3$  a  $y$  neplatí úplne presne, mohol prepísať do tvaru:

$$\begin{aligned} y_1 &= \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \beta_3 x_{13} + u_1 \\ y_2 &= \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \beta_3 x_{23} + u_2 \\ &\vdots \\ y_n &= \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \beta_3 x_{n3} + u_n \end{aligned} \quad (3.82)$$

Kým v prípade, ak sme predpokladali, že vzťahy medzi veličinami platia presne bolo možné odhadnúť veľkosť vzorky potrebnej na presné stanovenie koeficientov z Frobeniovej vety, zahrnutie poruchového člena nám situáciu dramaticky mení. V skutočnosti totiž často potrebujeme okrem koeficientov  $\beta_0, \beta_1, \beta_2$  a  $\beta_3$  popísať aj vlastnosti poruchových členov. Pokiaľ hodnoty  $y$  v sebe obsahujú aj náhodný prvok daný poruchovým členom, ako si môžeme byť istí, že poznatky zo vzorky nám skutočne vypovedajú o premenných  $x_1, x_2, x_3$  a  $y$ , a nie sú výsledkom náhody – poruchových členov  $u$ ? Toto je kľúčová otázka, ktorou sa zaoberá klasická ekonometria.

Uvedme ešte jeden alternatívny zápis skúmaného modelu. Doterajší výklad bol pomerne jednoduchý – zaoberali sme sa len popisom základného modelu. V ďalšom texte bude potrebné často popisovať operácie, ktoré sú pomerne zložité, a v tvare sústavy rovníc je ich zápis dosť neprehľadný. Veľmi efektívne sa však dá s modelmi pracovať vtedy, ak sú vyjadrené pomocou operácií s maticami. Ukážme si takýto zápis na sústave rovníc, ktorú sme si popísali.

Definujeme najprv tri vektory pre hodnoty  $\mathbf{y}$ ,  $\mathbf{u}$  a hľadané koeficienty  $\boldsymbol{\beta}$ .

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad \mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{pmatrix} \quad (3.83)$$

Zachováme konvenciu, že vektory budeme vždy vytvárať ako stĺpcové, ak nepovieme inak. Vektory  $\mathbf{y}$  a  $\mathbf{u}$  sú  $n$ -prvkovými stĺpcovými vektormi (ak by sme ich brali ako matice, mali by rozmer  $n \times 1$ ). Vektor  $\boldsymbol{\beta}$  je stĺpcovým vektorom s  $k + 1$  prvkami, pričom  $k \in \mathbb{N}$

definujeme ako počet premenných  $x$ . V našom prípade by bolo  $k = 3$ , pretože máme premenné  $x_1, x_2, x_3$ .

Okrem spomínaných vektorov definujeme maticu  $\mathbf{X}$  rádu  $n \times (k + 1)$  nasledovne:

$$\mathbf{X} = \begin{pmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{pmatrix} \quad (3.84)$$

Potom sú nasledovné zápisy ekvivalentné:

$$\begin{aligned} y_1 &= \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \cdots + \beta_k x_{1k} + u_1 \\ y_2 &= \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \cdots + \beta_k x_{2k} + u_2 \\ &\vdots \end{aligned} \quad (3.85)$$

$$\begin{aligned} y_n &= \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \cdots + \beta_k x_{nk} + u_n \\ y_i &= \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + u_i, \quad i = 1, 2, \dots, n \end{aligned} \quad (3.86)$$

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix} + \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \quad (3.87)$$

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \quad (3.88)$$

Z uvedeného je zrejmé, že maticový zápis umožňuje značne zjednodušiť zápis celého modelu, pričom stále zostáva ľahko čitateľný.

Je možné ešte vysvetliť, prečo matica  $\mathbf{X}$  obsahuje okrem hodnôt jednotlivých premenných aj prvý stĺpec, ktorý obsahuje samé jednotky. V prípade maticového násobenia daného posledným vzťahom musíme násobiť prvky matice  $\mathbf{X}$  aj koeficientom  $\beta_0$ . Keďže tomu neprislúcha žiadna z premenných, doplnením stĺpca jednotiek umožníme vznik súčinu  $1 \cdot \beta_0 = \beta_0$ .

V ďalšom texte budeme premennú  $y$  nazývať vysvetľovanou veličinou (alebo aj závislou premennou) a premenné  $x_1, x_2, \dots, x_k$  vysvetľujúce (nezávislé) premenné.

### 3.2.3 Aplikácie lineárnych modelov

Na prvý pohľad sa zdá, že model popísaný v predchádzajúcej časti predstavuje najjednoduchšiu, a teda aj najmenej zaujímavú alternatívu, ktorou popisujeme vzťahy medzi skúmanými veličinami. Je pravdou, že lineárna funkcia patrí medzi najjednoduchšie –

jednoduchšia je snád' už len konštantná funkcia. Napriek tomu je model predstavený v predchádzajúcej časti využiteľný na modelovanie veľkej skupiny vzťahov. Rôznym možnostiam využitia lineárnych modelov je venovaná táto podkapitola.

Pripomeňme si, že všeobecný lineárny model skúmame v tvare:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + u_i, \quad i = 1, 2, \dots, n \quad (3.89)$$

V najjednoduchšom prípade pracujeme len s jednou vysvetľujúcou premennou a model má tvar:

$$y_i = \beta_0 + \beta_1 x_i + u_i, \quad i = 1, 2, \dots, n \quad (3.90)$$

Ako sme spomínali vyššie, grafom funkcie  $y_i = \beta_0 + \beta_1 x_i$  je priamka. Čo v prípade, ak by sme mali dôvod domnievať sa, že pre popísanie vzťahu medzi  $y_i$  a  $x_i$  je vhodná parabola? Zrejme by sme hľadali model v tvare:

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + u_i, \quad i = 1, 2, \dots, n \quad (3.91)$$

Grafom funkcie  $y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2$  by bola parabola, ktorej vlastnosti sú dané koeficientmi  $\beta_0$ ,  $\beta_1$  a  $\beta_2$ .

Z hľadiska klasickej ekonometrie je tento model lineárnym modelom, napriek tomu, že skúmame nelineárnu závislosť. Ak hovoríme o lineárnom modeli, myslí sa tým **linearita v parametroch**. Uvažujeme o nej vo vzťahu ku koeficientom  $\beta_0$ ,  $\beta_1$  a  $\beta_2$ , a nie k premenným, ktorý do modelu dosadzujeme. Aby sme túto skutočnosť ilustrovali, zaveďme substitúciu:

$$z_i = x_i^2, \quad i = 1, 2, \dots, n \quad (3.92)$$

S využitím tejto substitúcie môžeme model zapísať v tvare:

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 z_i + u_i, \quad i = 1, 2, \dots, n \quad (3.93)$$

Tento zápis už na prvý pohľad predstavuje lineárny model. Všetky modely, ktoré je možné substitúciou previesť na tvar všeobecného lineárneho modelu nazývame lineárnymi modelmi a môžeme pri nich využiť všetky poznatky, ktoré uvedieme v nasledujúcich podkapitolách.

Predchádzajúce úvahy nám umožňujú pristupovať k mnohým modelom ako k lineárnym – kým na začiatku sa zdalo, že obmedzenie sa na lineárne modely je veľmi reštriktívne, z uvedeného je zrejmé, že v skutočnosti máme k dispozícii veľmi bohatú skupinu modelov, ktoré vieme považovať za lineárne v parametroch. Uveďme si niekoľko ďalších príkladov:

$$y_i = \beta_0 + \beta_1 x_i + u_i \quad (3.94)$$

$$y_i = \beta_0 + \beta_1 \ln(x_i) + u_i \quad (3.95)$$

$$y_i = \beta_0 + \beta_1 \cos(x_i) + u_i \quad (3.96)$$

$$y_i = \beta_0 + \beta_1 \ln(x_{i1}) + \beta_2 e^{x_{i2}} + \beta_3 \cos(\ln(x_{i3}^2)) + u_i \quad (3.97)$$

Nie každý model sa však dá odhadnúť ako lineárny – uvedme si niekoľko príkladov aj pre túto skupinu (pre odhad nelineárnych modelov odporúčame napr. publikáciu Davidson – MacKinnon, 2003).

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_1 \beta_2 x_{i3} + u_i \quad (3.98)$$

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_1 x_{i3} / \beta_2 + u_i \quad (3.99)$$

$$y_i = \beta_0 + \beta_1 \ln(x_{i1}) + \beta_2 e^{x_{i2}} + \beta_1 \cos(\ln(\beta_2 x_{i3}^2)) + u_i \quad (3.100)$$

Špecifické miesto medzi modelmi, ktoré je možné previesť na lineárny tvar majú modely, ktoré súvisia s logaritmickejšími transformáciami. Ide o nasledovné tri prípady:

$$y_i = \beta_0 + \beta_1 \ln(x_i) + u_i \quad (3.101)$$

$$\ln(y_i) = \beta_0 + \beta_1 x_i + u_i \quad (3.102)$$

$$\ln(y_i) = \beta_0 + \beta_1 \ln(x_i) + u_i \quad (3.103)$$

Model (3.101) sa niekedy nazýva lineárno-logaritmickejší model, (3.102) logaritmickejší-lineárny (alebo log-lineárny) model a nakoniec (3.103) ako logaritmickejší-logaritmickejší (log-log) model.

Dôvodom pre logaritmovanie premenných môže byť niekoľko. V prípade, ak je vzťah medzi skúmanými veličinami  $y_i$  a  $x_i$  nelineárny, je možné pomocou logaritmovania premenných získať transformované premenné, pre ktoré je lineárna aproximácia vyhovujúca.

Druhým dôvodom je transformácia v prípadoch, ak potrebujeme pracovať s premennými, ktoré majú výrazné zošikmenie. Premenné, ktoré majú vysokú šikmosť spravidla nie sú normálne rozdelené – toto rozdelenie je symetrické. Mnoho testov v štatistike, ale aj modelov ktoré budeme preberať v ďalších častiach je založených na predpoklade normality dát. V rámci indukčívnej štatistiky preto existujú testy normality, ktorých cieľom je posúdiť, či daná premenná má normálne rozdelenie. Jeden z tejto skupiny testov, tzv. Jarque-Berov test je priamo založený na koeficientoch šikmosti a špicatosti. Je zrejmé, že v prípade premenných s vysokou šikmosťou zrejme dôjde k porušeniu predpokladu o normalite. Pomerne často je však možné premenné s vysokou šikmosťou logaritmickejší transformáciou previesť na premenné, pri ktorých normalitu nevyvraciam. Získavame tak premenné, ktoré spĺňajú predpoklady na ďalšie testovanie alebo modelovanie. Z tohto dôvodu je logaritmickejší transformácia premenných veľmi často využívaná pri tvorbe ekonomických modelov. V oblasti financií je napríklad známa aplikácia tejto transformácie pri modeloch na

oceňovanie opcií, ktoré priamo používajú lognormálne rozdelenie – pravdepodobnostné rozdelenie, pri ktorom má logaritmus skúmanej premennej normálne rozdelenie.

V prípade *lineárno-logaritmickeho* modelu je závislá premenná nezmenená, ale namiesto vysvetľujúcej premennej  $x_i$  používame jej logaritmicke transformáciu. V predchádzajúcej podkapitole sme si vysvetlili, ako je možné interpretovať odhadované koeficienty pri vysvetľujúcich premenných v modeli. Tým, že v lineárno-logaritmicke modeli nahrádzame premennú  $x_i$  jej logaritmom, mení sa aj interpretácia regresného koeficientu.

V modeli obsahujúcom premennú  $x$  predstavuje koeficient  $\beta_1$  hodnotu, o ktorú čakáme že sa zmení premenná  $y_i$  ak sa premenná  $x_i$  zmení o jednotku. V prípade lineárno-logaritmickeho modelu predstavuje koeficient  $\beta_1$  hodnotu, o ktorú sa zmení premenná  $y$ , ak sa zmení o jednotku  $\ln(x_i)$ . Túto zmenu môžeme zapísať ako:

$$\ln(x_i) + 1 = \ln(x_i) + \ln(e) = \ln(ex_i) \quad (3.104)$$

V predchádzajúcom vzťahu predstavuje  $e$  základ prirodzeného logaritmu, s hodnotou približne 2.718. Interpretácia, ktorú dostávame, je teda nasledovná: koeficient  $\beta_1$  vyjadruje, o koľko sa zmení premenná  $y$  ak sa premenná  $x_i$  zmení 2.718-násobne. Tá istá skutočnosť sa dá vyjadriť aj inak - koeficient  $\beta_1$  vyjadruje, o koľko sa zmení premenná  $y_i$  ak sa premenná  $x_i$  zmení o 171.8 %, pretože:

$$2.718 * 100 - 100 = 171.8 \quad (3.105)$$

Najčastejšie používaným logaritmom je prirodzený logaritmus – ako je vidno, ten však niekedy vedie k ťažšej interpretácii (prečo nás zaujíma práve 171,8 % nárast?). Namiesto prirodzeného logaritmu preto môžeme použiť logaritmus s iným základom. Napríklad pri základe 2 dostávame interpretáciu, o koľko sa zmení premenná  $y$  ak sa premenná  $x$  zdvojnásobí.

Môžeme to otočiť aj naopak – ak sa premenná  $x_i$  zmení o  $a \in \mathbb{R}$  percent, ako sa zmení premenná  $y$ ?

Pôvodná hodnota je  $x_i$ , po zvýšení o  $a$  percent dostávame  $(1+a/100)x_i$ . Potom:

$$\beta_1 \ln[(1 + a/100) x_i] = \beta_1 \ln(1 + a/100) + \beta_1 \ln(x_i) \quad (3.106)$$

Premenná  $y_i$  sa pri  $a$  percentnej zmene premennej  $x_i$  zmení o  $\beta_1 \ln(1 + a/100)$ .

Pre malé zmeny  $a$  platí približný vzťah:

$$\ln(1 + a/100) \approx a/100 \quad (3.107)$$

Preto je možné koeficient  $\beta_1/100$  približne interpretovať ako zmenu premennej  $y_i$  pri zmene premennej  $x_i$  o jedno percento.

O *log-lineárnom modeli* hovoríme, ak je v tvare:

$$\ln(y_i) = \beta_0 + \beta_1 x_i + u_i \quad (3.108)$$

V tomto prípade nie je logaritmovaná nezávislá, ale závislá premenná. Predchádzajúci výraz je možné prepísať ako:

$$y_i = e^{\beta_0 + \beta_1 x_i + u_i} = e^{\beta_0} e^{\beta_1 x_i} e^{u_i} \quad (3.109)$$

Ak by sme chceli poznať interpretáciu koeficientu  $\beta_1$ , skúsme zmeniť hodnotu premennej  $x_i$  o jednotku.

$$e^{\beta_0} e^{\beta_1(x_i+1)} e^{u_i} = e^{\beta_0} e^{\beta_1 x_i + \beta_1} e^{u_i} = e^{\beta_1} e^{\beta_0 + \beta_1 x_i + u_i} = e^{\beta_1} y_i \quad (3.110)$$

Znamená to, že zmena hodnoty  $x_i$  o jednotku je sprevádzaná zmenou  $y_i$   $e^{\beta_1}$ -násobne.

Podobne ako v predchádzajúcom modeli, aj tu si pre rýchly odhad môžeme pomôcť aproximáciou – pre malé hodnoty  $\beta_1$  je výraz  $e^{\beta_1}$  približne rovný  $1 + \beta_1$ . Ak by sme napríklad mali  $\beta_1 = 0.02$ , potom by jednotková zmena  $x_i$  bola sprevádzaná  $e^{0.02}$  násobnou zmenou  $y_i$ , pričom  $e^{0.02} \approx 1.02$ , čiže by šlo o zmenu o približne dve percentá (skutočná hodnota je približne 1.0202).

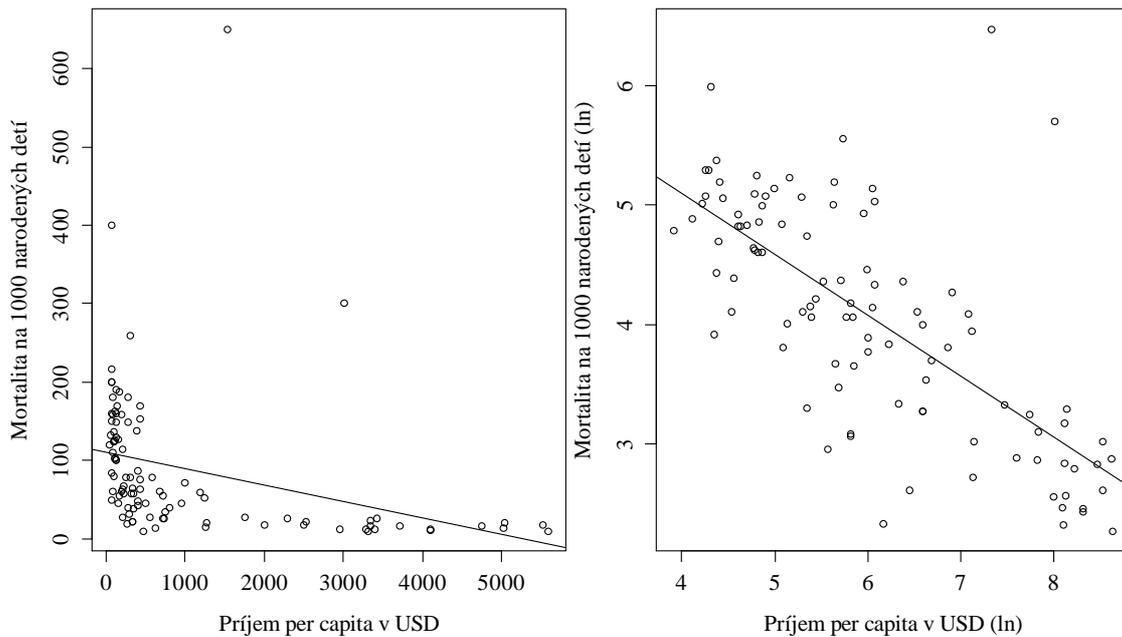
*Log-log* model má tvar:

$$\ln(y_i) = \beta_0 + \beta_1 \ln(x_i) + u_i \quad (3.111)$$

V tomto prípade sú logaritmované obe premenné – závislé, aj nezávislá. Tento model je možné ekvivalentne prepísať do tvaru:

$$y_i = e^{\beta_0 + \beta_1 \ln(x_i) + u_i} = e^{\beta_0 + u_i + \ln(x_i^{\beta_1})} = e^{\beta_0 + u_i} e^{\ln(x_i^{\beta_1})} = e^{\beta_0 + u_i} x_i^{\beta_1} \quad (3.112)$$

Keďže v tomto modeli máme logaritmované obidve premenné, pričom pre malé zmeny je možné logaritmy aproximovať percentuálnymi zmenami, koeficient  $\beta_1$  je možné (približne) interpretovať ako elasticitu  $y_i$  voči  $x_i$  – vyjadruje približne, o koľko percent sa zmení premenná  $y_i$  ak sa  $x_i$  zmení o percento.



Obrázok 6: Príklad lineárneho a log-log modelu

*Zdroj: vlastné spracovanie*

Obrázok 6 zobrazuje príklad, v ktorom je logaritmickej transformácii oboch premenných užitočná. Použité údaje vychádzajú z databázy Leinhardt v knižnici `car`. Databáza obsahuje údaje o príjme per capita v USD a mortalite detí z obdobia rokov v období okolo roku 1970. Ľavá časť obrázku predstavuje vzťah medzi oboma veličinami v pôvodnom tvare. Z obrázku je zrejماً značná nelinearita. Pravá časť obrázku predstavuje vzťah oboch premenných po ich logaritmickej transformácii. I keď sa vo vzorke naďalej vyskytujú odľahlé (extrémne) hodnoty, ktoré by analýzu mohli ovplyvniť, je zrejماً, že logaritmickej transformácia vedie k hodnotám, ktoré je možné lepšie popísať lineárnym modelom.

Obrázok môžeme v programe R vytvoriť nasledovne:

```
> library(car)
> data(Leinhardt)
> par(family="serif")
> par(mar=c(5, 4, 0.3, 0.3))
> par(mfrow=c(1, 2))
> plot(Leinhardt$income, Leinhardt$infant, xlab="Príjem per
capita v USD", ylab="Mortalita na 1000 narodených detí")
> abline(lm(Leinhardt$infant~Leinhardt$income))
> lnIncome <- log(Leinhardt$income)
> lnInfant <- log(Leinhardt$infant)
> plot(lnIncome, lnInfant, xlab="Príjem per capita v USD (ln)",
ylab="Mortalita na 1000 narodených detí (ln)")
> abline(lm(lnInfant~lnIncome))
```

Na tomto mieste pripomeňme ešte jeden spôsob, ako je možné špecifikovať ekonometrický model. Ide o použitie takzvaných indikátorových premenných (angl. *dummy variables*). Uvažujme nad modelom:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + u_i \quad (3.113)$$

kde premenná  $x_{i2}$  je binárna, takže nadobúda len dve hodnoty – 0 a 1. Hodnota 0 väčšinou charakterizuje neprítomnosť nejakého znaku a hodnota 1 jeho prítomnosť (odtiaľ názov indikátorová premenná). Pozrime sa, aké má dôsledky zahrnutie takejto premennej do modelu.

Bežnou indikátorovou premennou býva pohlavie. Povedzme, že by sme chceli kvantifikovať model, v ktorom modelujeme ako závislú premennú mesačný príjem zamestnanca, premenná  $x_1$  zodpovedá počtu rokov jeho praxe a premenná  $x_2$  je binárna premenná kódujúca pohlavie: položíme  $x_{i2} = 0$  v prípade, ak respondent  $i$  je muž a  $x_{i2} = 1$  v prípade, ak je respondentom  $i$  žena. Premennú použijeme ako bežnú vysvetľujúcu premennú a model kvantifikujeme bežným spôsobom. Ako interpretujeme výsledok?

Ak by sme zobrali do úvahy model, ktorý dostávame pre mužov ( $x_{i2} = 0$ ), tak zistíme, že je v tvare:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + u_i \quad (3.114)$$

$$y_i = \beta_0 + \beta_1 x_i + u_i \quad (3.115)$$

Naproti tomu v prípade, ak je respondentom  $i$  žena ( $x_{i2} = 1$ ), nadobudne model tvar:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + u_i \quad (3.116)$$

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 + u_i \quad (3.117)$$

Oproti predchádzajúcemu modelu pre mužských respondentov je rozdiel len v dodatočnej konštante  $\beta_2$ . Podľa znamienka (a významnosti) tohto koeficientu by bolo možné určiť, či v rámci skúmanej skupiny zamestnancov je priemerný plat žien nižší, ako je plat mužov. Indikátorová premenná potom umožňuje kvantifikáciu efektu príslušnosti k skupine.

### 3.3 Odhad lineárneho modelu

Predmetom tejto kapitoly je dohad lineárneho modelu, ktorý sme si predstavili v predchádzajúcej časti. V ďalšom texte už budeme pracovať takmer výhradne s maticovým zápisom – pripomeňme preto, že základný model zapisujeme nasledovne:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \quad (3.118)$$

Vo väčšine prípadov pri empirickom výskume vychádzame zo situácie, v ktorej hľadáme vzťah medzi premennými tvoriacimi maticu  $\mathbf{X}$  a premennou, ktorej hodnoty sa nachádzajú vo vektore  $\mathbf{y}$ . Odhad samotný má niekoľko krokov.

*Prvým krokom* je voľba premenných modelu. Okrem závislej premennej, ktorú chceme modelovať, je potrebné zvoliť nezávislé premenné, pomocou ktorých ju budeme vysvetľovať. Platí, že ak do modelu nezarádime nejakú relevantnú premennú, jej efekt sa čiastočne stane súčasťou náhodného poruchového člena  $\mathbf{u}$ . Keďže každý model je len abstrakciou reality, je našou snahou identifikovať relevantné premenné pre daný druh analýzy a všetky ostatné faktory zanedbávame. Keďže ich efekt nedokážeme posúdiť, pristupujeme k nim ako k náhodnej premennej (resp. k náhodnému vektoru).

Ako je možné zvoliť relevantné premenné? Možností je niekoľko. V prvom rade je hlavným zdrojom pre výber premenných dedukcia. Ak vytvárame ekonometrický model, spravidla je to za nejakým konkrétnym účelom – skúmame nejaký problém, a prostredníctvom kvantitatívnej analýzy na vzorke o ňom chceme rozhodnúť. Väčšina problémov vychádza z ekonomickej teórie: správanie sa spotrebiteľov môže byť na makroekonomickej úrovni popísané teóriami spotreby, výrobný proces v podniku má svoje zákonitosti, nákupné správanie v maloobchode má svoje zákonitosti. Väčšina problémov, ktoré sa v praxi vyskytujú už bola v minulosti skúmaná buď v rovnakých, alebo podobných podmienkach, a preto je možné tieto informácie využiť.

Druhou možnosťou ako identifikovať vhodný model je pomocou exploratívnych techník – vyskúšaním rôznych špecifikácií. V prípade, ak kvantifikujeme model, existujú postupy, ktorými je možné overiť adekvátnosť špecifikácie modelu. Ak má poruchový člen obsahovať všetky efekty, ktoré v modeli explicitne nevystupujú, a model je správne špecifikovaný, tzn. žiadnu relevantnú premennú sme nevynechali, potom by náhodný člen nemal obsahovať žiadnu využiteľnú informáciu, ktorou by bolo možné model vylepšiť. Ak je náhodný člen súhrnom faktorov, ktoré neobsahujú systematickú informáciu vo vzťahu k  $\mathbf{y}$ , potom je model asi špecifikovaný správne. Adekvátnosť špecifikácie teda môžeme vyšetriť na základe skúmania náhodných odchýlok.

Ekonometrická teória ďalej definuje ukazovatele, ktoré popisujú kvalitu modelu (napr. koeficient determinácie a informačné kritériá). S ich pomocou je možné porovnávať rôzne modely a vybrať ten, v ktorom je skúmaná premenná  $\mathbf{y}$  vysvetlená lepšie.

Kým prístup k výberu premenných na základe ekonomickej teórie bol deduktívny, prístup založený na porovnávaní modelov a skúmaní náhodných odchýlok je empirický. Bez

ohľadu na to, ako výber premenných realizujeme, je potrebné v tejto fáze zabezpečiť údaje za všetky zvolené premenné.

*Druhým krokom* je funkčná špecifikácia modelu, ktorá výberom premenných nekončí. Musíme sa zaoberať otázkou, aký tvar závislosti medzi závislou vysvetľovanou a vysvetľujúcou premennou chceme modelovať. V predchádzajúcej kapitole sme si uviedli niekoľko možností, ako rôzne môžu vyzeráť modely, ktoré sú lineárne v parametroch – možností je nekonečne veľa.

Výber funkčnej špecifikácie (typu funkcie pre vyjadrenie  $y$ ) je tiež možné realizovať podobne ako výber premenných – na základe teórie, alebo na základe empirického testovania. Príkladom pre prvú kategóriu je napríklad voľba spotrebnej funkcie – existujú teórie, ktoré predpokladajú funkčnú závislosť na disponibilnom dôchodku, alebo permanentnom dôchodku a podobne. V závislosti od teórie môžu v modeli vystupovať diskontované budúce príjmy a iné vzťahy. Príklad na výber funkčného tvaru založený na empirickom skúmaní sme videli na konci predchádzajúcej podkapitoly o log-log modeli. V tomto prípade sa veľmi často využíva grafické zobrazenie závislostí dvoch premenných, napríklad vo forme  $x$ - $y$  grafu (angl. *scatterplot*). V mnohých prípadoch je vhodné v tejto fáze využiť niektorú z vizualizačných techník, aby bolo možné lepšie porozumieť charakteru údajov, ktoré skúmame – týmto spôsobom často vieme rozhodnúť o voľbe vhodnej funkčnej špecifikácie.

*Tretím krokom* je kvantifikácia modelu – ak máme k dispozícii údaje a funkčnú špecifikáciu, môžeme pristúpiť k odhadu parametrov modelu  $\beta$ . Odhad parametrov je možné realizovať viacerými metódami – najčastejšie metódou najmenších štvorcov, metódou maximálnej vierohodnosti alebo metódou momentov. V tejto publikácii sa budeme primárne zaoberať metódou najmenších štvorcov.

*Štvrtým krokom* je verifikácia modelu – potom, ako kvantifikujeme model je potrebné sa presvedčiť, či je model korektný. To zahŕňa kontrolu adekvátnosti špecifikácie, prípadne posúdenie kvality modelu tak, ako to bolo spomínané vyššie – je totiž možné, že aj keď je model špecifikovaný správne, jeho schopnosť vysvetliť hodnoty  $y$  môže byť nízka. Verifikácia modelu je obzvlášť dôležitá v prípade, ak zostavujeme model na základe teoretických modelov – je možné, že pri verifikácii zistíme, že teóriou navrhovaný model sa na popisovanie vzťahov skúmaných premenných nehodí, resp. nie je adekvátny, čo má samo osebe praktickú aj vedeckú hodnotu (teoretický model potom zrejme nezodpovedá realite a vzniká priestor na tvorbu nového).

*Piatym krokom* je využitie kvantifikovaného modelu. Modely je možné využiť rôzne – jednou z možných aplikácií je predikcia, kde na základe vysvetľujúcich premenných dokážeme predikovať hodnotu závislej premennej. Predikciu je možné realizovať tak na údajoch v rámci vzorky (angl. *within sample*) alebo na údajoch mimo vzorky (angl. *out of sample forecast*). Príklad na predikciu si je možné predstaviť takto: povedzme, že zostavíme model spotrebiteľského správania pri nákupe jogurtov. Ak na základe vzorky odhadneme model predikujúci nákupné správanie sa zákazníka, získame model, ktorý je použiteľný aj na osoby, ktoré neboli súčasťou vzorky. Ak by sme do takého modelu dosadili charakteristiky nového spotrebiteľa, ktorý sa na vzorke nepodieľal, mohli by sme pomocou modelu napríklad predikovať, či si kúpi jogurt, ktorého predaj sme analyzovali.

Druhou aplikáciou pre využitie kvantifikovaného modelu je testovanie štatistických hypotéz. Hlavne z hľadiska empirického výskumu je veľmi často kvantifikovaný model ani nie kvôli tomu, aby sme poznali presné hodnoty koeficientov  $\beta$ , ale aby sme pomocou nich overili existenciu alebo neexistenciu vzťahu medzi premennými. V prípade spomínaných spotrebných teórií by nás napríklad mohlo zaujímať, či je preukázateľné, že spotreba závisí od disponibilného príjmu. Teória predpokladá, že áno – je tomu tak ale aj naozaj? Na overenie tohto tvrdenia je možné špecifikovať a kvantifikovať ekonometrický model, ktorý by takéto tvrdenie mohol vyvrátiť.

### **3.3.1 Vzťah ekonometrického modelu a jeho odhadu**

Kým pristúpime k popisu estimačných procedúr, ktoré umožňujú kvantifikovať špecifikovaný ekonometrický model, je potrebné vysvetliť vzťah medzi ekonometrickým modelom a jeho odhadom. Rozdiel medzi nimi je dôležitý a je nutné vedieť medzi nimi jasne rozlišovať. Z dôvodu významu tohto členenia venujeme vysvetleniu samostatnú podkapitolu.

Zatiaľ sme sa rozdielom medzi modelom a jeho odhadom zámerne nevenovali – model bol pre nás vždy len jeden. V skutočnosti si pod definíciou uvedenou vyššie predstavujeme špecifikáciu modelu, ktorý platí pri údajoch za celú populáciu.

Podobne ako v induktívnej štatistike rozlišujeme vzorku a populáciu. Chápanie pojmu vzorka je jednoduché – ak máme k dispozícii údaje od respondentov pri prieskume, respondenti tvoria vzorku. Ak modelujeme vzťah medzi nastavením vstupných parametrov výrobného procesu a jeho výstupmi, potom môže vzorku predstavovať séria experimentov, ktoré vykonáme s rôznymi vstupnými parametrami.

Definovanie populácie môže byť o niečo náročnejšie. Niektoré príklady sú jednoduché – ak skúmame preferencie spotrebiteľov pri nákupe jogurtov v rámci určitého územia (Slovenska), populáciu môžu tvoriť všetci kupujúci.

V prípade časových radov je situácia o niečo ťažšia. Ak by sme realizovali prieskum na 100 respondentoch, ktorí by tvorili vzorku, tak je jasné, že vyberáme len časť z populácie a snažíme sa zovšeobecniť výsledky na celú populáciu. Ak však máme napríklad časový rad týkajúci sa HDP na obyvateľa, a od vzniku SR máme len 20 ročných údajov, môže vyvstať otázka: ide o vzorku, alebo populáciu? Kým pri štúdiu o jogurtoch je jasné, že 100 respondentov je len malá časť celku (populácie), v akom zmysle je vzorkou 20 údajov o HDP – čo presne je populáciou, ak existuje len tých 20 údajov?

V prípade časových radov existuje rozsiahla rozpracovaná teória o ich analýze a modelovaní, ktorá je dosť náročná. Zjednodušene si to preto predstavme tak, že ak modelujeme podobné veličiny, predpokladáme, že ak sa premenné riadili nejakým rozdelením v minulosti, budú sa ním riadiť aj v budúcnosti. Z tohto pohľadu máme k dispozícii tiež len vzorku – 20 údajov z populácie, ktorá zahŕňa minulé aj budúce pozorovania. Tento predpoklad je pomerne reštriktívny, ale nateraz si s ním vystačíme – presnejšia odpoveď by si vyžadovala ďalšie vedomosti z teórie o časových radoch. Zatiaľ je dôležité len pochopiť, že aj v tomto prípade je možné na hodnoty premenných nazerať ako na vzorku.

Model v tvare, ako sme si ho definovali vyššie, t. j. v tvare:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \quad (3.119)$$

budeme považovať za hypotetický model, o ktorom sme presvedčení, že platí v rámci skúmanej populácie. Tento model považujeme za „pravý“ model, ktorého parametre  $\boldsymbol{\beta}$  sa snažíme odhadnúť. Tento model je „pravý“ v tom zmysle, že  $\boldsymbol{\beta}$  sú dané a  $\mathbf{X}$  obsahuje práve tie premenné, ktoré majú vo vzťahu vystupovať (špecifikácia je správna). Tento model v praxi prakticky nikdy nepoznáme.

Všimnime si, že v tomto – nazvime ho populačnom – modeli stále existuje náhodný poruchový člen. Predpokladá sa preto, že aj keby sme mali k dispozícii všetky jednotky z populácie, vzťah medzi  $\mathbf{y}$  a  $\mathbf{X}$  nie je presný – vždy existujú medzi  $\mathbf{y}$  a  $\mathbf{X}\boldsymbol{\beta}$  nejaké odchýlky.

Ako už bolo povedané, tento model prakticky nikdy nepoznáme. Spravidla je to dané tým, že nemáme k dispozícii celú populáciu, ale len vzorku. Hneď prvou chybou, ktorej sa môžeme dopustiť pri ekonometrickom modelovaní je chyba špecifikácie modelu (nemusíme do matice  $\mathbf{X}$  správne identifikovať všetky premenné, resp. nesprávne určíme funkčný tvar

závislosti). Keďže tomuto problému sme sa už venovali, predpokladajme ďalej, že model máme špecifikovaný správne.

Povedzme, že na základe vzorky odhadneme parametre  $\beta$ . Aby sme rozlíšili skutočné a odhadované hodnoty, označme odhadnuté hodnoty symbolom  $\hat{\beta}$ . Model, ktorý takto dostaneme, budeme zapisovať nasledovne:

$$\mathbf{y} = \mathbf{X}\hat{\beta} + \mathbf{e} \quad (3.120)$$

V zápise si všimnime niekoľko skutočností. V prvom rade vidíme, že vektor  $\mathbf{y}$  a matica  $\mathbf{X}$  zostali nezmenené – nenahrádzali sme ich symbolmi so strieškou. Je to preto, že skúmaním vzorky získavame skutočné hodnoty  $\mathbf{y}$ , ktorým prislúchajú skutočné hodnoty  $\mathbf{X}$ . Oproti populačnému modelu ich je však menej – ide o vzorku. Technicky by bolo asi možné označiť tieto matice inak, ale vo väčšine učebníc ekonometrie sa tak nerobí, prikloníme sa preto k zažitej konvencii. Predstaviť si to môžeme aj tak, že ak by sme mali tie isté údaje, ktoré sa nachádzajú vo vzorke, potom podľa populačného modelu by pre ne platilo (3.119). My však len na základe vzorky nedokážeme odhadnúť populačný model presne – dostávame preto (3.120).

Ako už bolo povedané, nahrádzame neznámy vektor  $\beta$  naším odhadom  $\hat{\beta}$ . Keďže pri našom odhade pravdepodobne  $\beta \neq \hat{\beta}$ , rozdiely  $\mathbf{y} - \mathbf{X}\hat{\beta}$  nebudú rovné  $\mathbf{u}$ . Označujeme ich preto  $\mathbf{e}$ . Prvky vektora  $\mathbf{e}$  nazývame *rezíduá* a vektor  $\mathbf{e}$  nazývame vektorom rezíduí. Hodnoty vektora  $\mathbf{X}\hat{\beta}$  nazývame *vyrovnané hodnoty* (angl. *fitted values, predicted values*) a označujeme ich aj  $\hat{\mathbf{y}}$ . Predstavujú náš odhad hodnôt  $\mathbf{y}$ , rezíduá  $\mathbf{e}$  predstavujú chybu nášho odhadu. Model (3.120) nazývame regresným modelom. Pomocou koeficientu  $\hat{\beta}$  je možné definovať nadrovinu (priamku v rovine, rovinu v trojrozmernom priestore a pod.), pričom vyrovnané hodnoty  $\hat{\mathbf{y}} = \mathbf{X}\hat{\beta}$  všetky ležia v tejto tzv. regresnej nadrovine. Rezíduá  $\mathbf{e}$  je možné považovať aj za odchýlky skutočných hodnôt od regresnej nadroviny.

### 3.3.2 Odhad modelu metódou najmenších štvorcov

Z predchádzajúcej podkapitoly vieme, že naším cieľom pri ekonometrickom modelovaní bude odhadnúť koeficienty  $\beta$  pomocou vektora odhadov  $\hat{\beta}$ . Namiesto pravého populačného modelu:

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{u} \quad (3.121)$$

máme model:

$$\mathbf{y} = \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{e} \quad (3.122)$$

Vyrovnané hodnoty sme si označili ako:

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}} \quad (3.123)$$

Stále sme si však nevysvetlili, ako vypočítať odhad  $\hat{\boldsymbol{\beta}}$ .

Ideálne by bolo, ak by sme vedeli vypočítať priamo vektor  $\boldsymbol{\beta}$ . Tak ako to ale pri výberovom skúmaní býva, je len ťažko možné očakávať, že dostaneme zhodou okolností takú vzorku, že náš odhad bude úplne presný. Nech by sme zo vzorky počítali akúkoľvek charakteristiku, väčšinou aj pridaním jedného pozorovania do vzorky sa výberové charakteristiky budú meniť. Musíme teda akceptovať, že náš odhad presný nebude.

Podobne ako tomu je v štatistike, je možné tú istú populačnú charakteristiku (v našom prípade regresné parametre  $\boldsymbol{\beta}$ ) odhadovať rôznymi bodovými odhadmi. Otázkou je, ako definovať kritérium pre výber odhadu – ktorý odhad považujeme za dobrý?

*Metóda najmenších štvorcov* vychádza z porovnania skutočných a vyrovnaných hodnôt závislej premennej. Skutočné hodnoty tvoria zložky vektora  $\mathbf{y}$ . Odhad týchto hodnôt spočítaný na základe modelu je  $\hat{\mathbf{y}}$ . Ak by náš model presne popisoval závislú premennú, odchýlky medzi skutočnými a vyrovnanými hodnotami by mali byť malé, ideálne nulové.

Pri voľbe odhadu by sme preto mohli postupovať tak, že by sme hľadali taký vektor odhadov  $\hat{\boldsymbol{\beta}}$ , pri ktorom budú odchýlky:

$$\mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{e} \quad (3.124)$$

čo najmenšie. Pripomeňme, že spomínané odchýlky skutočných a vyrovnaných hodnôt označujeme ako rezíduá. Našou snahou by potom mohlo byť minimalizovať rezíduá.

Musíme však byť presnejší – zrejme ak by sme sa snažili napríklad o minimalizáciu súčtu prvkov  $\mathbf{e}$ , alebo minimalizáciu priemeru rezíduí, model by asi nespĺňal naše predstavy – ak by bola priemerná odchýlka  $-10^{400}$ , čo je veľmi veľké záporné číslo, je zrejme, že chyba odhadu by bola obrovská, aj keď odchýlka by bola záporná (teda „malá“ – bola by menšia ako nula).

Taktiež si uvedomme, aký by malo dôsledok, ak by sme za vyrovnané hodnoty dosadili priemernú hodnotu zložiek  $\mathbf{y}$ . Súčet rezíduí sa dá maticovo zapísať aj takto:

$$\sum_{i=1}^n e_i = \mathbf{1}^T \mathbf{e} \quad (3.125)$$

V tomto vzťahu označujeme vektor rezíduí  $\mathbf{e} = (e_1, e_2, \dots, e_n)$ ,  $e_i \in \mathbb{R}$  pre  $i = 1, 2, \dots, n$ , počet pozorovaní je  $n \in \mathbb{N}$  a vektor  $\mathbf{1}$  je vektor s  $n$  zložkami, všetkými rovný jednej (keďže pojmom jednotkový vektor sa v lineárnej algebre myslí iný typ vektora, budeme  $\mathbf{1}$  nazývať vektorom jednotiek).

Odtiaľ priemerná hodnota rezíduí je:

$$\frac{1}{n} \sum_{i=1}^n e_i = \frac{1}{n} \mathbf{1}^T \mathbf{e} \quad (3.126)$$

Ak by sme za  $\hat{\mathbf{y}}$  zvolili vektor:

$$\bar{\mathbf{y}} = \bar{y} \mathbf{1} = \left( \frac{1}{n} \sum_{i=1}^n y_i \right) \mathbf{1} \quad (3.127)$$

teda vektor, ktorý má  $n$  zložiek a každá z nich je rovná priemernej hodnote zložiek vektora  $\mathbf{y}$ , potom by sme pre súčet odchýlok rezíduí dostali:

$$\mathbf{1}^T \mathbf{e} = \mathbf{1}^T (\mathbf{y} - \hat{\mathbf{y}}) \quad (3.128)$$

$$= \mathbf{1}^T (\mathbf{y} - \bar{\mathbf{y}}) \quad (3.129)$$

$$= \mathbf{1}^T (\mathbf{y} - \bar{y} \mathbf{1}) \quad (3.130)$$

V druhom riadku sme za  $\hat{\mathbf{y}}$  dosadili  $\bar{\mathbf{y}}$ , za čo sme zasa v treťom riadku dosadili  $\bar{y} \mathbf{1}$ . Po roznásobení dostávame:

$$\mathbf{1}^T (\mathbf{y} - \bar{y} \mathbf{1}) = \mathbf{1}^T \mathbf{y} - \mathbf{1}^T \bar{y} \mathbf{1} \quad (3.131)$$

$$= \mathbf{1}^T \mathbf{y} - \bar{y} \mathbf{1}^T \mathbf{1} \quad (3.132)$$

$$= \mathbf{1}^T \mathbf{y} - \bar{y} n \quad (3.133)$$

V druhom riadku sme využili fakt, že násobenie matíc reálnym číslom je asociatívne a komutatívne. V poslednom riadku sme využili skutočnosť, že:

$$\mathbf{1}^T \mathbf{1} = (1, 1, \dots, 1) \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = n \quad (3.134)$$

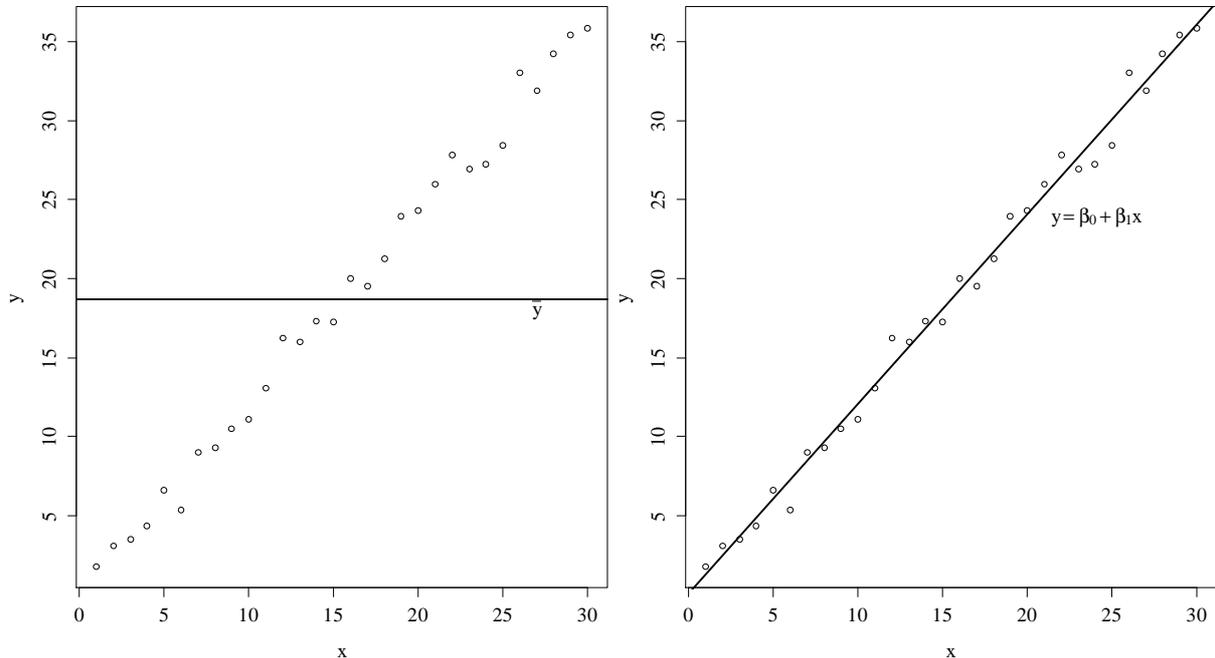
Keďže  $\mathbf{1}^T \mathbf{y} = n \bar{y}$ , dostávame:

$$\mathbf{1}^T \mathbf{y} - \bar{y} n = n \bar{y} - \bar{y} n = 0 \quad (3.135)$$

potom ale:

$$\mathbf{1}^T \mathbf{e} = \mathbf{1}^T (\mathbf{y} - \hat{\mathbf{y}}) = 0 \quad (3.136)$$

Voľba  $\hat{y} = \bar{y}$  teda vedie k nulovým priemerným odchýlkam, nulovým priemerným rezíduám. Takýto výsledok znie pomerne atraktívne – nulové priemerné odchýlky znamenajú, že v priemere dostávame „správne výsledky“.



Obrázok 7: Odhad pomocou priemeru (vľavo) a pomocou regresnej priamky (vpravo)

*Zdroj: vlastné spracovanie*

```

> set.seed(12345)
> par(family = "serif")
> par(mar = c(5, 4, 0.3, 0.3))
> par(mfrow = c(1, 2))
> x <- 1:30
> y <- 1.2*x + rnorm(length(x), 0, 1)
> plot(x, y)
> abline(h = mean(y), lwd = 2)
> text(27, 18, expression(bar("y")))
> plot(x, y)
> abline(lm(y ~ x), lwd = 2)
> text(24, 24, expression(y == beta[0] + beta[1] * x))

```

Obrázok 7 nám zobrazuje, ako by ale takáto voľba mohla graficky vyzerat'. V ľavej časti obrázku je zobrazený x-y graf údajov, spolu s konštantnou priamkou zodpovedajúcou  $\bar{y}$ . Je vidieť, že priemerná hodnota nevyjadruje závislosť medzi premennými x a y. Pravá časť obrázku popisuje rovnaké údaje, ale tentoraz je využitá regresná priamka (určená metódou najmenších štvorcov, ktorej sa budeme venovať za chvíľu), nie konštanta. V tomto prípade je vyrovnanie o dosť zmyslupnejšie – našim cieľom preto nebude uspokojiť sa s priemernou

hodnotou, ale pokúsiť sa odhadnúť regresnú priamku, ktorá by popisovala vzťah medzi premennými lepšie.

Nulové priemerné rezíduá sme dostali v prípade, ak sme použili na vyrovnanie priemer závislej premennej z toho dôvodu, že ako z obrázku vidno, časť údajov je pod priamkou zodpovedajúcou priemeru a časť je nad ňou. Dochádza k tomu, že odchýlky  $y - \bar{y}$  sú v ľavej časti grafu záporné a v pravej časti kladné. Pre malé  $x$  sa hodnoty  $y$  nachádzajú pod priemerom, pre vysoké nad ním. Z tohto dôvodu sa pri výpočte priemernej odchýlky kompenzujú kladné odchýlky so zápornými, a priemerná odchýlka je nulová.

Namiesto stratégie, pri ktorej požadujeme malú priemernú odchýlku by sme mohli požadovať, aby sa akákoľvek odchýlka (kladná či záporná) považovala za neželanú odchýlku – aby sa odchýlky medzi sebou nemohli kompenzovať. Aby sme vylúčili podobný prípad, mohli by sme minimalizovať súčet absolútnych hodnôt rezíduí:

$$\sum_{i=1}^n |e_i| \quad (3.137)$$

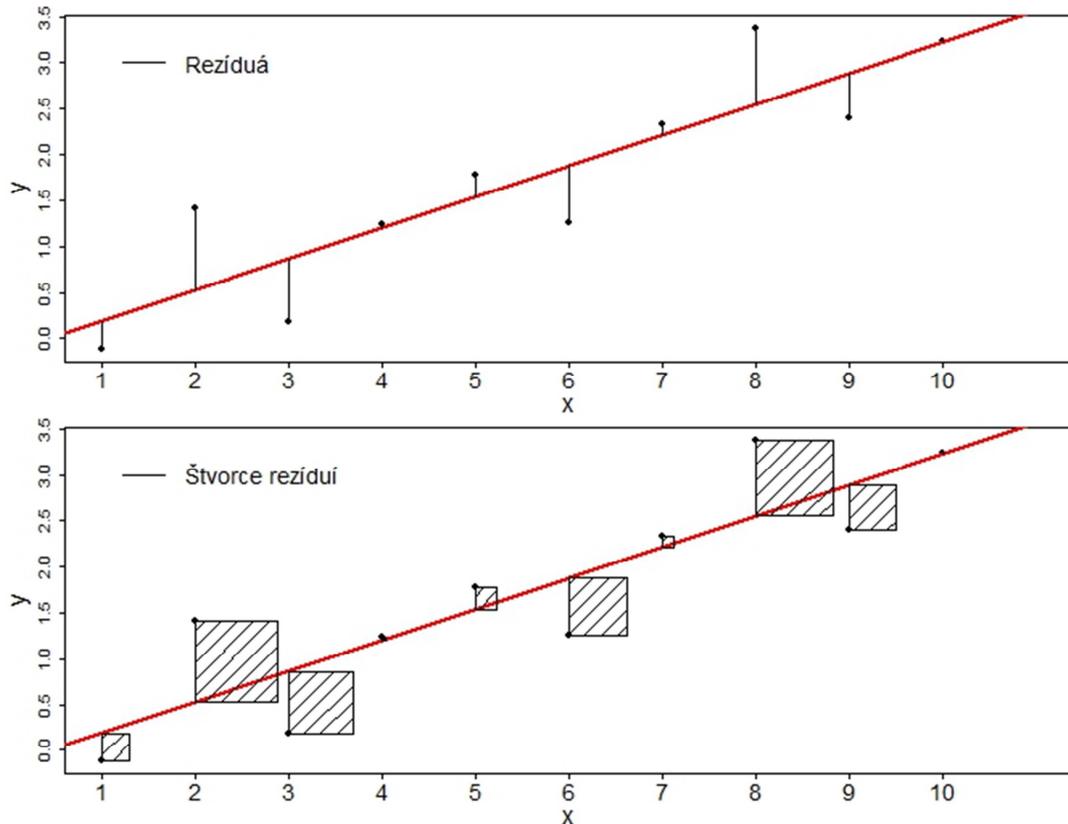
Tento spôsob – odhad regresných koeficientov  $\beta$  založený na minimalizácii súčtu (alebo priemeru) absolútnych odchýlok rezíduí – sa naozaj zvykne využívať, a používa sa preň označenie *robustná regresia* (angl. *robust regression*). Je využiteľný hlavne v prípade, ak sú porušené predpoklady modelu (konkrétne tzv. heteroskedasticita), alebo v prípade, ak sú v skúmanom súbore prítomné extrémne hodnoty, ktoré z nejakých dôvodov nie je vhodné odstrániť a pokračovať v analýze bez nich.

Najčastejšie využívaným kritériom je však minimalizácia súčtu štvorcov (druhých mocnín) rezíduí – metóda sa preto volá metóda najmenších štvorcov (angl. *ordinary least squares*, OLS). Kritérium v tomto prípade môžeme vyjadriť ako:

$$\sum_{i=1}^n e_i^2 = (e_1, e_2, \dots, e_n) \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix} = \mathbf{e}^T \mathbf{e} \quad (3.138)$$

Pri metóde najmenších štvorcov požadujeme, aby sme vybrali model, pri ktorom je súčet štvorcov rezíduí najmenší. Ako taký model nájdeme? Uvedomme si, že na to, aby sme popísali model v požadovanom tvare nám stačí nájsť vektor  $\hat{\beta}$  - týmto vektorom je úplne daný vzťah vysvetľujúcich a vysvetľovaných premenných. V skutočnosti je vektor rezíduí  $\mathbf{e}$  funkciou neznámeho vektora  $\hat{\beta}$  – v závislosti od toho, aké koeficienty zvolíme, dostaneme

veľké alebo malé rezíduá. Podmienka o minimalizácii štvorcov rezíduí je preto podmienkou týkajúcou sa odhadu regresných parametrov. Názorne si rezíduá a štvorce rezíduí môžeme ukázať na nasledujúcom obrázku (Obrázok 8).



Obrázok 8 Vizualizácia štvorcov rezíduí

Zdroj: vlastné spracovanie v programe R

```
> x <- 1:10
> y <- x*0.3 + round(rnorm(10, 0, 0.5), 2)
> dev.off()
> par(mfrow = c(2,1))
> par(cex = 0.6)
> par(mar = c(2.4, 0, 1.2, 0.1), oma = c(2, 4.0, 0.5, 0.5))
> par(tcl = -0.25)
> par(mgp = c(2, 0.6, 0))
> model <- lm(y~x)
> plot(y~x, pch = 19, xlim = c(1, 11), xaxt = "n")
> abline(model, col = "red", lwd = 2)
> segments(x0 = 1:10, y0 = fitted(model), x1 = 1:10, y1 = y, col = "black", lty = 1, lwd = 1)
> legend(y = 3.0, x = 1, yjust = 0.5, cex = 1.5, legend = c("Rezíduá"), col = "black", lwd = 1.5, bty = "n")
> mtext("y", side = 2, line = 1.9)
> mtext("x", side = 1, line = 1.9)
> axis(1, cex.axis = 1.5, at = 1:10)
> plot(y~x, pch = 19, xlim = c(1, 11), xaxt = "n")
> abline(model, col = "red", lwd = 2)
```

```

> rect(1:10, fitted(model), 1:10+abs(resid(model)), y, density
= 10, angle = 45, col = NULL, border = TRUE, lty = 1, lwd = 1)
> legend(y = 3.0, x = 1, yjust = 0.5, cex = 1.5, legend =
c("Štvorce rezíduí"), col = "black", lwd = 1.5, bty = "n")
> mtext("y", side = 2, line = 1.9)
> mtext("x", side = 1, line = 1.9)
> axis(1, cex.axis = 1.5, at = 1:10)

```

Názorne si to môžeme ukázať ak využijeme definíciu rezíduí z rovnice  $\mathbf{y} = \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{e}$ .

Odtiaľ:

$$\mathbf{e} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}} \quad (3.139)$$

Po dosadení do účelovej funkcie dostávame:

$$\mathbf{e}^T \mathbf{e} = (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})^T (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \quad (3.140)$$

Ďalej použijeme pravidlá pre operácie s maticami, ktoré sme popísali v podkapitolách 1.3.1 až 1.3.3. Najprv transponujeme matice v prvej zátvorke a dostaneme:

$$(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})^T (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) = (\mathbf{y}^T - \hat{\boldsymbol{\beta}}^T \mathbf{X}^T)(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \quad (3.141)$$

Výrazy v zátvorkách roznásobíme, pričom je treba dávať pozor na to, že násobenie matíc nie je komutatívne – rozlišujeme násobenie sprava a násobenie zľava, poradie matíc preto nemôžeme ľubovoľne meniť:

$$(\mathbf{y}^T - \hat{\boldsymbol{\beta}}^T \mathbf{X}^T)(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) = \mathbf{y}^T \mathbf{y} - \mathbf{y}^T \mathbf{X}\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}^T \mathbf{X}^T \mathbf{y} + \hat{\boldsymbol{\beta}}^T \mathbf{X}^T \mathbf{X}\hat{\boldsymbol{\beta}} \quad (3.142)$$

Všimnime si teraz matice  $\mathbf{y}^T \mathbf{X}\hat{\boldsymbol{\beta}}$  a  $\hat{\boldsymbol{\beta}}^T \mathbf{X}^T \mathbf{y}$ . Výrazy nie len že obsahujú tie isté vektory, ale môžeme si všimnúť že platí:

$$(\mathbf{y}^T \mathbf{X}\hat{\boldsymbol{\beta}})^T = \hat{\boldsymbol{\beta}}^T \mathbf{X}^T \mathbf{y} \quad (3.143)$$

Matice dané predchádzajúcimi súčinnami sú navzájom transponované. V skutočnosti toho majú spoločného ešte viac – aby sme sa o tom presvedčili, overme rozmery (dimenziu) oboch matíc.

V predchádzajúcich podkapitolách sme si zaviedli maticový zápis pre špecifikáciu modelu. Ukázali sme si, že matica  $\mathbf{X}$  je rozmeru  $n \times (k + 1)$ , vektor  $\mathbf{y}$  má rozmer  $n \times 1$  a vektor  $\hat{\boldsymbol{\beta}}$  má rozmer  $(k + 1) \times 1$ . Potom súčin  $\mathbf{y}^T \mathbf{X}\hat{\boldsymbol{\beta}}$  je súčinom matíc/vektorov s rozmermi  $1 \times n$ ,  $n \times (k + 1)$  a  $(k + 1) \times 1$ . Po vynásobení matíc dostávame „maticu“ rozmeru  $1 \times 1$ , čiže obyčajné reálne číslo.

V prípade súčiny  $\hat{\boldsymbol{\beta}}^T \mathbf{X}^T \mathbf{y}$  násobíme matice/vektory o rozmeroch  $1 \times (k + 1)$ ,  $(k + 1) \times n$  a  $n \times 1$ . Výsledkom je opäť reálne číslo. Keďže obidva súčiny majú za výsledok reálne číslo, a sú jeden k druhému transponované, ide v skutočnosti o to isté číslo. Platí preto:

$$\mathbf{y}^T \mathbf{X} \hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}^T \mathbf{X}^T \mathbf{y} \quad (3.144)$$

Ak sa vrátíme ku vzťahu (3.142) a dosadíme za výraz  $\mathbf{y}^T \mathbf{X} \hat{\boldsymbol{\beta}}$  výraz  $\hat{\boldsymbol{\beta}}^T \mathbf{X}^T \mathbf{y}$ , dostaneme:

$$\mathbf{y}^T \mathbf{y} - \mathbf{y}^T \mathbf{X} \hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}^T \mathbf{X}^T \mathbf{y} + \hat{\boldsymbol{\beta}}^T \mathbf{X}^T \mathbf{X} \hat{\boldsymbol{\beta}} = \mathbf{y}^T \mathbf{y} - 2\hat{\boldsymbol{\beta}}^T \mathbf{X}^T \mathbf{y} + \hat{\boldsymbol{\beta}}^T \mathbf{X}^T \mathbf{X} \hat{\boldsymbol{\beta}} \quad (3.145)$$

Sériou ekvivalentných úprav sme dostali rovnosť:

$$\mathbf{e}^T \mathbf{e} = \mathbf{y}^T \mathbf{y} - 2\hat{\boldsymbol{\beta}}^T \mathbf{X}^T \mathbf{y} + \hat{\boldsymbol{\beta}}^T \mathbf{X}^T \mathbf{X} \hat{\boldsymbol{\beta}} \quad (3.146)$$

Naším cieľom bolo minimalizovať  $\mathbf{e}^T \mathbf{e}$ , ak meníme  $\hat{\boldsymbol{\beta}}$ . Pri hľadaní extrémov sa bežne používajú prvé derivácie, ktoré kladieme rovné nule (nutná podmienka existencie extrému). V našom prípade by išlo o parciálne derivácie – celkovo  $k + 1$  parciálnych derivácií, ktoré by sme museli položiť rovné nule a riešiť ako systém rovníc. Môžeme však efektívne využiť poznatky o maticových deriváciách, ktoré sme uvádzali v kapitole 1.3.6. Chceme vypočítať:

$$\frac{\partial \mathbf{e}^T \mathbf{e}}{\partial \hat{\boldsymbol{\beta}}} \quad (3.147)$$

Môžeme postupovať po jednotlivých výrazoch. Derivácia  $\mathbf{y}^T \mathbf{y}$  podľa  $\hat{\boldsymbol{\beta}}$  bude nulovým vektorom – hodnoty  $\mathbf{y}$  sú hodnoty získané zo vzorky a nie sú funkciou  $\hat{\boldsymbol{\beta}}$ . Ide o situáciu podobnú tej, keď derivujeme konštantu – výsledok je nulový.

Druhým výrazom je  $-2\hat{\boldsymbol{\beta}}^T \mathbf{X}^T \mathbf{y}$ . Podľa kapitoly 1.3.6. platí:

$$\frac{\partial (-2\hat{\boldsymbol{\beta}}^T \mathbf{X}^T \mathbf{y})}{\partial \hat{\boldsymbol{\beta}}} = -2\mathbf{X}^T \mathbf{y} \quad (3.148)$$

Posledným výrazom je  $\hat{\boldsymbol{\beta}}^T \mathbf{X}^T \mathbf{X} \hat{\boldsymbol{\beta}}$ , čo je vlastne kvadratická forma. Matica, ktorú dostaneme súčinom  $\mathbf{X}^T \mathbf{X}$  je symetrická matica, čo môžeme jednoducho overiť:

$$(\mathbf{X}^T \mathbf{X})^T = \mathbf{X}^T \mathbf{X} \quad (3.149)$$

Pre deriváciu potom platí:

$$\frac{\partial (\hat{\boldsymbol{\beta}}^T \mathbf{X}^T \mathbf{X} \hat{\boldsymbol{\beta}})}{\partial \hat{\boldsymbol{\beta}}} = 2\mathbf{X}^T \mathbf{X} \hat{\boldsymbol{\beta}} \quad (3.150)$$

Pre deriváciu  $\mathbf{e}^T \mathbf{e}$  potom dostávame:

$$\frac{\partial \mathbf{e}^T \mathbf{e}}{\partial \hat{\boldsymbol{\beta}}} = -2\mathbf{X}^T \mathbf{y} + 2\mathbf{X}^T \mathbf{X} \hat{\boldsymbol{\beta}} \quad (3.151)$$

Pre splnenie nutnej podmienky existencie extrémumu majú byť parciálne derivácie rovné nulám, čiže:

$$-2\mathbf{X}^T \mathbf{y} + 2\mathbf{X}^T \mathbf{X} \hat{\boldsymbol{\beta}} = \mathbf{0} \quad (3.152)$$

Odtiaľ:

$$\mathbf{X}^T \mathbf{X} \hat{\boldsymbol{\beta}} = \mathbf{X}^T \mathbf{y} \quad (3.153)$$

V poslednom kroku sme dostali maticovú rovnicu, v ktorej by sme potrebovali osamostatniť vektor  $\hat{\boldsymbol{\beta}}$ . Pokiaľ je matica  $\mathbf{X}^T \mathbf{X}$  regulárna, existuje k nej inverzná matica, ktorou by sme celú rovnicu mohli prenásovať zľava. Toto je jeden z kľúčových krokov pri postupe stanovenia odhadu regresného modelu – pokračovať môžeme len v prípade, ak matica  $\mathbf{X}^T \mathbf{X}$  nie je singularná. Tento predpoklad budeme uvádzať aj neskôr pri popisovaní podmienok, ktoré na model kladieme. V texte sa budeme zaoberať len prípadmi, ak inverzná matica existuje. Na tomto mieste to pripomínáme hlavne preto, že tu môžeme jasne vidieť dôsledky nesplnenia tohto predpokladu – ak je matica  $\mathbf{X}^T \mathbf{X}$  singularná, nedokážeme jednoznačne vypočítať hodnotu odhadov regresných koeficientov, o čom nás môže informovať aj výstup zo štatistického software.

Pokiaľ je matica  $\mathbf{X}^T \mathbf{X}$  regulárna, môžeme prenásovať zľava celú rovnicu maticou k nej inverznou:

$$(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{X} \hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \quad (3.154)$$

Súčinom navzájom inverzných matic  $(\mathbf{X}^T \mathbf{X})^{-1}$  a  $\mathbf{X}^T \mathbf{X}$  bude jednotková matica. Akého bude rádu? Keďže matica  $\mathbf{X}$  je rádu  $n \times (k + 1)$ , potom je výsledok súčinu  $\mathbf{X}^T \mathbf{X}$  štvorcovou maticou rádu  $(k + 1) \times (k + 1)$ . Inverzná matica k  $\mathbf{X}^T \mathbf{X}$  bude tiež štvorcová matica rovnakého rádu. Budeme teda mať  $(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{X} = \mathbf{I}_{k+1}$ . Keďže z vlastností jednotkovej matice  $\mathbf{I}_{k+1} \hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}$ , z rovnice dostávame:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \quad (3.155)$$

Uvedený vzťah je konečne výrazom, ktorý vieme využiť na výpočet odhadu regresných koeficientov  $\hat{\boldsymbol{\beta}}$ . Umožňuje nám vypočítať koeficienty  $\hat{\boldsymbol{\beta}}$ , ktoré majú vlastnosť, že vedú k najmenšej sume štvorcov rezíduí modelu. Obrázok 7 uvedený vyššie obsahuje v pravej časti zobrazenie priamky, ktorá bola vypočítaná práve pomocou metódy najmenších štvorcov.

### 3.3.3 Predpoklady lineárneho modelu

V predchádzajúcej podkapitole sme si ukázali odvodenie vzťahu, ktorý je možné použiť na odhad regresných koeficientov lineárneho modelu. Ako bolo naznačené v kapitole venovanej formulovaniu ekonometrických modelov, veľmi často odhadujeme ekonometrické modely nie len za účelom výpočtu ich koeficientov, ale aj za účelom testovania hypotéz o existencii, resp. neexistencii vzťahov medzi vysvetľovanou premennou a vysvetľujúcimi premennými. Odhad pomocou metódy najmenších štvorcov nekládol na výpočet žiadne zásadné podmienky – jediným predpokladom, s ktorým sme sa zatiaľ stretli, bola regulárnosť matice  $\mathbf{X}^T\mathbf{X}$ . V prípade, ak chceme ekonometrický model použiť aj na účely indukčnej štatistiky, t. j. na overovanie hypotéz a zovšeobecňovanie výsledkov zo vzorky na populáciu (čo sa deje vo väčšine prípadov), je potrebné na model klásť dodatočné predpoklady.

*Prvým predpokladom* sú nulové stredné hodnoty zložiek vektora náhodných porúch  $\mathbf{u}$ . Tento predpoklad môžeme zapísať nasledovne:

$$E(\mathbf{u}) = \mathbf{0} \quad (3.156)$$

Požiadavka na nulovú strednú hodnotu zložiek vektora náhodných porúch sa dá vysvetliť tak, že očakávame, že hodnoty  $\mathbf{u}$  systematickým spôsobom neovplyvňujú hodnoty  $\mathbf{y}$ . Niekedy sú hodnoty poruchového člena kladné, niekedy záporné, ale v priemere je ich hodnota nulová. Dôsledky porušenia tohto predpokladu nebývajú veľmi vážne – ak špecifikujeme ekonometrický model tak, že obsahuje aj konštantu ( $\beta_0$ ), jej odhad bude skreslený. V ekonometrických modeloch veľmi často konštantu  $\beta_0$  nemá významnú úlohu. Väčšinou analyzujeme vzťahy vysvetľujúcich veličín a vysvetľovanej, pričom konštantný člen nemá kľúčovú interpretáciu.

Ďalším dôvodom, pre ktorý niekedy pre nás konštantný člen nie je veľmi smerodajný je ten, že v prípade, ak nadobúdajú vysvetľujúce premenné veľké hodnoty, nebýva odhad konštantného člena veľmi presný. Uvedme si hypotetický príklad pre takúto situáciu.

Povedzme, že skúmame ziskovosť stavebných firiem (závislá premenná) od počtu kilometrov diaľnic, ktoré firmy postavili za posledných 10 rokov. Nech v našom príklade firma s najmenším počtom kilometrov postavila 80 km diaľnic a naopak maximum nech bolo 120 km. O čom by vypovedal konštantný člen?

Pripomeňme, že interpretácia konštantného člena súvisí z geometrického hľadiska s priesečníkom s osou  $y$  v prípade, ak je vysvetľujúca premenná nulová. V našom príklade by to znamenalo, že koeficient  $\beta_0$  predstavuje odhad ziskovosti stavebnej firmy, ktorá nepostavila

za posledných 10 rokov ani kilometer diaľnic. Uvedomme si, vzorka pre takúto firmu nie je reprezentatívna – vo vzorke máme len firmy, ktoré diaľnice stavajú. O firmách, ktoré ich nestavajú, zo vzorky nevieme nič. Aj náš odhad zisku firiem, ktoré nestavajú diaľnice, preto na základe našej vzorky asi nebude veľmi spoľahlivý. Ak sme si tejto skutočnosti vedomí vopred, nebudeme zrejme konštantnému členu prikladať veľkú váhu. Prvý predpoklad preto nie je kritický.

*Druhý predpoklad* kladený na lineárny model je možné vyjadriť algebraicky nasledovne:

$$E(\mathbf{u}\mathbf{u}^T) = \sigma^2 \mathbf{I}_n \quad (3.157)$$

kde  $\sigma^2 \in \mathbb{R}^+$ .

Kým prvý predpoklad nebol z hľadiska využívania lineárnych modelov zásadný, druhý predpoklad je kľúčový a budeme mu venovať značný priestor. Kým sa dostaneme k vysvetleniu jeho dôležitosti, musíme najprv vysvetliť, čo vlastne hovorí.

Pripomeňme (kapitola 3.1.3), že variančno-kovariančnú maticu náhodného vektora  $\mathbf{u}$  môžeme zapísať takto:

$$\text{var}(\mathbf{u}) = E\left((\mathbf{u} - E(\mathbf{u}))(\mathbf{u} - E(\mathbf{u}))^T\right) \quad (3.158)$$

Keďže podľa prvého predpokladu pre lineárny model uvedeného vyššie máme  $E(\mathbf{u}) = \mathbf{0}$ , platí:

$$\text{var}(\mathbf{u}) = E\left((\mathbf{u} - E(\mathbf{u}))(\mathbf{u} - E(\mathbf{u}))^T\right) = E\left((\mathbf{u} - \mathbf{0})(\mathbf{u} - \mathbf{0})^T\right) = E(\mathbf{u}\mathbf{u}^T) \quad (3.159)$$

Z toho vidíme, že druhý predpoklad lineárneho modelu je vlastne výrokom o variančno-kovariančnej matici  $\text{var}(\mathbf{u})$ . Druhý predpoklad hovorí, že variančno-kovariančná matica vektora  $\mathbf{u}$  má tento tvar:

$$\text{var}(\mathbf{u}) = \begin{pmatrix} \sigma^2 & 0 & \dots & 0 \\ 0 & \sigma^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma^2 \end{pmatrix} \quad (3.160)$$

Pripomeňme, že prvky hlavnej diagonály variančno-kovariančnej matice vektora  $\mathbf{u}$  predstavujú rozptyly zložiek tohto vektora. Keďže prvky na hlavnej diagonále sú rovnaké, druhý predpoklad požaduje, aby poruchové členy mali rovnaký rozptyl. Túto požiadavku môžeme považovať za bežnú pri využívaní indukčnej štatistiky – keďže štatistika vie skúmať len hromadné javy, je potrebné, aby všetky prvky  $\mathbf{u}$  mali rovnaký rozptyl.

V predchádzajúcej časti sme si vysvetlili, že pri kvantifikovaní ekonometrických modelov potrebujeme využívať aparát induktívnej štatistiky, pretože pri skúmaní na vzorke potrebujeme rozlíšiť dva efekty: efekt, ktorý je daný vzťahom vysvetľovanej a vysvetľujúcich premenných, a efekt náhodného člena. Určite by bolo nežiaduce prehlásiť, že medzi premennými existuje vzťah, keď v skutočnosti ide o náhodu. Aby sme posúdili vplyv náhodného poruchového člena, potrebujeme sa o ňom „dozvedieť viac“ – pokiaľ má každá zložka náhodného člena ten istý rozptyl, dáva nám to možnosť odhadnúť jeho veľkosť a s prihliadnutím k nemu posúdiť, či výsledky dosiahnuté kvantifikáciou modelu môžu byť výsledkom náhody. Požiadavka na konštantný rozptyl prvkov  $\mathbf{u}$  je dôležitá hlavne z hľadiska možnosti posúdenia významnosti výsledkov modelu. Požiadavka na konštantný rozptyl zložiek vektora  $\mathbf{u}$  sa označuje aj ako podmienka homoskedasticity.

V druhom predpoklade je však zahrnutá ešte jedna skutočnosť. Pri definícii variančno-kovariančnej matice sme uviedli, že mimodiagonálne prvky variančno-kovariančnej matice predstavujú kovariancie medzi zložkami náhodného vektora. Keďže podľa druhého predpokladu sú mimodiagonálne prvky nulové, znamená to, že požadujeme, aby boli nekorelované.

Požiadavka na nekorelovanosť má rôzny význam pri prierezových údajoch a časových radoch. Pri prierezových údajoch skúmame hodnoty za rôzne štatistické jednotky v tom istom čase. Nekorelovanosť pri prierezových údajoch požaduje, aby náhodná chyba, ktorá sa vyskytne u jednej štatistickej jednotky nebola korelovaná s chybou, ktorá vznikne pri inej štatistickej jednotke.<sup>1</sup> Pri časových radoch je situácia odlišná – tam jednotlivé zložky vektora  $\mathbf{u}$  súvisia s vývojom v čase. Tu je ľahké si predstaviť, že ak existuje v nejakom čase náhodná chyba, chyba v nasledujúcom pozorovaní môže veľmi ľahko súvisieť s predchádzajúcou, čo by bolo porušením nášho predpokladu. Pri definícii náhodného poruchového člena sme uviedli, že ho vytvárajú všetky ostatné faktory, ktoré sme nezahrnuli do modelu. Ak by sme do modelu nepridali nejakú premennú, ktorá je podstatná, jej efekt by sa mohol prejaviť v poruchovom člene. Napríklad by sa mohlo stať, že záporné hodnoty poruchového člena sú sprevádzané aj následnými zápornými hodnotami, a kladné hodnoty systematicky ďalšími kladnými hodnotami. Kým autokorelácia je pri ekonomických časových radoch veľmi bežný jav, lineárny model tak, ako sme si ho opísali, predpokladá jej absenciu pri poruchovom člene. Z toho vyplýva, že tento model nie je veľmi vhodný pre väčšinu časových radov – toto

---

<sup>1</sup> Tomuto problému sa hovorí aj priestorová korelácia, avšak v tejto publikácii sa tomu problému nebudeme bližšie venovať.

je jeden z hlavných dôvodov, prečo predstavuje analýza časových radov pomerne veľký samostatný celok v rámci ekonometrie. Vznikol z potreby pracovať s modelmi, pri ktorých je tento predpoklad často porušený.

*Tretím predpokladom* je požiadavka, aby bola matica  $\mathbf{X}$  deterministická. Doposiaľ sme sa pri popise modelu zaoberali len jednou stochastickou veličinou, ktorou bol náhodný poruchový člen  $\mathbf{u}$  (v skutočnosti bol v dôsledku náhodnosti  $\mathbf{u}$  náhodným aj vektor  $\mathbf{y}$ ,  $\mathbf{e}$  a  $\hat{\mathbf{y}}$ ). Tretí predpoklad požaduje, aby náhodnou nebola matica  $\mathbf{X}$ . Čo to znamená, ak matica  $\mathbf{X}$  nie je náhodná, si môžeme vysvetliť na veľmi zjednodušenom príklade. Povedzme, že vo vzorke máme niekoľko podnikov (alebo respondentov). Matica  $\mathbf{X}$  obsahuje hodnoty premenných, ktoré na nich vo vzorke sledujeme – ich vybrané charakteristiky. Predpoklad o tom, že matica  $\mathbf{X}$  nie je náhodná znamená aj to, že ak by sme sa konkrétneho respondenta opýtali na určitú jeho charakteristiku (napríklad výsledok hospodárenia za rok 2012), dostaneme tú istú odpoveď, a to bez ohľadu na to, koľko krát by sme sa pýtali – odpoveď by bola stále rovnaká, konštantná (deterministická). Ak sa respondenta opýtame na rok jeho narodenia, predpokladáme, že ani tento sa nebude meniť, aj keby sme sa naň pýtali opakovane.

Tretí predpoklad hovorí, že pri modeli neuvažujeme o situáciách, kedy by sa tá istá charakteristika u toho istého respondenta mohla meniť – bola by náhodná, keďže by sme nevedeli, aká bude odpoveď, ak by sme sa pýtali opakovane. Samozrejme, aj takéto situácie sa v realite stávajú a nie je ťažké vymyslieť kontrapríklad. Model, ktorým sa budeme zaoberať ale vychádza z deterministickej matice  $\mathbf{X}$ .<sup>2</sup>

*Štvrtý predpoklad* sa opäť týka matice  $\mathbf{X}$ . Môžeme ho vyjadriť nasledovne:

$$h(\mathbf{X}) = k + 1 \quad \wedge \quad k + 1 \leq n \quad (3.161)$$

Predpoklad má opäť dve časti: prvá časť hovorí, že hodnosť matice  $\mathbf{X}$  je  $k + 1$ , kde  $k$  je počet vysvetľujúcich premenných. S dôvodom pre tento predpoklad sme sa už stretli – ak je hodnosť matice  $\mathbf{X}$  rovná  $k + 1$ , bude súčin  $\mathbf{X}^T \mathbf{X}$  regulárnou maticou, tým pádom bude k  $\mathbf{X}^T \mathbf{X}$  existovať inverzná matica a následne bude možné vypočítať odhad vektora  $\hat{\boldsymbol{\beta}}$  ako  $\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}$ . Podobne druhá časť predpokladu hovorí, že počet odhadovaných parametrov nemôže byť väčší ako je počet pozorovaní. Tento predpoklad je možné vždy ľahko overiť a je na nás, aby sme analýzu realizovali tak, aby predpoklad nebol porušený.

---

<sup>2</sup> V skutočnosti ekonometria rozoberá aj situáciu, ak je  $\mathbf{X}$  náhodná matica – pokiaľ sú splnené určité podmienky, výsledky sú takmer totožné s tými, ktoré získame pre deterministickú maticu  $\mathbf{X}$ . V podobnej forme môže platiť napríklad aj tzv. Gauss-Markovova podmienka. Rozoberať preto budeme len prípad s konštantnou maticou  $\mathbf{X}$ .

Uvedme si jeden spôsob, ako môže dôjsť k porušeniu tohto predpokladu. Povedzme, že do modelu by sme chceli dať tri premenné za každého respondenta – počet synov, počet dcér a počet detí celkom. V takomto prípade  $k = 3$  a hodnosť matice  $\mathbf{X}$  by mala byť  $k + 1 = 4$ . My však vieme, že hodnosť matice bude menšia – je zrejmé, že stĺpec matice zodpovedajúci počtom detí spolu je súčtom predchádzajúcich dvoch stĺpcov (počet synov a dcér) a teda tieto tri stĺpce sú lineárne závislé. Štvrtý predpoklad takto nie je splnený – matica  $\mathbf{X}^T\mathbf{X}$  bude singulárna a odhad  $\hat{\boldsymbol{\beta}}$  nebude možné vypočítať.

Podobný príklad vzniká, ak by sme chceli na vzorke podnikov použiť ako vysvetľujúce premenné výšku výnosov, výšku nákladov a výšku zisku. Keďže zisk získame ako výnosy mínus náklady, stĺpce matice  $\mathbf{X}$  budú lineárne závislé a model opäť nebude možné vypočítať.

Kým predchádzajúce predpoklady predstavujú skutočnosti, nad ktorými máme pri výskume malú kontrolu, splnenie tohto predpokladu jednoznačne spadá na tvorcu modelu. Keďže nám prípadné porušenie modelu štatistický softvér vždy oznámi, je jednoduché chybu opraviť – spravidla je potrebné vynechať niektorú z premenných, ktoré sú lineárne závislé.

Problém s druhou časťou predpokladu je tiež v kompetencii tvorcu modelu – ak máme 10 pozorovaní, asi nebude možné odhadnúť hodnotu 20 koeficientov.

*Piatym predpokladom* je predpoklad o pravdepodobnostnom rozdelení vektora náhodných porúch  $\mathbf{u}$ . Keďže sa ekonometrické modely využívajú aj na účely induktívnej štatistiky, nie je prekvapivé, že vzniká potreba poznať pravdepodobnostné rozdelenie člena, ktorý spôsobuje stochastický charakter ostatných veličín. Piaty predpoklad hovorí, že:

$$\mathbf{u} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_n) \quad (3.162)$$

Náhodný vektor  $\mathbf{u}$  má podľa tohto predpokladu normálne rozdelenie s nulovým vektorom stredných hodnôt a variančno-kovariančnou maticou  $\sigma^2 \mathbf{I}_n$ . Tento posledný predpoklad využíva prvý a druhý predpoklad o vlastnostiach vektora  $\mathbf{u}$ , a pridáva nový o normálnom rozdelení.

Tabuľka 4 uvádza všetkých päť predpokladov spoločne. V ďalšom texte sa budeme na predpoklady modelu odkazovať číslami v tabuľke, teda P1 až P5.

Tabuľka 4: Predpoklady lineárneho modelu

Označenie	Predpoklad	Charakteristika
P1	$E(\mathbf{u}) = \mathbf{0}$	Nulový vektor stredných hodnôt $\mathbf{u}$
P2	$E(\mathbf{u}\mathbf{u}^T) = \sigma^2\mathbf{I}_n$	Nekorelovanosť a homoskedasticita $\mathbf{u}$
P3	$\mathbf{X}$ je deterministická	Deterministickosť $\mathbf{X}$
P4	$h(\mathbf{X}) = k + 1 \wedge k + 1 \leq n$	Hodnosť matice $\mathbf{X}$
P5	$\mathbf{u} \sim N(\mathbf{0}, \sigma^2\mathbf{I}_n)$	Normalita $\mathbf{u}$

Zdroj: vlastné spracovanie

### 3.3.4 Odhad lineárneho modelu metódou maximálnej vierohodnosti

Metóda najmenších štvorcov, ktorú sme si definovali v podkapitole 3.3.2 nevyžadovala žiadne zásadné predpoklady – v podstate sme pri nej museli urobiť iba rozhodnutie týkajúce sa funkcie, ktorú budeme minimalizovať ( $\mathbf{e}^T\mathbf{e}$ ). Z piatich predpokladov uvedených vyššie sme potrebovali len jeden – konkrétne predpoklad P4 o hodnosti matice  $\mathbf{X}$ .

Alternatívne je možné parametre modelu  $\hat{\boldsymbol{\beta}}$  odhadnúť aj inými metódami – napríklad metódou maximálnej vierohodnosti (angl. *maximum likelihood estimation*, MLE).

Odhadovať budeme rovnaký model ako doposiaľ, teda:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \quad (3.163)$$

Podľa predpokladu P5 vieme, že:

$$\mathbf{u} \sim N(\mathbf{0}, \sigma^2\mathbf{I}_n) \quad (3.164)$$

Keďže podľa predpokladu P3 je matica  $\mathbf{X}$  deterministická, a vektor  $\boldsymbol{\beta}$  je vektorom konštant, a nie náhodných premenných, poznáme aj pravdepodobnostné rozdelenie vektora  $\mathbf{y}$  – je ním rozdelenie:

$$\mathbf{y} \sim N(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I}_n) \quad (3.165)$$

Odhad metódou maximálnej vierohodnosti je založený práve na fakte, že poznáme typ pravdepodobnostného rozdelenia  $\mathbf{y}$ . V skutočnosti ak pracujeme so vzorkou, poznáme len typ rozdelenia, ale nie jeho parametre. Stále nepoznáme vektor  $\boldsymbol{\beta}$  a rozptyl chybového člena  $\sigma^2$ . Naším cieľom je ich zo vzorky odhadnúť. Odhadneme preto rozdelenie:

$$\mathbf{y} \sim N(\mathbf{X}\hat{\boldsymbol{\beta}}, \hat{\sigma}^2\mathbf{I}_n) \quad (3.166)$$

Ako však vybrať vhodné hodnoty  $\hat{\boldsymbol{\beta}}$  a  $\hat{\sigma}^2$ ? Nachvíľu zanedbajme, že normálne rozdelenie je spojité rozdelenie a uvažujme, ako by sme postupovali v prípade diskrétného rozdelenia. Pri diskrétnych rozdeleniach je možné pracovať s funkciou udávajúcou

pravdepodobnosť javov, s ktorou môžu nastať. V našom prípade je javom, ktorý nastal to, že vo vzorke sme pozorovali hodnoty  $\mathbf{y}$ , ktorým prislúchali určité hodnoty  $\mathbf{X}$ . Mohli by sme sa pýtať nasledovne: aké by museli byť hodnoty  $\hat{\boldsymbol{\beta}}$  a  $\hat{\sigma}^2$ , aby bola pravdepodobnosť javu, že vo vzorke získame dané hodnoty  $\mathbf{y}$  a  $\mathbf{X}$  najvyššia? Inak povedané, pri akých hodnotách  $\hat{\boldsymbol{\beta}}$  a  $\hat{\sigma}^2$  máme najväčšiu pravdepodobnosť, že dostaneme také pozorovania, aké sa objavili v našej vzorke?

Myšlienka odhadu pomocou maximálnej vierohodnosti je jednoduchá a zároveň veľmi elegantná. Ak by sme chceli postupovať podľa tejto metódy, musíme upustiť od nášho zjednodušujúceho predpokladu a čeliť skutočnosti, že musíme pracovať s normálnym rozdelením, ktoré nie je diskrétné. Pre spojité rozdelenia nemáme pravdepodobnosti, ako je tomu v prípade diskrétnych rozdelení. V ich prípade však môžeme pracovať s funkciou hustoty pravdepodobnosti, ktorá je jej určitou analógiou (táto analógia je trochu nepresná – našim cieľom je na tomto mieste skôr sprostredkovať princíp metódy maximálnej vierohodnosti, ako vysvetľovať rozdiely medzi diskrétnymi a spojitými rozdeleniami. Pre hlbšie pochopenie rozdielov odporúčame niektorú z učebníc matematickej štatistiky (Anděl, 1985). Hustota pravdepodobnosti normálneho rozdelenia má pre pozorovanie  $i$ , kde  $i = 1, 2, \dots, n$ ,  $n \in \mathbb{N}$  je počet pozorovaní a  $\mathbf{X}_i$  je  $i$ -tý riadok matice  $\mathbf{X}$  tvar:

$$f_i(y_i, \hat{\boldsymbol{\beta}}, \hat{\sigma}) = \frac{1}{\hat{\sigma}\sqrt{2\pi}} e^{-\frac{(y_i - \mathbf{X}_i \hat{\boldsymbol{\beta}})^2}{2\hat{\sigma}^2}} \quad (3.167)$$

Funkciu vierohodnosti dostaneme, ak vynásobíme hustoty za každé pozorovanie. Ak by sme pokračovali s našou (nepresnou) analógiou – ak by sme mali pravdepodobnosť toho, že sa v našej vzorke objavia hodnoty  $y_i$  a  $\mathbf{X}_i$  pre každé  $i$  a pozorovania boli na sebe nezávislé – potom pravdepodobnosť, že vo vzorke pozorujeme  $\mathbf{y}$  a  $\mathbf{X}$  je daná súčinom pravdepodobností jednotlivých pozorovaní. Keďže pracujeme so spojitým rozdelením, násobíme hustoty pravdepodobnosti. Funkcia vierohodnosti má potom tvar:

$$L(\mathbf{y}, \hat{\boldsymbol{\beta}}, \hat{\sigma}) = \prod_{i=1}^n \frac{1}{\hat{\sigma}\sqrt{2\pi}} e^{-\frac{(y_i - \mathbf{X}_i \hat{\boldsymbol{\beta}})^2}{2\hat{\sigma}^2}} \quad (3.168)$$

Keďže s funkciou v tomto tvare sa ťažko pracuje, väčšinou sa používa jej logaritmus. Keďže logaritmus je na svojom definičnom obore rastúcou funkciou, a teda aj prostou, hľadané maximum dosiahneme pri tých istých hodnotách  $\hat{\boldsymbol{\beta}}$  a  $\hat{\sigma}^2$  ako pri pôvodnej funkcii.

Po logaritmovaní funkcie hustoty dostávame funkciu, ktorú označíme  $l_i$ :

$$l_i(y_i, \hat{\boldsymbol{\beta}}, \hat{\sigma}^2) = -\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln(\hat{\sigma}^2) - \frac{1}{2\hat{\sigma}^2} (y_i - \mathbf{X}_i \hat{\boldsymbol{\beta}})^2 \quad (3.169)$$

Logaritmom funkcie vierohodnosti (angl. log-likelihood) je potom funkcia:

$$\begin{aligned} \ln(L(\mathbf{y}, \hat{\boldsymbol{\beta}}, \hat{\sigma}^2)) &= -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\hat{\sigma}^2) - \frac{1}{2\hat{\sigma}^2} \sum_{i=1}^n (y_i - \mathbf{X}_i \hat{\boldsymbol{\beta}})^2 \\ &= -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\hat{\sigma}^2) - \frac{1}{2\hat{\sigma}^2} (\mathbf{y} - \mathbf{X} \hat{\boldsymbol{\beta}})^T (\mathbf{y} - \mathbf{X} \hat{\boldsymbol{\beta}}) \end{aligned} \quad (3.170)$$

Keďže funkciu vierohodnosti (alebo aj jej logaritmus) chceme maximalizovať, ďalším krokom by bolo zderivovať posledný výraz podľa jednotlivých hľadaných parametrov.

Zderivujme najprv logaritmus funkcie vierohodnosti podľa  $\hat{\sigma}^2$ :

$$\frac{\partial \ln(L(\mathbf{y}, \hat{\boldsymbol{\beta}}, \hat{\sigma}^2))}{\partial \hat{\sigma}^2} = -\frac{n}{2} \frac{1}{\hat{\sigma}^2} + \frac{1}{2} \mathbf{e}^T \mathbf{e} \frac{1}{(\hat{\sigma}^2)^2} \quad (3.171)$$

V predchádzajúcom výraze sme využili definíciu vektora rezíduí, daný výrazom  $\mathbf{e} = \mathbf{y} - \mathbf{X} \hat{\boldsymbol{\beta}}$ . Keďže hľadáme extrém funkcie, z nutnej podmienky pre extrém položíme túto deriváciu rovnú nule:

$$-\frac{n}{2} \frac{1}{\hat{\sigma}^2} + \frac{1}{2} \mathbf{e}^T \mathbf{e} \frac{1}{(\hat{\sigma}^2)^2} = 0 \quad (3.172)$$

Keďže je zrejmé  $\hat{\sigma}^2 > 0$ , prenásobíme rovnicu výrazom  $2\hat{\sigma}^2$  a dostaneme:

$$-n\hat{\sigma}^2 + \mathbf{e}^T \mathbf{e} = 0 \quad (3.173)$$

Odtiaľ nakoniec dostávame:

$$\hat{\sigma}^2 = \frac{\mathbf{e}^T \mathbf{e}}{n} \quad (3.174)$$

Odhadom rozptylu poruchového člena  $\hat{\sigma}^2$  je teda priemer štvorcov rezíduí. Zostáva nám ešte odhadnúť vektor  $\hat{\boldsymbol{\beta}}$ . Aby sme ho zistili, musíme pracovať s maticovou deriváciou:

$$\frac{\partial \ln(L(\mathbf{y}, \hat{\boldsymbol{\beta}}, \hat{\sigma}^2))}{\partial \hat{\boldsymbol{\beta}}} = -\frac{1}{2\hat{\sigma}^2} \frac{\partial (\mathbf{y} - \mathbf{X} \hat{\boldsymbol{\beta}})^T (\mathbf{y} - \mathbf{X} \hat{\boldsymbol{\beta}})}{\partial \hat{\boldsymbol{\beta}}} \quad (3.175)$$

Keďže pracujeme s parciálnymi deriváciami, pozeráme sa na  $\hat{\sigma}^2$  ako na konštantu. Zároveň máme položiť deriváciu rovnú nulovému vektoru, čím dostávame:

$$-\frac{1}{2\hat{\sigma}^2} \frac{\partial (\mathbf{y} - \mathbf{X} \hat{\boldsymbol{\beta}})^T (\mathbf{y} - \mathbf{X} \hat{\boldsymbol{\beta}})}{\partial \hat{\boldsymbol{\beta}}} = \mathbf{0} \quad (3.176)$$

$$\frac{\partial(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})^T (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})}{\partial \hat{\boldsymbol{\beta}}} = \frac{\partial \mathbf{e}^T \mathbf{e}}{\partial \hat{\boldsymbol{\beta}}} = \mathbf{0}$$

Posledný výraz však už poznáme – je to presne ten istý výraz, s ktorým sme počítali aj pri metóde najmenších štvorcov. Odhad vektora  $\hat{\boldsymbol{\beta}}$  je potom nutne rovnaký:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \quad (3.177)$$

### Príklad 3.7

Koncom januára 2012 sme z internetovej stránky [www.reality.sk](http://www.reality.sk) stiahli ceny a rozlohy posledných pridaných bytov do tejto databázy z mesta Košice. Zaujímá nás, či existuje štatisticky významný vzťah medzi cenou bytu a jeho rozlohou. Údaje sú uvedené v nasledujúcej tabuľke.

#### Úlohy:

Zobrazte graficky skúmanú funkčnú závislosť.

1. Kvantifikujte regresný model, na základe ktorého by bolo možné overiť tvrdenie, že cena bytu závisí od jeho rozlohy.
2. Ak by sa rozloha bytu zvýšila o  $1 \text{ m}^2$ , o koľko by sa zvýšila jeho cena?
3. Za akú cenu by ste predali (podľa daného modelu) byt o rozlohe  $85 \text{ m}^2$ ?
4. Určte, ktoré byty je možné podľa daného modelu považovať za nadhodnotené.

#### Riešenie:

Keďže v tvrdení, ktoré máme overiť (či cena bytu závisí od jeho rozlohy) je presne dané, ktorá premenná je vysvetľovaná, kvantifikovať budeme nasledujúci model:

Tabuľka 5: Cena a rozloha vybraných bytov v Košiciach

Byt	Cena	m <sup>2</sup>
1	113500	96
2	64000	68
3	36500	43
4	63000	56
5	69000	51
6	48400	37
7	66900	67
8	86000	92
9	61900	68
10	63000	64
11	81500	74
12	72500	71
13	109800	99.6
14	89900	80
15	52700	42

Zdroj: [www.reality.sk](http://www.reality.sk)

$$Cena_i = \beta_0 + \beta_1 Rozloha_i + u_i \quad (3.178)$$

kde:

$i = 1, 2, \dots, n$

$Cena_i$  – cena  $i$ -teho bytu

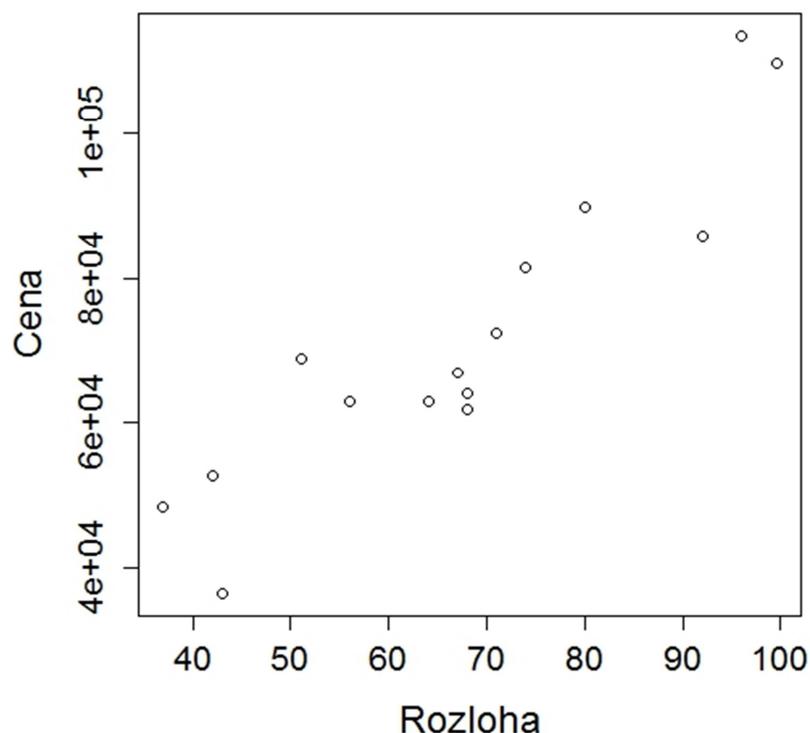
$u_i$  – náhodná chyba

$n$  – počet pozorovaní

$Rozloha_i$  – rozloha  $i$ -teho bytu

$\beta_0, \beta_1$  – regresné koeficienty

```
> Cena <- c(113500, 64000, 36500, 63000, 69000, 48400, 66900,
86000, 61900, 63000, 81500, 72500, 109800, 89900, 52700)
> Rozloha <- c(96, 68, 43, 56, 51, 37, 67, 92, 68, 64, 74, 71,
99.6, 80, 42)
> plot(Cena ~ Rozloha, cex.axis = 1.3, cex.lab = 1.5)
```



Obrázok 9: x-y graf vzťahu ceny a rozlohy bytov

Zdroj: výstup zo softvéru R

Z uvedeného obrázku môžeme vidieť, že zrejme naozaj existuje lineárna závislosť medzi cenou bytu a jeho rozlohou. V tomto bode je potrebné zdôrazniť, že k dispozícii máme prierezové údaje.

```
> Reg.Model = lm(Dataset$Cena ~ Dataset$Rozloha)
> summary(Reg.Model)

Call:
lm(formula = Dataset$Cena ~ Dataset$Rozloha)

Residuals:
    Min       1Q   Median       3Q      Max
-11006  -7158   2408   5760  13441

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    4220.8     8404.5   0.502   0.624
Dataset$Rozloha 1006.6     120.4   8.359 1.38e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 8726 on 13 degrees of freedom
Multiple R-squared:  0.8431,    Adjusted R-squared:  0.8311
F-statistic: 69.88 on 1 and 13 DF,  p-value: 1.377e-06
```

Výsledky z kvantifikácie modelu poukazujú na štatisticky významný vzťah medzi cenou a rozlohou skúmaných bytov. Na 1 % hladine významnosti môžeme zamietnuť nulovú hypotézu, že koeficient pri premennej „Rozloha“ je rovný nule. Daným modelom sa nám podarilo popísať 83.11 % celkovej variability ceny bytov, čo je možné považovať za dôkaz silného vzťahu. Najmä ak vezmeme do úvahy, že ceny bytov závisia samozrejme od množstva iných faktorov. Model by sme mohli rozšíriť o výhodnosť polohy bytu (meranou napr. vzdialenosťou od centra), ale aj o rôzne indikátorové premenné (napr. 0 = byt nemá balkón, 1 = byt má balkón; 0 = byt je panelový, 1 = byt je tehlový; 0 = ide o starý byt, 1 = ide o novostavbu; a podobne).

Na zodpovedanie otázky „o koľko by sa zvýšila cena bytu, ak by sa zvýšila jeho rozloha o 1 m<sup>2</sup>?“, stačí pozrieť odhad koeficientu  $\beta_1$ . V našom prípade, ak sa rozloha zvýšila o jednu jednotku, cena bytu by sa na základe modelu v priemere zvýšila o 1006.6 EUR.

Ak by sme sa rozhodovali o predaji bytu s rozlohou 85m<sup>2</sup> na základe daného modelu, by sme priemernú (ponúkanú) cenu mohli stanoviť na 89781.80 EUR. Túto cenu sme vypočítali dosadením danej rozlohy (85m<sup>2</sup>) do odhadnutého lineárneho modelu:

$$Cena = 4220.8 + 1006.6 \times 85m^2.$$

Poslednou úlohou bolo určiť, ktoré byty je podľa daného modelu možné považovať za nadhodnotené. Pod týmto pojmom zrejme každý intuitívne rozumie fakt, že daný byt sa predáva za vyššiu cenu, ako je jeho hodnota. My túto hodnotu samozrejme nepoznáme, avšak predikované hodnoty z nášho modelu je možné použiť na jej odhad. Potom rezíduá z modelu naznačujú či je byt nad- alebo pod-hodnotený.

```
> Reg.Model$residuals
      1          2          3          4          5
12642.619 -8671.706 -11005.925  2407.869 13441.025
      6          7          8          9         10
 6933.862 -4765.075 -10830.856 -10771.706 -5645.181
     11         12         13         14         15
2788.506 -3191.600  5318.746  5148.719  6200.706
```

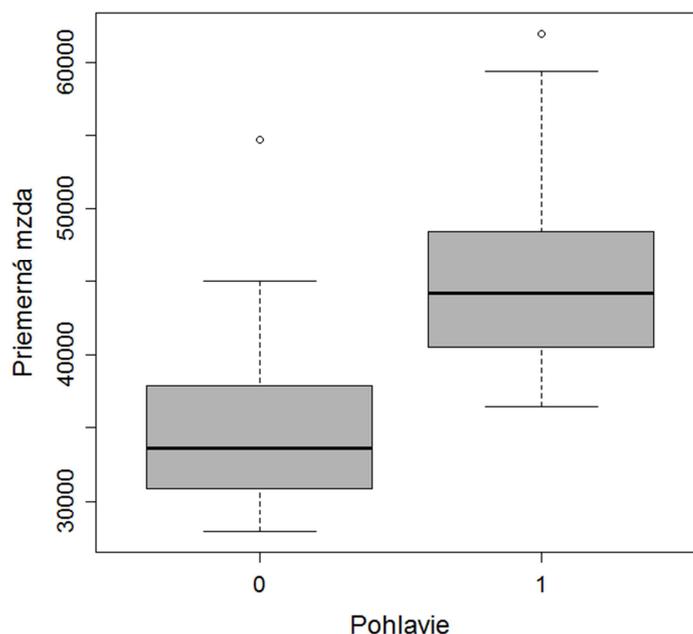
Všetky kladné rezíduá znamenajú, že dané pozorovania sa nachádzajú nad regresnou priamkou (byt je nadhodnotený) a všetky záporné rezíduá znamenajú, že dané pozorovania sa nachádzajú pod regresnou priamkou (byt je podhodnotený). V našom prípade môžeme za nadhodnotené považovať byty č.1 (12642.619), č.4 (2407.869), č.5 (13441.025), č.6



```
> numSummary(Mzda, groups = Pohlavie)
      mean      sd    0%    25%   50%    75%  100%   n
0 34681.69 5410.388 27855 30834.0 33611 37933.0 54698 51
1 45204.47 5915.428 36465 40530.5 44174 48440.5 61993 51
```

Ako môžeme vidieť, v danom súbore máme vytvorenú tzv. umelú premennú („Pohlavie“), ktorá nadobúda hodnotu 1, pokiaľ ide o mzdu mužov a hodnotu 0, pokiaľ ide o mzdu žien. Z uvedených základných charakteristík to vyzerá tak, že muži v jednotlivých štátoch USA zarábajú v priemere za rok viac ako ženy. Pre lepšiu prehľadnosť je veľmi dobrým nástrojom vizualizácie dát box-plot. Graf box-plot vytvoríme v softvéri R pomocou nasledujúceho príkazu:

```
> boxplot (Mzda ~ Pohlavie, col = grey (0.7), density = 10, xlab =
= c("Pohlavie"), ylab = c("Priemerná mzda"), cex.axis = 1.3,
cex.lab = 1.5)
```



Obrázok 10: Box-plot priemerných miezd mužov (1) a žien (0)

*Zdroj: výstup zo softvéru R*

Na prvý pohľad z box-plotu vidíme, že priemerné mzdy mužov v jednotlivých štátoch USA sú vyššie, ako priemerné mzdy žien. Dobrým indikátorom významného rozdielu je situácia ak je dolný kvartil jedného boxplotu vyšší, ako horný kvartil druhého boxplotu. Inak povedané, aspoň 75 % mužov zarába viac ako 75 % žien. Naše podozrenie si overíme pomocou kvantifikácie nasledujúceho regresného modelu:

$$Mzda_i = \beta_0 + \beta_1 Pohlavie_i + u_i \quad (3.179)$$

kde:

$i = 1, 2, \dots, n$        $Mzda_i$  – mzda v  $i$ -tom štáte       $u_i$  – náhodná chyba  
 $n$  – počet pozorovaní       $Pohlavie_i$  – umelá premenná       $\beta_0, \beta_1$  – regresné koeficienty

```
> Reg.Model = lm(Mzda ~ Pohlavie)
> summary (Reg.Model)

Call:
lm(formula = Dataset$Mzda ~ Dataset$Pohlavie)

Residuals:
    Min       1Q   Median       3Q      Max
-8739  -4180  -1068   3262  20016

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)    34681.7      793.8   43.693 < 2e-16 ***
Dataset$Pohlavie 10522.8     1122.5   9.374 2.33e-15 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5669 on 100 degrees of freedom
Multiple R-squared:  0.4677,    Adjusted R-squared:  0.4624
F-statistic: 87.87 on 1 and 100 DF,  p-value: 2.332e-15
```

### 3.3.5 Geometrická interpretácia metódy najmenších štvorcov

V kapitole 2 sme si prezentovali vybraný výsledky teórie lineárnej algebr. S ich pomocou si môžeme priblížiť aj geometrickú interpretácia metódy najmenších štvorcov, a lineárnej regresie.

V predchádzajúcich častiach sme si ukázali, že odhadovaný populačný model predpokladáme v tvare:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \quad (3.180)$$

Odhad vektora  $\boldsymbol{\beta}$  dostaneme v tvare:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \quad (3.181)$$

V odhadnutom modeli

$$\mathbf{y} = \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{e} \quad (3.182)$$

predstavuje vektor  $\mathbf{e}$  rezíduá modelu, a vektor

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}} \quad (3.183)$$

nazývame vektorom vyrovnaných hodnôt. Môžeme si všimnúť, že vektor  $\hat{\mathbf{y}}$  dostávame ako súčin  $\mathbf{X}\hat{\boldsymbol{\beta}}$  - čo si vzhľadom na definíciu lineárneho modelu v časti 3.3.2 môžeme rozpísať aj nasledovne:

$$\hat{\mathbf{y}} = \mathbf{x}_1\hat{\beta}_0 + \mathbf{x}_2\hat{\beta}_1 + \dots + \mathbf{x}_{k+1}\hat{\beta}_k = \sum_{i=0}^k \mathbf{x}_{i+1}\hat{\beta}_i \quad (3.184)$$

V predchádzajúcom výraze predstavujú vektory  $\mathbf{x}_i$  pre  $i = 1, 2, \dots, k + 1$  stĺpce matice  $\mathbf{X}$ . Zápis možno pôsobí trochu neštandardne, ale treba si uvedomiť, že matica  $\mathbf{X}$  je rozmeru  $n \times (k + 1)$  a pre regresné koeficienty sme zaviedli označenie  $\beta_0$  až  $\beta_k$  - stĺpce matice označujeme číslami počnúc číslom 1, ale regresné koeficienty označujeme od nuly ( $\beta_0$  je konštantný člen).

V predchádzajúcom výraze vidíme, že vektor  $\hat{\mathbf{y}}$  v skutočnosti vyjadrujeme ako lineárnu kombináciu stĺpcových vektorov matice  $\mathbf{X}$ . Každý stĺpec matice  $\mathbf{X}$  pritom zodpovedá hodnotám jednej premennej.

V súlade s definíciami 2.2 by sme mohli definovať lineárny obal stĺpcových vektorov matice  $\mathbf{X}$  ako  $[\mathbf{X}]$ . Ide o vektorový priestor, ktorého prvky je možné vyjadriť ako lineárne kombinácie stĺpcových vektorov tejto matice. Je zrejmé, že  $\hat{\mathbf{y}} \in [\mathbf{X}]$ , keďže podľa predchádzajúcej diskusie získame vyrovnané hodnoty ako lineárnu kombináciu vektorov tvoriacich  $\mathbf{X}$  s koeficientmi  $\beta_0$  až  $\beta_k$ .

Spôsob, ako k vektoru  $\mathbf{y}$  môžeme získať vyrovnané hodnoty  $\hat{\mathbf{y}}$  je jednoduchý. Platí:

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} = \mathbf{P}\mathbf{y} \quad (3.185)$$

kde sme označili  $\mathbf{P} = \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T$ . Výraz  $\hat{\mathbf{y}} = \mathbf{P}\mathbf{y}$  sa nápadne podobá na výrazy v kapitole 2.5, kde sme si definovali lineárne zobrazenia. V našom prípade je maticou lineárnej transformácie matica  $\mathbf{P}$ . Túto maticu nazývame aj projekčnou maticou.

Všimnime si ďalej, že vektor  $\mathbf{y}$  sme definovali ako vektor rozmeru  $n \times 1$  a matica  $\mathbf{X}$  má podľa predpokladu P4 uvedenom v kapitole 3.3.3 hodnotu  $(k + 1)$ , a  $k + 1 < n$ . Z toho vyplýva, že  $[\mathbf{X}]$  nemôže obsahovať všetky vektory  $\mathbf{y} \in \mathbb{R}^n$ . Vektorový priestor  $[\mathbf{X}]$  má rozmer  $k + 1$ , zatiaľ čo  $\mathbb{R}^n$  je rádu  $n$ .

Pre niektoré vektory preto môže platiť  $\mathbf{y} \in [\mathbf{X}]$  (ležia v regresnej nadrovine), ale niektoré vektory ležia mimo tohto vektorového priestoru. V ich prípade sme definovali:

$$\mathbf{e} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{y} - \hat{\mathbf{y}} \quad (3.186)$$

Podľa predchádzajúceho vzťahu však

$$\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \mathbf{P}\mathbf{y} = (\mathbf{I}_n - \mathbf{P})\mathbf{y} \quad (3.187)$$

Označme maticou  $\mathbf{M}$  maticu

$$\mathbf{M} = \mathbf{I}_n - \mathbf{P} = \mathbf{I}_n - \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X} \quad (3.188)$$

Rezíduá modelu potom získame pomocou matice  $\mathbf{M}$  ako:

$$\mathbf{e} = \mathbf{M}\mathbf{y} \quad (3.189)$$

Aj matica  $\mathbf{M}$  je maticou lineárnej transformácie, teda projekčnou maticou. Kým pomocou matice  $\mathbf{P}$  z vektorov  $\mathbf{y}$  získavame vyrovnané hodnoty  $\hat{\mathbf{y}}$ , pomocou matice  $\mathbf{M}$  získavame vektor rezíduí  $\mathbf{e}$ . Zrejme potom platí:

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{e} = \mathbf{P}\mathbf{y} + \mathbf{M}\mathbf{y} = (\mathbf{P} + \mathbf{M})\mathbf{y} \quad (3.190)$$

Ukážme si niekoľko vlastností projekčných matíc  $\mathbf{P}$  a  $\mathbf{M}$ .

Matica  $\mathbf{P} = \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T$  je na prvý pohľad symetrickou maticou – vieme, že matica  $\mathbf{X}^T \mathbf{X}$  je symetrická vždy, rovnako aj k nej inverzná matica  $(\mathbf{X}^T \mathbf{X})^{-1}$ . Keďže  $\mathbf{M} = \mathbf{I}_n - \mathbf{P}$  je rozdielom dvoch symetrických matíc (jednotková matica je symetrická), aj  $\mathbf{M}$  je symetrická matica.

Matice  $\mathbf{P}$  aj  $\mathbf{M}$  sú idempotentné matice, t.j. matice, ktorých súčin samých so sebou dáva pôvodnú maticu. Platí totiž:

$$\mathbf{P}\mathbf{P} = \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T = \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{I}_{k+1} \mathbf{X}^T = \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T = \mathbf{P} \quad (3.191)$$

Podobne pre maticu  $\mathbf{M}$  s využitím idempotentnosti  $\mathbf{P}$  platí:

$$\mathbf{M}\mathbf{M} = (\mathbf{I}_n - \mathbf{P})(\mathbf{I}_n - \mathbf{P}) = \mathbf{I}_n - \mathbf{P} - \mathbf{P} + \mathbf{P}\mathbf{P} = \mathbf{I}_n - \mathbf{P} - \mathbf{P} + \mathbf{P} = \mathbf{I}_n - \mathbf{P} = \mathbf{M} \quad (3.192)$$

Matice  $\mathbf{P}$  a  $\mathbf{M}$  zároveň definujú navzájom ortogonálne vektorové priestory. To vidíme, ak si uvedomíme, že platí:

$$\mathbf{P}\mathbf{M} = \mathbf{P}(\mathbf{I}_n - \mathbf{P}) = \mathbf{P} - \mathbf{P}\mathbf{P} = \mathbf{P} - \mathbf{P} = \mathbf{0} \quad (3.193)$$

Zrejme rovnaký výsledok platí aj pre súčin  $\mathbf{M}\mathbf{P}$ . Tento výsledok je veľmi zaujímavý – akýkoľvek vektor, ktorý získame ako projekciu maticou  $\mathbf{P}$  bude ortogonálny vzhľadom na akýkoľvek iný vektor, ktorý získame ako projekciu pomocou matice  $\mathbf{M}$ . Keďže vieme, že  $\hat{\mathbf{y}} = \mathbf{P}\mathbf{y}$  a  $\mathbf{e} = \mathbf{M}\mathbf{y}$ , z toho potom nutne vyplýva, že vektor vyrovnaných hodnôt a rezíduí sú navzájom kolmé (ortogonálne). Ako sme uviedli v kapitole 2.7, vektor  $\hat{\mathbf{y}}$  môžeme chápať ako ortogonálnu projekciu vektora  $\mathbf{y}$  do vektorového priestoru  $[\mathbf{X}]$ .

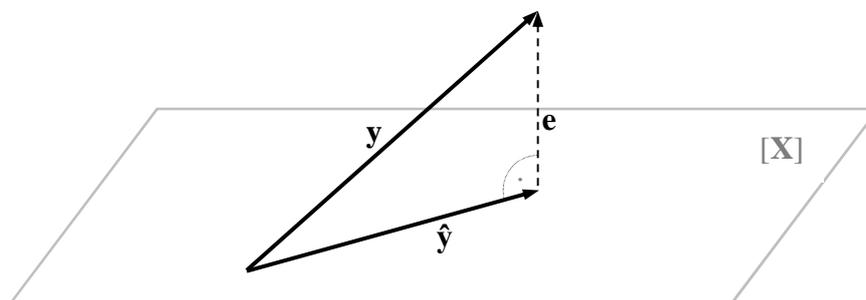
Z vlastností ortogonálnej projekcie vyplýva, že pri  $\hat{\mathbf{y}}$  ide o vektor, pri ktorom je vzdialenosť  $\|\mathbf{y} - \hat{\mathbf{y}}\|$  najmenšia pre  $\hat{\mathbf{y}} \in [\mathbf{X}]$ . Toto kritérium je ekvivalentné (vzhľadom k

definícii  $\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}}$ ) kritériu minimalizácie  $\|\mathbf{e}\|$ . Vzhľadom na vzťah medzi normou vektora a skalárnym súčinom, ktorý sme definovali v časti 2.6, ale platí:

$$\|\mathbf{e}\| = \mathbf{e}^T \mathbf{e} = \sum_{i=1}^n e_i^2 \quad (3.194)$$

Posledný výraz predstavuje súčet štvorcov rezíduí – máme minimalizovať presne ten výraz, ktorý sme využili na výpočet odhadov v podkapitole 3.3.2 metódou najmenších štvorcov.

Z uvedeného dôvodu sa na lineárny regresný model môžeme geometricky pozeráť aj ako na ortogonálnu projekciu vektora  $\mathbf{y}$  do do vektorového priestoru  $[\mathbf{X}]$ .



Obrázok 11: Lineárny model ako ortogonálna projekcia

Zdroj: vlastné spracovanie

### 3.4 Vlastnosti odhadu lineárneho modelu

Predmetom tejto kapitoly je popísanie vybraných vlastností odhadu lineárneho modelu, ktorý sme si priblížili v predchádzajúcej podkapitole. Keďže odhad regresných koeficientov pomocou metódy najmenších štvorcov a maximálnej vierohodnosti je zhodný, veľa skutočností platí pre obidva modely. Odlišné sú vlastnosti pri odhade rozptylu prvkov poruchového člena  $\mathbf{u}$  – na túto výnimku ale na príslušnom mieste upozorníme.

#### 3.4.1 Základné vzťahy pre odhadované regresné koeficienty

V tejto časti popíšeme niektoré vzťahy, ktoré platia pri odhade vektora regresných koeficientov  $\hat{\boldsymbol{\beta}}$ . Tieto vlastnosti sa ďalej využívajú v nasledujúcich podkapitolách – sami o sebe nie sú možno až tak zaujímavé, ale je dobré ich uviesť, aby sme ich mohli neskôr využiť. Podrobnejší popis uvádza Hatrák (2007).

Z predchádzajúcej kapitoly vieme, že odhad  $\hat{\boldsymbol{\beta}}$  je daný vzťahom:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \quad (3.195)$$

Ak by sme postupovali späťne oproti postupu, ako sme tento vzťah odvodili, dostávame:

$$\mathbf{X}^T \mathbf{X} \hat{\boldsymbol{\beta}} = \mathbf{X}^T \mathbf{y} \quad (3.196)$$

$$\mathbf{X}^T \mathbf{y} - \mathbf{X}^T \mathbf{X} \hat{\boldsymbol{\beta}} = \mathbf{0} \quad (3.197)$$

$$\mathbf{X}^T (\mathbf{y} - \mathbf{X} \hat{\boldsymbol{\beta}}) = \mathbf{0} \quad (3.198)$$

$$\mathbf{X}^T \mathbf{e} = \mathbf{0} \quad (3.199)$$

Prvý vzťah sme dostali pre násobení rovnice  $\mathbf{X}^T \mathbf{X}$  zľava. V poslednej rovnici sme využili definíciu vektora rezíduí ako rozdielu skutočných a vyrovnaných hodnôt. Posledný vzťah sa objavuje v mnohých výrazoch pri neskoršom výklade – je to prvý zo série dôležitých vzťahov.

S jeho pomocou môžeme dokázať jednu zaujímavú vlastnosť vektora rezíduí. Ak si vzťah (3.199) rozpíšeme v maticovej podobe, dostávame:

$$\mathbf{X}^T \mathbf{e} = \mathbf{0} \quad (3.200)$$

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_{11} & x_{21} & \cdots & x_{n1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1k} & x_{2k} & \cdots & x_{nk} \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad (3.201)$$

Z prvého riadku môžeme vidieť, že platí:

$$\mathbf{1}^T \mathbf{e} = 0 \quad (3.202)$$

Tento vzťah môžeme zapísať aj inak:

$$\mathbf{1}^T \mathbf{e} = e_1 + e_2 + \cdots + e_n = \sum_{i=1}^n e_i = 0 \quad (3.203)$$

Súčet rezíduí je pre odhad pomocou metódy najmenších štvorcov vždy nulový. Z toho ale vyplýva, že aj priemerná odchýlka skutočných a modelovaných hodnôt je nulová, pretože ju vypočítame ako:

$$\frac{1}{n} \mathbf{1}^T \mathbf{e} = 0 \quad (3.204)$$

Pokiaľ je priemer rezíduí nulový, bolo by možno očakávať, že priemer vyrovnaných hodnôt by mal byť rovný priemeru skutočných hodnôt  $\mathbf{y}$ . Priemer skutočných hodnôt je daný vzťahom:

$$\frac{1}{n} \mathbf{1}^T \mathbf{y} \quad (3.205)$$

Môžeme ho ale rozpísať nasledovne:

$$\frac{1}{n} \mathbf{1}^T \mathbf{y} = \frac{1}{n} \mathbf{1}^T (\mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{e}) \quad (3.206)$$

$$= \frac{1}{n} \mathbf{1}^T \mathbf{X}\hat{\boldsymbol{\beta}} + \frac{1}{n} \mathbf{1}^T \mathbf{e} \quad (3.207)$$

$$= \frac{1}{n} \mathbf{1}^T \mathbf{X}\hat{\boldsymbol{\beta}} \quad (3.208)$$

$$= \frac{1}{n} \mathbf{1}^T \hat{\mathbf{y}} \quad (3.209)$$

Zjavne teda platí:

$$\frac{1}{n} \mathbf{1}^T \mathbf{y} = \frac{1}{n} \mathbf{1}^T \hat{\mathbf{y}} \quad (3.210)$$

$$\mathbf{1}^T \mathbf{y} = \mathbf{1}^T \hat{\mathbf{y}} \quad (3.211)$$

Priemer skutočných hodnôt  $\mathbf{y}$  a vyrovnaných hodnôt  $\hat{\mathbf{y}}$  je rovnaký.

Ďalšou vlastnosťou, ktorú budeme potrebovať, je rovnosť (Hatrák, 2007):

$$\hat{\mathbf{y}}^T \mathbf{e} = 0 \quad (3.212)$$

Tento vzťah platí, pretože:

$$\hat{\mathbf{y}}^T \mathbf{e} = (\mathbf{X}\hat{\boldsymbol{\beta}})^T \mathbf{e} = \hat{\boldsymbol{\beta}}^T \mathbf{X}^T \mathbf{e} = \hat{\boldsymbol{\beta}}^T (\mathbf{X}^T \mathbf{e}) = \hat{\boldsymbol{\beta}}^T \mathbf{0} = 0 \quad (3.213)$$

Využili sme pritom prvú vlastnosť v tejto podkapitole, a to že  $\mathbf{X}^T \mathbf{e} = \mathbf{0}$ .

Ďalšou vlastnosťou, ktorú si ukážeme, je vzťah medzi odhadom, skutočnou hodnotou hľadaných parametrov a náhodným poruchovým členom. Vieme, že:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \quad (3.214)$$

Keďže pre skutočné hodnoty  $\mathbf{y}$  podľa populačného modelu platí  $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$ , po dosadení dostávame:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T (\mathbf{X}\boldsymbol{\beta} + \mathbf{u}) \quad (3.215)$$

Po roznásobení dostaneme:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{X} \boldsymbol{\beta} + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{u} \quad (3.216)$$

Keďže  $(\mathbf{X}^T \mathbf{X})^{-1}$  a  $\mathbf{X}^T \mathbf{X}$  sú navzájom inverzné matice rádu  $k + 1 \times k + 1$ , ich súčinom je jednotková matica  $\mathbf{I}_{k+1}$ , a preto:

$$\hat{\boldsymbol{\beta}} = \boldsymbol{\beta} + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{u} \quad (3.217)$$

Odtiaľ vidíme, že odhad regresných parametrov je funkciou tak skutočnej hodnoty parametra  $\boldsymbol{\beta}$ , ako aj poruchového člena  $\mathbf{u}$ .

Posledným vzťahom, ktorý je pri analýze lineárnych modelov užitočný je vzťah vyjadrujúci rezíduá  $\mathbf{e}$  ako funkciu náhodných porúch  $\mathbf{u}$ . Z definície rezíduí vieme, že:

$$\mathbf{e} = \mathbf{y} - \mathbf{X} \hat{\boldsymbol{\beta}} \quad (3.218)$$

Keďže  $\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}$ , máme:

$$\mathbf{y} - \mathbf{X} \hat{\boldsymbol{\beta}} = \mathbf{y} - \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} = (\mathbf{I}_n - \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T) \mathbf{y} \quad (3.219)$$

V poslednom výraze sme vybrali vektor  $\mathbf{y}$  za zátvorku. Zároveň však platí  $\mathbf{y} = \mathbf{X} \boldsymbol{\beta} + \mathbf{u}$ , takže:

$$\begin{aligned} \mathbf{y} - \mathbf{X} \hat{\boldsymbol{\beta}} &= \mathbf{y} - \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} = (\mathbf{I}_n - \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T) (\mathbf{X} \boldsymbol{\beta} + \mathbf{u}) \\ &= \mathbf{X} \boldsymbol{\beta} + \mathbf{u} + \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{X} \boldsymbol{\beta} + \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{u} \\ &= \mathbf{X} \boldsymbol{\beta} + \mathbf{u} - \mathbf{X} \boldsymbol{\beta} - \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{u} \\ &= (\mathbf{I}_n - \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T) \mathbf{u} \\ &= \mathbf{M} \mathbf{u} \end{aligned} \quad (3.220)$$

V poslednom kroku sme zaviedli substitúciu  $\mathbf{M} = \mathbf{I}_n - \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T$ . Matica  $\mathbf{M}$  je symetrickou maticou, pretože:

$$\mathbf{M}^T = (\mathbf{I}_n - \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T)^T = \mathbf{I}_n - \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T = \mathbf{M} \quad (3.221)$$

Zároveň je aj idempotentnou maticou, keďže:

$$\mathbf{M} \mathbf{M} = (\mathbf{I}_n - \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T) (\mathbf{I}_n - \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T) \quad (3.222)$$

$$= \mathbf{I}_n - \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T - \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T + \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \quad (3.223)$$

$$= \mathbf{I}_n - 2\mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T + \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \quad (3.224)$$

$$= \mathbf{I}_n - \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T = \mathbf{M} \quad (3.225)$$

Vzťah medzi  $\mathbf{e}$  a  $\mathbf{u}$  je preto daný výrazom:

$$\mathbf{e} = \mathbf{M}\mathbf{u} \quad (3.226)$$

Aby sme vlastnosti uviedli aj v prehľadnejšej forme, vypíšeme si ich aj všetky na jednom mieste:

$$\mathbf{X}^T \mathbf{e} = \mathbf{0} \quad (3.227)$$

$$\mathbf{1}^T \mathbf{e} = 0 \quad (3.228)$$

$$\mathbf{1}^T \mathbf{y} = \mathbf{1}^T \hat{\mathbf{y}} \quad (3.229)$$

$$\hat{\mathbf{y}}^T \mathbf{e} = 0 \quad (3.230)$$

$$\hat{\boldsymbol{\beta}} = \boldsymbol{\beta} + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{u} \quad (3.231)$$

$$\mathbf{e} = \mathbf{M}\mathbf{u} = (\mathbf{I}_n - \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T) \mathbf{u} \quad (3.232)$$

### 3.4.2 Vlastnosti bodového odhadu – neskreslenosť

V predchádzajúcej kapitole sme predstavili niektoré vlastnosti odhadu najmenších štvorcov, ktoré by sa dali označiť za jeho algebraické vlastnosti. V tejto kapitole sa budeme venovať niektorým štatistickým vlastnostiam tohto odhadu.

Ukázali sme si dve metódy odhadu parametrov  $\hat{\boldsymbol{\beta}}$ , ktoré viedli k rovnakým výsledkom. Otázkou zostáva, či je tento odhad naozaj „dobrým“ odhadom. Skôr ako vysvetlíme, čo pod týmto označením myslíme, môžeme zauvažovať, prečo má skúmanie vlastností estimátorov vôbec zmysel.

Ako sme už spomínali, pravý populačný model v skutočnosti nepoznáme. Spôsobov, akými by bolo možné odhadnúť populačný parameter ( $\boldsymbol{\beta}$ ) je nekonečne veľa. Ak máme k dispozícii len jednu vzorku, na základe ktorej uskutočňujeme svoj odhad, radi by sme mali istotu, že využívame informáciu, ktorú nám vzorka poskytuje v maximálnej možnej miere.

Uveďme jednoduchý príklad. Ak chceme zistiť priemernú výšku Slovákov, môžeme vytvoriť náhodným výberom vzorku a zmerať 100 obyvateľov. Naším odhadom priemernej výšky (strednej hodnoty) Slovákov môže byť priemer vypočítaný zo vzorky, ale aj výška dvadsiateho tretieho respondenta. V oboch prípadoch dostaneme ako odhad jedno číslo, v oboch prípadoch je to číslo získané zo vzorky, v oboch prípadoch odhadujeme parameter tej istej populácie. Jednou vlastnosťou, ktorá sa u bodových odhadov sleduje je ich neskreslenosť – či použitím konkrétneho odhadu systematicky nenadhodnocujeme alebo nepodhodnocujeme

skutočnú odhadovanú veličinu. Zhodou okolností sú obidva zvolené odhady neskreslené. Existuje však rozdiel v efektívnosti (výdatnosti) – priemer využíva údaje z celej vzorky, výška dvadsiateho tretieho respondenta používa len jeden údaj. Výberový priemer preto odhaduje populačnú charakteristiku s väčšou presnosťou (menším rozptylom) – v tomto ohľade je lepším estimátorom.

Analogicky v prípade lineárnej regresie pre nás môže byť dôležité vedieť, či sme si zvolili odhad, ktorý je skreslený a efektívny, alebo či existuje lepší spôsob, ako odhadovať hľadaný parameter  $\hat{\beta}$ .

Začnime preto s *neskreslenosťou*. Estimátor  $\hat{\theta}$  je neskresleným odhadom populačného parametra  $\theta$  vtedy, ak platí:

$$E(\hat{\theta}) = \theta \quad (3.233)$$

V našom prípade sa snažíme odhadnúť  $\beta$  pomocou  $\hat{\beta}$ . Otázkou je, čomu sa rovná  $E(\hat{\beta})$ ?

Začneme vzorcom pre odhad regresného parametra:

$$E(\hat{\beta}) = E((\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}) \quad (3.234)$$

Podľa vlastností uvedenej v kapitole 3.4.2 vieme, že za výraz vpravo môžeme dosadiť:

$$E((\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}) = E(\beta + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{u}) \quad (3.235)$$

Jedinou náhodnou veličinou v poslednom výraze na pravej strane je vektor  $\mathbf{u}$  – vektor  $\beta$  je konštantným populačným parametrom a matica  $\mathbf{X}$  je podľa predpokladu P3 deterministická. Odtiaľ z vlastností strednej hodnoty dostávame:

$$E(\beta + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{u}) = \beta + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T E(\mathbf{u}) \quad (3.236)$$

Podľa predpokladu P1 však platí  $E(\mathbf{u}) = \mathbf{0}$ , takže posledný sčítanec je rovný nulovému vektoru. Dostávame preto rovnosť:

$$E(\hat{\beta}) = \beta + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T E(\mathbf{u}) = \beta \quad (3.237)$$

Odhad  $\hat{\beta}$  získaný metódou najmenších štvorcov, alebo metódou maximálnej vierohodnosti je neskresleným odhadom vektora koeficientov  $\beta$ .

### 3.4.3 Vlastnosti bodového odhadu – efektívnosť

Efektívnosť, nazývaná aj výdatnosťou odhadu, sa zaoberá rozptylom estimátorov. Ak porovnáваме dva neskreslené odhady, potom hovoríme, že efektívnejší je ten, ktorý má

menší rozptyl. V prípade, ak k estimátoru neexistuje efektívnejší estimátor, nazývame ho efektívnym.

Ak by sme chceli posúdiť efektívnosť odhadu  $\hat{\beta}$ , musíme byť schopní najprv vypočítať jeho rozptyl. Ako sme uviedli v podkapitole 3.1.3, variančno-kovariančnú maticu pre estimátor  $\hat{\beta}$  môžeme napísať takto:

$$\text{var}(\hat{\beta}) = E\left(\left(\hat{\beta} - E(\hat{\beta})\right)\left(\hat{\beta} - E(\hat{\beta})\right)^T\right) \quad (3.238)$$

Vzhľadom na neskreslenosť  $\hat{\beta}$  môžeme napísať:

$$\text{var}(\hat{\beta}) = E\left(\left(\hat{\beta} - \beta\right)\left(\hat{\beta} - \beta\right)^T\right) \quad (3.239)$$

Keďže sme v kapitole 3.4.1 odvodili vzťah  $\hat{\beta} = \beta + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{u}$ , z neho vyplýva, že  $\hat{\beta} - \beta = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{u}$ . Do variančno-kovariančnej matice preto môžeme dosadiť tento výraz, čím dostaneme:

$$\text{var}(\hat{\beta}) = E\left(\left(\hat{\beta} - \beta\right)\left(\hat{\beta} - \beta\right)^T\right) = E\left(\left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T \mathbf{u} \mathbf{u}^T \mathbf{X} \left(\mathbf{X}^T \mathbf{X}\right)^{-1}\right) \quad (3.240)$$

Pri úprave sme pri poslednej inverznej matici využili jej symetrickosť: pracujeme s výrazom  $\left(\left(\mathbf{X}^T \mathbf{X}\right)^{-1}\right)^T$ . Keďže je však matica  $\mathbf{X}^T \mathbf{X}$  symetrická, je symetrická aj k nej inverzná matica  $\left(\mathbf{X}^T \mathbf{X}\right)^{-1}$  a preto  $\left(\left(\mathbf{X}^T \mathbf{X}\right)^{-1}\right)^T = \left(\mathbf{X}^T \mathbf{X}\right)^{-1}$ . Použili sme preto jednoduchší tvar.

V rovnici (3.240) sú všetky matice a vektory s výnimkou vektora  $\mathbf{u}$  deterministické -  $\beta$  je parameter populácie a  $\mathbf{X}$  je deterministická podľa predpokladu P3. Stredná hodnota je preto rovná:

$$E\left(\left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T \mathbf{u} \mathbf{u}^T \mathbf{X} \left(\mathbf{X}^T \mathbf{X}\right)^{-1}\right) = \left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T E\left(\mathbf{u} \mathbf{u}^T\right) \mathbf{X} \left(\mathbf{X}^T \mathbf{X}\right)^{-1} \quad (3.241)$$

Podľa predpokladu P2 však vieme, že  $E\left(\mathbf{u} \mathbf{u}^T\right) = \sigma^2 \mathbf{I}_n$ . Po dosadení:

$$\left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T E\left(\mathbf{u} \mathbf{u}^T\right) \mathbf{X} \left(\mathbf{X}^T \mathbf{X}\right)^{-1} = \left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T \sigma^2 \mathbf{I}_n \mathbf{X} \left(\mathbf{X}^T \mathbf{X}\right)^{-1} \quad (3.242)$$

$$= \sigma^2 \left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T \mathbf{I}_n \mathbf{X} \left(\mathbf{X}^T \mathbf{X}\right)^{-1} \quad (3.243)$$

$$= \sigma^2 \left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T \mathbf{X} \left(\mathbf{X}^T \mathbf{X}\right)^{-1} \quad (3.244)$$

$$= \sigma^2 \left(\mathbf{X}^T \mathbf{X}\right)^{-1} \quad (3.245)$$

V treťom riadku sme využili vlastnosti jednotkovej matice a nakoniec v poslednom riadku vlastnosti inverznej matice. Celkovo môžeme uzavrieť, že:

$$\text{var}(\hat{\boldsymbol{\beta}}) = \sigma^2 (\mathbf{X}^T \mathbf{X})^{-1} \quad (3.246)$$

Je dobré všimnúť si niekoľko skutočností. V prvom rade výsledok závisí zásadným spôsobom na platnosti predpokladu P2. Ak by tomu tak nebolo, nemohlo by dôjsť k zjednodušujúcemu kroku v rovnici (3.242). Na tomto mieste je zrejmé, aký dôležitý je predpoklad P2 – každý štatistický softvér totiž pre lineárny regresný model ráta variančno-kovariančnú maticu čiastočne podľa odvodeného vzorca, t. j. za predpokladu že P2 platí.

Druhou poznámkou je, že aj keď sme schopní zadať vzťah pre výpočet  $\text{var}(\hat{\boldsymbol{\beta}})$ , nie sme schopní ho reálne vypočítať – výraz stále obsahuje  $\sigma^2$ , čo je neznáma veličina. Odhad  $\sigma^2$  na chvíľku odložme – dokončíme najprv úvahu, či je možné, aby bol nejaký iný estimátor efektívnejší, ako  $\hat{\boldsymbol{\beta}}$ .

Označme  $\tilde{\boldsymbol{\beta}}$  iný estimátor, ktorý je možné vyjadriť v tvare  $\tilde{\boldsymbol{\beta}} = \mathbf{C}\mathbf{y}$ , kde  $\mathbf{C}$  je matica rádu  $(k+1) \times n$  (Greene, 2012, s. 60). Uvedený zápis znamená, že aj  $\tilde{\boldsymbol{\beta}}$  je lineárnym odhadom.

Aby sme mohli porovnávať efektívnosť  $\hat{\boldsymbol{\beta}}$  a  $\tilde{\boldsymbol{\beta}}$ , musia byť obidva odhady neskreslené. O  $\hat{\boldsymbol{\beta}}$  vieme, že je neskreslený. Ak má byť  $\tilde{\boldsymbol{\beta}}$  neskresleným odhadom, musí platiť:

$$E(\mathbf{C}\mathbf{y}) = E(\mathbf{C}(\mathbf{X}\boldsymbol{\beta} + \mathbf{u})) = E(\mathbf{C}\mathbf{X}\boldsymbol{\beta} + \mathbf{C}\mathbf{u}) = \boldsymbol{\beta} \quad (3.247)$$

Z vlastností strednej hodnoty vieme, že:

$$E(\mathbf{C}\mathbf{X}\boldsymbol{\beta} + \mathbf{C}\mathbf{u}) = \mathbf{C}\mathbf{X}\boldsymbol{\beta} + \mathbf{C}E(\mathbf{u}) \quad (3.248)$$

Podľa predpokladu P1 však  $E(\mathbf{u}) = \mathbf{0}$ . Odtiaľ:

$$\mathbf{C}\mathbf{X}\boldsymbol{\beta} + \mathbf{C}E(\mathbf{u}) = \mathbf{C}\mathbf{X}\boldsymbol{\beta} \quad (3.249)$$

Keďže odhad má byť neskreslený, musí platiť:

$$\mathbf{C}\mathbf{X}\boldsymbol{\beta} = \boldsymbol{\beta} \quad (3.250)$$

To sa však môže stať len vtedy, ak:

$$\mathbf{C}\mathbf{X} = \mathbf{I}_{k+1} \quad (3.251)$$

Definujme ďalej maticu  $\mathbf{D}$  ako:

$$\mathbf{D} = \mathbf{C} - (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \quad (3.252)$$

Motívom pre túto definíciu je ten, že  $\mathbf{D}\mathbf{y}$  predstavuje rozdiel v estimátoroch:

$$\mathbf{D}\mathbf{y} = \left( \mathbf{C} - (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \right) \mathbf{y} = \mathbf{C}\mathbf{y} - (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} = \tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}} \quad (3.253)$$

Jeden z dôsledkov definície  $\mathbf{D}$  je, že:

$$\mathbf{C} = \mathbf{D} + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \quad (3.254)$$

Pre  $\mathbf{CX}$  máme:

$$\mathbf{CX} = \left( \mathbf{D} + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \right) \mathbf{X} = \mathbf{DX} + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{X} = \mathbf{DX} + \mathbf{I}_{k+1} \quad (3.255)$$

Keďže však podľa (3.251) máme  $\mathbf{CX} = \mathbf{I}_{k+1}$ , musí platiť:

$$\mathbf{DX} + \mathbf{I}_{k+1} = \mathbf{I}_{k+1} \quad (3.256)$$

odkiaľ nutne:

$$\mathbf{DX} = \mathbf{0} \quad (3.257)$$

Ak by sme chceli vypočítať variančno-kovariančnú maticu  $\tilde{\boldsymbol{\beta}}$ , postupovali by sme podobne, ako v prípade  $\hat{\boldsymbol{\beta}}$  v rovnici (3.241), len namiesto  $(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T$  by sme používali maticu  $\mathbf{C}$ . Odhadom variančno-kovariančnej matice by bolo:

$$\text{var}(\tilde{\boldsymbol{\beta}}) = \sigma^2 \mathbf{C} \mathbf{C}^T \quad (3.258)$$

Ak dosadíme rovnicu (3.254), posledný vzťah bude mať tvar:

$$\text{var}(\tilde{\boldsymbol{\beta}}) = \sigma^2 \mathbf{C} \mathbf{C}^T = \sigma^2 \left( \mathbf{D} + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \right) \left( \mathbf{D} + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \right)^T \quad (3.259)$$

Po roznásobení:

$$\begin{aligned} & \sigma^2 \left( \mathbf{D} + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \right) \left( \mathbf{D} + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \right)^T = \\ & \sigma^2 \left( \mathbf{D} \mathbf{D}^T + \mathbf{D} \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{D}^T + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \right) \end{aligned} \quad (3.260)$$

Prostredné dva sčítance sú rovné nulovým maticiam, pretože  $\mathbf{DX} = \mathbf{X}^T \mathbf{D}^T = \mathbf{0}$ . Posledný výraz sa kvôli súčinu matice a matice k nej inverznej rovná  $(\mathbf{X}^T \mathbf{X})^{-1}$ , takže:

$$\text{var}(\tilde{\boldsymbol{\beta}}) = \sigma^2 \mathbf{D} \mathbf{D}^T + \sigma^2 (\mathbf{X}^T \mathbf{X})^{-1} = \sigma^2 \mathbf{D} \mathbf{D}^T + \text{var}(\hat{\boldsymbol{\beta}}) \quad (3.261)$$

Keďže  $\mathbf{D} \mathbf{D}^T$  je pozitívne definitná matica, rozptyly  $\tilde{\boldsymbol{\beta}}$  nemôžu byť menšie ako  $\hat{\boldsymbol{\beta}}$  a  $\tilde{\boldsymbol{\beta}}$  nemôže byť efektívnejším estimátorom ako  $\hat{\boldsymbol{\beta}}$ .  $\hat{\boldsymbol{\beta}}$  je preto efektívnym estimátorom  $\boldsymbol{\beta}$ .

V ekonometrickej teórii je známa tzv. Gauss-Markovova veta, podľa ktorej je odhad  $\hat{\boldsymbol{\beta}}$  najlepším lineárnym neskresleným odhadom  $\boldsymbol{\beta}$  (angl. *best linear unbiased estimator*, BLUE). Slovo „najlepší“ sa viaže na efektívnosť tohto estimátora – jeho neskreslenosť sme si ukázali v predchádzajúcej podkapitole.

### 3.4.4 Vlastnosti bodového odhadu – konzistentnosť

Poslednou vlastnosťou bodového odhadu, ktorá sa zvykne skúmať, je konzistentnosť. Konzistentnosť je možné zjednodušene vyjadriť ako vlastnosť, pri ktorej sa vypočítavaný odhad  $(\hat{\beta})$  pri zvyšujúcej sa vzorke blíži skutočnej hodnote odhadovaného parametra.

Konzistentný odhad má tú vlastnosť, že pre veľké vzorky by odhadnutá hodnota mala byť blízko odhadovanej – mali by sme dostávať pomerne presný výsledok. Presnosť tohto výsledku bude tým vyššia, čím väčšia je použitá vzorka.

Na prvý pohľad ide o vlastnosť, ktorá sa podobá na neskreslenosť. Pri neskreslenosti sme požadovali, aby  $E(\hat{\beta}) = \beta$ . Tento vzťah hovorí, že v prípade, ak použijeme  $\hat{\beta}$  na odhad  $\beta$ , nebudeme sa dopúšťať systematického skreslenia – nebudeme systematicky ani nadhodnocovať, ani systematicky podhodnocovať skutočný vektor hľadaných regresných koeficientov. Na druhej strane nám neskreslenosť nehovorí nič o tom, ako je to s presnosťou odhadu – vieme len, že „v priemere“ sa nedopúšťame systematického skreslenia. Naproti tomu konzistentnosť hovorí, že so zväčšujúcou sa vzorkou sa odhad spresňuje.

V skutočnosti môžeme mať prípad, v ktorom by sme využívali estimátor, ktorý by síce bol skreslený, ale bol aj konzistentný. To by znamenalo, že sa síce dopúšťame určitého skreslenia, ale že zväčšovaním vzorky je možné prísť k pomerne presnému odhadu (príklad – skutočnú hodnotu parametra populácie, ktorý nech je pre jednoduchosť číslo 5 by sme so zväčšujúcou sa vzorkou odhadovali povedzme ako 4.5, 4.8, 4.9, 4.999, 4.999999,... . Ako postupnosť naznačuje, skutočnú hodnotu vždy podhodnocujeme, odhad je zrejme skreslený. Na druhej strane sa odhad so zväčšovaním vzorky spresňuje a určite má svoj význam pre praktické použitie).

V prípade, ak by estimátor nebol konzistentný, nazývame ho nekonzistentným. Jedným z dôvodov nekonzistentnosti môže byť situácia, kedy nie je možné určiť, k čomu sa hodnota blíži pri zväčšovaní vzorky – ak hodnota nekonverguje. Horší prípad môže nastať, že sa odhad zväčšovaním vzorky blíži k určitej hodnote (konverguje), ale tá hodnota je iná ako hodnota hľadaného parametra. Tento druhý prípad je horší v tom, že pri zväčšovaní vzorky (ktoré sa spravidla realizuje s cieľom zvýšenia presnosti odhadu) sa s vyššou mierou istoty blížime k nesprávnemu číslu (alebo ako v našom prípade vektoru).

Aby sme koncepciu konzistentného odhadu uviedli formálnejšie, musíme definovať konvergenciu podľa pravdepodobnosti. Hovoríme, že estimátor  $\hat{\boldsymbol{\beta}}$  je konzistentným estimátorom  $\boldsymbol{\beta}$  vtedy, ak

$$\text{plim}_{n \rightarrow \infty}(\hat{\boldsymbol{\beta}}) = \boldsymbol{\beta} \quad (3.262)$$

Konvergencia podľa pravdepodobnosti (označovaná plim) je definovaná ako:

$$\lim_{n \rightarrow \infty}(\hat{\boldsymbol{\beta}}) = P\left(\left|\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right| < \varepsilon\right) = 1 \quad (3.263)$$

pre ľubovoľné  $\varepsilon \in \mathbb{R}^+$ .

Odhad regresných koeficientov vieme napísať v tvare:

$$\hat{\boldsymbol{\beta}} = \boldsymbol{\beta} + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{u} \quad (3.264)$$

Keďže vektor  $\boldsymbol{\beta}$  má rozmer  $(k+1) \times 1$  a nezávisí od vzorky a jej veľkosti  $n$ , bude:

$$\text{plim}_{n \rightarrow \infty}(\boldsymbol{\beta}) = \boldsymbol{\beta} \quad (3.265)$$

Zostáva výraz  $(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{u}$ . Ak by bolo  $\text{plim}_{n \rightarrow \infty}((\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{u}) = \mathbf{0}$ , estimátor  $\hat{\boldsymbol{\beta}}$  by bol

konzistentný. Problémom je, že matica  $\mathbf{X}$  má rozmery  $n \times (k+1)$  a vektor  $\mathbf{u}$  rozmery  $n \times 1$ . Súčin  $\mathbf{X}^T \mathbf{X}$  má potom síce rozmer  $(k+1) \times (k+1)$ , ale každý prvok tejto matice je súčtom  $n$  čísel. Keďže limitu počítame pre  $n \rightarrow \infty$ , je možné, že vo všeobecnosti tento súčet môže divergovať. Z tohto dôvodu sa namiesto  $\text{plim}_{n \rightarrow \infty}((\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{u}) = \mathbf{0}$  počíta:

$$\text{plim}_{n \rightarrow \infty}((\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{u}) = \text{plim}_{n \rightarrow \infty} \left( \frac{n}{n} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{u} \right) = \text{plim}_{n \rightarrow \infty} \left( \left( \frac{1}{n} \mathbf{X}^T \mathbf{X} \right)^{-1} \frac{1}{n} \mathbf{X}^T \mathbf{u} \right) \quad (3.266)$$

Za predpokladu existencie konečnej limity môžeme tento výraz ďalej napísať ako:

$$\text{plim}_{n \rightarrow \infty} \left( \left( \frac{1}{n} \mathbf{X}^T \mathbf{X} \right)^{-1} \frac{1}{n} \mathbf{X}^T \mathbf{u} \right) = \text{plim}_{n \rightarrow \infty} \left( \left( \frac{1}{n} \mathbf{X}^T \mathbf{X} \right)^{-1} \right) \text{plim}_{n \rightarrow \infty} \left( \frac{1}{n} \mathbf{X}^T \mathbf{u} \right) \quad (3.267)$$

Prvky matice  $\mathbf{X}^T \mathbf{X}/n$  už nebudú súčty, ale priemery – tie už pravdepodobne konečné byť môžu. Dá sa ukázať, že za pomerne jednoduchých predpokladov je  $\text{plim}_{n \rightarrow \infty} \left( \frac{1}{n} \mathbf{X}^T \mathbf{u} \right) = \mathbf{0}$  (Greene, 2012). Keďže  $\mathbf{X}^T \mathbf{X}/n$  je konečná matica, potom aj k nej inverzná matica obsahuje konečné prvky (predpokladáme regulárnosť  $\mathbf{X}^T \mathbf{X}$ ), a jej súčin s nulovým vektorom je nulový vektor. Potom dostávame:

$$\text{plim}_{n \rightarrow \infty}(\hat{\boldsymbol{\beta}}) = \text{plim}_{n \rightarrow \infty}(\boldsymbol{\beta} + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{u}) = \boldsymbol{\beta} \quad (3.268)$$

Estimátor  $\hat{\boldsymbol{\beta}}$  je preto konzistentným odhadom  $\boldsymbol{\beta}$ .

### 3.4.5 Odhad rozptylu poruchového člena

V predchádzajúcej časti sme si ukázali, že pre variančno-kovariančnú maticu odhadov regresných koeficientov platí vzťah  $\text{var}(\hat{\boldsymbol{\beta}}) = \sigma^2 (\mathbf{X}^T \mathbf{X})^{-1}$ . Nedostatkom je, že v ňom stále vystupuje neznámy parameter  $\sigma^2$  – empirickú variančno-kovariančnú maticu odhadov preto zatiaľ nedokážeme vypočítať. Je preto potrebné odhadnúť  $\sigma^2$ .

Pri odhade  $\sigma^2$  budeme vychádzať z výrazu  $E(\mathbf{e}^T \mathbf{e})$ . Keďže je rozptyl založený na súčte štvorcov odchýlok od priemeru, výraz  $\mathbf{e}^T \mathbf{e}$  je dobrým kandidátom – predstavuje súčet štvorcov rezíduí.

Zaujímá nás výraz  $E(\mathbf{e}^T \mathbf{e})$  – v kapitole 3.4.1 sme si však ukázali, že platí rovnosť:

$$\mathbf{e} = \mathbf{M}\mathbf{u} = (\mathbf{I}_n - \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T) \mathbf{u} \quad (3.269)$$

potom dostávame:

$$E(\mathbf{e}^T \mathbf{e}) = E(\mathbf{u}^T \mathbf{M}^T \mathbf{M} \mathbf{u}) \quad (3.270)$$

Keďže sme ukázali, že matica  $\mathbf{M}$  je symetrická a idempotentná, výraz je ekvivalentný:

$$E(\mathbf{u}^T \mathbf{M}^T \mathbf{M} \mathbf{u}) = E(\mathbf{u}^T \mathbf{M} \mathbf{M} \mathbf{u}) = E(\mathbf{u}^T \mathbf{M} \mathbf{u}) \quad (3.271)$$

V skutočnosti pri výpočte  $\mathbf{u}^T \mathbf{M} \mathbf{u}$  dostávame kvadratickú formu, ktorá predstavuje číslo – vektor  $\mathbf{u}^T$  má rozmer  $1 \times n$ , matica  $\mathbf{M}$  je rádu  $n \times n$ , a vektor  $\mathbf{u}$  má rozmer  $n \times 1$ . Výsledkom je naozaj číslo. V tomto prípade (ak pracujeme s maticou rádu  $1 \times 1$ , teda reálnym číslom) môžeme  $\mathbf{u}^T \mathbf{M} \mathbf{u}$  stotožniť so stopou matice  $\text{tr}(\mathbf{u}^T \mathbf{M} \mathbf{u})$  (Greene, 2012). Odtiaľ s využitím vlastností strednej hodnoty a vlastností stopy matice (hlavne skutočnosti, že stopa sa pri cyklickej permutácii činiteľov nemení) dostávame:

$$E(\mathbf{u}^T \mathbf{M} \mathbf{u}) = E(\text{tr}(\mathbf{u}^T \mathbf{M} \mathbf{u})) \quad (3.272)$$

$$= E(\text{tr}(\mathbf{M} \mathbf{u} \mathbf{u}^T)) \quad (3.273)$$

$$= \text{tr}(\mathbf{M} E(\mathbf{u} \mathbf{u}^T)) \quad (3.274)$$

Podľa predpokladu P2 je  $E(\mathbf{u} \mathbf{u}^T) = \sigma^2 \mathbf{I}_n$ . Potom:

$$\text{tr}(\mathbf{M} E(\mathbf{u} \mathbf{u}^T)) = \text{tr}(\mathbf{M} \sigma^2 \mathbf{I}_n) \quad (3.275)$$

$$= \sigma^2 \text{tr}(\mathbf{M}) \quad (3.276)$$

$$= \sigma^2 \text{tr}(\mathbf{I}_n - \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T) \quad (3.277)$$

$$= \sigma^2 \text{tr}(\mathbf{I}_n) - \sigma^2 \text{tr}(\mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T) \quad (3.278)$$

V predposlednom riadku sme dosadili z definície  $\mathbf{M}$ . V poslednom riadku sme zas využili, že stopa rozdielu matíc je rozdiel stôp.

$$\sigma^2 \text{tr}(\mathbf{I}_n) - \sigma^2 \text{tr}(\mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T) = \sigma^2 n - \sigma^2 \text{tr}((\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{X}) \quad (3.279)$$

Keďže stopa matice predstavuje súčet prvkov na jej hlavnej diagonále, stopa jednotkovej matice rádu  $n$  je číslo  $n$ . Zároveň si môžeme všimnúť, že v druhej stope bola opäť využitá cyklická permutácia. Opäť dostávame jednotkovú maticu, tentoraz rádu  $k + 1$ .

$$\sigma^2 n - \sigma^2 \text{tr}((\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{X}) = \sigma^2 n - \sigma^2 \text{tr}(\mathbf{I}_{k+1}) = \sigma^2 n - \sigma^2 (k + 1) = \sigma^2 (n - (k + 1)) \quad (3.280)$$

Došli sme k záveru, že:

$$E(\mathbf{e}^T \mathbf{e}) = \sigma^2 (n - (k + 1)) \quad (3.281)$$

preto:

$$E\left(\frac{\mathbf{e}^T \mathbf{e}}{n - (k + 1)}\right) = \sigma^2 \quad (3.282)$$

a výraz:

$$s^2 = \frac{\mathbf{e}^T \mathbf{e}}{n - (k + 1)} \quad (3.283)$$

je neskresleným odhadom  $\sigma^2$ .

Ak by sme chceli odhadnúť variančno-kovariančnú maticu  $\text{var}(\hat{\boldsymbol{\beta}})$ , stačí nám nahradiť  $\sigma^2$  výrazom pre  $s^2$ . Odhad variančno-kovariančnej matice vektora regresných koeficientov sa počíta ako:

$$s^2 (\mathbf{X}^T \mathbf{X})^{-1} = \frac{\mathbf{e}^T \mathbf{e}}{n - (k + 1)} (\mathbf{X}^T \mathbf{X})^{-1} \quad (3.284)$$

Odhad variančno-kovariančnej matice poruchového člena je možné využiť pri testovaní štatistických hypotéz o regresných koeficientoch. Keďže predpoklad P5 lineárneho modelu predpokladá normalitu  $\mathbf{u}$ , potom aj odhad regresných koeficientov má normálne rozdelenie, s vektorom stredných hodnôt  $\boldsymbol{\beta}$  a variančno-kovariančnou maticou  $\text{var}(\hat{\boldsymbol{\beta}}) = \sigma^2 (\mathbf{X}^T \mathbf{X})^{-1}$ . Túto maticu vieme odhadnúť ako  $s^2 (\mathbf{X}^T \mathbf{X})^{-1}$ , takže vieme odhadnúť

celé pravdepodobnostné rozdelenie estimátora  $\hat{\boldsymbol{\beta}}$ . Využívajúc tieto vlastnosti môžeme napríklad testovať štatistickú hypotézu  $H_0: \hat{\beta}_i = 0$  pre  $i \in \{0, 1, \dots, k\}$  oproti alternatíve  $H_1: \hat{\beta}_i \neq 0$ . Ak by bola nulová hypotéza pravdivá, znamenalo by to, že regresný koeficient pri premennej  $i$  je nulový – bez ohľadu na to, akú nadobudne táto premenná hodnotu, nesúvisí so zmenami hodnoty závislej premennej. Ak túto hypotézu nevieme zamietnuť, nemáme dôvod tvrdiť, že existuje vzťah medzi závislou premennou a  $i$ -tou nezávislou premennou.

Za účelom testovania tejto hypotézy môžeme zostrojiť testovaciu štatistiku:

$$t = \frac{\hat{\beta}_i}{s^2 s_{\hat{\beta}_i}} \quad (3.285)$$

kde  $s_{\hat{\beta}_i}$  je  $i$ -tý prvok hlavnej diagonály matice  $(\mathbf{X}^T \mathbf{X})^{-1}$ . Testovacia štatistika má  $t$  rozdelenie s  $n - (k + 1)$  stupňami voľnosti.

### 3.4.6 Koeficient determinácie

Definujme maticu  $\mathbf{N}$  takto:

$$\mathbf{N} = \mathbf{I}_n - \frac{1}{n} \mathbf{u} \mathbf{u}^T \quad (3.286)$$

Matica  $\mathbf{N}$  je symetrická a idempotentná. Je symetrická, pretože:

$$\mathbf{N}^T = \left( \mathbf{I}_n - \frac{1}{n} \mathbf{u} \mathbf{u}^T \right)^T = \mathbf{I}_n^T - \frac{1}{n} \mathbf{u} \mathbf{u}^T = \mathbf{I}_n - \frac{1}{n} \mathbf{u} \mathbf{u}^T = \mathbf{N} \quad (3.287)$$

Je aj idempotentná, pretože:

$$\mathbf{N} \mathbf{N} = \left( \mathbf{I}_n - \frac{1}{n} \mathbf{u} \mathbf{u}^T \right) \left( \mathbf{I}_n - \frac{1}{n} \mathbf{u} \mathbf{u}^T \right) = \mathbf{I}_n - \frac{1}{n} \mathbf{u} \mathbf{u}^T - \frac{1}{n} \mathbf{u} \mathbf{u}^T + \frac{1}{n^2} \mathbf{u} \mathbf{u}^T \mathbf{u} \mathbf{u}^T \quad (3.288)$$

Avšak v poslednom výraze predstavuje súčin  $\mathbf{u}^T \mathbf{u} = n$ , takže:

$$\mathbf{I}_n - \frac{1}{n} \mathbf{u} \mathbf{u}^T - \frac{1}{n} \mathbf{u} \mathbf{u}^T + \frac{1}{n^2} \mathbf{u} \mathbf{u}^T \mathbf{u} \mathbf{u}^T = \mathbf{I}_n - \frac{1}{n} \mathbf{u} \mathbf{u}^T - \frac{1}{n} \mathbf{u} \mathbf{u}^T + \frac{n}{n^2} \mathbf{u} \mathbf{u}^T = \mathbf{I}_n - \frac{1}{n} \mathbf{u} \mathbf{u}^T = \mathbf{N} \quad (3.289)$$

A naozaj  $\mathbf{N} \mathbf{N} = \mathbf{N}$ , čo je podmienka idempotentnosti.

Matica  $\mathbf{N}$  má tú vlastnosť, že pre každý vektor  $\mathbf{z}$  rozmeru  $n \times 1$  predstavuje súčin:

$$\mathbf{z}^T \mathbf{N} \mathbf{N} \mathbf{z} = \mathbf{z}^T \mathbf{N} \mathbf{z} = \sum_{i=1}^n (z_i - \bar{z})^2 = n \sigma_z^2 \quad (3.290)$$

súčet štvorcov odchýlok zložiek vektora  $\mathbf{z}$  od ich priemeru. Presvedčme sa o tom nasledovne.

Zoberme výraz  $\mathbf{Nz}$ .

$$\mathbf{Nz} = \left( \mathbf{I}_n - \frac{1}{n} \mathbf{u}^T \right) \mathbf{z} = \mathbf{z} - \mathbf{u} \frac{1}{n} \mathbf{u}^T \mathbf{z} \quad (3.291)$$

Už sme v predchádzajúcich kapitolách videli, že pre ľubovoľný vektor  $\mathbf{z}$  je  $\frac{1}{n} \mathbf{u}^T \mathbf{z} = \bar{z}$ ,

keďže:

$$\frac{1}{n} \mathbf{u}^T \mathbf{z} = \frac{1}{n} (1, 1, \dots, 1) \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix} = \frac{1}{n} \sum_{i=1}^n z_i \quad (3.292)$$

Dostávame teda:

$$\mathbf{Nz} = \mathbf{z} - \mathbf{u} \frac{1}{n} \mathbf{u}^T \mathbf{z} = \mathbf{z} - \mathbf{u} \bar{z} = \begin{pmatrix} z_1 - \bar{z} \\ z_2 - \bar{z} \\ \vdots \\ z_n - \bar{z} \end{pmatrix} \quad (3.293)$$

Súčin  $\mathbf{z}^T \mathbf{N} \mathbf{N} \mathbf{z}$  potom predstavuje:

$$\mathbf{z}^T \mathbf{N} \mathbf{N} \mathbf{z} = (z_1 - \bar{z}, z_2 - \bar{z}, \dots, z_n - \bar{z}) \begin{pmatrix} z_1 - \bar{z} \\ z_2 - \bar{z} \\ \vdots \\ z_n - \bar{z} \end{pmatrix} = \sum_{i=1}^n (z_i - \bar{z})^2 = n \sigma_z^2 \quad (3.294)$$

Výraz  $\mathbf{z}^T \mathbf{N} \mathbf{N} \mathbf{z} / n = \mathbf{z}^T \mathbf{N} \mathbf{z} / n$  potom predstavuje rozptyl zložiek vektora  $\mathbf{z}$ .

Poslednou vlastnosťou matice  $\mathbf{N}$ , ktoré si ukážeme, je  $\mathbf{N} \mathbf{e} = \mathbf{e}$ . Platí totiž:

$$\mathbf{N} \mathbf{e} = \left( \mathbf{I}_n - \frac{1}{n} \mathbf{u}^T \right) \mathbf{e} = \mathbf{e} - \frac{1}{n} \mathbf{u}^T \mathbf{e} = \mathbf{e} \quad (3.295)$$

keďže  $\mathbf{u}^T \mathbf{e} = 0$ . Pripomeňme ešte, že v kapitole 3.4.1 sme dokázali aj vzťah  $\mathbf{X}^T \mathbf{e} = \mathbf{0}$ .

Pokúsme sa teraz vypočítať výraz  $\mathbf{y}^T \mathbf{N} \mathbf{y}$ , ktorý súvisí s rozptylom zložiek  $\mathbf{y}$  (chýba mu člen  $1/n$ ):

$$\mathbf{y}^T \mathbf{N} \mathbf{y} = (\hat{\mathbf{y}} + \mathbf{e})^T \mathbf{N} (\hat{\mathbf{y}} + \mathbf{e}) \quad (3.296)$$

$$= (\hat{\mathbf{y}}^T \mathbf{N} + \mathbf{e}^T \mathbf{N}) (\hat{\mathbf{y}} + \mathbf{e}) \quad (3.297)$$

$$= \hat{\mathbf{y}}^T \mathbf{N} \hat{\mathbf{y}} + \hat{\mathbf{y}}^T \mathbf{N} \mathbf{e} + \mathbf{e}^T \mathbf{N} \hat{\mathbf{y}} + \mathbf{e}^T \mathbf{N} \mathbf{e} \quad (3.298)$$

$$= \hat{\mathbf{y}}^T \mathbf{N} \hat{\mathbf{y}} + 2 \hat{\mathbf{y}}^T \mathbf{e} + \mathbf{e}^T \mathbf{e} \quad (3.299)$$

$$= \hat{\mathbf{y}}^T \mathbf{N} \hat{\mathbf{y}} + 2\hat{\boldsymbol{\beta}}^T \mathbf{X}^T \mathbf{e} + \mathbf{e}^T \mathbf{e} \quad (3.300)$$

$$= \hat{\mathbf{y}}^T \mathbf{N} \hat{\mathbf{y}} + \mathbf{e}^T \mathbf{e} \quad (3.301)$$

V štvrtom riadku sme využili fakt, že  $\hat{\mathbf{y}}^T \mathbf{N} \mathbf{e}$  je reálne číslo ( $\hat{\mathbf{y}}^T$  má rozmer  $1 \times n$ , matica  $\mathbf{N}$  má rozmer  $n \times n$  a vektor  $\mathbf{e}$  je rozmeru  $n \times 1$ ). Zároveň  $\hat{\mathbf{y}}^T \mathbf{N} \mathbf{e} = (\mathbf{e}^T \mathbf{N} \hat{\mathbf{y}})^T$ , ale keďže matica rádu  $1 \times 1$  sa vždy rovná svojej transponovanej matici (je to len jedno číslo), platí  $\hat{\mathbf{y}}^T \mathbf{N} \mathbf{e} = \mathbf{e}^T \mathbf{N} \hat{\mathbf{y}}$  a následne  $\hat{\mathbf{y}}^T \mathbf{N} \mathbf{e} + \mathbf{e}^T \mathbf{N} \hat{\mathbf{y}} = 2\hat{\mathbf{y}}^T \mathbf{N} \mathbf{e}$ . V poslednej rovnici sme zas využili už spomínanú rovnosť  $\mathbf{X}^T \mathbf{e} = \mathbf{0}$ .

Ak by sme obidve strany vydělili číslom  $n$ , dostaneme:

$$\frac{1}{n} \mathbf{y}^T \mathbf{N} \mathbf{y} = \frac{1}{n} \hat{\mathbf{y}}^T \mathbf{N} \hat{\mathbf{y}} + \frac{1}{n} \mathbf{e}^T \mathbf{e} \quad (3.302)$$

$$\sigma_y^2 = \sigma_{\hat{\mathbf{y}}}^2 + \sigma_e^2$$

Výsledok, ktorý sme dostali znamená, že rozptyl vysvetľovanej premennej  $\mathbf{y}$ , a tým aj jeho variabilita, môže byť rozložený na dve časti: rozptyl vyrovnaných hodnôt, ktorý nazývame aj vysvetlenou variabilitou a rozptyl rezíduí, ktorý vysvetliť nevieme. Koncepciu vysvetlenej variability je možné pochopiť nasledovne. Ak existuje vzťah medzi vysvetľujúcimi a vysvetľovanou premennou, tak pri rôznych hodnotách vysvetľujúcich premenných prirodzene očakávame, že sa bude meniť aj vysvetľovaná premenná – je to dané ich vzájomným vzťahom. Variabilitu spôsobenú rezíduami si ale vysvetliť nevieme.

Uvedený rozklad nám môže byť užitočný pri konštrukcii ukazovateľa kvality modelu. Zrejme čím väčší je podiel vysvetlenej variability na celkovej variabilite  $\mathbf{y}$ , tým lepšie model vysvetľuje premennú  $\mathbf{y}$ .

Ukazovateľ dávajúci do pomeru spomínané premenné sa nazýva koeficient determinácie a je daný vzťahom:

$$R^2 = \frac{\hat{\mathbf{y}}^T \mathbf{N} \hat{\mathbf{y}}}{\mathbf{y}^T \mathbf{N} \mathbf{y}} = 1 - \frac{\mathbf{e}^T \mathbf{e}}{\mathbf{y}^T \mathbf{N} \mathbf{y}} \quad (3.303)$$

Keďže čitateľ v prvom zlomku je len časťou výrazu v menovateli a oba výrazy sú nezáporné, vždy platí  $0 \leq R^2 \leq 1$ . Hodnota 0 znamená, že model vôbec nevysvetľuje variabilitu závislej premennej, hodnota rovná 1 zasa znamená, že vzťah medzi nezávislými premennými a závislou premennou je možné exaktne popísať regresnou nadrovinou a rezíduá sú nulové.

Pri porovnávaní rôznych modelov pre tú istú závislú premennú  $\mathbf{y}$  je známy jeden nedostatok  $R^2$ . Koeficient determinácie je neklesajúcou funkciou počtu vysvetľujúcich premenných v modeli – to znamená, že by sme podľa neho do modelu mali pridávať premenné aj vtedy, keď je prírastok vysvetlenej variability veľmi malý. V praxi preferujeme skôr modely s čo najnižším počtom premenných. Aby bolo možné urobiť určitý kompromis medzi počtom premenných a konceptom koeficientu determinácie, často sa používa takzvaný korigovaný (angl. adjusted) koeficient determinácie. Korekcia sa realizuje na základe vzťahu veľkosti vzorky a počtu odhadovaných parametrov. O zaradení dodatočnej premennej do modelu môžeme rozhodovať na základe korigovaného koeficientu determinácie – ak by sa pridaním premennej jeho hodnota nezvýšila, premenná nepomáha vysvetliť závislú premennú dostatočne na to, aby sme ju v modeli ponechali. Korigovaný koeficient determinácie nám preto umožňuje nájsť kompromis medzi dvomi protichodnými cieľmi – kvalitou vyrovnaní a požiadavkou na čo najnižší počet vysvetľujúcich premenných.

$$R_a^2 = 1 - \frac{\frac{\mathbf{e}^T \mathbf{e}}{n - (k + 1)}}{\frac{\mathbf{y}^T \mathbf{N} \mathbf{y}}{n - 1}} = 1 - \frac{n - 1}{n - (k + 1)} \frac{\mathbf{e}^T \mathbf{e}}{\mathbf{y}^T \mathbf{N} \mathbf{y}} \quad (3.304)$$

Koeficient determinácie je možné využiť aj na testovanie štatistických hypotéz. Konkrétne je možné testovať lineárny model ako celok pomocou hypotézy:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0 \quad (3.305)$$

Oproti alternatívnej hypotéze:

$$H_1 : \exists \beta_i \neq 0, \quad i \in \{1, 2, \dots, k\} \quad (3.306)$$

Ak by sme nulovú hypotézu nevedeli vyvrátiť, znamenalo by to, že je možné, že neexistuje vzťah medzi ktoroukoľvek z vysvetľujúcich premenných a závislou premennou – inak povedané, model ako celok nedokáže vysvetliť, ani preukázať vzťah medzi premennými.

Testovacou štatistikou pre túto hypotézu je:

$$F = \frac{R^2}{1 - R^2} \frac{n - (k + 1)}{k} \quad (3.307)$$

Táto štatistika má  $F$  rozdelenie s  $k$  a  $n - (k + 1)$  stupňami voľnosti.

S pomocou koeficientu determinácie je možné realizovať aj ďalšie štatistické testy, napríklad test na lineárne reštrikcie. Povedzme napríklad, že by sme mali dve skupiny vysvetľujúcich premenných,  $\mathbf{X}_1$  a  $\mathbf{X}_2$ . Nech má matica  $\mathbf{X}_1$   $k_1$  stĺpcov a matica  $\mathbf{X}_2$   $k_2$  stĺpcov (premených), a  $k = k_1 + k_2$ . Jeden z modelov, ktorý by sme mohli kvantifikovať, je model:

$$\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \mathbf{u} = \begin{bmatrix} \mathbf{X}_1 & \mathbf{X}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{bmatrix} + \mathbf{u} \quad (3.308)$$

kde posledný výraz predstavuje zápis cez blokové matice, definované v Kapitole 1.3.7.

Alternatívnym modelom by mohol byť nasledovný model:

$$\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{u} \quad (3.309)$$

Tento model dostaneme z pôvodného, ak by platilo:

$$\boldsymbol{\beta}_2 = \mathbf{0} \quad (3.310)$$

Ak by sme chceli otestovať, či je posledná rovnosť pravdivá, mohli by sme vychádzať z kvality vyrovnaní oboma modelmi, t. j. z toho, ako dobre modely popisujú závislú premennú. Je dôležité, že závislá premenná je v oboch prípadoch rovnaká. Model v tvare (3.308) nazývame modelom bez reštrikcií a model (3.309) obmedzeným modelom. Koeficienty determinácie pre tieto modely budeme označovať  $R^2_{BR}$  a  $R^2_O$ .

Test je možné realizovať prostredníctvom štatistiky:

$$F = \frac{R^2_{BR} - R^2_O}{1 - R^2_{BR}} \frac{n - k}{k_2} \quad (3.311)$$

Ak je štatistika  $F$  väčšia, ako kritická hodnota  $F$  rozdelenia s  $k_2$  a  $n - k$  stupňami voľnosti, hypotézu o  $\boldsymbol{\beta}_2 = \mathbf{0}$  zamietame.

Sú známe aj ďalšie testy, ktoré testujú podobnú hypotézu, ako sme si práve ukázali: patrí medzi ne tzv. Waldov test, LM test a LR test. Tieto testy sú asymptoticky ekvivalentné, no pri malých vzorkách nemajú lepšie vlastnosti, ako spomínaný  $F$  test.

### Príklad 3.9

V tomto príklade sa pokúsime o kvantifikáciu jedného zo základných tvrdení vyučovaných v rámci makroekonómie: „spotreba je funkciou dôchodku“ (John Maynard Keynes). Vychádzať budeme z údajov uvedených v nasledujúcej tabuľke.

Tabuľka 6: HDP a spotreba za obdobie 1995 – 2010 v SR

Rok	HDP	CONS
1995	7.0	5.1
1996	7.7	5.9
1997	8.3	6.3
1998	8.8	6.8
1999	9.0	6.9
2000	9.5	7.3
2001	10.4	8.1
2002	11.1	8.7
2003	11.5	8.9
2004	12.3	9.4
2005	13.5	10.3
2006	15.0	11.4
2007	16.9	12.4
2008	18.2	13.6
2009	17.0	13.8
2010	18.0	14.0

Zdroj: Eurostat

Pozn.: údaje sú v parite kúpnej sily na jedného obyvateľa v trhových cenách

Premenná CONS predstavuje spotrebu a dôchodok bude reprezentovaný premennou HDP (hrubý domáci produkt).

**Úlohy:**

1. Zobrazte graficky skúmanú funkčnú závislosť.
2. Kvantifikujte regresný model, na základe ktorého by bolo možné overiť tvrdenie, že spotreba je funkciou dôchodku.
3. Rezíduá z modelu zobrazte graficky.
4. Overte, či platí  $\mathbf{1}^T \mathbf{e} = 0$ .
5. Interpretujte získané výsledky a vyjadrite sa k hraničnému sklonu k spotrebe.

**Riešenie:**

Keďže v tvrdení, ktoré máme overiť (či spotreba je funkciou dôchodku) je presne dané, ktorá premenná je vysvetľovaná, kvantifikovať budeme nasledujúci model:

$$CONS_t = \beta_0 + \beta_1 HDP_t + u_t \tag{3.312}$$

kde:

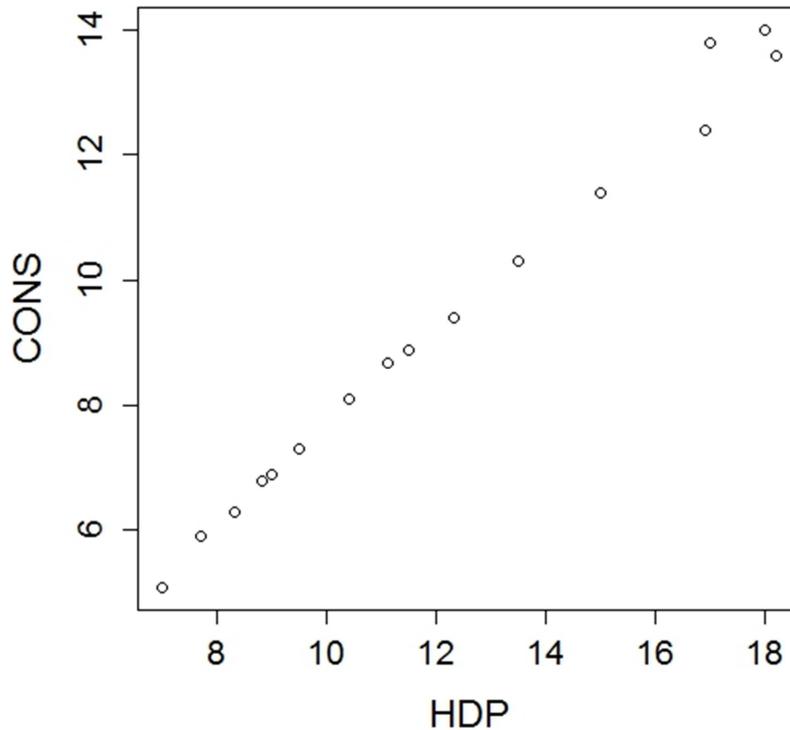
$t = 1, 2, \dots, T$

$CONS_t$  – spotreba v  $t$ -tom roku

$u_t$  – náhodná chyba

$T$  – počet pozorování     $HDP_t$  – dôchodok v  $t$ -tom roku     $\beta_0, \beta_1$  – regresné koeficienty

```
> HDP <- c(7.0, 7.7, 8.3, 8.8, 9.0, 9.5, 10.4, 11.1, 11.5, 12.3,
13.5, 15.0, 16.9, 18.2, 17.0, 18.0)
> CONS <- c(5.1, 5.9, 6.3, 6.8, 6.9, 7.3, 8.1, 8.7, 8.9, 9.4,
10.3, 11.4, 12.4, 13.6, 13.8, 14.0)
> plot(CONS ~ HDP, cex.axis=1.3, cex.lab=1.5)
```



Obrázok 12:  $x$ - $y$  graf vzťahu spotreby a dôchodku

*Zdroj: výstup zo softvéru R*

Už z daného grafu môžeme vidieť, že existuje silná lineárna závislosť medzi skúmanými premennými. V záujme korektnosti a exaktnosti je však nutné daný vzťah kvantifikovať<sup>4</sup>.

```
> Reg.Model = lm(CONS ~ HDP)
> summary (Reg.Model)

Call:
lm(formula = CONS ~ HDP)

Residuals:
    Min       1Q   Median       3Q      Max
```

<sup>4</sup> Pri tomto príklade sa dopúšťame nepresnosti, pretože do úvahy neberieme (ne)stacionárnosť časových radov a nebudeme riešiť ani predpoklady regresného modelu. Našich cieľom je len matematická kvantifikácia modelu a interpretácie koeficientov.

```

-0.57224 -0.06871 0.00923 0.09616 0.75078

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.03674    0.25126  -0.146   0.886
HDP          0.76976    0.01979  38.887 1.15e-15 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2941 on 14 degrees of freedom
Multiple R-squared: 0.9908,    Adjusted R-squared: 0.9902
F-statistic: 1512 on 1 and 14 DF,  p-value: 1.149e-15

```

Tento Koeficient pri HDP je možné interpretovať ako hraničným sklon k spotrebe. Hodnota koeficientu je kladná, čo naznačuje priamy vzťah medzi úrovňou HDP a spotrebou. Tento výsledok sa dal očakávať a je v súlade so základnou ekonomickou teóriou. Pri väčších hodnotách HDP pozorujeme väčšie hodnoty spotreby. Odhad regresného koeficientu pri HDP má hodnotu  $\approx 0.77$ , čo naznačuje, že s prírastkom HDP (nezávislá premenná) o jednu jednotku spotreba (závislá premenná) narastá o 0.77 jednotiek. Táto hodnota je pritom pomerne dosť vysoká a naznačuje, že pomerne veľkú časť produktu vynakladajú na spotrebu. Tento stav je skôr typické pre menej rozvinuté krajiny.

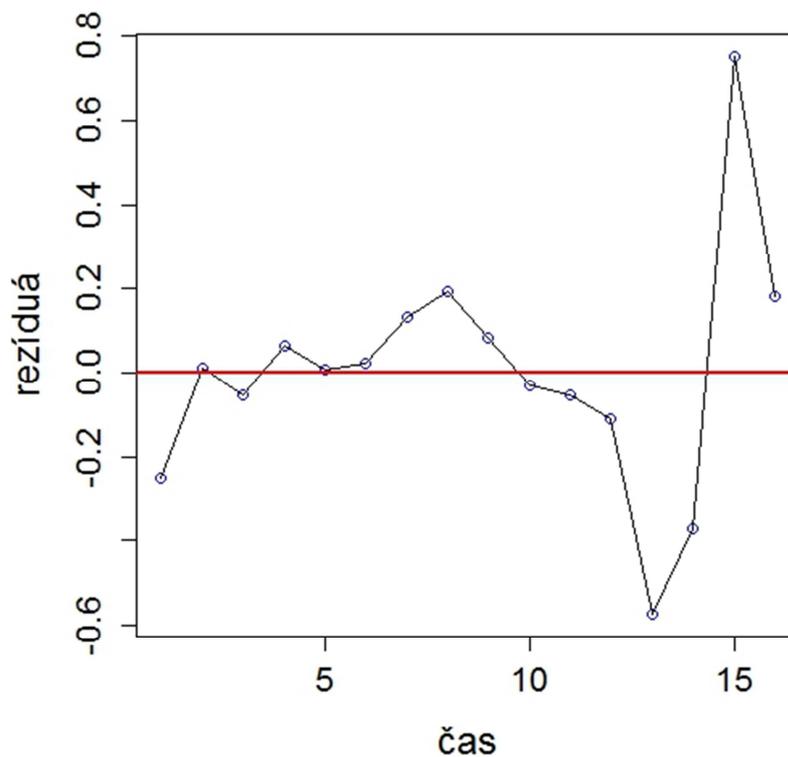
Hodnotiť významnosti koeficientov je problematické, keďže ide o časové rady a neoverili sme predpoklad stacionarity. Z tohto dôvodu by sme nemali dôverovať výsledkom z regresnej analýzy. Taktiež si všimnime pomerne vysokú hodnotu koeficientu determinácie vieme povedať, že model popisuje 99.02 % variability. V spoločenských vedách, najmä pri analýze časových radov, sú takéto výsledky mimoriadne ojedinelé a skôr naznačujú, že tento model je zrejme chybný. Ako sa neskôr ukáže problematický je tu práve predpoklad o stacionarite. Ak majú obe premenné rovnaký deterministický trend, alebo rovnaký charakter stochastického trendu, potom sa medzi nimi môže identifikovať silný vzťah (ako je to aj v našom prípade), ale v skutočnosti je to len spoločný trend.

Ostala nám ešte vizualizácia rezíduí a kontrola predpokladu o ich nulovej strednej hodnote (resp.  $\mathbf{1}^T \mathbf{u} = 0$ ), ktorú vykonáme napríklad pomocou výberového aritmetického priemeru.

```

> plot(Reg.Model$residuals, col = "blue", cex.axis = 1.3,
      cex.lab = 1.5)
> lines(Reg.Model$residuals)
#Úloha 3 => zobrazí rezíduá z modelu v grafe

```



Obrázok 13: Rezíduá z regresného modelu

*Zdroj: výstup zo softvéru R*

```
> mean(Reg.Model$residuals)
#Úloha 4 => zobrazí priemer vypočítaných rezíduí
[1] 8.239937e-18
```

Obrázok rezíduá nám naznačuje, či existujú určité intervaly systematického nad- alebo pod-hodnocovania predikovaných hodnôt spotreby. Napríklad od štvrtého až po deviate pozorovanie model mierne podhodnocoval spotrebu, čo vidno z toho, že rezíduá sú väčšie ako 0. Zároveň vidíme, že model relatívne stratil svoju presnosť v posledných rokoch 2008, 2009 a 2010, čo je práve obdobie, ktoré bolo charakteristické prudkým a náhlým spomalením (dokonca poklesom) ekonomického rastu. Takýto obrázok teda naznačuje aj prítomnosť heteroskedasticity, keďže rozptyl rezíduí sa na konci sledovaného obdobia javí byť väčším ako na začiatku.

### 3.5 Porušenie predpokladov lineárneho modelu

Výsledky v predchádzajúcich kapitolách nám umožňujú pracovať s lineárnymi modelmi – hovorili sme o ich špecifikácii, o ich odhade a vlastnostiach. Celý výklad sa

doposiaľ zásadným spôsobom opieral o predpoklady, ktoré sme na model kládli. V podkapitole 3.3.3 sme si uviedli všetky predpoklady aj s popisom dôsledkov, ak nie sú dodržané. V časti o efektívnosti bodového odhadu regresných koeficientov sme videli, že štatistické vlastnosti estimátora metódy najmenších štvorcov záviseli na predpoklade P2.

Nie je pravdou, že v prípade, ak sú predpoklady modelu porušené, je celý model zbytočný. Podľa toho, o ktorý predpoklad ide, a akým spôsobom je porušený si model môže zachovávať niektoré dobré vlastnosti, ale niektoré môže strácať – je preto dôležité poznať dôsledky porušenia predpokladov.

V tejto kapitole sa budeme zaoberať hlavne porušením predpokladu P2 – tento predpoklad je porušený najčastejšie, a má asi aj najväčšie dôsledky. V krátkosti ale spomeňme aj ostatné predpoklady.

Predpoklad P1 o nulovej strednej hodnote vektora  $\mathbf{u}$  sme už podrobnejšie popísali – jeho nedodržanie sa prejaví hlavne v skreslenom odhade konštanty  $\beta_0$ . Predpoklad P3 o deterministickom charaktere matice  $\mathbf{X}$  nie je (za splnení niektorých dodatočných predpokladov) reštriktívnym – väčšina výsledkov platí aj pre stochastickú maticu  $\mathbf{X}$ . Predpoklad P4 hovorí o hodnosti matice  $\mathbf{X}$  – na porušenie tohto predpokladu nás okamžite upozorní štatistický softvér, pretože v prípade jeho nastania nie je možné vypočítať odhad regresných koeficientov. Predpoklad P5 je o normalite vektora  $\mathbf{u}$  – jeho porušenie ovplyvňuje výsledky štatistických testov, ale v niektorých prípadoch tiež nie je kritické, ak sa je možné pri väčších vzorkách spoľahnúť na centrálnu limitnú vetu, ktorá hovorí o asymptoticky normálnom pravdepodobnostnom rozdelení priemerov. Aj z uvedeného zoznamu možných porušení predpokladov vidíme, že predpoklad P2 má medzi predpokladmi významné postavenie – jeho porušenie má vážnejšie dôsledky.

### 3.5.1 Dôsledky porušenia predpokladov o homoskedasticite a autokorelácii

Doposiaľ sme sa zaoberali prípadmi, ak platil predpoklad P2, teda platilo:

$$E(\mathbf{u}\mathbf{u}^T) = \sigma^2 \mathbf{I}_n = \begin{pmatrix} \sigma^2 & 0 & \dots & 0 \\ 0 & \sigma^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma^2 \end{pmatrix} \quad (3.313)$$

pre  $\sigma^2 \in \mathbb{R}^+$ . Pripomeňme, že ide o predpoklad o tvare variančno-kovariančnej matice vektora náhodných porúch  $\mathbf{u}$ . Podľa predpokladu majú zložky vektora  $\mathbf{u}$  ten istý rozptyl  $\sigma^2$  (homoskedasticita) a sú vzájomne nekorelované.

Z uvedeného vyplýva, že k porušeniu predpokladu môže prísť tromi spôsobmi – buď sú zložky  $\mathbf{u}$  heteroskedastické (majú rôzny rozptyl) – vtedy by boli prvky matice  $E(\mathbf{u}\mathbf{u}^T)$  na hlavnej diagonále rôzne. Druhou možnosťou je, že sú zložky  $\mathbf{u}$  korelované – to by znamenalo, že mimodiagonálne prvky matice  $E(\mathbf{u}\mathbf{u}^T)$  by neboli nulové. Tretou možnosťou je kombinácia predchádzajúcich – k porušeniu predpokladov môže dochádzať tak na diagonálnych, ako aj na mimodiagonálnych prvkoch.

Uvažujme preto všeobecný prípad, kedy pripúšťame všetky alternatívy. V takom prípade zavedme maticu  $\mathbf{\Omega}$ , ktorá má rozmer  $n \times n$ . Jej bližšiu štruktúru bližšie nešpecifikujeme. Predpokladajme tiež, že ako jediný je porušený predpoklad o variančno-kovariančnej matici vektora  $\mathbf{u}$ , ostatné predpoklady nech sú splnené. V rozpore s predpokladom P2 nech:

$$E(\mathbf{u}\mathbf{u}^T) = \mathbf{\Omega} \quad (3.314)$$

Ak by sme chceli posúdiť dôsledky porušenia P2, stačí nám zistiť, pri ktorých výsledkoch bol tento predpoklad použitý.

Pri odhade metódou najmenších štvorcov sme získali odhad  $\hat{\boldsymbol{\beta}}$  minimalizáciou sumy štvorcov rezíduí  $\mathbf{e}^T\mathbf{e}$ . V tomto kroku sme predpoklad nepoužívali, takže zatiaľ k zmenám nedochádza.

Ďalej sme posudzovali vlastnosti bodového odhadu  $\hat{\boldsymbol{\beta}}$ . Zistili sme, že išlo o neskreslený odhad. Využili sme pritom skutočnosť, že strednú hodnotu  $\hat{\boldsymbol{\beta}}$  je možné zapísať v tvare:

$$E(\hat{\boldsymbol{\beta}}) = \boldsymbol{\beta} + (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T E(\mathbf{u}) = \boldsymbol{\beta} \quad (3.315)$$

pričom posledná rovnosť vyplývala z predpokladu P1,  $E(\mathbf{u}) = \mathbf{0}$ . Keďže teraz uvažujeme o porušení predpokladu P2, predpokladáme, že predpoklad P1 platí a preto aj v tejto situácii zostáva  $\hat{\boldsymbol{\beta}}$  neskresleným odhadom.

Ďalšou vlastnosťou, ktorú sme skúmali, bola efektívnosť odhadu  $\hat{\boldsymbol{\beta}}$ . Je pravdou, že má tento estimátor zo všetkých neskreslených odhadov najmenší rozptyl? V podkapitole 3.4.3 sme ukázali, že pre variančno-kovariančnú maticu  $\hat{\boldsymbol{\beta}}$  platí:

$$\text{var}(\hat{\boldsymbol{\beta}}) = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T E(\mathbf{u}\mathbf{u}^T) \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \quad (3.316)$$

Ak sme podľa predpokladu P2 mohli za  $E(\mathbf{u}\mathbf{u}^T)$  dosadiť  $\sigma^2 \mathbf{I}_n$ , výraz sa zjednodušil na:

$$\text{var}(\hat{\boldsymbol{\beta}}) = \sigma^2 (\mathbf{X}^T \mathbf{X})^{-1} \quad (3.317)$$

Každý štatistický softvér sa pri výpočte štatistických testov riadi poslednou rovnosťou.

Ak je však predpoklad P2 porušený, skutočná variančno-kovariančná matica  $\hat{\boldsymbol{\beta}}$  má tvar:

$$\text{var}(\hat{\boldsymbol{\beta}}) = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \boldsymbol{\Omega} \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \quad (3.318)$$

Softvér v takom prípade používa nesprávny vzorec. V dôsledku toho budú nesprávne odhadnuté rozptyly  $\hat{\boldsymbol{\beta}}$  a všetky realizované štatistické testy ( $t$ -test,  $F$ -test) budú nesprávne. Napriek tomu, že by výstup z programu indikoval štatisticky významné výsledky, tieto nebudú smerodajné, keďže sú vypočítané nesprávne.

Ďalším dôsledkom je, že nie je isté, či variančno-kovariančná matica (3.318) predstavuje najmenší možný rozptyl pre odhad  $\hat{\boldsymbol{\beta}}$  už nemusí byť efektívnym odhadom.

Dôsledky porušenia predpokladu P2 sú teda rozsiahle: síce môžeme vypočítať odhad regresných koeficientov metódou najmenších štvorcov a tento odhad bude neskreslený, je ale možné, že bude neefektívny (informácia zo vzorky by mohla byť využitá aj na presnejší odhad, ako dostávame), a čo je horšie, zlyhávajú testy v rámci indukčnej štatistiky.

V nasledujúcej časti sa budeme zaoberať všeobecným riešením problému porušenia predpokladu P2 zovšeobecnenou metódou najmenších štvorcov, a následne si ukážeme konkrétne spôsoby, ako zistiť prítomnosť heteroskedasticity a autokorelácie, prípadne ako si počínať v prípade, ak nami skúmaný model vykazuje tieto porušenia predpokladov.

### 3.5.2 Zovšeobecnená metóda najmenších štvorcov

Ekonometrickému modelu, pri ktorom vychádzame z populačného modelu:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \quad (3.319)$$

a o vektore poruchových členov predpokladáme:

$$E(\mathbf{u}\mathbf{u}^T) = \boldsymbol{\Omega} \quad (3.320)$$

hovoríme zovšeobecnený lineárny model (angl. *generalised linear model*).

Problémom je v našom prípade matica  $\boldsymbol{\Omega}$ . Podobne ako v prípade bežných testov indukčnej štatistiky aj tentoraz môžeme najprv uvažovať so situáciou, kedy by sme maticu  $\boldsymbol{\Omega}$  považovali za známu. Uvažujeme teda o prípade, ak je porušený predpoklad P2, ale my

vieme, ako *presne* je porušený. Nasledovný postup realizujeme podľa dôkazu uvedeného v Davidson a MacKinnon (2003, s. 256).

K matici  $\Omega$  je užitočné nájsť maticu  $\Psi$  rozmeru  $n \times n$ , ktorá má vlastnosť:

$$\Omega^{-1} = \Psi \Psi^T \quad (3.321)$$

Ku každej symetrickej pozitívne definitnej matici je vždy možné takúto maticu nájsť – keďže je  $\Omega$  variančno-kovariančnou maticou, musí byť aj symetrická, aj pozitívne definitná. Podmienku teda splňa a vždy k nej bude možné nájsť vhodnú maticu  $\Psi$ . Matica  $\Psi$  dokonca nebude jedinečná – väčšinou sa na ňu kladú ešte dodatočné podmienky, ako napríklad požiadavka na symetrickosť, alebo aby bola trojuholníkovou maticou. Vhodnú maticu je možné nájsť napríklad Choleského dekompozíciou matice  $\Omega^{-1}$ .

Matica  $\Psi$  sa využíva na špecifikáciu transformovaného modelu:

$$\Psi^T \mathbf{y} = \Psi^T \mathbf{X} \boldsymbol{\beta} + \Psi^T \mathbf{u} \quad (3.322)$$

Ak by sme označili  $\mathbf{y}' = \Psi^T \mathbf{y}$ ,  $\mathbf{X}' = \Psi^T \mathbf{X}$  a  $\mathbf{u}' = \Psi^T \mathbf{u}$ . Po zavedení substitúcií dostávame:

$$\mathbf{y}' = \mathbf{X}' \boldsymbol{\beta} + \mathbf{u}' \quad (3.323)$$

Už vieme, že odhad pomocou metódy najmenších štvorcov bude mať tvar:

$$\begin{aligned} \hat{\boldsymbol{\beta}} &= (\mathbf{X}'^T \mathbf{X}')^{-1} \mathbf{X}'^T \mathbf{y}' & (3.324) \\ &= (\mathbf{X}^T \Psi \Psi^T \mathbf{X})^{-1} \mathbf{X}^T \Psi \Psi^T \mathbf{y} \\ &= (\mathbf{X}^T \Omega^{-1} \mathbf{X})^{-1} \mathbf{X}^T \Omega^{-1} \mathbf{y} \end{aligned}$$

Tento odhad bude neskreslený, pretože:

$$E(\hat{\boldsymbol{\beta}}) = E\left((\mathbf{X}^T \Omega^{-1} \mathbf{X})^{-1} \mathbf{X}^T \Omega^{-1} \mathbf{y}\right) \quad (3.325)$$

$$= E\left((\mathbf{X}^T \Omega^{-1} \mathbf{X})^{-1} \mathbf{X}^T \Omega^{-1} (\mathbf{X} \boldsymbol{\beta} + \mathbf{u})\right) \quad (3.326)$$

$$= E\left((\mathbf{X}^T \Omega^{-1} \mathbf{X})^{-1} \mathbf{X}^T \Omega^{-1} \mathbf{X} \boldsymbol{\beta} + (\mathbf{X}^T \Omega^{-1} \mathbf{X})^{-1} \mathbf{X}^T \Omega^{-1} \mathbf{u}\right) \quad (3.327)$$

$$= E\left(\boldsymbol{\beta} + (\mathbf{X}^T \Omega^{-1} \mathbf{X})^{-1} \mathbf{X}^T \Omega^{-1} \mathbf{u}\right) \quad (3.328)$$

$$= \boldsymbol{\beta} + (\mathbf{X}^T \Omega^{-1} \mathbf{X})^{-1} \mathbf{X}^T \Omega^{-1} E(\mathbf{u}) \quad (3.329)$$

$$= \boldsymbol{\beta} \quad (3.330)$$

Ako je to s predpokladmi modelu? Pre variančno-kovariančnú maticu poruchových členov máme:

$$E(\mathbf{u}' \mathbf{u}'^T) = E(\Psi^T \mathbf{u} \mathbf{u}^T \Psi) \quad (3.331)$$

$$= \Psi^T E(\mathbf{u} \mathbf{u}^T) \Psi \quad (3.332)$$

$$= \Psi^T \Omega \Psi \quad (3.333)$$

Na začiatku sme definovali  $\Omega^{-1} = \Psi \Psi^T$ , takže  $\Omega = (\Psi \Psi^T)^{-1}$ . Pre posledný výraz z vlastností inverznej matice dostávame  $(\Psi \Psi^T)^{-1} = (\Psi^T)^{-1} \Psi^{-1}$ . Po dosadení dostaneme:

$$\Psi^T \Omega \Psi = \Psi^T (\Psi^T)^{-1} \Psi^{-1} \Psi = \mathbf{I}_n \quad (3.334)$$

Transformovaný vektor poruchových členov spĺňa predpoklad P2, pretože:

$$E(\mathbf{u}'\mathbf{u}'^T) = \mathbf{I}_n \quad (3.335)$$

Variančno-kovariančná matica odhadov regresných koeficientov má tvar:

$$\text{var}(\hat{\boldsymbol{\beta}}) = (\mathbf{X}^T \Psi \Psi^T \mathbf{X})^{-1} \quad (3.336)$$

Dá sa ukázať (Davidson – MacKinnon, 2003), že estimátor získaný zovšeobecnenou metódou najmenších štvorcov (angl. *generalized least squares*, GLS) je neskreslený a efektívny odhad  $\boldsymbol{\beta}$ . Aj keď má dobré vlastnosti, použitie tohto estimátora je stále problematické, keďže predpokladá, že poznáme variančno-kovariančnú maticu  $\Omega$ .

Pomerne známou je aj metóda tzv. vážených najmenších štvorcov (angl. *weighted least squares*). Je špeciálnym prípadom vyššie popísanej zovšeobecnenej metódy najmenších štvorcov pre prípad, kedy je matica  $\Omega$  diagonálna. Môžeme ju potom napísať v tvare:

$$\Omega = \begin{pmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_n^2 \end{pmatrix} \quad (3.337)$$

kde  $\sigma_i^2 \in \mathbb{R}^+$  pre  $i = 1, 2, \dots, n$ . Pri takomto porušení predpokladu P2 (označovanom aj ako heteroskedasticita) môžeme v prípade známej matice  $\Omega$  odhadnúť model tak, že každú regresnú rovnicu vydělíme  $\sigma_i$ . Transformácia by potom vyzerala takto:

$$y_i/\sigma_i = \beta_0/\sigma_i + \beta_1 x_{i1}/\sigma_i + \beta_2 x_{i2}/\sigma_i + \dots + \beta_k x_{ik}/\sigma_i + u_i/\sigma_i, \quad i = 1, 2, \dots, n \quad (3.338)$$

Ide o špeciálny prípad zovšeobecneneho lineárneho modelu, pretože inverznou maticou k  $\Omega$  definovanej vyššie je matica:

$$\Omega^{-1} = \begin{pmatrix} \frac{1}{\sigma_1^2} & 0 & \dots & 0 \\ 0 & \frac{1}{\sigma_2^2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\sigma_n^2} \end{pmatrix} \quad (3.339)$$

Odtiaľ vidíme, že pre  $\Psi$  stačí zvoliť:

$$\Psi = \begin{pmatrix} 1 & 0 & \dots & 0 \\ \sigma_1 & & & \\ 0 & \frac{1}{\sigma_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\sigma_n} \end{pmatrix} \quad (3.340)$$

### 3.5.3 Heteroskedasticita

O heteroskedasticite hovoríme, ak poruchové členy  $\mathbf{u}$  nemajú rovnaký rozptyl, čo je v rozpore s predpokladom P2. Variančno-kovariančná matica prvkov  $\mathbf{u}$  potom vyzerá tak, ako sme si to popísali v predchádzajúcej podkapitole pri metóde vážených najmenších štvorcov.

$$\Omega = \begin{pmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_n^2 \end{pmatrix} \quad (3.341)$$

Príčiny heteroskedasticity môžu byť rôzne – niekedy súvisí heteroskedasticita s niektorou z vysvetľujúcich premenných – napr. čím je niektorá vysvetľujúca premenná vyššia, tým vyšší môže byť aj rozptyl poruchového člena.

Heteroskedasticitu je možné identifikovať niekoľkými spôsobmi. V niektorých situáciách je možné nadobudnúť podozrenie o prítomnosti heteroskedasticity aj vizuálnym posúdením rezíduí. Za týmto účelom je vhodné zobrazíť si priebeh rezíduí, neraz ako  $x$ - $y$  graf, kde na horizontálnej osi zvolíme niektorú z vysvetľujúcich premenných a na zvislú os zobrazíme rezíduá.

Vizuálne posúdenie je však stále len subjektívnou metódou, takže má len orientačný charakter. Napriek tomu sa dá odporúčať, keďže vizualizácia dát spravidla značne napomáha pri rozhodovaní sa o spôsobe modelovania skúmanej premennej.

Druhou skupinou metód, ktoré slúžia na testovanie heteroskedasticity sú postupy založené na štatistických testoch, realizovaných spravidla na pomocných regresiach. Ak máme podozrenie, že niektorá premenná súvisí s meniacim sa rozptylom poruchových členov, môže byť logickou voľbou pokúsiť sa tento vzťah modelovať. Ak by sme vedeli preukázať štatisticky významný vzťah medzi rozptylom rezíduí a niektorou z premenných, rozptyl zrejme nebude konštantný.

Zostáva otázka, čo by v tomto prípade predstavovalo závislú premennú. Rozptyl náhodných porúch  $\sigma_i^2$  pre  $i = 1, 2, \dots, n$  je pre nás neznámy. Väčšinou sa preto za závislú premennú vyberajú štvorce rezíduí,  $e_i^2$ .

Tabuľka 7: Prístupy k testovaniu heteroskedasticity

Metóda	Test
Park	$\ln e_i^2 = \beta_0 + \beta_1 \ln x_{ih} + v_i$
Glejser	$ e_i  = \beta_1 x_{ih} + v_i$
	$ e_i  = \beta_1 (\sqrt{x_{ih}})^{-1} + v_i$
	$ e_i  = \beta_1 (\sqrt{x_{ih}})^1 + v_i$
	$ e_i  = \beta_0 + \beta_1 \sqrt{x_{ih}} + v_i$
	$ e_i  = \sqrt{\beta_0 + \beta_1 x_{ih}} + v_i$
	$ e_i  = \sqrt{\beta_0 + \beta_1 x_{ih}^2} + v_i$
Goldfeld–Quandt	$F$ -test, $\mathbf{e}^T \mathbf{e}$ pre dva modely
White	$e_i^2 = \gamma_0 + \sum_{j=1}^k \gamma_j x_{ij} + \sum_{j=1}^k \gamma_{k+i} x_{ij}^2$ $+ \sum_{j=1}^{k-1} \sum_{l=j+1}^k \gamma_{2k+(2k-j)(j-1)/2+l-j} x_j x_k + v_i$

Zdroj: vlastné spracovanie

Tabuľka 7 prezentuje rôzne prístupy k testovaniu heteroskedasticity na základe pomocných regresíí. V prípade testov, ktoré navrhli Park a Glejser je potrebné najprv identifikovať premennú, o ktorej si myslíme, že súvisí s heteroskedasticitou zložiek  $\mathbf{u}$ . V tabuľke predpokladáme, že zvolená premenná má index  $h \in \{1, 2, \dots, n\}$ . Vo väčšine prípadov (okrem Goldfeld-Quandtovho a Whiteovho testu) sa testuje významnosť  $\beta_1$ .

Goldfeld-Quandtov test porovnáva pomocou  $F$ -testu súčet štvorcov rezíduí pre dva modely totožné s pôvodným modelom, vypočítané na základe reštrikcie vzoriek – pozorovania sa zoradia podľa hodnôt premennej, o ktorej si myslíme, že spôsobuje heteroskedasticitu, a následne kvantifikujeme dva modely, jeden na prvej pätine (alebo tretine) vzorky, a druhý model na základe poslednej pätiny (tretiny) vzorky. Pozorovania v strede vynechávame. Zmyslom testu je porovnať, či rezíduá rovnako špecifikovaného modelu majú na podvzorkách rôzny rozptyl – ak áno, indikuje to heteroskedasticitu.

Test ktorý navrhol White vychádza z predpokladu, že nevieme vopred identifikovať premennú, ktorá by mohla súvisieť s heteroskedasticitou. Pomocná regresia preto zahŕňa všetky pôvodné regresory, ich druhé mocniny, ale aj súčiny regresorov. Je zrejmé, že tento test je možné použiť len na väčších vzorkách – v teste vystupuje veľmi veľa regresorov.

### Príklad 3.10

Overte, či regresný model závislosti cien bytov a ich rozlohy (Príklad 3.7) nie je zaťažený problémom heteroskedasticity rezíduí.

Aj v tomto príklade využijeme knižnicu `lmtest`, v ktorej pomocou príkazu `bptest` vieme použiť na testovanie heteroskedasticity klasický Breusch – Paganov test. V rámci neho sa testuje združená nulová hypotéza o konštantných rozptyloch rezíduí (homoskedasticita).

```
> Reg.Model = lm(Dataset$Cena ~ Dataset$Rozloha)
> library(lmtest)
> bptest(Reg.Model)

studentized Breusch-Pagan test

data:  Reg.Model
BP = 0.0165, df = 1, p-value = 0.8979
```

V danom regresnom modeli sa heteroskedasticita nepreukázala, keďže  $p$ -hodnota je väčšia ako hladina významnosti 5 %. Opäť však treba pripomenúť, že ide o regresný model kvantifikovaný na nízkom počte pozorovaní, preto aj Breusch – Paganov test môže mať nízku silu.

V prípade, ak pomocou testov zistíme, že sú poruchové členy heteroskedastické, máme niekoľko možností, ako s problémom naložiť. Ak máme informáciu, ktorá premenná súvisí s heteroskedasticitou  $\mathbf{u}$ , môžeme využiť metódu vážených najmenších štvorcov, resp. zovšeobecnenú metódu najmenších štvorcov.

Iným prístupom je použitie tzv. estimátorov variančno-kovariančnej matice konzistentných aj v prítomnosti heteroskedasticity (angl. heteroskedasticity-consistent covariance matrix estimator). Podstatou týchto estimátorov je realizácia konzistentného odhadu variančno-kovariančnej matice odhadov regresných koeficientov. Pripomeňme, že táto matica má tvar:

$$\text{var}(\hat{\boldsymbol{\beta}}) = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \boldsymbol{\Omega} \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \quad (3.342)$$

White (1980) sa zaoberal odhadom podobných matíc, aj keď v trochu modifikovanom tvare. Pracoval s asymptotickou variančno-kovariančnou maticou Davidson – MacKinnon (2003):

$$\lim_{n \rightarrow \infty} \left( \frac{1}{n} \mathbf{X}^T \mathbf{X} \right)^{-1} \lim_{n \rightarrow \infty} \left( \frac{1}{n} \mathbf{X}^T \mathbf{\Omega} \mathbf{X} \right)^{-1} \lim_{n \rightarrow \infty} \left( \frac{1}{n} \mathbf{X}^T \mathbf{X} \right)^{-1} \quad (3.343)$$

Kým odhadnúť prvú a poslednú limitu nebol problém, White dokázal, že je možné konzistentne odhadnúť prostrednú limitu s využitím nekonzistentného odhadu matice  $\mathbf{\Omega}$ . Teda dokázal, aj nekonzistentný odhad  $\mathbf{\Omega}$  môže viesť ku konzistentnému odhadu variančno-kovariančnej matice odhadov regresných koeficientov. Tento odhad sa realizuje pomocou výrazu:

$$(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \hat{\mathbf{\Omega}} \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \quad (3.344)$$

kde  $\hat{\mathbf{\Omega}}$  je diagonálna matica. Existuje niekoľko prístupov, ako túto maticu vytvoriť, najbežnejšie sa označujú  $HC_0$  až  $HC_3$ .

Estimátor  $HC_0$  navrhol vo svojom článku sám White, ide o najjednoduchší prípad, keď diagonálne prvky  $\hat{\mathbf{\Omega}}$  vypočítame podľa vzorca:

$$\hat{\omega}_{ii} = e_i^2 \quad (3.345)$$

Na diagonále sú teda štvorce rezíduí.

Estimátor  $HC_1$  uskutočňuje korekciu vychádzajúcu z počtu pozorovaní a odhadovaných parametrov:

$$\hat{\omega}_{ii} = e_i^2 \frac{n}{n-k} \quad (3.346)$$

Pre estimátor  $HC_2$  sa definuje premenná  $h_i = \mathbf{X}_i (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}_i^T$ , kde  $\mathbf{X}_i$  je  $i$ -ty riadok matice  $\mathbf{X}$ . Pre estimátor  $HC_2$  potom kladieme:

$$\hat{\omega}_{ii} = \frac{e_i^2}{1-h_i} \quad (3.347)$$

Ďalšou možnosťou je estimátor  $HC_3$ , pri ktorom:

$$\hat{\omega}_{ii} = \frac{e_i^2}{(1-h_i)^2} \quad (3.348)$$

Poslednou možnosťou je estimátor  $HC_4$ , pri ktorom:

$$\hat{\omega}_{ii} = \frac{e_i^2}{(1-h_i)^{\delta_i}} \quad (3.349)$$

kde

$$\delta_i = \min \left\{ 4, \frac{nh_i}{k+1} \right\}, \quad i = 1, 2, \dots, n \quad (3.350)$$

Výber estimátora nie je jednoznačný a názory sa rozchádzajú aj v publikovaných článkoch. Panuje pomerne značná zhoda, že najhorším je zrejme odhad pomocou  $HC_0$ . Existujú štúdie, ktoré uprednostňujú odhad pomocou  $HC_3$ . Odhady pomocou  $HC_2$  a  $HC_3$  sú odporúčané aj vtedy, ak sa vo vzorke nachádzajú odľahlé hodnoty v rámci vysvetľujúcich premenných.

Využitie  $HC$  odhadov pri regresnom modeli má svoje výhody. V prvom rade na rozdiel od iných metód nie je potrebné transformovať vysvetľujúce a vysvetľovanú premennú. Transformácie často menia interpretáciu koeficientov v regresnom modeli – videli sme, že pri interpretácii regresného modelu zodpovedá regresný koeficient marginálnej zmene závislej premennej, kým napr. pri transformovanom log-log modeli už ide o elasticitu. Ak vytvárame ekonometrický model v určitom tvare z deduktívnych dôvodov (ako dôsledok testovania konkrétnej teórie), transformovaný model nemusí byť vyhovujúci.

Ďalšou nevýhodou transformovaných modelov je, že aj v prípade, ak sa vieme spätnou transformáciou dostať k odhadom pôvodných regresných koeficientov v originálnom modeli, strácame informáciu o testoch na štatistickú významnosť a koeficiente determinácie. Hlavne v prípade nelineárnych transformácií (akou je aj spomínaná frekventovaná logaritmická transformácia) sa už potom k vlastnostiam v pôvodnom, netransformovanom modeli nevieme vyjadriť.

Aj z tohto dôvodu je využitie estimátorov variančno-kovariančnej matice konzistentných aj v prítomnosti heteroskedasticity veľmi atraktívne – nedochádza k žiadnej transformácii modelu, interpretácia a štatistické testy zostávajú v pôvodnom tvare. V skutočnosti sú pri použití týchto estimátorov, ako aj a bez nich (pri bežnom odhade variančno-kovariančnej matice odhadov regresných koeficientov) odhadnuté regresné koeficienty, a preto aj rezíduá rovnaké. Musia byť, pretože estimátory  $HC$  odhadujú len spomínanú variančno-kovariančnú maticu, a nie vektor regresných koeficientov.

Bežnou chybou býva výpočet modelu s použitím  $HC$  odhadov, a následné testovanie heteroskedasticity pomocou niektorého z testov. Je pomerne bežné, že v takom prípade test identifikuje heteroskedasticitu – čo je aj logické, pretože testy na heteroskedasticitu vychádzajú z rezíduí, ktoré sa použitím  $HC$  estimátora nemenia. Vo výstupe štatistického software sa preto treba pri ich použití riadiť len výsledkami odhadu regresných koeficientov,

variančno-kovariančnej matici (štandardným chybám z nej vypočítanej) a signifikanciou. Výpočet testov na heteroskedasticitu je po použití *HC* estimátora už potom zbytočné.

Treba si však uvedomiť, že výsledky pre tieto estimátory sú len asymptotické, v malých vzorkách teda môžu viesť aj k nepresným výsledkom. Pre nami prezentované *HC* estimátory sa neodporúča používať ich na vzorkách, pri ktorých je menej ako 100 pozorovaní.

Okrem estimátorov *HC*, ktoré riešia problém heteroskedasticity existujú aj estimátory variančno-kovariančnej matice *HAC*, ktoré sú okrem heteroskedasticity konzistentné aj v prípade, ak je prítomná autokorelácia. Ich použitie a obmedzenia sú prakticky rovnaké, avšak ich výpočet je náročnejší – v tejto publikácii sa nimi bližšie zaoberať nebudeme.

### **3.5.4 Autokorelácia**

Kým heteroskedasticita sa týkala rozptylov, a teda hlavnej diagonály variančno-kovariančnej matice poruchových členov, autokorelácia sa týka jej mimodiagonálnych prvkov. Podľa predpokladu by mali byť nulové, a poruchové členy by mali byť nekorelované – ak tomu tak nie je, dôsledky sú značné, ako sme uvádzali v časti 3.5.1.

Autokoreláciou sa má zmysel zaoberať hlavne v prípade časových radov. Pri nich je možné očakávať, že susedné pozorovania môžu mať korelované náhodné členy. Môže sa to stať napríklad z dôvodu, že máme zle špecifikovaný model – napríklad sme mohli vynechať relevantnú premennú, efekt ktorej sa potom stane súčasťou poruchového člena, prípadne sme zvolili zľú funkčnú závislosť modelu. Autokorelácia má menší význam v prípade priereзовých dát – na týchto údajoch, ktoré sú získané za rôzne štatistické jednotky v tom istom čase nie je spravidla jasne definované poradie. Ak by sme napríklad v databáze, s ktorou pracujeme vymenili jedno pozorovanie za iné, nič sa nemení – hodnoty zostávajú tie isté, bez ohľadu na poradie, v akom ich uvádzame.<sup>5</sup> Pri časových radoch ale poradie definované je – tu má „susednosť“ pozorovaní zmysel. Niektoré ekonometrické programy (ako napríklad Gretl) dokonca ani neumožňujú výpočet testov na autokoreláciu v prípade, ak je súbor dát definovaný ako priereзовý.

---

<sup>5</sup> Pri niektorých zaujímavých aplikáciách je možné hovoriť o takzvanej priestorovej korelácií. Uvažujme o jednoduchom prípade, kde našou vysvetľovanou premennou sú (štatistickou jednotkou) sú okresy a štatistickým znakom nezamestnanosť. Po odhade tohto jednoduchého modelu by sme zrejme získali určitú koreláciu medzi rezíduami vzájomne susediacich okresov (prípadne by rezíduá záviseli od vzdialenosti medzi okresmi).

Dôsledky autokorelácie si demonštrujeme na príklade najjednoduchšieho možného spôsobu, ako môže byť v tomto smere predpoklad P2 porušený – pomocou tzv. autoregresného procesu prvého rádu.

V tomto prípade majú byť poruchové členy korelované – jednoduchý spôsob, ako to dosiahnuť, je položiť:

$$u_t = \rho u_{t-1} + \varepsilon_t \quad (3.351)$$

V predchádzajúcom výraze  $t = 1, 2, \dots, n$  – doposiaľ sme pozorovania označovali indexom  $i$ , avšak pri práci s časovými radmi sa väčšinou používa index  $t$ . Z tohto dôvodu zachováme obvyklú konvenciu.

Vidíme, že poruchový člen  $u_t$  je funkciou predchádzajúceho,  $u_{t-1}$ . Pre konštantu  $\rho$  na teraz platí obmedzenie,  $-1 < \rho < 1$ . Definujeme aj nové náhodné premenné  $\varepsilon_t$  (budeme ich nazývať inovácie), o ktorých predpokladáme že sú navzájom nezávislé, majú rovnaké rozdelenie pravdepodobnosti s nulovou strednou hodnotou a konštantným rozptylom  $\sigma_\varepsilon^2$ . Taktiež platí, že  $\varepsilon_t$  sú nezávislé na  $u_{t-1}$ .

Rekurentná definícia  $u_t$  znamená, že platí:

$$u_t = \rho u_{t-1} + \varepsilon_t \quad (3.352)$$

$$= \rho(\rho u_{t-2} + \varepsilon_{t-1}) + \varepsilon_t = \rho^2 u_{t-2} + \rho \varepsilon_{t-1} + \varepsilon_t \quad (3.353)$$

$$= \rho^2(\rho u_{t-3} + \varepsilon_{t-2}) + \rho \varepsilon_{t-1} + \varepsilon_t = \rho^3 u_{t-3} + \rho^2 \varepsilon_{t-2} + \rho \varepsilon_{t-1} + \varepsilon_t \quad (3.354)$$

$$= \rho^p u_{t-p} + \sum_{i=0}^{p-1} \rho^i \varepsilon_{t-i}, \quad p \in \{1, 2, \dots, t-1\} \quad (3.355)$$

Náhodný člen  $u_t$  teda môžeme vždy zapísať pomocou oneskorených hodnôt  $u_{t-p}$  a náhodných inovácií.

Od náhodného procesu, akým je aj  $u_t$  niekedy požadujeme, aby bol stacionárny. Stacionarita sa dá definovať rôzne, ale v našom prípade použijeme definíciu tzv. slabej, alebo kovariančnej stacionarity. Aby sme tak proces  $u_t$  mohli nazývať slabo stacionárnym, musí spĺňať nasledovné podmienky:

- nepodmienená stredná hodnota  $E(u_t)$  a rozptyl  $\sigma_{u_t}^2$  existujú a sú v čase konštantné,
- kovariancie  $\text{cov}(u_t, u_{t-p})$  sú nezávislé od  $t$  a závisia len od  $p$  (teda kovariancia medzi dvomi poruchovými členmi závisí len na ich vzájomnom oneskorení a nie na čase, v ktorom sa realizovali).

Podmienky slabej stacionarity vlastne hovoria o stabilite, nemennosti niektorých štatistických vlastností  $u_t$  v čase. Skúsme sa preto pozrieť, či je proces  $u_t$  tak, ako sme ho navrhli v rovnici (3.351) stacionárny.

Podľa (3.355) môžeme pri oneskorení o  $p$  hodnôt napísať člen  $u_t$  ako:

$$u_t = \rho^p u_{t-p} + \sum_{i=0}^{p-1} \rho^i \varepsilon_{t-i} \quad (3.356)$$

Často sa v tomto bode navrhuje teoretická abstrakcia: čo by sa stalo, ak by sme počítali limitu pre  $p$  idúce do nekonečna? Samozrejme, tým by sme opustili našu vzorku, ktorá nie je nekonečná. Je však zaujímavé vidieť, ako by sa poruchový člen správal asymptoticky. Približne si to môžeme predstaviť tak, ako keby sme mali veľmi veľa pozorovaní. Pre limitu by platilo:

$$u_t = \sum_{i=0}^{\infty} \rho^i \varepsilon_{t-i} \quad (3.357)$$

Aká by bola v tomto limitnom prípade stredná hodnota  $u_t$ ?

$$E(u_t) = E\left(\sum_{i=0}^{\infty} \rho^i \varepsilon_{t-i}\right) = \sum_{i=0}^{\infty} \rho^i E(\varepsilon_{t-i}) = 0 \quad (3.358)$$

Keďže  $\varepsilon_t$  má pre všetky  $t$  nulovú strednú hodnotu, je aj stredná hodnota  $u_t$  rovná nule.

Vypočítajme teraz rozptyl  $u_t$ . Tu môžeme využiť, že  $D(\varepsilon_t) = \sigma_\varepsilon^2$ , ako aj to, že  $\varepsilon_t$  a  $\varepsilon_{t-p}$  sú pre ľubovoľné prirodzené  $p$  nezávislé, teda aj nekorelované. Potom:

$$\sigma_{u_t}^2 = D(u_t) = D\left(\sum_{i=0}^{\infty} \rho^i \varepsilon_{t-i}\right) = \sum_{i=0}^{\infty} \rho^{2i} D(\varepsilon_{t-i}) = \sigma_\varepsilon^2 \sum_{i=0}^{\infty} \rho^{2i} \quad (3.359)$$

V prípade, ak platí  $-1 < \rho < 1$ , má nekonečný rad v poslednom výraze konečný súčet:

$$D(u_t) = \sigma_\varepsilon^2 \sum_{i=0}^{\infty} \rho^{2i} = \frac{\sigma_\varepsilon^2}{1-\rho^2} \quad (3.360)$$

Vidíme, že tak stredná hodnota, ako aj rozptyl  $u_t$  nezávisí od indexu  $t$  – v čase sa teda nemenia. Zároveň je zřejmé, že podmienka  $-1 < \rho < 1$  je kľúčová – ak by tomu tak nebolo, nekonečný rad pre rozptyl by nebol konvergentný. Podmienka  $-1 < \rho < 1$  sa preto nazýva aj podmienkou stacionarity.

Zatiaľ sme pracovali s  $u_t$  v tvare nekonečného radu – čo je tvar, ktorý je ťažké si predstaviť, keďže pracujeme s konečnými vzorkami. Je možné aj iné vyjadrenie – položme pre  $u_1$  strednú hodnotu  $E(u_1) = 0$  a pre rozptyl  $D(u_1) = \sigma_\varepsilon^2/(1-\rho^2)$ .

Potom pre  $u_2$  dostávame v poslednom výraze konečný súčet:

$$E(u_2) = E(\rho u_1 + \varepsilon_2) = \rho E(u_1) + E(\varepsilon_2) = 0 + 0 = 0 \quad (3.361)$$

Pre rozptyl dostaneme:

$$D(u_2) = D(\rho u_1 + \varepsilon_2) \quad (3.362)$$

$$= \rho^2 D(u_1) + D(\varepsilon_2) \quad (3.363)$$

$$= \frac{\rho^2 \sigma_\varepsilon^2}{1 - \rho^2} + \sigma_\varepsilon^2 \quad (3.364)$$

$$= \frac{\rho^2 \sigma_\varepsilon^2 + \sigma_\varepsilon^2 - \rho^2 \sigma_\varepsilon^2}{1 - \rho^2} \quad (3.365)$$

$$= \frac{\sigma_\varepsilon^2}{1 - \rho^2} \quad (3.366)$$

Indukciou by sme mohli pokračovať, ale zistili by sme, že pre všetky  $t$  je  $E(u_t) = 0$  a  $D(u_t) = \sigma_\varepsilon^2 / (1 - \rho^2)$ .

Z hľadiska stacionarity je ešte potrebné overiť kovariancie  $\text{cov}(u_t, u_{t-p})$  pre  $p \in \{1, 2, \dots, t-1\}$ . Platí:

$$\text{cov}(u_t, u_{t-p}) = \text{cov}\left(\rho^p u_{t-p} + \sum_{i=0}^{p-1} \rho^i \varepsilon_{t-i}, u_{t-p}\right) \quad (3.367)$$

$$= \text{cov}(\rho^p u_{t-p}, u_{t-p}) + \sum_{i=0}^{p-1} \text{cov}(\rho^i \varepsilon_{t-i}, u_{t-p}) \quad (3.368)$$

$$= \rho^p \text{cov}(u_{t-p}, u_{t-p}) + \sum_{i=0}^{p-1} \rho^i \text{cov}(\varepsilon_{t-i}, u_{t-p}) \quad (3.369)$$

$$= \rho^p D(u_{t-p}) \quad (3.370)$$

$$= \rho^p \frac{\sigma_\varepsilon^2}{1 - \rho^2} \quad (3.371)$$

Vidíme, že kovariancia nezávisí na časovom indexe  $t$ , ale závisí len na oneskorení  $p$ . Proces  $u_t$  tak, ako sme ho definovali, je slabo stacionárny. Tento proces voláme aj autoregresívny proces prvého rádu (poruchový člen je priamo funkciou svojej jednej predchádzajúcej hodnoty  $u_{t-1}$ ).

Variančno-kovariančnú maticu  $\text{var}(\mathbf{u})$  môžeme zapísať nasledovne:

$$\text{var}(\mathbf{u}) = E(\mathbf{u}\mathbf{u}^T) = \mathbf{\Omega} = \frac{\sigma_\varepsilon^2}{1-\rho^2} \begin{pmatrix} 1 & \rho & \rho^2 & \dots & \rho^{n-1} \\ \rho & 1 & \rho & \dots & \rho^{n-2} \\ \rho^2 & \rho & 1 & \dots & \rho^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \dots & 1 \end{pmatrix} \quad (3.372)$$

Z uvedeného výrazu je vidieť, že aj ten najjednoduchší autoregresívny proces generuje pomerne komplikovanú variančno-kovariančnú maticu. Kým pri heteroskedasticite došlo len k zmenám na hlavnej diagonále, pri autokorelácii sa mení celá matica. Taktiež si môžeme všimnúť, že najväčšie hodnoty v matici sa sústreďujú okolo hlavnej diagonály – keďže  $|\rho| < 1$ , vyššie mocniny  $\rho$  predstavujú v absolútnej hodnote menšie čísla. Je tomu tak preto, že čím je oneskorenie ( $p$ ) väčšie, tým menšia bude (za predpokladu homoskedasticity) hodnota kovariancie.

Predchádzajúci vzťah ukazuje, že autokorelácia má značné dôsledky z hľadiska splnenia predpokladu P2. Ako je možné prítomnosť autokorelácie identifikovať?

Historicky asi najznámejším testom autokorelácie je tzv. Durbin-Watsonov test autokorelácie. Testuje najjednoduchšiu formu autokorelácie, proces AR(1), o ktorom sme hovorili v predchádzajúcej časti. Tento test je založený na testovacej štatistike:

$$d = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\mathbf{e}^T \mathbf{e}} \quad (3.373)$$

Štatistika  $d$  nadobúda hodnoty od 0 do 4. Hodnota  $d = 2$  indikuje neprítomnosť autokorelácie. Hodnoty menšie ako 2 hovoria o pozitívnej autokorelácii, hodnoty väčšie ako 2 o zápornej autokorelácii. Jednou z nevýhod Durbin-Watsonovej testovacej štatistiky je to, že nemá jednoduché známe pravdepodobnostné rozdelenie. Existujú preňho rôzne aproximácie, avšak vzhľadom na to, že ide len o aproximácie, ako aj fakt, že je možné testovať len najjednoduchšiu formu autokorelácie spravidla odporúčame realizovať iný test na autokoreláciu. Durbin-Watsonova štatistika je napriek tomu veľmi známou a väčšina štatistických programov ju pri modelovaní uvádza – môže preto slúžiť ako rýchla informácia o tom, či môžeme očakávať porušenie predpokladu P2 z dôvodu prítomnosti autokorelácie. Hodnoty menšie ako 1, resp. väčšie ako 3 nás nabádajú k bližšiemu preskúmaniu rezíduí.

### Príklad 3.11

Overte, či regresný model závislosti spotreby od dôchodku v SR (Príklad 3.9) nie je zaťažený problémom autokorelácie rezíduí.

Danú úlohu môžeme overiť pomocou testu Durbin – Watson. Nainštalovať potrebujeme knižnicu `lmtest` a následne pomocou príkazu `dwtest` vieme overiť nulovú hypotézu  $H_0: \rho = 0$  (autokorelácia sa nevyskytuje).

```
> Reg.Model = lm(CONS~HDP, data = Dataset)
> library(lmtest)
> dwtest(Reg.Model)

Durbin-Watson test

data:  Reg.Model
DW = 1.6293, p-value = 0.1419
alternative hypothesis: true autocorrelation is greater than 0
-----
> dwtest (Reg.Model, alternative = "two.sided")

Durbin-Watson test

data:  Reg.Model
DW = 1.6293, p-value = 0.2839
alternative hypothesis: true autocorelation is not 0
```

Tento test má pôvodne nastavenú alternatívnu hypotézu, že autokorelácia je väčšia ako 0. Alternatívnu hypotézu si však môžeme zvoliť, teda či ide o kladnú autokoreláciu ( $H_1: \rho > 0$ ; `alternative = „greater“`) alebo zápornú ( $H_1: \rho < 0$  `alternative = „less“`). Taktiež si môžeme zvoliť obojstranný test (`alternative = "two.sided"`), teda voči nulovej hypotéze postaviť alternatívnu  $H_1: \rho \neq 0$ .

V našom príklade sa nulovú hypotézu nepodarilo zamietnuť, teda môžeme považovať výskyt autokorelácie za nepreukázaný. Implikácia takeého zistenia spočíva v skutočnosti, že získaným výsledkom z daného regresného modelu môžeme veriť. Treba však poukázať aj na nízky počet pozorovaní v danom modeli, čo sa môže prejaviť aj na sile testu.

Pomerne často využívaným testom autokorelácie je Ljung-Boxov  $Q$  test. Tento test je založený na autokorelácií rezíduí, presnejšie na takzvanej autokorelačnej funkcii. Pre rezíduá by sme mohli definovať odhad rozptylu:

$$s_e^2 = \frac{1}{n-1} \sum_{i=1}^n (e_i - \bar{e})^2 = \frac{1}{n-1} \sum_{i=1}^n e_i^2 = \frac{1}{n-1} \mathbf{e}^T \mathbf{e} \quad (3.374)$$

V predchádzajúcom vzťahu sme využili vlastnosť, ktorú sme dokázali v Kapitole 3.4.1, ktorá hovorí, že:

$$\bar{e} = \frac{1}{n} \mathbf{1}^T \mathbf{e} = \frac{1}{n} \mathbf{0} = 0 \quad (3.375)$$

Ďalej pre  $p \in \{1, 2, \dots, n-1\}$  môžeme definovať výberovú autokovarianciu:

$$\text{cov}(e_i, e_{i-p}) = \frac{1}{n-1} \sum_{i=p+1}^n (e_i - \bar{e})(e_{i-p} - \bar{e}) = \frac{1}{n-1} \sum_{i=1}^n e_i e_{i-p} \quad (3.376)$$

S pomocou výrazov pre výberový rozptyl rezíduí a autokovarianciu je možné definovať autokorelačnú funkciu rezíduí  $\rho(p)$ :

$$\rho(p) = \frac{\text{cov}(e_i, e_{i-p})}{s_e^2} \quad (3.377)$$

Ljung-Boxov test je potom založený na štatistike (Greene, 2012):

$$Q(p) = T(T+2) \sum_{j=1}^p \frac{\rho(j)}{n-j} \quad (3.378)$$

Posledným testom, ktorý spomenieme, navrhli Breusch a Godfrey. Tento test, ako aj niektoré ďalšie pokročilé spôsoby testovania porušenia predpokladu P2 si ukážeme v nasledujúcej kapitole.

### 3.5.5 Multikolarita

O multikolarite hovoríme vo vzťahu k predpokladu P4, ktorý hovorí o hodnosti matice  $\mathbf{X}$ . Podľa P4 má platiť:

$$h(\mathbf{X}) = k+1 \quad \wedge \quad k+1 \leq n \quad (3.379)$$

Predpoklad P4 sa zavádza z toho dôvodu, aby bola matica  $\mathbf{X}^T \mathbf{X}$  regulárna, a bolo možné odhadnúť  $\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}$ .

V kapitole 3.3.3 o predpokladoch lineárneho modelu sme si vysvetlili, že o porušení tohto predpokladu sa s určitosťou dozvieme, a teda že tento problém je jasne identifikovateľný a riešiteľný. Identifikovateľný je preto, lebo jeho nastanie každý štatistický program ohlásí, keďže je nemožné vypočítať odhad regresných koeficientov. Problém riešiteľný, pretože vzniká ako dôsledok zaradenia nesprávnej kombinácie premenných do modelu – konkrétne ide o lineárne závislé premenné. V takom prípade nie je možné jednoznačne určiť koeficienty lineárnej kombinácie regresorov. Riešenie spravidla spočíva v úprave vysvetľujúcich premenných – spravidla sa vyberie nejaká vhodná vlastná podmnožina vysvetľujúcich premenných z pôvodného modelu.

Problém s multikolinearitou nastáva pomerne často pri použití indikátorových premenných (angl. *dummy variables*), o ktorých sme hovorili v kapitole 3.2.3. Vráťme sa k modelu, v ktorom modelujeme ako závislú premennú mesačný príjem zamestnanca, premenná  $x_1$  zodpovedá počtu rokov jeho praxe a premenná  $x_2$  je binárna premenná, kde položíme  $x_{i2} = 0$  v prípade, ak respondent  $i$  je muž a  $x_{i2} = 1$  v prípade, ak je respondentom  $i$  žena. Premenná  $x_3$  nech je binárna premenná, kde položíme  $x_{i3} = 0$  v prípade, ak respondent  $i$  je žena a  $x_{i3} = 1$  v prípade, ak je respondentom  $i$  muž.

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + u_i \quad (3.380)$$

pre  $i = 1, 2, \dots, n$  a  $n \in \mathbb{N}$ . Dôvodom, prečo by sme mohli chcieť špecifikovať model týmto spôsobom by mohla byť snaha získať samostatný regresný koeficient pre každé pohlavie samostatne –  $\beta_2$  pre ženy a  $\beta_3$  pre mužov.

Takto špecifikovaný model nie je možné kvantifikovať, nakoľko dochádza k exaktnej multikolinearite – stĺpce matice  $\mathbf{X}$  sú určite lineárne závislé a  $h(\mathbf{X}) < k + 1$ , čo je v rozpore s predpokladom P4 (v našom prípade  $k = 3$ ). Túto skutočnosť môžeme ľahko ukázať, ak si uvedomíme, že matica  $\mathbf{X}$  má tvar:

$$\mathbf{X} = \begin{pmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{pmatrix} \quad (3.381)$$

Prvý stĺpec matice  $\mathbf{X}$  obsahuje vektor jednotiek. Pre vyššie definované premenné zrejme platí:

$$x_{i2} + x_{i3} = 1 \quad (3.382)$$

Odtiaľ je zřejmé, že prvý stĺpec matice je lineárnou kombináciou tretieho a štvrtého stĺpca. Matica  $\mathbf{X}$  je potom singularná a nie je možné vypočítať odhad regresných koeficientov. Ak by sme premennú  $x_3$  vynechali, model by bolo možné odhadnúť, a mal by interpretáciu, o ktorej sme hovorili v podkapitole 3.2.3.

Porušenie predpokladu P4, ktoré sme si opísali, je veľmi jednoduché identifikovať a preto je aj jednoduché sa mu vyvarovať. O dosť menej zreteľným je prípad multikolinearity, pri ktorom nehovoríme o exaktnej multikolinearite – stĺpce matice  $\mathbf{X}$  nemusia byť lineárne závislé, stačí, ak sú vysoko korelované.

Ak by boli stĺpce matice  $\mathbf{X}$ , a teda vysvetľujúce premenné, ktoré do modelu vkladáme vysoko korelované, k porušeniu predpokladu P4 nedochádza – hodnosť matice  $\mathbf{X}$  môže mať hodnotu  $k + 1$  a je možné vypočítať odhad regresných koeficientov.

V prípade vysoko korelovaných vysvetľujúcich premenných sa problém prejavuje inak. Dá sa ukázať (Greene, 2012), že:

$$D(\beta_i) = \frac{\sigma^2}{\sum_{j=1}^n (x_{ji} - \bar{x}_i)^2} \frac{1}{1 - R_i^2} \quad (3.383)$$

V predchádzajúcom vzťahu  $i = 1, 2, \dots, n$ ,  $\bar{x}_i = \frac{1}{n} \sum_{j=1}^n x_{ji}$ ,  $\sigma^2$  je rozptyl prvkov vektora  $\mathbf{u}$  a  $R_i^2$  je koeficient determinácie regresného modelu, pri ktorom je závislou premennou  $x_i$  a vysvetľujúcimi všetky ostatné premenné. Je zrejmé, že ak sú premenné tvoriace vector  $\mathbf{X}$  korelované, koeficienty determinácie  $R_i^2$  budú tiež vysoké. Podľa predchádzajúceho vzťahu však potom budú vysoké aj rozptyly regresných koeficientov  $D(\beta_i)$ .

Keďže aj v prípade o korelovaných vysvetľujúcich premenných hovoríme o multikolinearite, môžeme sa vyjadriť o dôsledkoch multikolinearity. Sú ním vysoké odhady rozptylov regresných koeficientov. Dôležitosť tejto skutočnosti sa prejaví pri štatistickom testovaní parametrov modelu – vysoké rozptyly znamenajú široké intervalové odhady regresných koeficientov. To má za dôsledok to, že je veľmi ťažké zamietnuť nulovú hypotézu pri  $t$ -testoch skúmajúcich významnosť regresných koeficientov. Typickým prípadom, keď je vhodné preveriť prítomnosť multikolinearity je situácia, v ktorej nie sú žiadne, resp. je len málo vysvetľujúcich premenných významných.

Štatistické programy často umožňujú výpočet charakteristiky, ktorá slúži na posúdenie vplyvu a existencie multikolinearity. Ide vlastne o posledný člen predchádzajúceho vzťahu, ktorý sa označuje ako  $VIF_i$ , a nazýva sa inflačný faktor rozptylu (angl. *variance inflation factor*). Zjednodušene povedané sa  $VIF_i$  snaží charakterizovať, aký má korelácia medzi vysvetľujúcimi premennými efekt na rozptyl regresného koeficientu  $\beta_i$ :

$$VIF_i = \frac{1}{1 - R_i^2} \quad (3.384)$$

Nasledujúca tabuľka zobrazuje vzťah medzi  $R_i^2$  a  $VIF_i$ .

Tabuľka 8: Inflačné faktory rozptylu  $VIF_i$

$R_i^2$	0.500	0.600	0.700	0.800	0.900	0.920	0.940	0.960	0.980	0.990	0.999
$VIF_i$	1.33	1.56	1.96	2.78	5.26	6.51	8.59	12.76	25.25	50.25	500.25

Zdroj: vlastné spracovanie

## 4 Simulačné metódy testovania predpokladov lineárneho modelu

V Kapitole 3.3.3 sme vymedzili predpoklady lineárneho regresného modelu. V Kapitole 3.4 a 3.5 sme špecifikovali, na čo má v našej ekonometrickej analýze vplyv porušenie týchto predpokladov. V tejto kapitole si prezentujeme niektoré vybrané štatistické testy, ktoré môžu slúžiť na identifikáciu porušenia niektorého zo spomínaných predpokladov. Na rozdiel od tradičných testov, sme sa aj vzhľadom na podporu programu R, rozhodli prezentovať moderné simulačné metódy testovania predpokladov lineárneho regresného modelu.

### 4.1 Testovanie normality rezíduí

Jedným zo štandardných predpokladov lineárneho regresného modelu je normalita rezíduí (P5 v Kapitole 3.3.3). Z aplikačného hľadiska nás teda zaujíma, či rezíduá môžu pochádzať z normálneho rozdelenia pravdepodobnosti. Pre tento účel je možné použiť ľubovoľný z tradičných testov, napr.: Kolmogorov – Smirnovov, Shapiro – Wilkov, Pearsonov  $\chi^2$ , Anderson – Darlingov, Jarque – Berov test normality. Opis niektorých z týchto testov môžeme nájsť v Lyócsa et al. (2013). Keďže tento predpoklad sa vyskytuje v indukčnej štatistike pomerne často, existuje mnoho prác, ktoré skúmajú vlastnosti týchto testov. Napríklad Yazici – Yolacan (2007) zistili, že chyba I. druhu<sup>6</sup> Anderson – Darlingovho testu je blízko k nominálne stanovenej hladine významnosti  $\alpha = 0.05$ , a to bez ohľadu na veľkosť vzorky. Pri menších vzorkách mal Jarque – Berov test (často používaný práve v ekonometrii) túto chybu na úrovni až  $\approx 0.192$  ( $n = 20$ ) a  $\approx 0.071$  ( $n = 30$ ), kým pri väčších vzorkách táto chyba výrazne klesala aj pod<sup>7</sup> nominálnych 0.05 (0.033,  $n = 40$ ; 0.017,  $n = 50$ ). Pri skúmaní chyby II. druhu<sup>8</sup> boli najpriaznivejšie výsledky namerané znovu pri Anderson – Darlingovom a Jarque – Berovom teste. Pozorovania sa generovali z rôznych rozdelení, pričom pri vzorkách  $n = 20, 30$  dosahoval Anderson – Darlingov test veľmi dobré výsledky,

---

<sup>6</sup> Pravdepodobnosť, že chybné zamietneme správnu nulovú hypotézu, v tomto prípade pravdepodobnosť, že zamietneme hypotézu o normalite rezíduí, ak v skutočnosti rezíduá pochádzajú z normálneho rozdelenia pravdepodobnosti.

<sup>7</sup> Čo môže mať negatívny dopad na silu testu.

<sup>8</sup> Pravdepodobnosť, že chybné nezamietneme nesprávnu nulovú hypotézu. V našom prípade pravdepodobnosť, že nezamietneme hypotézu o normalite rezíduí, ak v skutočnosti rezíduá nepochádzajú z normálneho rozdelenia pravdepodobnosti.

kým Jarque – Berov test dosahoval dobré výsledky pri vzorkách  $n = 40, 50$  (bližšie pozri Tabuľku 5, s. 181 – 182, Yazici – Yolacan, 2007).

Pomerne rozsiahlu štúdiu sily testov nájdeme ďalej v Romão et al. (2010). Táto štúdia je rozsiahla nie len do počtu porovnávaných testov (33 rôznych testov normality), ale aj typov nie normálnych rozdelení pravdepodobnosti (tri kategórie: symetrické, asymetrické, odvodené od normálneho), ktoré sa používali pri skúmaní sily testov. Z tejto štúdie sa javilo ako vhodných hneď niekoľko testov, avšak pre naše potreby len pripomenieme, že tak Anderson – Darlingov ako aj Jarque – Berov test dosahovali pomerne dobré výsledky, pričom medzi týmito dvoma testami, mal vyššiu silu vo väčšine prípadov Anderson – Darlingov test (všetky vzorky do  $n = 100$ ), avšak rozdiely v sile testov sa s rastúcou veľkosťou vzorky zmenšovali.

#### 4.1.1 Monte – Carlo Jarque – Berov test

Romão et al. (2010) ukázali, že v prípade Jarque – Berovho testu, nie je vhodné používať kritické hodnoty odvodené od asymptotického rozdelenia pravdepodobnosti testovacej štatistiky (a to aj pre väčšie vzorky)<sup>9</sup>. Keďže Jarque – Berov test patrí v ekonometrii k najpoužívanejším a zároveň sa vo väčšine prípadov stretávame so vzorkami väčšími ako 40 pozorovaní, budeme v ďalšej práci prezentovať tzv. Monte – Carlo Jarque – Berov test (pre iné testy, pozri bližšie Dufour et al., 1998). Použijeme pritom algoritmus prezentovaný v Godfrey (2009).

Uvažujme o lineárnom regresnom modeli:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \quad (4.1)$$

Vektor  $\mathbf{y}$  predstavuje závislú premennú, matica  $\mathbf{X}$  zahrňuje striktno exogénne regresory<sup>10</sup> (v stĺpcoch matice  $\mathbf{X}$ ). Chceme overiť, či rezíduá  $\mathbf{u}$  je možné považovať za realizácie z viacrozmerneho normálneho rozdelenia pravdepodobnosti. Predpokladáme pritom, že tieto rezíduá sú *iid* a sú homoskedastické. Zjavne je teda vhodné pred použitím tohto testu overiť predpoklad o homoskedasticite rezíduí<sup>11</sup>. V empirickými prácach sme len zriedkakedy svedkami homoskedastických rezíduí a ešte zriedkavejšie rezíduí, ktoré môžeme zároveň považovať za realizácie z normálneho rozdelenia pravdepodobnosti. Avšak kým

---

<sup>9</sup> Pojmy testovacia štatistika a charakteristika budeme voľne zamieňať tak, ako v Lyócsa et al. (2013).

<sup>10</sup> Pod striktnou exogenitou regresorov (vysvetľujúcich premenných) budeme rozumieť nasledujúcu vlastnosť: (náhodné) regresory sú nezávislé od rezíduí.

<sup>11</sup> Ak rezíduá nemôžeme považovať za homoskedastické, nemali by byť ani realizáciami z rovnakého rozdelenia pravdepodobnosti.

porušenie homoskedasticity môže mať výrazné dopady na nami prijaté závery z analýzy lineárnych regresných modelov, tak ako bolo spomenuté vyššie (pozri Kapitolu 3.5), porušenie predpokladu o normalite sa spravidla nepovažuje za závažný nedostatok regresného modelu.

K overovaniu tohto predpokladu použijeme odhady rezíduí (t. j. najprv odhadneme model (4.1)) a Jarque – Berovu testovaciu štatistiku vypočítame nasledovne ( $n$  je rozsah empirického štatistického súboru):

$$S = \frac{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^3}{\left( \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \right)^{3/2}} = \frac{\frac{1}{n} \sum_{i=1}^n (\hat{u}_i - \bar{\hat{u}})^3}{\left( \frac{1}{n} \sum_{i=1}^n (\hat{u}_i - \bar{\hat{u}})^2 \right)^{3/2}} = \frac{\frac{1}{n} \sum_{i=1}^n (\hat{u}_i)^3}{\left( \frac{1}{n} \sum_{i=1}^n (\hat{u}_i)^2 \right)^{3/2}} \quad (4.2)$$

$$K = \frac{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^4}{\left( \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \right)^2} = \frac{\frac{1}{n} \sum_{i=1}^n (\hat{u}_i - \bar{\hat{u}})^4}{\left( \frac{1}{n} \sum_{i=1}^n (\hat{u}_i - \bar{\hat{u}})^2 \right)^2} = \frac{\frac{1}{n} \sum_{i=1}^n (\hat{u}_i)^4}{\left( \frac{1}{n} \sum_{i=1}^n (\hat{u}_i)^2 \right)^2} \quad (4.3)$$

Testovacia charakteristika má potom tvar:

$$JB = \frac{n}{6} \left( S^2 + \frac{(K-3)^3}{4} \right) \quad (4.4)$$

Algoritmus Monte – Carlo Jarque – Berovho testu normality rezíduí je potom nasledovný:

Krok 1: Odhadnite regresný model (4.1) pomocou metódy najmenších štvorcov. Výstupom z tohto kroku sú odhadované rezíduá  $\hat{\mathbf{u}}$ .

Krok 2: Vypočítajte testovaciu štatistiku  $JB$  pomocou (4.2) - (4.4).

Krok 3: Pomocou generátora pseudonáhodných čísel, generujte  $n$  pozorovaní z normovaného normálneho rozdelenia pravdepodobnosti, ktoré si označíme ako  $\mathbf{y}_b^*$ , kde index  $b$  značí, že ide o  $b$ -tú iteráciu a hviezdička značí, že ide o pseudo-hodnoty,  $b = 1, \dots, B$ .

Krok 4: Odhadnite nasledujúci regresný model:

$$\mathbf{y}_b^* = \mathbf{X}\boldsymbol{\beta}^* + \mathbf{u}_b^* \quad (4.5)$$

Výstupom z tohto kroku sú odhadované rezíduá  $\hat{\mathbf{u}}_b^*$ .

Krok 5: Použite  $\hat{\mathbf{u}}_b^*$  a vypočítajte testovaciu štatistiku  $JB_b^*$  pomocou (4.2) - (4.4). Táto testovacia štatistika je vypočítaná za predpokladu platnosti nulovej hypotézy.

Krok 6: Krok 3 – 5 opakujte  $B$  krát, pričom  $B$  zvolíte tak, aby platilo, že  $(B + 1)\alpha_d$  je celé číslo, kde  $\alpha_d$  je želaná hladina významnosti (spravidla 0.05). Teda v prípade ak  $\alpha_d = 0.05$ , potom sú bežnou voľbou  $B = 999, 9999, \dots$ . Potom  $p$ -hodnotu, na základe ktorej môžeme rozhodnúť o nulovej hypotéze normality rezíduí, si môžeme vypočítať nasledovne:

$$JB_{PVAL} = \frac{\left| \left\{ JB_b^*; JB_b^* \geq JB \right\} \right| + 1}{(B + 1)} \quad (4.6)$$

Celý postup je v programe R možné pomerne ľahko automatizovať. Vytvorili sme funkciu, ktorá okrem vyššie spomínaného algoritmu ponúka aj dve ďalšie verzie predchádzajúceho algoritmu. Druhá verzia nevykonáva test z Kroku 5 na  $\hat{\mathbf{u}}_b^*$ , ale priamo na  $\mathbf{y}_b^*$ . Túto možnosť sme si vo funkcii označili ako  $T$ . Tretia verzia testu v Kroku 3 označuje generované hodnoty ako  $\mathbf{u}_b^*$  a ďalej platí,  $\mathbf{y}_b^* = \hat{\mathbf{y}} + \mathbf{u}_b^*$ , kde  $\hat{\mathbf{y}}$  sú predikované hodnoty z modelu (4.1). Ostatné kroky sú rovnaké ako v prvom algoritme. Túto možnosť sme si vo funkcii označili ako  $S$ .

```
JBPVAL <- function(DEP = NULL, X = NULL, B = 1000, algorithm =
  c("G", "T", "S")) {
  library(lawstat)
  n <- length(DEP)
  # Krok 1
  MOD_1 <- lm(DEP ~ X)
  u_hat <- residuals(MOD_1)
  # Krok 2
  JB <- rjb.test(u_hat, option = "JB")$statistic
  y_predict <- predict(MOD_1)
  JB_star <- c()
  for (i in 1:B) {
    # Krok 3
    temp <- rnorm(n)
    y_star <- switch(algorithm, G = temp, T = temp, S =
  y_predict + temp);
    # Krok 4
    MOD_2 <- lm(y_star ~ X)
    # Krok 5
    u_star_hat <- residuals(MOD_2)
    temp <- rjb.test(u_star_hat, option = "JB")$statistic
    JB_star[i] <- switch(algorithm, G = temp, T =
  rjb.test(y_star, option = "JB")$statistic, S = temp)
  }
  # Krok 6
```

```

pval <- round((sum(JB_star >= JB) + 1)/(B + 1), 8)
results <- list()
results[["JB-Statistics"]] <- JB
results[["MC-PValue"]] <- pval
return(results)
}

```

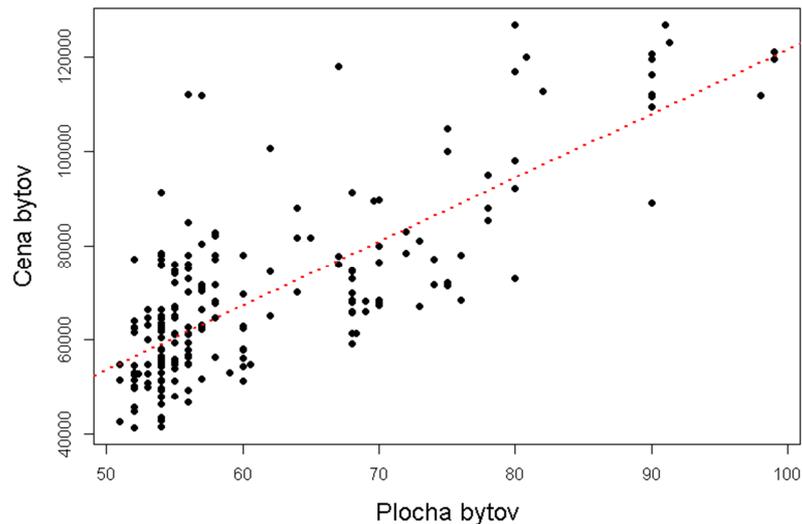
#### Príklad 4.1

Použijeme databázu `data_prices.csv` (pozri [www.econometrics.sk](http://www.econometrics.sk)). Ide o databázu cien bytov v krajských mestách na Slovensku z roku 2010. Databázu vytvorila D. Pančurová z EUBA PHF. Databáza okrem cien bytov obsahuje nasledujúce premenené: počet izieb, plocha bytu, výťah (1 – áno, 0 – nie), nadzemné podlažie, konštrukcia bytu (1 – panelový byt, 0 – tehlový byt), balkón/lodžia (1 – áno, 0 – nie) a premenné charakterizujúce lokalitu bytu. Údaje si najprv upravíme tak, aby sme získali iba byty s kompletnými údajmi, ktorých plocha je väčšia ako 50, ale menšia ako 100 metrov štvorcových, ktorých cena je viac ako 5000,- EUR, a ktoré sa nachádzajú v mestskej časti Košice – Sever. Zaujíma sa iba o jednoduchý model vzťahu medzi cenou bytu (závislá premenná) a podlahovou plochou bytu (nezávislá premenná). Nasledujúce kódy predstavujú zmeny v databáze a tvorbu  $x$ - $y$  obrázku.

```

> data <- read.csv(file = "D:\\... cesta ...\\data_prices.csv",
  header = T, dec = ".", sep = ";")
> data <- na.omit(data[,c(1, 3, 4)]); dim(data);
[1] 26407      3
> data <- subset(data, subset = data$Plocha > 50 & data$Plocha <
  100 & data$Cena > 5000 & data$KESE == 1)
> summary(data)
      Cena          Plocha          KESE
Min.   : 41400   Min.   :51.00   Min.   :1
1st Qu.: 55131   1st Qu.:54.00   1st Qu.:1
Median : 64350   Median :56.00   Median :1
Mean   : 68186   Mean   :60.61   Mean   :1
3rd Qu.: 75330   3rd Qu.:67.00   3rd Qu.:1
Max.   :126700   Max.   :99.00   Max.   :1
> plot(data$Plocha, data$Cena, type = "p", pch = 19, xlab =
  "Plocha bytov", ylab = "Cena bytov", cex.lab = 1.5)
> abline(lm(data$Cena ~ data$Plocha), col = "red", lty = 3, lwd
  = 2.5)

```



Obrázok 14: x-y graf vzťahu cien bytov a ich plochy v mestskej časti Košice - Sever

Zdroj: vlastné spracovanie

V pôvodnej databáze sa nachádzajú údaje, ktoré sú zrejme výsledkom chybných záznamov, zavádzania, prípadne ide o extrémne byty. Príkladom sú záznamy, kde byty mali veľkosť  $1\text{ m}^2$ , prípadne predajnú cenu 1,- EUR. Úpravou databázy bolo našou snahou získať homogénnejšiu vzorku, ktorá lepšie zodpovedá realite na trhu. Z obrázku (pozri Obrázok 14) je zrejмый tzv. priamy (a nie prekvapujúci) vzťah medzi cenou a plochou bytu. Tento vzťah sme odhadli pomocou klasického jednoduchého lineárneho regresného modelu, ktorý bol prezentovaný v Kapitole 3. Výsledky naznačujú, že v priemere zaplatíme za každý  $\text{m}^2$  navyše 1353.98,- EUR. Celkovo tento jednoduchý model vysvetľuje až  $\approx 59\%$  variability cien bytov.

```
> cena_plocha <- lm(Cena ~ Plocha, data = data)
> summary(cena_plocha)

Call:
lm(formula = Cena ~ Plocha, data = data)

Residuals:
    Min       1Q   Median       3Q      Max
-21447  -7841  -1835   6063  50049

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -13871.42   4372.18  -3.173  0.0017 **
Plocha       1353.98    71.13  19.036 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

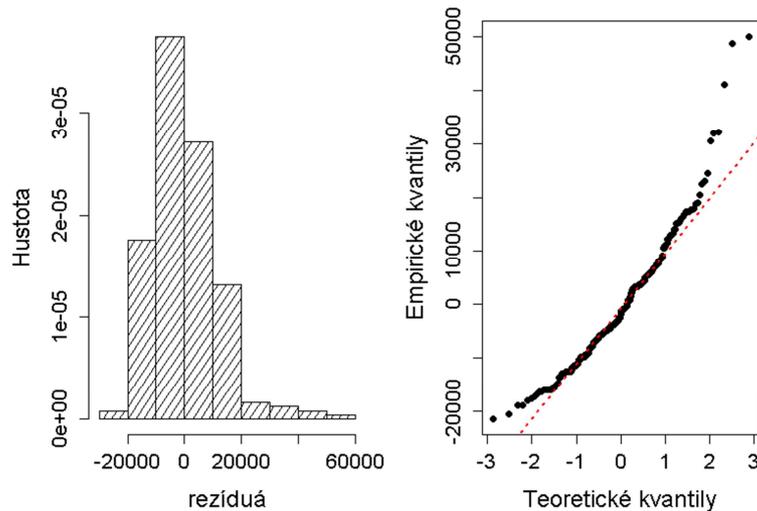
Residual standard error: 11590 on 249 degrees of freedom
```

```
Multiple R-squared: 0.5927,    Adjusted R-squared: 0.5911  
F-statistic: 362.4 on 1 and 249 DF,  p-value: < 2.2e-16
```

Zaujímá nás overovanie predpokladu o normalite rezíduí (ako uvidíme v ďalšej časti, rezíduá z tohto modelu môžeme považovať za homoskedastické).

Použitím postupu od Godfrey (2009) sme dospeli k záveru, že rezíduá nie je možné považovať za realizácie z normálneho rozdelenia pravdepodobnosti. K záveru pripájame aj histogram rezíduí a tzv. kvantil – kvantilový graf (tzv. Q-Q graf). Zjednodušene môžeme tento obrázok interpretovať tak, že pokiaľ empiricky namerané hodnoty ležia na priamke, potom ide o rozdelenie, ktoré zodpovedá normálnemu rozdeleniu pravdepodobnosti. Z obrázku (pozri Obrázok 15) môžeme vidieť, že väčšina rezíduí („v strede rozpätia hodnôt“) sa správa podľa normálneho rozdelenia až na niekoľko extrémnych rezíduí (či už záporné alebo kladné), ktoré sa vyskytujú.

```
> JBPVAL(DEP = data$Cena, X = data$Plocha, B = 9999, algorithm =  
"G")  
$`JB-Statistics`  
X-squared  
 97.93565  
  
$MCPValue  
[1] 1e-04  
> resids <- residuals(cena_plocha)  
> par(mfrow = c(1, 2))  
> hist(resids, density = 10, main = NA, col = "black", cex.lab =  
 1.5, cex.axis = 1.3, freq = FALSE, ylab = "Hustota", xlab =  
  "rezíduá")  
> qqnorm(resids, main = NA, ylab = "Empirické kvantily", pch =  
 19, xlab = "Teoretické kvantily", cex.lab = 1.5, cex.axis =  
 1.3)  
> qqline(resids, col = "red", lty = 3, lwd = 2.5)
```



Obrázok 15: Vizualizácia normality rezíduí

Zdroj: vlastné spracovanie

Keďže používanie Jarque – Berovho testu je v ekonometrii pomerne rozšírené, ilustrujeme si správanie sa testu pri väčšej vzorke. Cieľom je zistiť, (i) či test správne nezamietne nulovú hypotézu v situácii, kde budú rezíduá realizáciami z normálneho rozdelenia pravdepodobnosti, (ii) či test správne zamietne nulovú hypotézu v situácii, kde budú rezíduá realizáciami z iného ako normálneho rozdelenia pravdepodobnosti. Podobne ako Dufour et al. (1998) testovať budeme nasledujúce rezíduá:

$$\hat{\mathbf{u}} = \mathbf{M}\mathbf{u} \quad (4.7)$$

kde platí:

$$\mathbf{M} = \mathbf{I}_n - \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \quad (4.8)$$

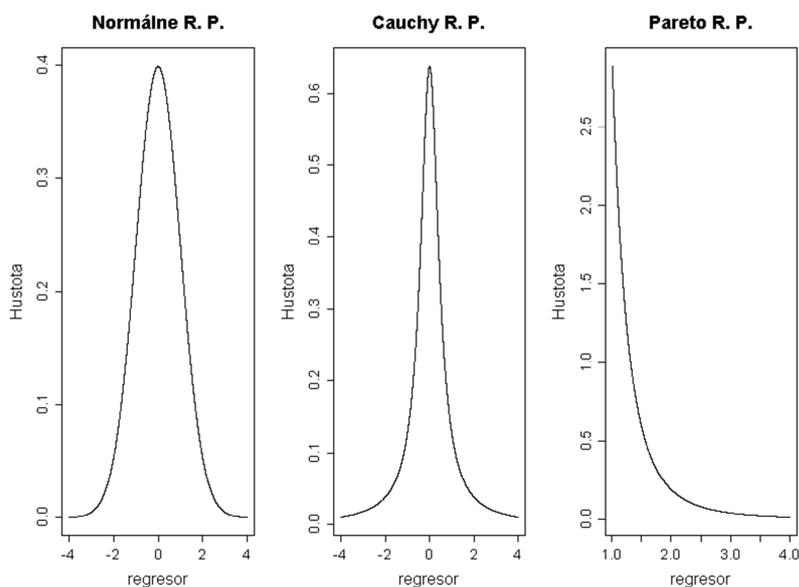
pričom  $\mathbf{I}_n$  je jednotková matica a rezíduá  $\mathbf{u}$  sú modelované rôznymi rozdeleniami pravdepodobnosti, čím sa vytvára rezíduum s požadovanými vlastnosťami. Všimnime si, že vo vzťahoch (4.7) a (4.8) vystupuje matica regresorov. Zjavne sa tak predpokladá, že výsledky ovplyvňujú aj regresory. Rozhodli sme sa uvažovať o štyroch maticiach regresorov (Tabuľka 10 obsahuje rovnaké označenia matíc):

- $\mathbf{X}_1$ , je tvorená iba konštantou (LS).
- $\mathbf{X}_2$ , je tvorená konštantou, štyrmi indikátorovými premennými a piatimi regresormi z normovaného normálneho rozdelenia (budeme označovať ako: C-4D-5N, C – konštanta, D – indikátorová premenná, N – premenná z normálneho rozdelenia, kde čísla označujú počet regresorov s požadovanou vlastnosťou).
- $\mathbf{X}_3$ , je tvorená konštantou, štyrmi indikátorovými premennými a piatimi regresormi z Cauchyho rozdelenia pravdepodobnosti s charakteristikou polohy 0 a škály 1,

(budeme označovať ako: C-4D-5C, C – konštanta, D – indikátorová premenná, 5C – päť náhodných premenných z Cauchiho rozdelenia).

- $\mathbf{X}_4$ , je tvorená konštantou, štyrmi indikátorovými premennými a piatimi regresormi z Paretoho rozdelenia s charakteristikou polohy 1 a tvaru 3, (budeme označovať ako: C-4D-5P, kde P – značí Pareto rozdelenie pravdepodobnosti).

Týmto spôsobom sme chceli uvažovať jednak so základným symetrickým rozdelením nezávislých premenných (Normálne rozdelenie), ako aj so špicatejším, ale symetrickým rozdelením (Cauchy rozdelenie) a výrazne pravostranným rozdelením (Pareto rozdelenie). Tieto rozdelenia sú na nasledujúcom obrázku (pozri Obrázok 16).



Obrázok 16: Rozdelenia regresorov použitých v experimente<sup>12</sup>

*Zdroj: vlastné spracovanie*

Pri modelovaní  $\mathbf{u}$ , sme uvažovali o troch skupinách rozdelení: symetrické rozdelenia, asymetrické rozdelenia a mix rozdelenia (tzv. zmiešané rozdelenia). Týmto spôsobom sme sa snažili napodobniť štúdie od Romão et al. (2010) a Dufour et al. (1998). Budeme však uvažovať iba o jednej veľkosti vzorky  $n = 200$ .

Nasledujúca tabuľka predstavuje zoznam použitých rozdelení vrátane skratiek, ktoré sa použili v tabuľke výsledkov (Tabuľka 10). Pre ilustráciu uvádzame aj vizualizáciu symetrických a asymetrických rozdelení (pozri Obrázok 17).

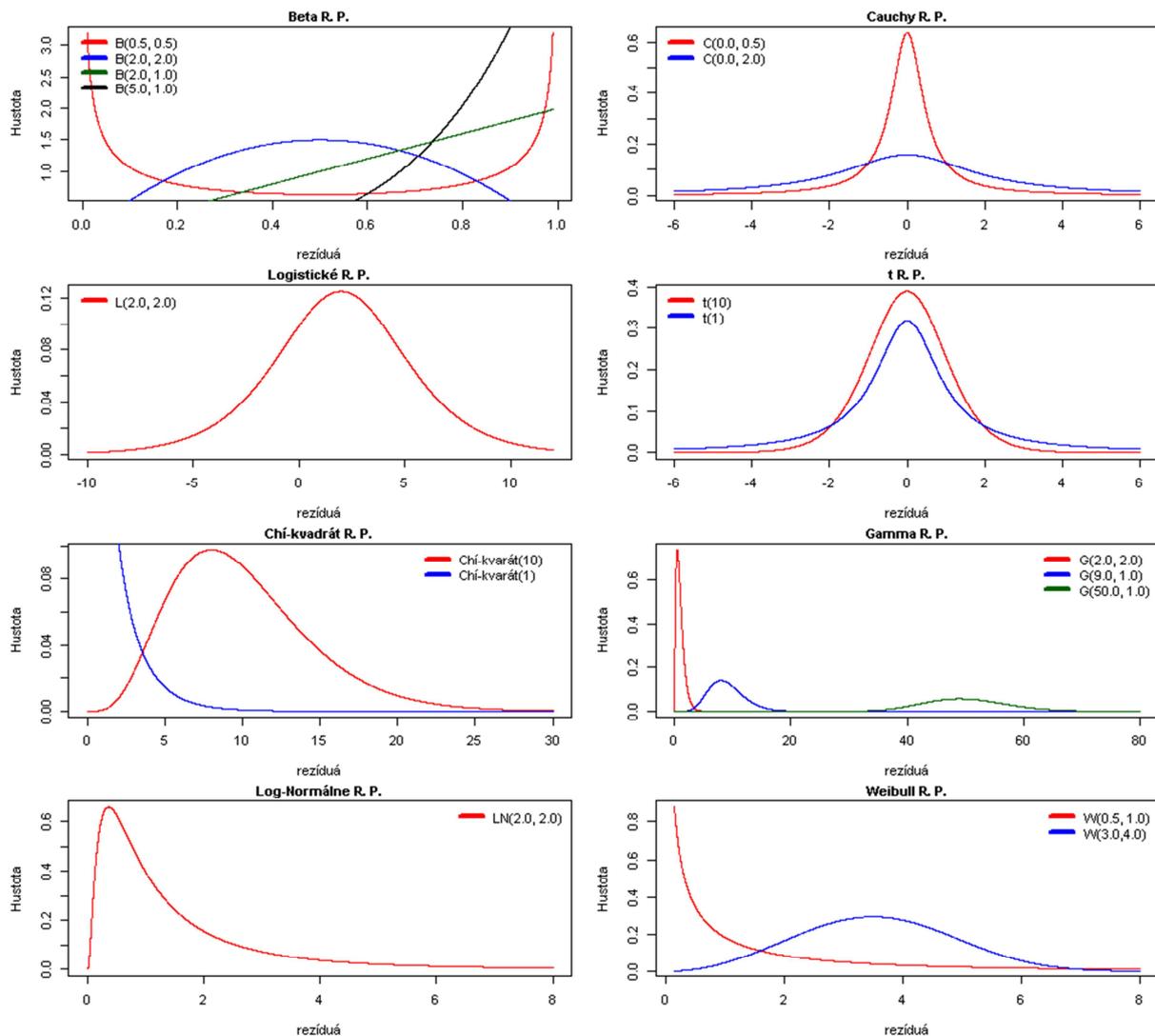
<sup>12</sup> R.P. je skratka pre rozdelenie pravdepodobnosti, používame ju aj v niektorých nasledujúcich tabuľkách a obrázkoch.

Tabuľka 9: Rozdelenia rezíduí

Symetrické rozdelenia				Mix normálne rozdelenia				
<b>Normálne</b>	$N(0, 1)$	$\mu = 0.0$	$\sigma = 1.0$	<b>NN1</b>	30% $n$	$\mu = 1.0$	$\sigma = 0.5$	
<b>Beta</b>	$B(0.5, 0.5)$	$\alpha = 0.5$	$\beta = 0.5$	<b>Mix dvoch normálnych</b>	70% $n$	$\mu = 0.0$	$\sigma = 1.0$	
	$B(2.0, 2.0)$	$\alpha = 2.0$	$\beta = 2.0$		<b>NN2</b>	50% $n$	$\mu = 1.0$	$\sigma = 0.5$
<b>Cauchy</b>	$C(0.0, 0.5)$	$x_0 = 0.0$	$\gamma = 0.5$	<b>NN3</b>	50% $n$	$\mu = 0.0$	$\sigma = 1.0$	
	$C(0.0, 2.0)$	$x_0 = 0.0$	$\gamma = 2.0$		30% $n$	$\mu = 3.0$	$\sigma = 0.5$	
<b>Logistické</b>	$L(2.0, 2.0)$	$\mu = 2.0$	$s = 2.0$	<b>NN4</b>	70% $n$	$\mu = 0.0$	$\sigma = 1.0$	
<b>Studentovo t</b>	$T(1)$	$df = 1$			50% $n$	$\mu = 3.0$	$\sigma = 0.5$	
	$T(10)$	$df = 10$		<b>NN5</b>	50% $n$	$\mu = 0.0$	$\sigma = 1.0$	
<b>Asymetrické rozdelenia</b>				<b>NN6</b>	30% $n$	$\mu = 1.0$	$\sigma = 4.0$	
	<b>Beta</b>	$B(2.0, 1.0)$	$\alpha = 2.0$		$\beta = 1.0$	70% $n$	$\mu = 0.0$	$\sigma = 1.0$
	$B(5.0, 1.0)$	$\alpha = 5.0$	$\beta = 1.0$		50% $n$	$\mu = 1.0$	$\sigma = 4.0$	
$\chi^2$	$\chi^2(1)$	$df = 1$		<b>NO1</b>	$\mu = 0.0$	$\sigma = 1$	$k = 2$	
	$\chi^2(10)$	$df = 10$			$x(\max) = x^{Q3} + k \cdot IQR$			
<b>Gama</b>	$G(2.0, 2.0)$	$\alpha = 2.0$	$\beta = 2.0$	<b>Normálne s extrém. hodnot.</b>	<b>NO2</b>	$\mu = 0.0$	$\sigma = 1$	$k = 2$
	$G(9.0, 1.0)$	$\alpha = 9.0$	$\beta = 1.0$		$x(\max) = x^{Q3} + k \cdot IQR$			
<b>Log-normálne</b>	$G(100.0, 1.0)$	$\alpha = 50.0$	$\beta = 1.0$		$x(\min) = x^{Q1} - k \cdot IQR$			
	$LN(0.0, 1.0)$	$\mu^* = 0.0$	$\sigma^* = 1.0$	<b>NO3</b>	$\mu = 0.0$	$\sigma = 1$	$k = 3$	
<b>Weibull.</b>	$W(0.5, 1.0)$	$\lambda = 0.5$	$\kappa = 1.0$		$x(\max) = x^{Q3} + k \cdot IQR$			
	$W(3.0, 4.0)$	$\lambda = 3.0$	$\kappa = 4.0$		$x(\min) = x^{Q1} - k \cdot IQR$			

Zdroj: vlastné spracovanie

Pozn.:\* ide o strednú hodnotu a smerodajnú odchýlku na logaritmickú škálu, t. j. hodnoty z lognormálneho rozdelenia pravdepodobnosti sú logaritmované a ich stredná hodnota a smerodajná odchýlka je uvedená v tabuľke.  $x(\max)$  predstavuje maximálnu hodnotu súboru,  $x(\min)$  predstavuje minimálnu hodnotu súboru,  $IQR$  predstavuje medzikvartilové rozpätie,  $xQ3$  je horný kvartil a  $xQ1$  je dolný kvartil. Rozdelenia NN1 sa skladajú z dvoch normálnych rozdelení, kde 30% $n$  značí podiel prvého rozdelenia. Rozdelenia NO1 majú v sebe jednu (NO1), resp. dve extrémne hodnoty (NO2, NO3).



Obrázok 17: Rozdelenia rezíduí použitých v experimente

*Zdroj: vlastné spracovanie*

Pre každú kombináciu regresorov a rozdelení sme vykonali 1000 iterácií a v každej iterácii Monte – Carlo Jarque – Berovho test s  $B = 999$  a postupom od Godfrey (2009). Počet iterácií, ako aj hodnota parametra  $B$  sú relatívne menšie. Dôvod spočíva vo vysokej časovej náročnosti celej simulácie. Výsledky zo simulácií sú uvedené v nasledujúcej tabuľke (Tabuľka 10).

Panel A zobrazuje výsledky chyby prvého druhu. Je zrejmé, že na regresoroch záleží. V prípade, ak je v regresii iba konštanta, test správne nezamietal nulovú hypotézu vo všetkých prípadoch simulácie. Ak sa v modeloch vyskytli konštanty aj regresory, chyba bola väčšia ako nominálnych 5 %. K podobným záverom dospeli aj Dufour et al. (1998), kde pri iných ako modeloch s konštantou bola empirická chyba I. druhu spravidla vyššia ako nominálnych 5 %. Hodnoty sa pritom pohybovali v podobnom intervale ako u nás, t. j. od 5 % do 10 %. Čo je

menej priaznivé, Dufour et al. (1998) konštatujú, že s rastúcou vzorkou sa chyba testu neznižovala (týka sa to aj ostatných, bežných testov normality).

Tabuľka 10: Empirická chyba I. druhu a sila Monte – Carlo Jarque – Berovho testu

Panel A: Chyba I. druhu					Panel C: Asymetrické rozdelenia				
	LS	C-4D-5N	C-4D-5C	C-4D-5P		LS	C-4D-5N	C-4D-5C	C-4D-5P
<b>N(0,1)</b>	0.000	0.098	0.064	0.088	B(2,1)	1.000	1.000	1.000	1.000
<b>Panel B: Symetrické rozdelenia</b>					B(5,1)	1.000	1.000	1.000	1.000
	LS	C-4D-5N	C-4D-5C	C-4D-5P	CHISQ(1)	1.000	1.000	1.000	1.000
<b>B(0.5,0.5)</b>	1.000	1.000	1.000	1.000	CHISQ(10)	1.000	1.000	1.000	1.000
<b>B(2.0,2.0)</b>	0.000	0.008	0.002	0.007	GAMMA(2, 2)	1.000	1.000	1.000	1.000
<b>C(0.0,0.5)</b>	1.000	1.000	1.000	1.000	GAMMA(9, 1)	1.000	0.981	0.985	0.985
<b>C(0.0,2.0)</b>	1.000	1.000	1.000	1.000	GAMMA(50, 1)	0.000	0.000	0.000	0.001
<b>L(2.0,2.0)</b>	1.000	0.994	0.975	0.974	LOGNORMAL(0,1)	1.000	1.000	1.000	1.000
<b>T(1)</b>	1.000	1.000	1.000	1.000	WEIBULL(0.5,1)	1.000	1.000	1.000	1.000
<b>T(10)</b>	0.025	0.252	0.330	0.291	WEIBULL(3,4)	0.000	0.000	0.000	0.000
<b>Panel D: Mix rozdelení</b>									
	LS	C-4D-5N	C-4D-5C	C-4D-5P					
<b>NN1</b>	0.000	0.000	0.000	0.000					
<b>NN2</b>	1.000	0.999	0.998	0.999					
<b>NN3</b>	1.000	0.961	0.957	0.973					
<b>NN4</b>	1.000	1.000	1.000	1.000					
<b>NN5</b>	1.000	1.000	1.000	1.000					
<b>NN6</b>	1.000	0.999	0.999	0.997					
<b>NO1</b>	0.000	0.081	0.070	0.073					
<b>NO2</b>	0.000	0.003	0.000	0.000					
<b>NO3</b>	1.000	0.971	0.957	0.949					

Zdroj: vlastné spracovanie

Pri symetrických, nie normálnych rozdeleniach (pozri panel B, Tabuľka 10) bola sila Jarque - Berovho testu veľmi nízka pre rozdelenia  $B(2;2)$  a  $t(10)$ . Pre Studentovo  $t$  rozdelenie výsledok nie je veľmi prekvapivý, keďže s rastúcim počtom stupňov voľnosti sa rozdelenie približuje k normálnemu. Efekt regresorov sa pritom nezdá byť zásadný. Pri asymetrických rozdeleniach bola sila testu vo väčšine prípadov veľká, > 98 % správnych zamietnutí nulovej hypotézy. Nízka sila bola iba pri mierne asymetrických rozdeleniach, ako Weibull(3,4) a Gamma(50,1).

Pri kombinácii dvoch rozdelení (panel D, modely  $NN$ ) je sila testov adekvátna a to bez rozdielu na použitých regresoroch. Výnimku tvorí iba  $NN1$  test, kde v empirickom rozdelení dominuje štandardizované rozdelenie a druhé, normálne rozdelenie má pomerne nízky rozptyl a malý posun strednej hodnoty na  $\mu = 1$ , v dôsledku ktorých zrejme test v mnohých prípadoch nebol schopný zamietnuť normálne rozdelenie. Ak sme do rozdelení pridali extrémnu hodnotu, potom sila testu bola adekvátna až v situácii, keď extrémna hodnota bola výrazne väčšia ( $k = 3$ ,  $NO3$ ) ako horný, resp. dolný kvartil empirického súboru.

## 4.2 Testovanie heteroskedasticity rezíduí

V podkapitole 3.3.3 sme si predstavili predpoklady lineárneho modelu. Výsledky, ktoré sme pre tieto modely odvodili môžu byť správne len vtedy, ak sú predpoklady využívané pri odvodení splnené. V kapitole 3.5.1 sme si ukázali, aké dôsledky má porušenie predpokladov lineárneho modelu. Veľmi závažné je hlavne porušenie predpokladu P2, kde predpokladáme absenciu autokorelácie a homoskedasticitu. V tejto kapitole uvedieme niekoľko možností, ako testovať predpoklad o homoskedasticite. Okrem tradičných testov (Breusch – Pagan, White) si ukážeme aj to, ako je možné realizovať Whiteov test s použitím metódy bootstrap.

### 4.2.1 Breusch – Paganov test

Regresná analýza sa zriedkakedy zaobíde bez overovania predpokladu o homoskedasticite rezíduí. Jedným z najpopulárnejších testov je zrejme Breusch – Paganov (1979) LM test. Popularita testu sa zakladá na jeho jednoduchosti. Pomerne veľkým nedostatkom pôvodnej verzie tohto testu je, že pre správnu inferenciu je potrebné, aby boli splnené nasledujúce predpoklady:

- rezíduá pochádzajú z lineárneho regresného modelu (správna špecifikácia modelu):

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \quad (4.9)$$

- v matici regresorov  $\mathbf{X}$  sú všetky premenné striktno exogénne. Pod striktnou exogenitou regresorov budeme na teraz rozumieť nasledujúcu vlastnosť: (náhodné) regresory sú nezávislé od rezíduí.
- rezíduá  $\mathbf{u}$  sú *iid* a pochádzajú z normálneho rozdelenia pravdepodobnosti so strednou hodnotou 0, avšak pripúšťa sa rôznych rozptyl.
- rozptyl rezíduí  $\sigma_i^2$  je možné modelovať pomocou nasledujúceho vzťahu:

$$\sigma_i^2 = h\left(\gamma_0 + \sum_{j=1}^q \gamma_j x_{i,j}\right) \quad (4.10)$$

kde  $z_{i,j}$  sú spravidla striktno exogénne premenné a  $h(\cdot)$  je ľubovoľná „hladká“ funkcia<sup>13</sup>, pre ktorú platí  $h'(\gamma_0) \neq 0$ . Testovaciu štatistiku  $nR^2$  (Koenker, 1981) je potom možné

---

<sup>13</sup> Zrejme najčastejšie sa uvažuje iba o funkcii  $h(r) = x$  (Zaman, 2000).

vypočítať použitím tzv. pomocnej regresie (pôvodná verzia testu uvažovalo o odlišnej špecifikácii a inej štatistike):

$$\hat{u}_i^2 = \gamma_0 + \sum_{j=1}^q \gamma_j x_{i,j} + e_i \quad (4.11)$$

Pri tejto verzii testu môžeme upustiť od predpokladu o normalite rezíduí  $\mathbf{u}$ . Testovacia charakteristika  $nR^2$  sa za predpokladu platnosti nulovej hypotézy o homoskedasticite riadi  $\chi^2(q)$  rozdelením pravdepodobnosti.

Ďalším dôvodom popularity tohto testu je skutočnosť, že pomocná regresia spravidla obsahuje menší počet regresorov ( $q$ ) ako Whiteov test prezentovaný v ďalšej kapitole 4.2.2 a 4.2.3. Všimnime si, že vo vzťahu (4.11) nás vlastne zaujíma nakoľko lineárna kombinácia regresorov pomáha vysvetliť rozptyl rezíduí. Ak výrazne, potom zjavne rozptyl rezíduí závisí od určitého stavu regresorov, t. j. v určitej situácii majú rezíduá väčší (menší) rozptyl ako v inej. Potom nemôžeme hovoriť o homoskedasticite.

#### 4.2.2 Whiteov test

Pokiaľ máme dostatočný počet pozorovaní, lepšou alternatívou je tzv. Whiteov test (1980). Preto práve tento test popíšeme podrobne aj s príslušnými bootstrap verziami testu. Podobne ako v predošlom prípade, uvažujeme o modeli (4.9), kde matica  $\mathbf{X}$  má rozmer  $n \times p$ , kde jednotlivé stĺpce predstavujú regresory, vrátane jednotkového vektora (teda konštanty – uvažujeme iba o modeloch, ktoré zahŕňajú konštantu). Regresory sú striktne exogénne, pričom variančno-kovariančná matica rezíduí  $\mathbf{u}$  je diagonálna matica  $\mathbf{\Omega} = \text{diag}(\sigma_1^2, \dots, \sigma_n^2)$ . Podobne ako v predošlom prípade, nepredpokladáme nutne rovnaký rozptyl (koniec koncov testujeme, či je možné rozptyly rezíduí považovať za homoskedastické). Všimnime si, že matica  $\mathbf{\Omega}$  je diagonálna, čo v praxi predpokladá nekorelovanosť rezíduí. Aj keď nás v tejto časti nezaujíma analýza časových radov, je to dôležitý prvok v prípade, ak uvažujeme o modeloch s časovými radmi. Potom je vhodné, aby sme najprv získali model, kde je rozumné predpokladať nekorelovanosť rezíduí (čo sa vyšetruje samostatným testom – bližšie pozri Kapitolu 4.3) a potom overovať predpoklad o heteroskedasticite.

Na testovanie môžeme použiť dve štatistiky (White, 1980). Prvá štatistika  $W$  je daná nasledujúcimi vzťahmi:

$$W = n \left( n^{-1} \sum_{i=1}^n (\hat{u}_i^2 - \hat{\sigma}^2) \boldsymbol{\psi}_i^T \right) \hat{\mathbf{D}}^{-1} \left( n^{-1} \sum_{i=1}^n (\hat{u}_i^2 - \hat{\sigma}^2) \boldsymbol{\psi}_i \right) \quad (4.12)$$

$$\hat{\sigma}^2 = n^{-1} \sum_{i=1}^n \hat{u}_i^2 \quad (4.13)$$

$$\hat{\mathbf{D}} = n^{-1} \sum_{i=1}^n (\hat{u}_i^2 - \hat{\sigma}^2)^2 (\boldsymbol{\psi}_i - \bar{\boldsymbol{\psi}})(\boldsymbol{\psi}_i - \bar{\boldsymbol{\psi}})^T \quad (4.14)$$

$$\bar{\boldsymbol{\psi}} = n^{-1} \sum_{i=1}^n \boldsymbol{\psi}_i \quad (4.15)$$

kde  $\boldsymbol{\psi}_i$  je  $p(p+1)/2 \times 1$  vektor, ktorého prvky pozostávajú z nasledujúcich kombinácií  $p$  regresorov:  $\mathbf{x}_{k=1,\dots,p}, x_{i,k}x_{i,m}$  ( $i = 1, \dots, n; k = 1, \dots, p; m = 1, \dots, k$ ). Ide vlastne o všetky možné dvojice regresorov (vrátane konštanty). Ich celkový počet je  $p(p+1)/2$ . Ak napríklad uvažujeme o jednoduchom lineárnom regresnom modeli s konštantou  $x_{i,0}$  a jednou nezávislou premennou  $x_{i,1}$ , potom má vektor tvar:  $\boldsymbol{\psi}_i = (1, x_{i,1}, x_{i,1}^2)^T$ . Testovacia štatistika  $W$  sa v prípade platnosti nulovej hypotézy riadi  $\chi^2(p)$  rozdelením pravdepodobnosti s  $p$  stupňami voľnosti.

Pri tomto teste potrebujeme podstatne viac parametrov ako v prípade Breusch – Paganovho testu (čo je dané kombináciami regresorov). Zároveň pri (4.12) je potrebné vypočítať inverznú maticu z  $\hat{\mathbf{D}}$ , kde aj v prípade, ak táto matica existuje, v prípade veľkých rezíduí môžeme naraziť na technické hranice výpočtovej techniky. V praxi sa používa oveľa jednoduchšia verzia tohto testu, kde testovacia štatistika má tvar  $W^* = nR^2$ , pričom koeficient determinácie sa vypočíta z nasledujúceho modelu:

$$\hat{u}_i^2 = \gamma_0 + \sum_{k=1}^p \sum_{m=1}^k \gamma_{0.5k(k-1)+m} x_{i,k} x_{i,m} + e_i \quad (4.16)$$

príčom ak sa medzi členmi  $x_{i,k}x_{i,m}$  nachádza konštanta, tak sa vzťah (4.16) odhaduje bez tejto konštanty (v programe R sa táto nadbytočná konštanta vylúči). Podobne ako v predošlom prípade, za predpokladu platnosti nulovej hypotézy sa  $W^*$  riadi  $\chi^2(p)$  rozdelením pravdepodobnosti s  $p$  stupňami voľnosti.

Hypotéza o homoskedasticite rezíduí sa overuje iba v prípade, ak lineárny regresný model (4.9) považujeme za správne špecifikovaný. White (1980, s. 823) k problematike združenej nulovej hypotézy konštatoval: „Nulová hypotéza nezahŕňa len homoskedasticitu rezíduí, ale aj nezávislosť rezíduí a regresorov a skutočnosť, že model je správne špecifikovaný. Ak je testovacia štatistika nevýznamná je to obzvlášť dobrá správa, keďže to

implikuje, že nie len že špecifikácia rozptylu modelu je správna, ale že aj lineárna špecifikácia modelu je správna“. White (1980) preto tento test do určitej miery považoval za všeobecný test špecifikácie modelu (4.9). Samozrejme, významná štatistika môže znamenať porušenie ľubovoľného z vyššie spomínaných predpokladov testu. Môže ísť o zlú špecifikáciu modelu, o nezávislosť regresorov a rezíduí, prípadne o heteroskedasticitu, ktorá sa spravidla prejavuje určitou funkčnou závislosťou rozptylu rezíduí a regresorov. Preto správny záver z použitia tohto testu do značnej miery závisí od našej dôvery v správnosť špecifikácie modelu a striktnej exogenite regresorov.

### 4.2.3 Whiteov test – Bootstrap

V predošlej časti sme napísali, že testovacia štatistika  $W^*$  (aj  $W$ ) sa asymptoticky riadia príslušným  $\chi^2(p)$  rozdelením pravdepodobnosti s  $p$  stupňami voľnosti. Ide teda o asymptotickú vlastnosť, ktorá v praxi nemusí byť vždy veľmi presná. Jednou z možností, ako vylepšiť presnosť tohto testu, je použiť bootstrap. Hodoshima – Ando (2007) skúmali osem spôsobov bootstrapovania testovacej štatistiky  $W^*$ . Nasledujúcu verziu bootstrapovania vyhodnotili ako najlepšiu (najvyššia sila testu pri očakávanej chybe I. druhu), a to aj pre prípad stochastických regresorov (čo je najčastejší prípad v empirickom výskume).

Krok 1: Odhadnite regresný model (4.9) pomocou OLS. Výstupom z tohto kroku sú odhadované rezíduá  $\hat{\mathbf{u}}$  a predikované hodnoty  $\hat{\mathbf{y}}$ .

Krok 2: Vypočítajte testovaciu štatistiku  $W^*$ .

Krok 3: Vytvorte vzorku  $\mathbf{u}^* = (u_1^*, \dots, u_n^*)$  náhodným výberom s opakovaním, zo vzorky  $\hat{\mathbf{u}} = (\hat{u}_1, \dots, \hat{u}_n)$ .

Krok 4: Vytvorte bootstrap závislú premennú nasledovne:

$$\mathbf{y}^* = \hat{\mathbf{y}} + \sqrt{\frac{n}{n-p}} \mathbf{u}^* \quad (4.17)$$

Krok 5: Na pseudo vzorke  $\mathbf{y}^*$  a  $\mathbf{X}$  odhadnite regresný model (4.9) a vypočítajte štatistiku  $W_b^*$ ,  $b = 1, \dots, B$ .

Krok 6: Kroky 3 – 5 opakujte  $B$  krát. Vypočítajte  $p$ -hodnotu nasledovne:

$$W^*BPVAL = \frac{|\{W_b^*; W_b^* \geq W^*\}|}{B} \quad (4.18)$$

Podobne ako pri teste normality, symbolom # označujeme počet. Hodnotu  $B$  zvolíme dostatočne veľkú, napr. Hodoshima – Ando (2007) pri simuláciách používali hodnotu 10000.

Zostrojíme si funkciu, ktorá bude implementovať obe verzie testovania významnosti  $W^*$  štatistiky. Argumentom funkcie je objekt funkcie `lm()` a hodnota počtu vzoriek, ktoré sa pri bootstrape použijú ( $B$ ). Výstupom je hodnota testovacej štatistiky  $W^*$ , 1 %, 5 %, 10 % asymptotické a bootstrap kritické hodnoty a  $p$ -hodnoty.

```
White.test <- function(model, B = 10000) {
  temp <- dim(model$model)[2]; DEP <- model$model[,1]; X <-
  as.matrix(model$model[,2:temp])
  PSI <- function(X) {
    X <- cbind(X, rep(1,dim(X)[1]))
    n <- dim(X)[1]; p <- dim(X)[2]
    PSI <- matrix(nrow = n, ncol = (p*(p+1)/2))
    t <- 1
    for (k in 1:p) {
      for (l in 1:k) {
        PSI[,t] <- X[,k]*X[,l]
        t <- t + 1
      }
    }
    results <- list()
    results[["PSI"]] <- PSI
    return(results)
  }

  W_star <- function(DEP = NULL, X = NULL) {
    psi <- PSI(X)$PSI;
    n <- dim(X)[1]; k <- dim(X)[2];
    model <- lm(DEP ~ X[,1:k]);
    u_hat <- residuals(model);
    u_hat2 <- u_hat^2
    model <- lm(u_hat2 ~ psi);
    W_star <- summary(model)$r.squared*n
    return(W_star)
  }

  W <- W_star(DEP, X)
  p <- dim(X)[2]
  A01_crit <- round(qchisq(0.99, (p*(p+1))/2), 5)
  A05_crit <- round(qchisq(0.95, (p*(p+1))/2), 5)
  A10_crit <- round(qchisq(0.90, (p*(p+1))/2), 5)
  A_pval <- round(1 - pchisq(W, (p*(p+1))/2), 5)

  n <- dim(X)[1];
  model <- lm(DEP ~ X);
  u_hat <- residuals(model);
  y_predict <- predict(model);
  Wb <- c()
  for (i in 1:B) {
    u_star <- u_hat[sample(1:n, size = n, replace = T)]
    y_star <- y_predict + sqrt(n/(n-p))*u_star
    Wb[i] <- W_star(DEP = y_star, X)
  }
}
```

```

B01_crit <- round(quantile(Wb, prob = 0.99), 5)
B05_crit <- round(quantile(Wb, prob = 0.95), 5)
B10_crit <- round(quantile(Wb, prob = 0.90), 5)
B_pval <- round(sum(Wb > W)/B, 5)
infer <- cbind(c(A01_crit, A05_crit, A10_crit), c(B01_crit,
B05_crit, B10_crit))
colnames(infer) <- c("ASYMPTOTIC", "BOOTSTRAP")
pvals <- cbind(A_pval, B_pval)
colnames(pvals) <- c("ASYMPTOTIC", "BOOTSTRAP")
results <- list()
results[["White"]] <- W
results[["crit"]] <- infer
results[["pvals"]] <- pvals
return(results)
}

```

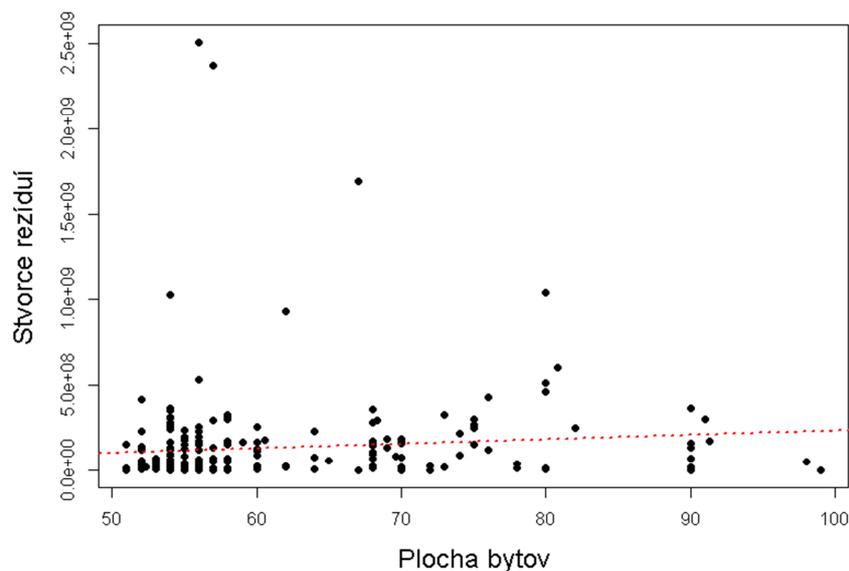
#### Príklad 4.2

Budeme voľne pokračovať v predošlom príklade, kde sme odhadovali jednoduchý lineárny regresný model so závislou premennou „cena bytu“ a nezávislou premennou „plocha bytu“. Ďalej bez dodatočného testovania predpokladajme, že lineárna špecifikácia modelu je správna. Taktiež považujeme plochu za striktno exogénnu premennú. Začnime jednoduchou vizualizáciou štvorcov rezíduí (odhad rozptylu rezíduí) vo vzťahu k ploche bytov.

```

> cena_plocha <- lm(Cena ~ Plocha, data = data)
> plot(data$Plocha, residuals(cena_plocha)^2, type = "p", pch =
  19, xlab = "Plocha bytov", ylab = "Štvorce rezíduí", cex.lab =
  1.5)
> abline(lm(residuals(cena_plocha)^2 ~ data$Plocha), col =
  "red", lty = 3, lwd = 2.5)

```



Obrázok 18:  $x$ - $y$  graf štvorcov rezíduí a plochy bytov

*Zdroj: vlastné spracovanie*

Z obrázku (pozri Obrázok 18) sa zdá, že by mohol existovať vzťah medzi štvorcami rezíduí a plochou bytov. Nie je však zrejmé, či tento mierne rastúci trend je spôsobený extrémnymi hodnotami (medzi 50 a 60  $m^2$ ), alebo je to systematický efekt. Pri overovaní prítomnosti heteroskedasticity pomocou Whiteovho testu sa naše podozrenie nepotvrdilo a nulovú hypotézu sme nezamietli (použitím oboch prístupov ku kritickým hodnotám – asymptotických ako aj bootstrap).

```
$White
[1] 2.141528

$crit
  ASYMPTOTIC BOOTSTRAP
99%   16.81189   27.16518
95%   12.59159   13.27134
90%   10.64464    7.88251

$pvals
  ASYMPTOTIC BOOTSTRAP
[1,]   0.90623    0.541
```

### 4.3 Testovanie autokorelácie rezíduí

Ak uvažujeme o klasických regresných modeloch ako (4.9), mnohé testy sprevádzajúce prácu s týmito testami predpokladajú, že variančno-kovariančná matica rezíduí  $\mathbf{u}$  je diagonálna. To znamená, že rezíduá nie sú navzájom korelované. Ak uvažujeme

o lineárnom regresnom modeli na prierezných údajoch, potom je tento predpoklad spravidla opodstatnený<sup>14</sup>. Pri analýze ekonomických časových radov je predpoklad o nezávislosti rezíduí otázný a vyžaduje si formálne overovanie.

Preto ak v regresnom modeli vystupujú časové rady, jedným z prvých krokov v diagnostike je overovanie predpokladu o nulovej autokorelácii rezíduí. Z predchádzajúcich dvoch testov sme mohli vidieť, že nezávislosť rezíduí<sup>15</sup> je nutným predpokladom k testovaniu hypotéz o normalite a homoskedasticite rezíduí. Z toho vyplýva určité poradie, v ktorom by sa diagnostika mala uskutočňovať. Zároveň však platí, že test, ktorým pri overovaní autokorelácie rezíduí použijeme, by nemal byť ovplyvnený prípadnou heteroskedasticitou a ne-normalitou rezíduí. Štandardné testy v takej situácii nie je možné použiť (bližšie pozri Godfrey – Tremayne, 2005). V nasledujúcej časti si odvodíme test, ktorý je vhodné použiť aj v situáciách, kde regresory zahŕňajú nielen exogénne premenné, ale aj oneskorené premenné závislej premennej. Tento test nazveme HR robust LM Breusch – Godfreyov test, ktorý bol prezentovaný v práci Godfrey (1997) a Godfrey – Tremayne (2005). Následne v ďalšej podkapitole použijeme bootstrap na odvodenie kritických hodnôt, ktoré ako bolo ukázané v práci Godfrey – Tremayne (2005), vedú k lepším štatistickým vlastnostiam testu ako asymptotické kritické hodnoty.

#### 4.3.1 HR robust LM Breusch - Godfreyov test

Uvažujme o regresnom modeli s exogénnymi a endogénnymi premennými v nasledujúcom tvar:

$$\mathbf{y} = \mathbf{Y}\boldsymbol{\alpha} + \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \quad (4.19)$$

kde ak si počet pozorovaní označíme ako  $T$ , potom rozmery vektorov a matic zo vzťahu (4.19) sú nasledujúce:  $\mathbf{y} - T \times 1$ ,  $\mathbf{Y} - T \times P$ ,  $\boldsymbol{\alpha} - P \times 1$ ,  $\mathbf{X} - T \times L$ ,  $\boldsymbol{\beta} - L \times 1$ ,  $\mathbf{u} - T \times 1$ . Pričom stĺpce matice  $\mathbf{Y}$  predstavujú oneskorené hodnoty závislej premennej  $y_{t-1}, \dots, y_{t-P}$

<sup>14</sup> Jednou z výnimiek sú už spomínané priestorové regresné modely. Uvažujme o situácii, v ktorej je našou snahou vysvetliť objem mesačných výdavkov jednej domácnosti na energie. Máme priemerné údaje za jednotlivé okresy. Potom je možné, že v ľubovoľnom regresnom modeli (ako nezávislú premennú by sme mohli použiť hustotu obyvateľstva na  $1 \text{ km}^2$ ) budú rezíduá modelu závislé, keďže objem mesačných výdavkov, ako aj hustota obyvateľstva sa spravidla nemení náhodne od okresu k okresu, ale postupne. Alebo inak: je pravdepodobnejšie, že v okrese A bude podobný objem mesačných výdavkov a hustoty obyvateľstva ako v susediacich okresoch. Týmto inak zaujímavým modelom sa v tejto publikácii venovať nebudeme.

<sup>15</sup> Korelácia je len určitá forma závislosti rezíduí, avšak je zrejme jedna z najčastejšie sa vyskytujúcich foriem závislosti. Preto pri overovaní predpokladu o nezávislosti rezíduí sa spravidla overuje autokorelácia rezíduí. Z tohto dôvodu píšeme text tak, ako keby sme testovaním autokorelácie riešili nezávislosť rezíduí. V skutočnosti však nevieme vylúčiť inú ako lineárnu formu závislosti.

a stĺpce matice  $\mathbf{X}$  zodpovedajú exogénnym premenným. Ďalej uvažujeme, že  $\mathbf{y}$  je slabo stacionárny proces.<sup>16</sup> Regresory v matici  $\mathbf{X}$  môžu byť predeterminované alebo stochastické, prípadne nemusia byť stacionárne (slabá forma stacionarity, tzv. kovariančne stacionárne)<sup>17</sup>. Našou snahou je overiť hypotézu, že rezíduá  $\mathbf{u}$  nie sú autokorelované, t. j. hodnota rezíduí v čase  $t$  nezávisí (lineárne) od predchádzajúcich  $Q$  rezíduí.

Označme si predikované hodnoty OLS modelu (4.19) a odhadované rezíduá nasledovne:  $\hat{\mathbf{y}} = \mathbf{Y}\hat{\mathbf{a}} + \mathbf{X}\hat{\mathbf{\beta}}$  a  $\hat{\mathbf{u}}$ . Potom si vytvoríme maticu  $\mathbf{U}$ , ktorej riadky predstavujú oneskorené hodnoty rezíduí  $\hat{u}_{t-1}, \dots, \hat{u}_{t-Q}$ . Všimnime si, že ak začíname analýzu s počtom pozorovaní  $R$ , potom ak v modeli (4.19) použijeme  $P = 1$ , budú mať jednotlivé premenné v modeli (4.19) časový rozmer  $R - 1 = T$ . Aby mohol byť rozmer matice  $\mathbf{U}$ ,  $T \times Q$ , budeme musieť zadefinovať, že v prípade, ak bude platiť  $t \leq Q$ , potom  $\hat{u}_t = 0$ . Efekt tejto úpravy dát bude pre dostatočne veľký počet pozorovaní zanedbateľný. Označme si variančno-kovariančnú maticu rezíduí  $\hat{\mathbf{\Omega}} = \text{diag}\{\hat{u}_1^2, \dots, \hat{u}_T^2\}$  a vektor rezíduí  $\hat{\mathbf{u}} = (\hat{u}_1, \dots, \hat{u}_T)^T$ , potom:

$$\mathbf{M}(\mathbf{W}) = \mathbf{I}_T - \mathbf{W}(\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T \quad (4.20)$$

$LM_{HR}$  štatistika má nasledujúci tvar:

$$LM_{HR} = \hat{\mathbf{u}}^T \mathbf{U} \{ \mathbf{U}^T \mathbf{M}(\mathbf{W}) \hat{\mathbf{\Omega}} \mathbf{M}(\mathbf{W}) \mathbf{U} \}^{-1} \mathbf{U}^T \hat{\mathbf{u}} \quad (4.21)$$

Za predpokladu platnosti nulovej hypotézy sa  $LM_{HR}$  štatistika riadi  $\chi^2$  rozdelením pravdepodobnosti s  $Q$  stupňami voľnosti. Okrem tejto štatistiky je možné použiť aj  $HA_{HR}$  štatistiku (prvú budeme preferovať pre lepšiu silu testu, pozri Godfrey – Tremayne, 2005). Jej odvodenie je však pomerne prácne:

$$\mathbf{A} = T^{-1} \begin{bmatrix} \mathbf{Y}^T \mathbf{Y} & \mathbf{Y}^T \mathbf{X} & \mathbf{Y}^T \mathbf{U} \\ \mathbf{X}^T \mathbf{Y} & \mathbf{X}^T \mathbf{X} & \mathbf{X}^T \mathbf{U} \\ \mathbf{U}^T \mathbf{Y} & \mathbf{U}^T \mathbf{X} & \mathbf{U}^T \mathbf{U} \end{bmatrix} \quad (4.22)$$

Vytvoríme si inverznú maticu k matici  $\mathbf{A}$  a príslušný kvadrant  $\mathbf{Y}^T \mathbf{U}$  tejto matice si označíme ako  $\mathbf{A}^{\mathbf{YU}}$  (jej rozmer je  $P \times Q$ ). Potom  $\mathbf{J} = \mathbf{U}(\mathbf{A}^{\mathbf{YU}})^T$ . A  $HA_{HR}$  má tvar:

$$HA_{HR} = \hat{\mathbf{u}}^T \mathbf{J} \{ \mathbf{J}^T \mathbf{M}(\mathbf{W}) \hat{\mathbf{\Omega}} \mathbf{M}(\mathbf{W}) \mathbf{J} \}^{-1} \mathbf{J}^T \hat{\mathbf{u}} \quad (4.23)$$

<sup>16</sup> Kapitola 5.3

<sup>17</sup> Vynechali sme ešte jeden predpoklad, ktorý hovorí o existencii zmysluplného vzťahu medzi regresormi navzájom (bližšie pozri Godfrey – Tremayne, 2005, s. 380).

Ak platí  $P < Q$ , potom za predpokladu platnosti nulovej hypotézy, sa  $HA_{HR}$  štatistika riadi  $\chi^2$  rozdelením pravdepodobnosti s  $P$  stupňami voľnosti. Ak  $P \geq Q$  sú (asymptoticky) rozdiely medzi štatistikou  $LM_{HR}$  a  $HA_{HR}$  minimálne. Preto aj vo funkcii, ktorú prezentujeme v ďalšej kapitole (pozri 4.3.2) sa pre prípad  $P \geq Q$  počíta iba  $LM_{HR}$  štatistika.

### 4.3.2 Bootstrap HR robust LM Breusch - Godfreyov test

V simuláciách, ktoré uskutočnili Godfrey – Tremayne (2005) boli asymptotické kritické hodnoty robustných testov pomerne nepresné. Ukázalo sa, že vhodnou alternatívou je použiť bootstrap metódu. Postup je pomerne priamočiary:

- Krok 1: Odhadnite regresný model (4.19) pomocou OLS. Výstupom z tohto kroku sú odhadované rezíduá  $\hat{\mathbf{u}}$  a predikované hodnoty  $\hat{\mathbf{y}}$ .
- Krok 2: Vypočítajte testovaciu štatistiku  $LM_{HR}$  (prípadne  $HA_{HR}$ ).
- Krok 3: Vypočítajte chybové členy  $\hat{\mathbf{u}}^*$  pomocou  $\hat{u}_i^* = \hat{u}_i Z_i$ , kde  $Z_i$  je náhodná premenná, ktorá nadobúda dve hodnoty s pravdepodobnosťou  $p$  a  $q$ :

$$Z_i = \begin{cases} -\frac{(\sqrt{5}-1)}{2}, & p = \frac{(\sqrt{5}+1)}{2\sqrt{5}} \\ \frac{(\sqrt{5}+1)}{2}, & q = 1-p \end{cases} \quad (4.24)$$

- Krok 4: Vypočítajte bootstrap závislé premenné  $\mathbf{y}^*$ , kde pri výpočte prvých  $P$  pozorovaní sa používajú pôvodné pozorovania zo vzorky:

$$\mathbf{y}^* = \mathbf{Y}^* \hat{\boldsymbol{\alpha}} + \mathbf{X} \hat{\boldsymbol{\beta}} + \hat{\mathbf{u}}^* \quad (4.25)$$

- Krok 5: Použitím pôvodných exogénnych regresorov  $\mathbf{X}$  a bootstrap údajov  $\mathbf{y}^*$ ,  $\mathbf{Y}^*$  vypočítajte testovaciu štatistiku  $LM_{HR}^*$  ( $HA_{HR}^*$ ).

- Krok 6: Krok 3 – 5 opakujte  $B$  krát. Následne vypočítajte empirickú  $p$ -hodnotu.

$$LM_{HR-pval} = \frac{\left\{ LM_{HR-pval}^*; LM_{HR-pval}^* \geq LM_{HR} \right\}}{B} \quad (4.26)$$

## 5 Úvod do analýzy časových radov

V prechádzajúcej kapitole sme sa zaoberali výsledkami klasickej ekonometrie. Prezentované závery sa veľmi často využívajú v prípade, ak analyzujeme prierezové údaje, pri ktorých skúmame vlastnosti rôznych štatistických jednotiek v rovnakom čase. Naproti tomu veľa ekonomických veličín má zmysel skúmať v podobe časových radov – teda ako vývoj vlastnosti jednej štatistickej jednotky v čase.

Skúmaniu časových radov by sa dala vymedziť samostatná časť ekonometrie. Mnoho problémov, ktoré rieši, sa pri prierezových údajoch nevyskytujú. Teória časových radov predstavuje veľmi bohatú a rozsiahlu časť ekonometrie. V tejto kapitole sa obmedzíme len na priblíženie základných modelov, ktoré sa najčastejšie používajú v praxi. Cieľom teda nie je vytvoriť celkový prehľad o danej problematike, ale poskytnúť prvotné informácie o danej oblasti.

V celej kapitole sa obmedzíme na popis tzv. slabo stacionárnych procesov. Časový rad si označíme podobne, ako tomu bolo v kapitole 3.5.4,  $u_t$ , pre  $t = 1, 2, \dots, n$ , kde  $n \in \mathbb{N}$ . V kapitole 5.1 a 5.2 budeme vychádzať z toho, že časový rad  $u_t$  je realizáciou slabou stacionárneho procesu. Pod slabo stacionárnym procesom budeme na teraz rozumieť, že nepodmienená stredná hodnota  $E(u_t)$  a rozptyl  $\sigma_u^2$  existujú a sú v čase konštantné, a kovariancie  $\text{cov}(u_t, u_{t-p})$  sú nezávislé od  $t$  a závisia len od  $p$ . Podrobnejšie sa definícii budeme venovať v Kapitole 5.3. Ak by sme chceli pri analýze konkrétneho časového radu overiť jeho stacionaritu, je potrebné vykonať niektorý z testov, ktorým sú venované Kapitoly 5.3.1 až 5.3.6.

### 5.1 Autoregresívne procesy a procesy kĺzavých priemerov

V podkapitole 3.5.4 sme definovali autoregresívny proces prvého rádu nasledovne:

$$u_t = \rho u_{t-1} + \varepsilon_t \quad (5.1)$$

Pre konštantu  $\rho$  platilo obmedzenie  $|\rho| < 1$ , náhodné premenné  $\varepsilon_t$  boli navzájom nezávislé, mali rovnaké rozdelenie pravdepodobnosti s nulovou strednou hodnotou a konštantným rozptylom  $\sigma_\varepsilon^2$ , a  $\varepsilon_t$  bolo nezávislé na  $u_{t-1}$ .

Pre tento proces sme si ukázali niekoľko vlastností, hlavne odvodili sme jeho variančno-kovariančnú maticu:

$$\text{var}(\mathbf{u}) = E(\mathbf{u}\mathbf{u}^T) = \mathbf{\Omega} = \frac{\sigma_\varepsilon^2}{1-\rho^2} \begin{pmatrix} 1 & \rho & \rho^2 & \dots & \rho^{n-1} \\ \rho & 1 & \rho & \dots & \rho^{n-2} \\ \rho^2 & \rho & 1 & \dots & \rho^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \dots & 1 \end{pmatrix} \quad (5.2)$$

Na tomto mieste môžeme formuláciu modelu zovšeobecniť tým, že budeme uvažovať o viac než o jednom oneskorení, teda pre  $u_t$  by sme mohli mať vzťah:

$$u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \dots + \rho_p u_{t-p} + \varepsilon_t \quad (5.3)$$

O takto definovanom  $u_t$  hovoríme, že predstavuje autoregresívny proces rádu  $p \in \mathbb{N}$ , kde  $\rho_1, \rho_2, \dots, \rho_p \in \mathbb{R}$ . Rád procesu je určený počtom oneskorení  $p$ . Tento proces označujeme aj ako proces AR( $p$ ), podobne proces definovaný (5.1) je procesom AR(1).

Autoregresívne procesy sa v literatúre často definujú aj využitím tzv. operátora oneskorenia (angl. *lag operator*), ktorý označujeme  $L$  (Hamilton, 1994, s. 25). Ak použijeme tento operátor na hodnotu časového radu, získavame predchádzajúcu hodnotu. Formálne:

$$Lu_t = u_{t-1} \quad (5.4)$$

Ďalej definujeme:

$$L^2 u_t = LLu_t = L(Lu_t) = Lu_{t-1} = u_{t-2} \quad (5.5)$$

Keď sa s operátorom oneskorenia stretávame prvý krát, môže na nás pôsobiť dosť neobvykle. V skutočnosti ale s podobnými operátormi pracujeme bežne a nespôsobuje nám to problémy. Zoberme si jednoduchý príklad: ak napíšeme výraz „3 – 2“, znamienko „–“ je binárnym operátorom predstavujúcim rozdiel dvoch čísel. Nazývame ho binárnym, pretože pre jeho aplikáciu potrebujeme dve čísla, ktorých rozdiel počítame. Nie každý si však uvedomí, že ide o iný operátor, ako keď napíšeme výraz „– 2“. V tomto prípade ide o unárny operátor (na jeho aplikáciu potrebujeme jedno číslo, nie dve ako predtým) a jeho význam nie je rozdielom: „– 2“ slovami znamená „číslo opačné k číslu 2“. Tento príklad ukazuje, že je pomerne bežné stretnúť sa s unárnymi operátormi. Ich argument nemusíme písať do zátvoriek, pokiaľ nechceme zvýrazniť poradie operácií. Tak ako nepíšeme vždy (– 2), podobne v prípade operátora oneskorenia nepíšeme  $L(u_t)$ . Operátoru oneskorenia dávame vyššiu prioritu ako násobeniu a deleniu, teda výraz  $Lu_t + u_{t-1}$  interpretujeme v zmysle výrazu  $L(u_t) + u_{t-1} = u_{t-1} + u_{t-1} = 2u_{t-1}$ , a nie v zmysle  $L(u_t + u_{t-1}) = u_{t-1} + u_{t-2}$ . Inak povedané, pri výraze  $Lu_t + u_{t-1}$  najskôr oneskorujeme ( $L$  má prioritu), a až potom sčítavame. Okrem

operátora oneskorenia môžeme definovať aj zápis pomocou tzv. polynómu oneskorenia (angl. *lag polynomial*). Výrazom  $\rho(L)$  budeme označovať výraz:

$$\rho(L) = \rho_1 L + \rho_2 L^2 + \dots + \rho_p L^p \quad (5.6)$$

Všimnime si, že polynóm oneskorenia nie je reálnym číslom – je ním definovaný operátor  $\rho(L)$ . Rád polynómu  $\rho(L)$  bude zrejmý z kontextu (ak skúmame proces AR( $p$ ), rádom je zrejme  $p$ ) no v prípade, ak by mohlo dôjsť k nedorozumeniu nášho explicitne upozorníme.

S pomocou polynómu oneskorenia môžeme napísať proces AR( $p$ ) ako:

$$u_t = \rho(L)u_t + \varepsilon_t \quad (5.7)$$

$$u_t - \rho(L)u_t = \varepsilon_t \quad (5.8)$$

$$(1 - \rho(L))u_t = \varepsilon_t \quad (5.9)$$

Okrem autoregresívneho procesu je druhým najčastejšie sa vyskytujúcim tzv. proces kľzavých priemerov (angl. *moving average process*). Proces kľzavých priemerov rádu  $q$ , označovaný MA( $q$ ) môžeme napísať v tvare:

$$u_t = \varepsilon_t + \alpha_1 \varepsilon_{t-1} + \alpha_2 \varepsilon_{t-2} + \dots + \alpha_q \varepsilon_{t-q} \quad (5.10)$$

To isté môžeme pomocou operátora oneskorenia definovať ako:

$$u_t = (1 + \alpha(L))\varepsilon_t \quad (5.11)$$

Proces kľzavých priemerov prvého rádu MA(1) je vyjadrený vzťahom:

$$u_t = \varepsilon_t + \alpha_1 \varepsilon_{t-1} \quad (5.12)$$

Keďže predpoklady o  $\varepsilon_t$  nemeníme, je hneď vidieť, že platí:

$$E(u_t) = E(\varepsilon_t + \alpha_1 \varepsilon_{t-1}) = E(\varepsilon_t) + \alpha_1 E(\varepsilon_{t-1}) = 0 \quad (5.13)$$

Pre rozptyl  $D(u_t)$  vďaka tomuto výsledku platí:

$$D(u_t) = E((u_t - E(u_t))^2) = E(u_t^2) \quad (5.14)$$

Preto:

$$D(u_t) = E(u_t^2) = E((\varepsilon_t + \alpha_1 \varepsilon_{t-1})^2) \quad (5.15)$$

$$= E(\varepsilon_t^2 + 2\alpha_1 \varepsilon_t \varepsilon_{t-1} + \alpha_1^2 \varepsilon_{t-1}^2) \quad (5.16)$$

$$= E(\varepsilon_t^2) + 2\alpha_1 E(\varepsilon_t \varepsilon_{t-1}) + \alpha_1^2 E(\varepsilon_{t-1}^2) \quad (5.17)$$

Z nezávislosti  $\varepsilon_t$  a  $\varepsilon_{t-1}$  vyplýva, že kovariancia  $E(\varepsilon_t, \varepsilon_{t-1})$  je nulová. Ďalej výraz  $E(\varepsilon_t^2)$  predstavuje rozptyl  $\varepsilon_t$ , keďže  $E(\varepsilon_t) = 0$ . Tento rozptyl označujeme podľa definície na začiatku

kapitoly  $\sigma_\varepsilon^2$ . Rozptyl  $\varepsilon_t$  je konštantný pre všetky indexy  $t$ , keďže  $\varepsilon_t$  sú podľa predpokladu homoskedastické. Odtiaľ:

$$D(u_t) = E(\varepsilon_t^2) + 2\alpha_1 E(\varepsilon_t \varepsilon_{t-1}) + \alpha_1^2 E(\varepsilon_{t-1}^2) = \sigma_\varepsilon^2 + \alpha_1^2 \sigma_\varepsilon^2 = \sigma_\varepsilon^2 (1 + \alpha_1^2) \quad (5.18)$$

Pre kovarianciu  $\text{cov}(u_t, u_{t-1})$  platí:

$$\text{cov}(u_t, u_{t-1}) = E(u_t u_{t-1}) = E((\varepsilon_t + \alpha_1 \varepsilon_{t-1})(\varepsilon_{t-1} + \alpha_1 \varepsilon_{t-2})) \quad (5.19)$$

$$= E(\varepsilon_t \varepsilon_{t-1} + \alpha_1 \varepsilon_t \varepsilon_{t-2} + \alpha_1 \varepsilon_{t-1}^2 + \alpha_1^2 \varepsilon_{t-1} \varepsilon_{t-2}) \quad (5.20)$$

$$= \alpha_1 E(\varepsilon_{t-1}^2) \quad (5.21)$$

$$= \alpha_1 \sigma_\varepsilon^2 \quad (5.22)$$

Pri úprave z druhého do tretieho riadku sme využili fakt, že z dôvodu nezávislosti  $\varepsilon_t$  pre rôzne hodnoty  $t$  dostávame nulové kovariancie  $E(\varepsilon_t \varepsilon_{t-1}) = E(\varepsilon_t \varepsilon_{t-2}) = E(\varepsilon_{t-1} \varepsilon_{t-2}) = 0$ . Z tohto dôvodu budú aj všetky kovariancie  $u_t$  a  $u_{t-p}$  pre  $p > 1$  nulové: dostávali by sme kovariancie  $E(\varepsilon_t \varepsilon_{t-p}) = E(\varepsilon_t \varepsilon_{t-p-1}) = E(\varepsilon_{t-1} \varepsilon_{t-p}) = E(\varepsilon_{t-1} \varepsilon_{t-p-1}) = 0$ .

Variančno-kovariančná matica pre proces MA(1) je tridiagonálna a vyzerá nasledovne:

$$\text{var}(\mathbf{u}) = E(\mathbf{u}\mathbf{u}^T) = \mathbf{\Omega} = \sigma_\varepsilon^2 \begin{pmatrix} 1 + \alpha_1^2 & \alpha_1 & 0 & \cdots & 0 \\ \alpha_1 & 1 + \alpha_1^2 & \alpha_1 & \cdots & 0 \\ 0 & \alpha_1 & 1 + \alpha_1^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 + \alpha_1^2 \end{pmatrix} \quad (5.23)$$

Z teoretického hľadiska je zaujímavé zaoberať sa aj procesmi, ktoré v praxi nevyužijeme. Ide hlavne o procesy AR( $\infty$ ) a MA( $\infty$ ). Ide o procesy, pri ktorých predpokladáme nekonečný počet oneskorených členov (a preto aj nekonečne veľký počet odhadovaných parametrov).

Uvažujme najprv o procese AR(1), ktorý by mal mať tvar:

$$(1 - \rho_1 L)u_t = \varepsilon_t \quad (5.24)$$

Keď tento vzťah berieme ako rovnicu, mohli by sme sa zaoberať myšlienkou, ako sa z nej dá osamostatniť  $u_t$ . Intuícia napovedá, že by to mohlo byť operáciou:

$$(1 - \rho_1 L)^{-1} (1 - \rho_1 L)u_t = (1 - \rho_1 L)^{-1} \varepsilon_t \quad (5.25)$$

Teda pre násobením rovnice  $(1 - \rho_1 L)^{-1}$ , pretože ak  $(1 - \rho_1 L)^{-1} (1 - \rho_1 L) = 1$ , dostaneme:

$$u_t = (1 - \rho_1 L)^{-1} \varepsilon_t \quad (5.26)$$

Otázkou ale je, čo presne je  $(1 - \rho_1 L)^{-1}$ ?

Aby sme to zistili, skúsme roznásobiť dva výrazy:  $(1 - \rho_1 L)$  a nekonečný rad:

$$1 + \rho_1 L + \rho_1^2 L^2 + \dots = \sum_{i=0}^{\infty} \rho_1^i L^i \quad (5.27)$$

pričom kladieme  $L^0 = 1$ . V súčine dostaneme:

$$(1 - \rho_1 L)(1 + \rho_1 L + \rho_1^2 L^2 + \dots) = (1 + \rho_1 L + \rho_1^2 L^2 + \rho_1^3 L^3 \dots) - (\rho_1 L + \rho_1^2 L^2 + \rho_1^3 L^3 \dots) \quad (5.28)$$

$$= 1 \quad (5.29)$$

Z toho ale vyplýva, že ak  $(1 - \rho_1 L)^{-1}(1 - \rho_1 L) = 1$ , a  $(1 + \rho_1 L + \rho_1^2 L^2 + \dots)(1 - \rho_1 L) = 1$ , potom môžeme položiť:

$$(1 - \rho_1 L)^{-1} = 1 + \rho_1 L + \rho_1^2 L^2 + \dots = \sum_{i=0}^{\infty} \rho_1^i L^i \quad (5.30)$$

Z uvedeného vyplýva, že proces AR(1) v tvare:

$$(1 - \rho_1 L)u_t = \varepsilon_t \quad (5.31)$$

Môžeme napísať ako proces MA( $\infty$ ):

$$u_t = (1 - \rho_1 L)^{-1} \varepsilon_t = \left( \sum_{i=0}^{\infty} \rho_1^i L^i \right) \varepsilon_t \quad (5.32)$$

Platí to aj naopak – proces MA(1) vieme napísať aj ako proces AR( $\infty$ ). Vzťah dokonca platí aj pre procesy AR( $p$ ) a MA( $q$ ) – stacionárny autoregresívny proces vieme vyjadriť ako proces kĺzavých priemerov nekonečného rádu, a stacionárny proces kĺzavých priemerov vieme vyjadriť ako autoregresívny proces nekonečného rádu.

Tvrdenie však neplatí o ľubovoľných dvoch procesoch – kľúčová podmienka je podmienka stacionarity. Procesy, pri ktorých vieme nájsť koeficienty polynómov oneskorenia splňajúce:

$$(1 - \rho(L))(1 + \alpha(L)) = 1 \quad (5.33)$$

nazývame invertovateľné (alebo invertibilné). Dá sa ukázať, že podmienkou pre invertovateľnosť procesu AR( $p$ ) je ekvivalentná podmienka jeho stacionarity. Ak by sme použili koeficienty polynómu oneskorenia a hľadali by sme korene polynómu:

$$1 - \rho(z) = 0 \quad (5.34)$$

pre  $z$  patriace do množiny komplexných čísel, podmienka stacionarity sa dá vyjadriť aj ako tvrdenie, že všetky korene (5.34) sú v absolútnej hodnote väčšie ako 1. Podľa základnej vety algebry bude mať takáto polynomická rovnica rádu  $p$  práve  $p$  riešení, pričom niektoré môžu byť viacnásobné, a niektoré môžu byť komplexné. Táto podmienka sa niekedy

formuluje aj tak, že korene rovnice musia ležať mimo jednotkovej kružnice (v obore komplexných čísel).

V prípade procesu MA( $q$ ) je podmienkou na invertovateľnosť tvrdenie, že všetky korene rovnice:

$$1 - \alpha(z) = 0 \quad (5.35)$$

Musia byť v absolútnej hodnote väčšie ako 1.

Vo vzťahu k invertovateľnosti si všimnime dve skutočnosti. Prvou je fakt, že proces MA(1):

$$u_t = \varepsilon_t + \alpha_1 \varepsilon_{t-1} \quad (5.36)$$

je stacionárny pre všetky  $\alpha_1 \in \mathbb{R}$ . Toto je zrejme aj z variančno-kovariančnej matice  $\mathbf{u}$  pre proces MA(1):

$$\text{var}(\mathbf{u}) = E(\mathbf{u}\mathbf{u}^T) = \mathbf{\Omega} = \sigma_\varepsilon^2 \begin{pmatrix} 1 + \alpha_1^2 & \alpha_1 & 0 & \dots & 0 \\ \alpha_1 & 1 + \alpha_1^2 & \alpha_1 & \dots & 0 \\ 0 & \alpha_1 & 1 + \alpha_1^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 + \alpha_1^2 \end{pmatrix} \quad (5.37)$$

Pre ľubovoľné  $\alpha_1 \in \mathbb{R}$  obsahuje táto matica konečné hodnoty. Avšak podmienka invertovateľnosti pre proces MA(1) hovorí, že:

$$1 - \alpha_1 z = 0 \quad (5.38)$$

Jediným koreňom je:

$$z = \frac{1}{\alpha_1} \quad (5.39)$$

Ak má byť tento koreň v absolútnej hodnote väčší ako 1, znamená to, že:

$$|z| = \frac{1}{|\alpha_1|} > 1 \quad (5.40)$$

$$|\alpha_1| < 1$$

Pre proces AR(1) sme dostali variančno-kovariančnú maticu:

$$\text{var}(\mathbf{u}) = E(\mathbf{u}\mathbf{u}^T) = \mathbf{\Omega} = \frac{\sigma_\varepsilon^2}{1 - \rho^2} \begin{pmatrix} 1 & \rho & \rho^2 & \dots & \rho^{n-1} \\ \rho & 1 & \rho & \dots & \rho^{n-2} \\ \rho^2 & \rho & 1 & \dots & \rho^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \dots & 1 \end{pmatrix} \quad (5.41)$$

Rozptyly prvkov  $\mathbf{u}$  sme odhadovali v rovnici (3.242) pomocou súčtu nekonečného radu.

$$D(u_t) = \sigma_\varepsilon^2 \sum_{i=0}^{\infty} \rho^{2i} = \frac{\sigma_\varepsilon^2}{1-\rho^2} \quad (5.42)$$

Tento rad má konečný súčet (a proces má konečný rozptyl) len v prípade, ak  $|\rho| < 1$ . Podmienka invertibility pre proces AR(1) je, aby koreň rovnice:

$$1 - \rho z = 0 \quad (5.43)$$

bol v absolútnej hodnote väčší ako nula, čo je ekvivalentné tvrdeniu:

$$|\rho| < 1 \quad (5.44)$$

Pre proces AR(1) je preto podmienka invertibility totožná s podmienkou stacionarity.

Pri pohľade na variančno-kovariančnú maticu procesov MA(1) a AR(1) vidíme hneď niekoľko skutočností. V prvom rade variančno-kovariančná matica vektora  $\mathbf{u}$  pre proces AR(1) obsahoval samé nenulové členy (ak  $\rho \neq 0$ ). V prípade procesu MA(1) sú nenulové kovariancie len pre susedné pozorovania (s oneskorením o jedno pozorovanie). Táto skutočnosť pre nás môže byť indikátorom, o aký proces ide – ak by sme videli variančno-kovariančnú maticu, je na základe charakteru tejto matice možné odhadnúť, či ide skôr o proces MA(1) alebo AR(1).

V prípade oboch procesov sú zložky  $\mathbf{u}$  homoskedastické – takže majú rovnaké rozptyly. Nižšie kovariancie mimo hlavnej diagonály potom v rámci tej istej matice sú potom ekvivalentné nižším koreláciám. Vo variančno-kovariančnej matici pre proces AR(1) vidíme, že korelácie klesajú s oneskorením – čím väčšia je vzdialenosť v čase (oneskorenie), tým menšia je korelácia. Pri procese MA(1) sú korelované len susedné hodnoty. Preto ak by sme hľadali vhodnú špecifikáciu pre časový rad, v ktorom sú korelácie aj medzi vzdialenejšími hodnotami, zrejme by sme nemali vybrať proces MA(1).

Z variančno-kovariančnej matice pre proces MA(1) ďalej vidíme, že nenulové korelácie medzi prvkami  $\mathbf{u}$  sú rovné:

$$\frac{\alpha_1}{1 + \alpha_1^2} \quad (5.45)$$

Z podmienky invertibility však máme, že  $|\alpha_1| < 1$ . Z toho vyplýva, že maximálna hodnota korelácie pri procese MA(1) musí byť menšia ako 0.5. Ak modelujeme proces, v ktorom je korelácia vyššia, zrejme nie je vhodné vybrať preň model MA(1).

## 5.2 Modely ARMA, ARCH a GARCH

V predchádzajúcej kapitole sme si predstavili základné procesy, o ktoré sa opiera teória modelovania časových radov. Definovali sme model  $AR(p)$ :

$$(1 - \rho(L))u_t = \varepsilon_t \quad (5.46)$$

ako aj model  $MA(q)$ :

$$u_t = (1 + \alpha(L))\varepsilon_t \quad (5.47)$$

Kombináciou uvedených dvoch modelov je možné vytvoriť proces, ktorý nazývame  $ARMA(p,q)$  (angl. *autoregressive moving average*):

$$(1 - \rho(L))u_t = (1 + \alpha(L))\varepsilon_t \quad (5.48)$$

Ide o model, ktorý obsahuje oneskorené hodnoty závislej premennej, ako aj oneskorené hodnoty inovácií. V prípade, ak sú splnené podmienky pre polynómy oneskorenia  $\rho(L)$  a  $\alpha(L)$  definované v predchádzajúcej časti, je aj proces  $ARMA(p,q)$  invertovateľný, a je ho možné vyjadriť ako proces  $AR(\infty)$  a  $MA(\infty)$ .

Model  $ARMA(p,q)$  môže sám osebe byť využívaný na modelovanie stacionárnych časových radov. Pri modelovaní je ale možné využiť aj exogénne regresory, ako sme to robili v klasickej ekonometrii. Model v tvare:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \quad (5.49)$$

kde zložky vektora  $\mathbf{u}$  predstavujú procesy  $ARMA(p,q)$ , pričom  $E(\mathbf{u}) = \mathbf{0}$  nazývame modelom  $ARMAX(p,q)$ .

Všimnime si, že ide o veľmi podobný model, ako sme rozoberali v kapitolách venovaných klasickej ekonometrii. Korelovanosť zložiek vektora náhodných porúch bola vtedy porušením predpokladu P2 – v prípade modelu  $ARMAX$  s autokoreláciou počítame pri špecifikácii modelu.

Odhad koeficientov modelu  $ARMA$ , ale aj  $ARMAX$  je možné realizovať pomocou odhadu metódou maximálnej vierohodnosti. Túto metódu sme si ukázali v Kapitole 3.3.4. Ukážme si preto len logaritmus funkcie vierohodnosti pre model  $ARMAX$  a inovácie s normálnym rozdelením pravdepodobnosti, ktorý je všeobecnejší (Davidson – MacKinnon, 2003, s. 558):

$$-\frac{n}{2}\ln(2\pi) - \frac{1}{2}\ln|\boldsymbol{\Omega}| - \frac{1}{2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T \boldsymbol{\Omega}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \quad (5.50)$$

Pripomeňme, že pre proces  $AR(1)$  sme mali variančno-kovariančnú maticu:

$$\mathbf{\Omega} = \frac{\sigma_\varepsilon^2}{1-\rho^2} \begin{pmatrix} 1 & \rho & \rho^2 & \dots & \rho^{n-1} \\ \rho & 1 & \rho & \dots & \rho^{n-2} \\ \rho^2 & \rho & 1 & \dots & \rho^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \dots & 1 \end{pmatrix} \quad (5.51)$$

Vidíme, že je vyjadrená pomocou neznámeho rozptylu  $\sigma_\varepsilon^2$  a korelácie  $\rho$ . Pre proces MA(1) obsahovali prvky matice  $\mathbf{\Omega}$  neznámy parameter  $\alpha_1$ . Variančno-kovariančná matica  $\mathbf{\Omega}$  bude preto aj v prípade zložitejšieho modelu ARMA( $p, q$ ), resp. ARMAX( $p, q$ ) obsahovať  $p$  autoregresných koeficientov a  $q$  koeficientov kľzavých priemerov. Okrem nich je potrebné odhadnúť ešte neznáme parametre  $\boldsymbol{\beta}$  a rozptyl  $\sigma_\varepsilon^2$ . Zámer je preto jednoduchý – odhadnúť všetky tieto parametre tak, aby sa maximalizovala funkcia vierohodnosti. Toto je optimalizačná úloha, riešenie ktorej umožňuje odhad modelu.

Ako sme videli v predchádzajúcich častiach, matica  $\mathbf{\Omega}$  môže byť pomerne komplikovaná – ukazovali sme si len prípady AR(1) a MA(1), ktoré sú najľahšie. Naším zámerom bolo vysvetliť princíp, na základe ktorého je možné hľadaný odhad získať.

Doteraz spomínané modely sa zaoberali problémom autokorelácie – všetky skúmané tvary variančno-kovariančných matíc viedli k tomu, že všetky rozptyly zložiek vektora  $\mathbf{u}$  boli rovnaké, teda zložky  $\mathbf{u}$  boli homoskedastické. Vystáva otázka, či by nebolo možné podobne, ako modely ARMA využívajú porušenie predpokladu P2 na rozšírenie špecifikácie modelu aj v prípade heteroskedasticity modelovať meniaci sa rozptyl. Vhodný model s akronymom ARCH navrhol v roku Engle (1982).

Nech  $\varepsilon_t$  pre  $t = 1, 2, \dots, n$  sú náhodné premenné s normovaným normálnym rozdelením pravdepodobnosti,  $q \in \mathbb{N}$  a  $\alpha_i \in \mathbb{R}^+$  pre  $i = 0, 1, \dots, q$ . Potom ARCH proces rádu  $q$ , označovaný aj ARCH( $q$ ) je daný nasledovne:

$$u_t = \sigma_t \varepsilon_t \quad (5.52)$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i u_{t-i}^2 \quad (5.53)$$

Dá sa ukázať (Hamilton, 1994, s. 659), že nepodmienенý rozptyl pre proces ARCH( $q$ ) je daný vzťahom:

$$\sigma^2 = \frac{\alpha_0}{1 - \sum_{i=1}^q \alpha_i} \quad (5.54)$$

Podmienkou stacionarity je potom:

$$\sum_{i=1}^q \alpha_i < 1 \quad (5.55)$$

Ak by podmienka neplatila, zároveň by bol menovateľ v predchádzajúcom vzťahu záporný – čo nedáva zmysel, keďže modelujeme nezáporný rozptyl. Z rovnakého dôvodu sa požaduje aj  $\alpha_i \in \mathbb{R}^+$ .

Zovšeobecnením modelov ARCH sú modely GARCH, ktoré navrhol Bollerslev (Bollerslev, 1986). Ich hlavnou výhodou je, že pri nich na modelovanie toho istého časového radu postačuje odhadovať spravidla oveľa menej parametrov, ako pri modeloch ARCH. Model GARCH( $p, q$ ) má tvar:

$$u_t = \sigma_t \varepsilon_t \quad (5.56)$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i u_{t-1}^2 + \sum_{j=1}^p \delta_j \sigma_{t-j}^2 \quad (5.57)$$

V tomto prípade pre nepodmienený rozptyl  $u_t$  platí:

$$\sigma^2 = \frac{\alpha_0}{1 - \sum_{i=1}^q \alpha_i - \sum_{j=1}^p \delta_j} \quad (5.58)$$

Podmienkou stacionarity je (Hamilton, 1994):

$$\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \delta_j < 1 \quad (5.59)$$

Označme  $\mathbf{\alpha}$  vektor parametrov  $\mathbf{\alpha} = (\alpha_0, \alpha_1, \dots, \alpha_q)$  a  $\mathbf{\delta}$  vektor parametrov  $\mathbf{\delta} = (\delta_1, \delta_2, \dots, \delta_p)$ . Podmienený rozptyl  $\sigma_t^2$  je potom možné abstraktne vyjadriť ako funkciu odhadovaných parametrov obsiahnutých v  $\mathbf{\alpha}$  a  $\mathbf{\delta}$ .

Odhad modelov ARCH a GARCH sa realizuje nasledovným spôsobom. Vychádzajme z modelu:

$$y_t = \mathbf{X}_t \boldsymbol{\beta} + u_t \quad (5.60)$$

$$y_t = \mathbf{X}_t \boldsymbol{\beta} + \sigma_t(\boldsymbol{\beta}, \mathbf{\alpha}, \mathbf{\delta}) \varepsilon_t \quad (5.61)$$

Hodnota závislej premennej  $y_t$  je modelovaná pomocou exogénnych premenných pre pozorovanie  $t$  obsiahnutých v  $\mathbf{X}_t$ .  $\mathbf{X}_t$  je  $t$ -ty riadok matice  $\mathbf{X}$ . Špecifikácia poruchového člena korešponduje s modelom GARCH. Člen  $\sigma_t(\boldsymbol{\beta}, \mathbf{\alpha}, \mathbf{\delta})$  vyjadruje, že podmienená štandardná odchýlka je funkciou odhadov vektorov  $\mathbf{\alpha}$ ,  $\boldsymbol{\beta}$  a  $\mathbf{\delta}$ . Poslednú rovnicu môžeme napísať v tvare:

$$\varepsilon_t = \frac{y_t - \mathbf{X}_t \boldsymbol{\beta}}{\sigma_t(\boldsymbol{\beta}, \mathbf{\alpha}, \mathbf{\delta})} \quad (5.62)$$

Keďže predpokladáme, že  $\varepsilon_t$  má normované normálne rozdelenie, môžeme k nemu nadefinovať logaritmus funkcie vierohodnosti založenej na hustote normálneho rozdelenia v tvare (Davidson – MacKinnon, 2003, s. 583):

$$-\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln(\sigma_t^2(\boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{\delta})) - \frac{1}{2} \frac{(y_t - \mathbf{X}_t\boldsymbol{\beta})^2}{\sigma_t(\boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{\delta})} \quad (5.63)$$

Odhad modelu potom môžeme realizovať metódou maximálnej vierohodnosti.

### 5.3 Stacionarita časových radov

Stacionarita je jednou z kľúčových vlastností časových radov. V niektorých prípadoch platí (ide o prípady, ktoré nás v ekonómii zaujímajú pomerne často), že ak je časový rad nestacionárny, potom efekt šoku sa s pribúdajúcim časom nestráca. Pokiaľ je časový rad stacionárny, tak akýkoľvek šok v čase  $t$  sa postupne vytráca, až úplne zmizne (pre  $t \rightarrow \infty$ ). Inými slovami môžeme povedať, že časový rad sa vracia k svojmu dlhodobému priemeru (v angl. *mean reversion*).

Vychádzajme z autoregresného procesu v nasledovnom tvare:

$$y_t = \rho y_{t-1} + \gamma D_t + e_t, \text{ iid } e_t \sim N(0,1) \quad (5.64)$$

Sústredme sa zatiaľ iba na prvú časť tohto vzťahu a uvažujme o špecifickom prípade,  $\rho = 1, \gamma = 0$ . Je zrejmé, že všetky budúce hodnoty budú v sebe zahŕňať aj hodnoty minulé. Ak  $\rho = 1$ , ide o nestacionárny časový rad. Minulé efekty sa pritom budú kumulatívne spočítavať:

$$y_t = y_{t-1} + e_t \quad (5.65)$$

$$y_t = y_{t-2} + e_{t-1} + e_t \quad (5.66)$$

$$y_t = y_{t-3} + e_{t-2} + e_{t-1} + e_t \quad (5.67)$$

$$y_t = y_{t=0} + \sum_{n=1}^t e_n \quad (5.68)$$

Na nasledujúcom obrázku (pozri Obrázok 19) sme znázornili niekoľko situácií. V každom časovom rade sme predpokladali, že k okamihu  $T/2$ , dôjde k šoku o veľkosti  $\gamma = 20$ , t. j.  $D_t = 1$  pre  $t = T/2$ . Pričom sme uvažovali o šiestich rôznych hodnotách  $\rho = 0.00, 0.50, 0.90, 0.95, 0.99, 1.00$ . Zostrojili sme si jednoduchú funkciu, ktorá takéto časové rady simuluje.

```
> arlsim <- function(TT, rho = c(rep(0, length(TT))), shock =
  c(rep(0, length(TT))), size = c(rep(0, length(TT))), errors =
  c("same", "different")) {
+   if (min(TT) < 3) stop("TT must be > 3")
```

```

+   if (max(abs(rho)) > 1) warning("abs(rho) should be <= 1")
+   results <- list()
+   if (errors == "same") es <- rnorm(max(TT))
+   for (n in 1:length(TT)) {
+     pos <- round(shock[n]*TT[n]);
+     if (errors == "same") {
+       et <- es
+       et[pos] <- et[pos] + size[n];
+     }
+     if (errors == "different") {
+       et <- rnorm(TT[n]); et[pos] <- et[pos] + size[n];
+     }
+     srs <- et[1]; srs[2] <- rho[n]*srs[1] + et[2];
+     for (t in 3:TT[n]) srs[t] <- rho[n]*srs[t-1] + et[t];
+     results[[n]] <- srs
+   }
+   return(results)
+ }
> dev.off()
> series <- arlsim(rep(200,6), rho = c(0.00, 0.50, 0.90, 0.95,
  0.99, 1.00), shock = rep(0.5, 6), size = rep(20, 6), errors =
  "same")
> rgb <- hcl(c(0), c = c(0), l = c(50), alpha = 0.25)
> par(mfrow = c(2, 3))
> par(cex = 0.6)
> par(mar = c(1.8, 0.2, 0.2, 0.2), oma = c(2.8, 2.0, 0.5, 0.5))
> par(tcl = -0.25)
> par(mgp = c(2, 0.6, 0))
> miny <- min(as.numeric(lapply(series, min)))
> maxy <- max(as.numeric(lapply(series, max)))
> plot(series[[1]], ylim = c(miny, maxy), main = NA, cex.axis =
  1.2, yaxt = "n", type = "l", lty = 1, lwd = 1.5)
> legend(x = 0, y = 25, legend = "rho = 0.00", lty = 1, bty =
  "n", cex = 1.5)
> axis(side = 2, cex.axis = 1.25)
> box(col = "grey60")
> plot(series[[2]], ylim = c(miny, maxy), main = NA, cex.axis =
  1.2, yaxt = "n", type = "l", lty = 1, lwd = 1.5)
> legend(x = 0, y = 25, legend = "rho = 0.50", lty = 1, bty =
  "n", cex = 1.5)
> box(col = "grey60")
> plot(series[[3]], ylim = c(miny, maxy), main = NA, cex.axis =
  1.2, yaxt = "n", type = "l", lty = 1, lwd = 1.5)
> legend(x = 0, y = 25, legend = "rho = 0.90", lty = 1, bty =
  "n", cex = 1.5)
> box(col = "grey60")
> plot(series[[4]], ylim = c(miny, maxy), main = NA, cex.axis =
  1.2, yaxt = "n", type = "l", lty = 1, lwd = 1.5)
> legend(x = 0, y = 25, legend = "rho = 0.95", lty = 1, bty =
  "n", cex = 1.5)
> axis(side = 2, cex.axis = 1.25)
> box(col = "grey60")
> plot(series[[5]], ylim = c(miny, maxy), main = NA, cex.axis =
  1.2, yaxt = "n", type = "l", lty = 1, lwd = 1.5)

```

```

> legend(x = 0, y = 25, legend = "rho = 0.99", lty = 1, bty =
  "n", cex = 1.5)
> box(col = "grey60")
> plot(series[[6]], ylim = c(miny, maxy), main = NA, cex.axis =
  1.2, yaxt = "n", type = "l", lty = 1, lwd = 1.5)
> legend(x = 0, y = 25, legend = "rho = 1.00", lty = 1, bty =
  "n", cex = 1.5)
> box(col = "grey60")
> series <- arlsim(rep(200,6), rho = c(0.00, 0.50, 0.90, 0.95,
  0.99, 1.00), shock = rep(0.5, 6), size = rep(20, 6), errors =
  "same")
> rgb <- hcl(c(0), c = c(0), l = c(50), alpha = 0.25)
> par(mfrow = c(2, 3))
> par(cex = 0.6)
> par(mar = c(1.8, 0.2, 0.2, 0.2), oma = c(2.8, 2.0, 0.5, 0.5))
> par(tcl = -0.25)
> par(mgp = c(2, 0.6, 0))
> miny <- min(as.numeric(lapply(series, min)))
> maxy <- max(as.numeric(lapply(series, max)))
> plot(series[[1]], ylim = c(miny, maxy), main = NA, cex.axis =
  1.2, yaxt = "n", type = "l", lty = 1, lwd = 1.5)
> legend(x = 0, y = 25, legend = "rho = 0.00", lty = 1, bty =
  "n", cex = 1.5)
> axis(side = 2, cex.axis = 1.25)
> box(col = "grey60")
> plot(series[[2]], ylim = c(miny, maxy), main = NA, cex.axis =
  1.2, yaxt = "n", type = "l", lty = 1, lwd = 1.5)
> legend(x = 0, y = 25, legend = "rho = 0.50", lty = 1, bty =
  "n", cex = 1.5)
> box(col = "grey60")
> plot(series[[3]], ylim = c(miny, maxy), main = NA, cex.axis =
  1.2, yaxt = "n", type = "l", lty = 1, lwd = 1.5)
> legend(x = 0, y = 25, legend = "rho = 0.90", lty = 1, bty =
  "n", cex = 1.5)
> box(col = "grey60")
> plot(series[[4]], ylim = c(miny, maxy), main = NA, cex.axis =
  1.2, yaxt = "n", type = "l", lty = 1, lwd = 1.5)
> legend(x = 0, y = 25, legend = "rho = 0.95", lty = 1, bty =
  "n", cex = 1.5)
> axis(side = 2, cex.axis = 1.25)
> box(col = "grey60")
> plot(series[[5]], ylim = c(miny, maxy), main = NA, cex.axis =
  1.2, yaxt = "n", type = "l", lty = 1, lwd = 1.5)
> legend(x = 0, y = 25, legend = "rho = 0.99", lty = 1, bty =
  "n", cex = 1.5)
> box(col = "grey60")
> plot(series[[6]], ylim = c(miny, maxy), main = NA, cex.axis =
  1.2, yaxt = "n", type = "l", lty = 1, lwd = 1.5)
> legend(x = 0, y = 25, legend = "rho = 1.00", lty = 1, bty =
  "n", cex = 1.5)
> box(col = "grey60")

```

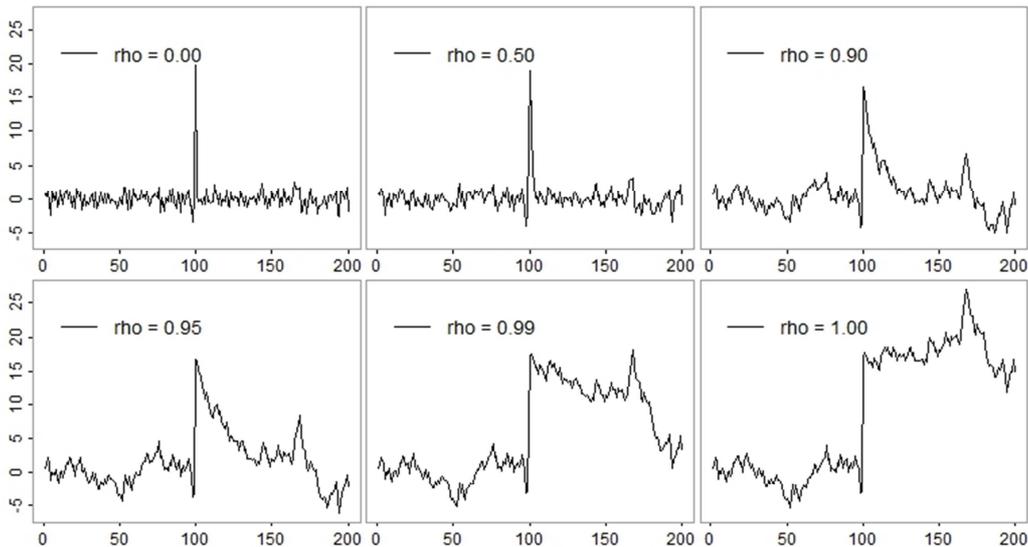
Z obrázku je pomerne jednoznačne vidno, ako sa dôsledky šoku s  $\rho$  blížiacemu sa k nule znižujú. Čím je  $\rho$  bližšie k 1 (zľava) o to perzistentnejší je proces a o to viac sa podobá na proces nestacionárny. V prípade, ak je  $\rho \geq 1$ , ide o nestacionárny časový rad, o tzv. explozívny proces.

Existuje niekoľko dôvodov, prečo ekonómov zaujíma práve existencia takýchto procesov.<sup>18</sup> Až do práce Nelson – Plosser (1982) mnoho ekonómov zastávalo názor, že ekonomické premenné ako je HDP alebo HNP oscilujú v okolí deterministického trendu. Inak povedané, existuje určitý dlhodobý deterministický trend a odchýlky od tohto trendu sú len dočasné. Nelson – Plosser (1982) však ukázali, že tomu tak nemusí byť, a že je celkom možné, že mnohé makroekonomické premenné majú skôr charakter nestacionárneho procesu v tvare podobnom ako v (5.68). Takýto výsledok mal pomerne závažné dôsledky. Ak by bol totiž celkový výstup z ekonomiky stacionárny, potom: 1) neistota ohľadom budúceho vývoja HDP a HNP je obmedzená, keďže môžeme poznať dlhodobý deterministický trend, 2) fluktuáciu okolo tohto trendu je možné interpretovať ako krátkodobé ekonomické cykly, 3) tieto fluktuácie je možné ovplyvňovať pomocou monetárnej a fiškálnej politiky a v neposlednom rade, 4) efekt intervencií nakoniec zoslabne, tak ako je to znázornené na nasledujúcom obrázku (pozri Obrázok 19). Na strane druhej, ak je celkový výstup z ekonomiky nestacionárny (v zmysle procesu definovanom vyššie (5.68)), potom: 1) neistota ohľadom budúceho vývoja v čase  $t + r$  rastie spolu s  $r$ , 2) reálne šoky do ekonomiky majú trvalý charakter, 3) celkový výstup z ekonomiky nemá tendenciu sa vrátiť k pôvodným hodnotám (resp. k pôvodnému deterministickému trendu).

Ďalším dôvodom prečo nás zaujíma, či sú časové rady stacionárne (presnejšie, či ich je možné charakterizovať prítomnosťou jednotkového koreňa, ktorý definujeme neskôr v tejto kapitole) je, že ak máme dva alebo viac nestacionárnych časových radov, môže byť ich lineárna kombinácia stacionárna. To by však znamenalo, že tieto časové rady sú v určitej dlhobodej rovnováhe. Hovoríme tomu aj kointegrácia. Tá však nie je predmetom tejto publikácie.

---

<sup>18</sup> V tejto časti voľne preberáme od Lyócsa et al. (2011).



Obrázok 19: Perzistentnosť šoku

Zdroj: výstup zo softvéru R

Doteraz sme pojem (ne)stacionarita používali pomerne vágne. Išlo nám najmä o intuitívne priblíženie tohto pojmu. Teraz zavedieme niektoré definície a zároveň bližšie špecifikujeme, čo budeme v empirickej práci s údajmi chápať pod pojmom stacionarita časového radu. Pri spracovaní tejto časti sme čerpali najmä z Burke – Hunter (2005), Poo (2003) a Hamilton (1994).<sup>19</sup> Keďže ide štandardnú teóriu, zdroje sme uviedli súhrne. Konkrétne zdroje uvádzame len na tom mieste, kde chceme zvýrazniť zdroj, z ktorého sme čerpali.

Stochastickým procesom budeme rozumieť usporiadanú postupnosť náhodných premenných  $Y_t$  v čase,  $\{Y_t\}_{-\infty}^{\infty}$ . Jeho realizácie sú potom  $\{y_1, y_2, \dots, y_T\} = \{y_t\}_1^T$ . Budeme ďalej predpokladať, že stochastický proces  $\{Y_t\}_{-\infty}^{\infty}$  je kovariančne stacionárny. To znamená, že pre všetky  $t$  a  $j$  platí:

$$E[Y_t] = \mu, |\mu| < \infty \quad (5.69)$$

$$\text{cov}(Y_t, Y_{t-j}) = E[(Y_t - \mu)(Y_{t-j} - \mu)] = \gamma_{YY}(j), |\gamma_{YY}(j)| < \infty \quad (5.70)$$

Ide o obmedzenia, ktoré sa týkajú prvého a druhého momentu. Prvé obmedzenie nedovolí, aby sa menila stredná hodnota v čase. Druhé obmedzenie tvrdí, že autokovariancia

<sup>19</sup> Kapitola 5.3 do značnej miery vychádza z interných materiálov, ktoré sa použili pri spracovaní metodológie ku kvalifikačnej práci Lyócsa (2013), ktorá však bola publikovaná neskôr ako tieto skriptá. Výnimku tvoria najmä príklady.

závisí len od oneskorenia  $j$  a nie od časového okamihu  $t$ . Pod druhou podmienkou sa skrýva aj podmienka, že nepodmienený rozptyl je konštantný (pre  $j = 0$ ). Na označenie kovariančne stacionárneho procesu sa zvykne používať aj výraz, *slabo stacionárny stochastický proces*. V literatúre sa môžeme stretnúť s rôznymi formami a definíciami stacionarity. Spravidla platí, že čím viac reštrikcií vytvoríme, o to prísnejšia (a pre empirický výskum o to menej praktická) je podmienka stacionarity stochastického procesu. Ďalej potrebujeme, aby pre kovariančne stacionárny proces platilo, že pre  $k \rightarrow \infty$ ,  $Y_t$  a  $Y_{t+k}$  boli nekorelované.

V praxi máme pre jeden stochastický proces  $\{Y_t\}_{-\infty}^{\infty}$  spravidla k dispozícii iba jednu realizáciu  $\{y_t\}_1^T$ . Ak chceme odhadnúť strednú hodnotu tohto procesu, na odhad môžeme použiť aritmetický priemer z nameraných hodnôt. Na mieste je otázka, či pomocou tohto postupu bude aritmetický priemer konvergovať ku skutočnej strednej hodnote stochastického procesu. Pri práci s časovými radmi budeme predpokladať, že tomu tak je. Tento problém sa týka odhadu všetkých momentov stochastického procesu.<sup>20</sup> Túto vlastnosť nazveme ergodicita. Zjednodušene si pod ergodicitou pritom môžeme predstaviť takú vlastnosť procesu, pri ktorej platí, že ak vyberieme určitú vzorku hodnôt z tohto procesu, potom bude výberový priemer na základe tejto vzorky skoro isto konvergovať k skutočnej strednej hodnote procesu a rovnako kovariancia k skutočnej kovariancii (Wang, 2009).

V úvode tejto kapitoly sme spomenuli, že nás zaujíma jeden špecifický prípad nestacionarity. Pritom podľa definície slabo stacionárneho procesu, je možné nájsť rôzne formy nestacionarity. Z definície vyššie vyplýva, že časový rad je nestacionárny, ak má deterministický trend:

$$y_t = \beta_0 + \beta_1 t + e_t, e_t \sim N(0,1), \beta_1 \neq 0 \quad (5.71)$$

Iným prípadom môže byť, ak sa mení rozptyl časového radu. Napríklad uvažujme o nasledujúcom empirickom procese:

$$y_t = \begin{cases} e_t \sim N(0,1), & t < T/2 \\ e_t \sim N(0,4), & t \geq T/2 \end{cases} \quad (5.72)$$

Stredná hodnota sa síce nemení, ale mení sa nepodmienený rozptyl. Preto podľa definície slabo stacionárneho procesu, by sme tento proces nemohli považovať za slabo stacionárny proces. Prípadne, ak sa v časovom rade vyskytujú štrukturálne zlomy v strednej hodnote:

---

<sup>20</sup> Táto podmienka nám do istej miery nahrádza podmienku náhodného výberu z populácie pri prierezovej štatistike.

$$y_t = \gamma D_t + e_t, e_t \sim N(0,1), \gamma \neq 0, k \in T : t < k \Rightarrow D_t = 0, t \geq k \Rightarrow D_t = 1 \quad (5.73)$$

Ako si ukážeme v Kapitole 5.4, analýza nestacionárnych časových radov vedie k netriviálnym problémom. Pokiaľ je však nestacionarita vyvolaná deterministickými zložkami, spravidla to pre modelovanie takéhoto časového radu nepredstavuje zásadný problém. Samozrejme v prípade, ak sú nám tieto deterministické zložky známe. Pri regresnej analýze môžeme tieto deterministické zložky do modelu zahrnúť vo forme vhodných regresorov, prípadne môžeme časový rad najprv upraviť tak, aby sa v ňom tieto deterministické zložky nevyskytovali. V ekonómii nás však zaujíma špecifický typ nestacionarity, tzv. stochastický trend. Uvažujme o nasledujúcom modeli s deterministickým trendom:

$$y_t = \beta_0 + \beta_1 t \quad (5.74)$$

kde  $\beta_0$  a  $\beta_1$  sú parametre modelu. V takom procese sú prírastky konštantné:

$$\Delta y_t = y_t - y_{t-1} = \beta_0 + \beta_1 t - \beta_0 - \beta_1 (t-1) = \beta_1 \quad (5.75)$$

Ak tieto prírastky ( $\beta_1$ ) nebudú konštantné, ale budú nezávislými realizáciami z normálneho rozdelenia pravdepodobnosti, povedzme  $\varepsilon_t \sim N(0, \sigma^2)$ , pôjde o určitú formu stochastického trendu. Takýto model môžeme zapísať ako:

$$\Delta y_t = \varepsilon_t \quad (5.76)$$

Toto je však špeciálny prípad, ktorý sme už riešili vyššie:

$$\Delta y_t = y_t - y_{t-1} = \varepsilon_t \quad (5.77)$$

$$y_t = y_{t-1} + \varepsilon_t = \rho y_{t-1} + \varepsilon_t, \rho = 1 \quad (5.78)$$

Keďže časové rady, ktoré generujú tento proces (nazývaný tiež náhodná prechádzka, tzv. random walk), sa podobajú na časové rady pozorované v ekonómii, v empirickom výskume sa venujeme odhaľovaniu práve tejto formy nestacionarity časového radu pomerne veľa priestoru. Pre zjednodušenie uvažujme o autoregresnom procese prvého rádu (AR(1)) tak, ako bol definovaný vyššie:

$$y_t = y_{t-1} + \varepsilon_t, \varepsilon_t \sim N(\mu, \sigma^2) \quad (5.79)$$

$$E[y_t] = 0 \quad (5.80)$$

$$D[y_t] = t\sigma^2 \quad (5.81)$$

$$\text{cov}(y_t, y_{t-j}) = (t-j)\sigma^2 \quad (5.82)$$

Z výrazu (5.81) a (5.82) vyplýva, že tento proces nie je kovariančne stacionárny. Rozptýl aj kovariancia závisia aj od polohy v čase  $t$ . Pomocou spätnej substitúcie môžeme

tento proces tiež zapísať ako (5.68). Takýto proces sa nazýva aj integrovaný proces. Burke – Hunter (2005) uvádzajú, že dôvod tohto názvu vychádza z výrazu (5.68), v ktorom sa vyskytuje suma náhodných členov. Ak by bol tento proces spojitý, pojem integrovaný (ide o analógiu k integrálu) by bol vhodný. V našom prípade je však možno vhodnejší výraz „sčítaný“ alebo „nasčítaný“ proces (Burke – Hunter, 2005, s. 204). Ostaneme však pri zaužívanom výraze – integrovaný proces.

Autoregresný proces môžeme vo všeobecnosti pomocou operátora oneskorenia zapísať ako:

$$\left(1 - \sum_{i=1}^p \phi_i L^i\right) y_t = \varepsilon_t \quad (5.83)$$

$$\phi(L) y_t = \varepsilon_t \quad (5.84)$$

Výraz:

$$\left(1 - \sum_{i=1}^p \phi_i L^i\right) = \phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p \quad (5.85)$$

sa nazýva polynóm stupňa  $p$  operátora  $L$ . Operátor  $L$  bol definovaný vyššie v Kapitole 5.1. K všeobecnému autoregresnému procesu vieme vytvoriť polynomickú funkciu stupňa  $p$ , tzv. charakteristickú rovnicu (kde  $z$  je komplexné číslo):

$$\left(1 - \sum_{i=1}^p \phi_i z^i\right) = \phi(z) = 0 \quad (5.86)$$

Pokiaľ sú všetky korene tejto rovnice v absolútnej hodnote väčšie ako 1, proces (5.83) je považovaný za kovariančne stacionárny (takéto silné tvrdenie súvisí s definíciou  $\varepsilon_t$ ). Pre model náhodnej prechádzky je riešením  $z = 1$ . Prítomnosť jednotkového koreňa je tak indikátorom určitej formy nestacionarity. Z (5.86) taktiež vyplýva, že ak je súčet regresných koeficientov rovný 1, potom je riešením rovnice 1. Zisťovanie (ne)prítomnosti jednotkového koreňa v časovom rade sa používa ako základný princíp pri mnohých testoch overujúcich (ne)stacionaritu.

Ako si neskôr ukážeme (kapitola 5.4 o klamlivej regresii), nestacionarita vie v mnohých prípadoch znamenať určitý problém pri spracovaní empirických dát. Hamilton (1994, s. 444) uvádza, že ak je nestacionarita časového radu spôsobená prítomnosťou jednotkového koreňa, potom môžeme tento stochastický trend odstrániť diferencovaním časového radu. Na druhej strane, ak je nestacionarita časového radu spôsobená prítomnosťou deterministického trendu, potom ju môžeme odstrániť detrendovaním. V oboch prípadoch sa

určitý typ informácie stráca. V prvom prípade hovoríme o zmenách pôvodnej premennej. Tomuto typu analýzy sa často hovorí aj krátkodobá analýza, lebo namiesto trendov nás zaujímajú zmeny. V druhom prípade hovoríme o odchýlkach od deterministického trendu, t. j. znovu krátkodobá analýza.

V empirickej analýze dát, sa často stretávame s potrebou určiť stupeň integrácie časového radu. Uvažujme o procese  $x_t$ , ktorý je nestacionárny proces, pričom nestacionarita je spôsobená len prítomnosťou tzv. jednotkového koreňa. Uvažujme ďalej o procese:

$$y_t = (1 - L)^d x_t \quad (5.87)$$

ktorý je možné vyjadriť ako stacionárny a invertibilný ARMA proces:

$$\left(1 - \sum_{i=1}^p \phi_i L^i\right) y_t = \left(1 - \sum_{i=1}^q \theta_i L^i\right) \varepsilon_t \quad (5.88)$$

ARMA model je stacionárny, ak všetky korene polynómu  $\phi(z)$  sú v absolútnej hodnote väčšie ako 1. ARMA model je invertibilný, ak všetky korene polynómu  $\phi(z)$  a  $\theta(z)$  sú v absolútnej hodnote väčšie ako 1. Stupeň integrácie procesu  $x_t$  môžeme chápať ako počet diferencií, ktoré je potrebné urobiť, aby sme mohli časový rad vyjadriť ako stacionárny a invertibilný ARMA proces. Táto vlastnosť procesu sa zvykne označovať ako  $x_t \sim I(d)$ . Substitúciou (5.87) do (5.88) vieme formulovať nasledujúci ARIMA( $p, d, q$ ) model (angl. *autoregressive integrated moving average*):

$$\left(1 - \sum_{i=1}^p \phi_i L^i\right) (1 - L)^d x_t = \left(1 - \sum_{i=1}^q \theta_i L^i\right) \varepsilon_t \quad (5.89)$$

$$\phi(L) \Delta^d x_t = \theta(L) \varepsilon_t \quad (5.90)$$

$$\pi(L) x_t = \theta(L) \varepsilon_t \quad (5.91)$$

Vo vzťahu (5.89) sme operátory na ľavej strane nahradili jedným operátorom (oba je možné vzájomne násobiť). Stupeň integrácie označuje potom počet jednotkových koreňov polynómu  $\pi(z)$ .

V ďalšom texte, pokiaľ nebude povedané inak, budeme pri pojme nestacionarita rozumieť prítomnosť jednotkového koreňa. Pri pojme stacionarita budeme rozumieť kovariančnú stacionaritu. V ďalších kapitolách prezentujeme niekoľko štandardných a aj menej štandardných testov, ktoré sa používajú na overovanie (ne)stacionarity časového radu. Pre empirické aplikácie testov pozri Lyócsa et al. (2011).

### 5.3.1 ADF-test

Vychádzajme z autoregresného modelu prvého rádu AR(1):

$$y_t = \alpha_0 + \rho y_{t-1} + w_t \quad (5.92)$$

pričom o náhodnej chybe  $w_t$  predpokladáme, že ide o biely šum (stacionárny proces, ktorý má konštantnú, nulovú strednú hodnotu a jeho časové oneskorenia sú nekorelované). Časový rad má jednotkový koreň, ak  $\rho_1 = 1$ . Vtedy hovoríme o už spomínanej náhodnej prechádzke. Pre procesy AR(1) platí postačujúca podmienka stacionarity, ktorá má tvar  $|\rho_1| < 1$ . V prípade, že máme k dispozícii nestacionárny časový rad (v tejto forme,  $\rho_1 = 1$ ), stacionaritu môžeme zabezpečiť jeho prvou diferenciou, značíme  $I(1)$ . Diferencovaním získame stacionárny  $I(0)$  proces. Prirodzeným testom na prítomnosť jednotkového koreňa je odhadnúť model (5.92) a testovať, či regresný koeficient  $\hat{\rho}$  je štatisticky významne odlišný od 1, povedzme pomocou  $t$ -testu.<sup>21</sup> Pod nulovou hypotézou sa však  $t$ -štatistika neriadi Studentovým- $t$  rozdelením pravdepodobnosti, Dickey – Fuller (1979). V praxi sa potom postupuje tak, že sa model (5.92) upraví na nasledovný tvar:

$$\Delta y_t = \alpha_0 + (\rho - 1)y_{t-1} + w_t = \alpha_0 + \delta y_{t-1} + w_t \quad (5.93)$$

Oproti výrazu (5.92) sa odpočítala oneskorená hodnota závislej premennej. Testovanie nulovej hypotézy je potom ekvivalentné testovaniu hypotézy  $\delta = 0$ , oproti alternatíve  $\delta < 0$ . Tomuto testu sa zvykne hovoriť Dickey – Fullero test, resp. DF test. Existujú rôzne špecifikácie tohto testu, pričom sa navzájom líšia zahrnutím deterministických zložiek:

$$\Delta y_t = \delta y_{t-1} + w_t \quad (5.94)$$

$$\Delta y_t = \alpha_0 + \delta y_{t-1} + w_t \quad (5.95)$$

$$\Delta y_t = \alpha_0 + \alpha_1 t + \delta y_{t-1} + w_t \quad (5.96)$$

Ide o test bez konštanty, test s konštantou a test s konštantou a deterministickým trendom. Neskôr naznačíme akým spôsobom zvoliť správnu špecifikáciu modelu.

Táto forma testu sa však už prakticky nevyskytuje. Dôvod spočíva v často nerealistickom predpoklade, že  $w_t$  sú nekorelované. Najbežnejšie sa na testovanie stacionarity využíva rozšírený Dickey–Fullerov test (angl. *Augmented Dickey Fuller* – ADF)<sup>22</sup>, ktorý je využiteľný aj na procesy iné ako AR(1), tzn. aj pre prípady, keď sú korelované oneskorenia

<sup>21</sup> Pre konštantu platí  $\alpha_0 = (1 - \rho)\mu$ , kde  $\mu$  je aritmetický priemer časového radu  $x_t$ , a teda ak je  $\rho = 1$ , potom  $\alpha_0 = 0$ . Je možné testovať spoločne obe reštrikcie.

<sup>22</sup> Bližšie k danému testu v pôvodnej práci Dickey – Fuller (1981).

vyššieho rádu – teda keď je porušený predpoklad o bielom šume. Autokorelácia náhodných chýb je korigovaná tým, že sa do modelu k vysvetľujúcim premenným pridá  $m$  oneskorených diferencií závislej premennej. Podobne ako v predošlom prípade, máme tri základné špecifikácie testu:

$$\Delta y_t = \delta y_{t-1} + \sum_{i=1}^m \beta_i \Delta y_{t-i} + w_t \quad (5.97)$$

$$\Delta y_t = \alpha_0 + \delta y_{t-1} + \sum_{i=1}^m \beta_i \Delta y_{t-i} + w_t \quad (5.98)$$

$$\Delta y_t = \alpha_0 + \alpha_1 t + \delta y_{t-1} + \sum_{i=1}^m \beta_i \Delta y_{t-i} + w_t \quad (5.99)$$

V modeli testujeme nulovú hypotézu  $\delta = 0$  (časový rad má jednotkový koreň a je nestacionárny) oproti alternatívnej  $\delta < 0$  (ide o stacionárny časový rad).

V softvéri R je možné využiť funkciu `ur.df()` (z knižnice `urca`) na výpočet rozšíreného Dickey – Fuller testu.<sup>23</sup> Ak nechceme využiť pôvodné asymptotické kritické hodnoty, tak pomocou funkcie `punitroot()` (knižnica `urca`) vieme vypočítať  $p$ -hodnoty a `qunitroot()` kritické hodnoty podľa simulácií od MacKinnon (1996).

Vo funkcii `ur.df()` si vieme potom vybrať, ktorú špecifikáciu ADF testu chceme aplikovať, `type = c("none", "drift", "trend")`. Ak zvolíme `type = "none"`, tak do rovnice nezahrnieme konštantu (koeficient  $\beta_0$ ) a ani trend (koeficient  $\beta_1$ ). Pri argumente `drift` je pridaná konštanta a pri argumente `trend` sú do testovanej regresie pridané aj konštanta a aj trend. Vo výsledkoch funkcie `ur.df()` sú podľa typu zvoleného modelu prezentované nasledujúce štatistiky<sup>24</sup>:

- `tau1` – pre model bez konštanty a bez trendu (`type = "none"`),
- `tau2` – pre model s konštantou (`type = "drift"`),
- `tau3` – pre model s konštantou a trendom (`type = "trend"`).

Vo výsledkoch máme k dispozícii ešte aj  $F$ -štatistiky (s neštandardným rozdelením), pri ktorých sa testujú nasledujúce združené hypotézy<sup>25</sup>:

- `phi1` :  $(\alpha_0, \delta) = (0, 0)$
- `phi2` :  $(\alpha_0, \alpha_1, \delta) = (0, 0, 0)$

<sup>23</sup> Taktiež môžeme využiť `adf.test` (knižnica `tseries`) alebo `adfTest` (knižnica `fUnitRoots`).

<sup>24</sup> Ide o  $t$ -štatistiky s neštandardným rozdelením. Bližšie <http://gremaq.univ-tlse1.fr/pig/Docs/PigTS.pdf>

<sup>25</sup> Pri modeli bez konštanty a trendu nemáme  $F$ -štatistiky, keďže v rovnici máme len jeden parameter ( $\delta$ ).

- $\text{phi3} : (\alpha_1, \delta) = (0, 0)$

Taktiež vieme nastaviť maximálny počet oneskorení (lags), ktoré chceme brať do úvahy ( $m$  oneskorení podľa vzťahu (5.97)) a kritérium, podľa ktorého má byť zvolený optimálny počet oneskorení (`selectlags = c("Fixed", "AIC", "BIC")`). Ak nastavíme argument `selectlags = "Fixed"`, tak počet oneskorení bude nastavený podľa počtu zadaného v kritériu lags. Môžeme však optimálny počet oneskorení vybrať podľa niektorého z informačných kritérií (AIC alebo BIC). Tejto problematike budeme ešte venovať priestor na konci tejto kapitoly.

Na ukážku budeme pracovať s nasledujúcimi časovými radmi<sup>26</sup>:

**Stacionárny časový rad (AR\_1):** 
$$y_t = 0.5y_{t-1} + w_t, \rho = 0.5 < 1 \quad (5.100)$$

**Nestacionárny časový rad (AR\_2):** 
$$y_t = 1.005y_{t-1} + w_t, \rho = 1.05 > 1 \quad (5.101)$$

**Jednotkový koreň (AR\_3):** 
$$y_t = y_{t-1} + w_t, \rho = 1 \quad (5.102)$$

```
> library(urca)
> x <- rnorm(500,0,1)
> AR_1 <- filter(x, filter = c(0.5), method = "recursive")
> AR_2 <- filter(x, filter = c(1.005), method = "recursive")
> AR_3 <- filter(x, filter = c(1), method = "recursive")
-----
> ADF_1 <- ur.df(AR_1, type="trend", lags=1); ADF_1
```

```
# Augmented Dickey-Fuller Test Unit Root / Cointegration Test #
#####

The value of the test statistic is: -11.5477 44.4514 66.6749

> ADF_2 <- ur.df(AR_2, type="trend", lags=1); ADF_2

#####
# Augmented Dickey-Fuller Test Unit Root / Cointegration Test #
#####

The value of the test statistic is: 0.2252 11.5994 3.9311

> ADF_3 <- ur.df(AR_3, type="trend", lags=1); ADF_3

#####
# Augmented Dickey-Fuller Test Unit Root / Cointegration Test #
#####
```

<sup>26</sup> Vygenerujeme väčší počet pozorovaní (500) a aby nám časový rad AR\_2 neexplodoval veľmi rýchlo, koeficient  $a$  nastavíme na 1.005. Pri väčšom počte pozorovaní dosiahneme lepšiu silu testu.

```
The value of the test statistic is: -2.1743 1.8519 2.3825
```

Vo výstupe funkcie teda môžeme vidieť (keďže sme zvolili model s konštantou aj s trendom) štatistiky tau3, phi2 a phi3. Podrobnejšie výsledky z testovanej regresie dostaneme použitím funkcie `summary()`. Zároveň tak vieme pozrieť, ktoré parametre boli významné a taktiež v závere výstupu máme uvedené kritické hodnoty pre štatistiky tau3, phi2 a phi3. Na ukážku uvidíme výsledky len pre vygenerovaný časový rad AR\_3.

```
> summary(ADF_3)

#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####

Test regression trend

Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)

Residuals:
    Min       1Q   Median       3Q      Max
-2.3135 -0.5719  0.1530  0.5936  2.2146

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.42197    0.33307  -1.267   0.2119
z.lag.1      -0.20000    0.09198  -2.174   0.0351 *
tt           -0.01114    0.01090  -1.022   0.3126
z.diff.lag   0.13132    0.14934   0.879   0.3840
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9643 on 44 degrees of freedom
Multiple R-squared:  0.09878,    Adjusted R-squared:  0.03733
F-statistic: 1.607 on 3 and 44 DF,  p-value: 0.2012

Value of test-statistic is: -2.1743 1.8519 2.3825

Critical values for test statistics:
      1pct  5pct 10pct
tau3 -4.15 -3.50 -3.18
phi2  7.02  5.13  4.31
phi3  9.31  6.73  5.61
```

Z uvedených výsledkov pre časový rad AR\_3 je zrejmé, že nulovú hypotézu  $\delta = 0$  nevieme zamietnuť ani na hladine významnosti 10 %. Časový rad teda môžeme považovať za

nestacionárny a vzhľadom na to, že poznáme proces generujúci tieto údaje, tak s týmito výsledkami môžeme byť spokojní (išlo o proces s jednotkovým koreňom, teda *unit root* proces).

V softvéri R je zrejme najjednoduchšie využiť pri ADF teste priamo funkciu `punitroot()` (knižnica `urca`), pomocou ktorej vieme vypočítať rovno  $p$ -hodnoty podľa tabelovaných kritických hodnôt z práce MacKinnon (1996) a nemusíme sledovať vypočítané testovacie štatistiky a kritické hodnoty. Pre model bez konštanty a bez trendu musíme zvoliť argument `trend = "nc"`, s konštantou ale bez trendu `trend = "c"` a pre model s trendom aj konštantou `trend = "ct"`. Na výber máme tiež výpočet  $t$ -štatistiky alebo rho štatistiky (normalizovaná štatistika) v argumente funkcie `statistic = c("t", "n")`.

```
> p_ADF_1 <- punitroot(ur.df(AR_1, type = "trend", lags =
  1)@teststat[1], N = length(AR_1), trend = "ct", statistic =
  "t"); p_ADF_1
[1] 1.830100e-19
> p_ADF_2 <- punitroot(ur.df(AR_2, type = "trend", lags =
  1)@teststat[1], N = length(AR_2), trend = "ct", statistic =
  "t"); p_ADF_2
[1] 0.9996988
> p_ADF_3 <- punitroot(ur.df(AR_3, type = "trend", lags =
  1)@teststat[1], N = length(AR_3), trend = "ct", statistic =
  "t"); p_ADF_3
[1] 0.4929388
```

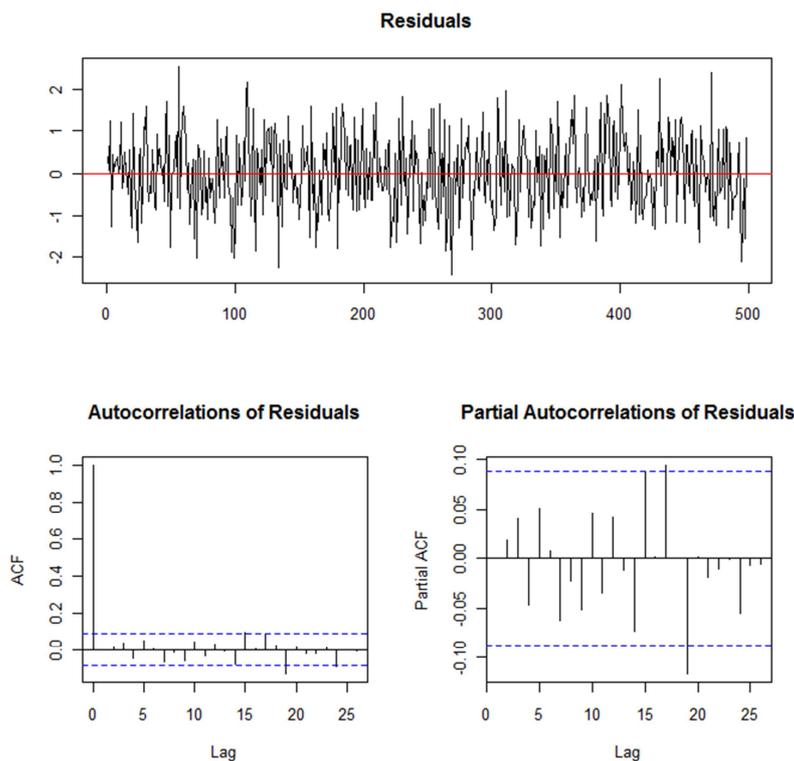
Z vygenerovaných časových radov by sme nulovú hypotézu vedeli zamietnuť len v prvom prípade (kde išlo o stacionárny časový rad) a pri zvyšných dvoch prípadoch by sme časové rady vyhodnotili ako nestacionárne.

S využitím funkcie `plot()` sa vieme tiež pozrieť na graf rezíduí a autokorelačné funkcie.<sup>27</sup>

```
> plot(ADF_3)
```

---

<sup>27</sup> Autokorelačná funkcia predstavuje závislosť medzi koreláciou a použitým oneskorením, teda  $\rho: \mathbb{Z} \rightarrow \langle 1, 1 \rangle$ , napr.  $\rho(2)$  bude značiť koreláciou medzi  $y_t$  a  $y_{t-2}$ .



Obrázok 20: Rezíduá a autokorelačné funkcie z ADF testu

*Zdroj: výstup zo softvéru R*

Pri doterajšej prezentácii testov sme sa vyhlili dvom dôležitým témam: (1) voľba špecifikácie testu, (2) voľba počtu oneskorení v pomocnej regresii. Najprv načrtujeme možnú stratégiu pre prvý problém. Elder – Kennedy (2001) prezentovali pomerne všeobecnú stratégiu voľby vhodnej špecifikácie modelu používaného pri zisťovaní (ne)stacionarity, ku ktorej sa v tejto publikácii prikláňame. Začnime s predpokladom, že časový rad z dlhodobého hľadiska obsahuje trend (rastúci alebo klesajúci). Tento trend je pritom dôsledkom buď stacionárneho procesu odvíjajúceho sa v okolí deterministického trendu, alebo dôsledkom prítomnosti jednotkového koreňa. To znamená, že rast nie je spôsobený prítomnosťou jednotkového koreňa a zároveň lineárnym trendom.

Za týchto predpokladov ďalej uvažujme o nasledujúcej situácii. O pôvodnom časovom rade vieme, že má rastúci alebo klesajúci trend. Napríklad z dlhodobého hľadiska by malo dochádzať k rastu reálnych cien akciových indexov ako aj celkovej produkcii ekonomiky. V oboch prípadoch v dôsledku technologického vývoja a rastu produktivity práce objem výroby v ekonomike narastá. To by sa malo odraziť aj v náraste cien akciových indexov. Je zrejme nepravdepodobné, aby dochádzalo k rastu ekonomiky a (ceteris paribus) systematickému poklesu reálnych cien akciových indexov. Ak by boli reálne akciové

indexy konštantné, potom nominálny nárast cien akcií by bol kompenzovaný rastom cenovej hladiny. Táto alternatíva je možná, je však reálnejšie pripustiť aspoň mierne rastúci trend v reálnych cenách akciových výnosov. V takomto prípade je vhodné použiť špecifikáciu s trendom a konštantou. V prípade, ak nevieme zamietnuť prítomnosť jednotkového koreňa, tak považujeme časový rad za nestacionárny a trend za dôsledok nenulovej konštanty.<sup>28</sup> V prípade, ak časový rad nerastie (napr. po diferencovaní), volí sa špecifikácia s konštantou. Ak sa potvrdí prítomnosť jednotkového koreňa, potom sa konštanta považuje priamo za nevýznamnú. V opačnom prípade by časový rad mal v sebe deterministický trend. Ak na základe testu sme toho názoru, že jednotkový koreň nie je prítomný, potom pozorovaný časový rad osciluje okolo konštantnej hodnoty.

Z použitia tejto stratégie testovania sa môže zdať, že ak analýza úrovňových premenných naznačuje prítomnosť jednotkového koreňa, proces je integrovaný rádu  $d > 1$ . Ignorujúc fakt, že príčiny nestacionarity môžu byť aj iné ako len vyšší stupeň integrácie, v ekonómii sme často svedkami tzv. štrukturálnych zlomov. Preto jednou z reálnych možností je, že napriek tomu, že nám štandardný test naznačuje prítomnosť aspoň jedného jednotkového koreňa, tak v skutočnosti ide o časový rad stacionárny okolo deterministického trendu, ktorý vykazuje tzv. štrukturálny zlom. V takomto prípade je náš model chybné špecifikovaný a naša analýza naznačuje prítomnosť  $I(1)$  procesu v situácii, kde v skutočnosti ide o  $I(0)$  nestacionárny proces. Perron (1989) ukázal, že ak časový rad má deterministický trend so štrukturálnymi zlomami (napr. posun v strednej hodnote, prípadne zmena tempa rastu), sila štandardných testov (s nulovou hypotézou „nestacionarita“) je nízka. Z tohto dôvodu nemusíme zamietnuť nulovú hypotézu o prítomnosti jednotkového koreňa napriek tomu, že časový rad je stacionárny okolo deterministického trendu so štrukturálnym zlomom.

Ak existuje podozrenie pre existenciu takéhoto štrukturálneho zlomu, potom vyššie uvedené špecifikácie nie sú správne. Z tohto dôvodu budeme prezentovať aj iné typy testov, ktoré tieto štrukturálne zlomy zohľadňujú.

Ostalo nám určiť správny počet oneskorení v ADF teste, prípadne v podobných testoch. V súčasnosti už existuje mnoho prístupov a na nešťastie, aspoň podľa našej skúsenosti, na voľbe vhodného počtu oneskorení do značnej miery záleží. Spravidla sa vyberá

---

<sup>28</sup> Významnosť konštanty pritom nie je podstatná. Vychádzame z nášho predpokladu a z neho vyplýva, že časový rad rastie/klesá a tento trend v prípade prítomnosti jednotkového koreňa musí byť spôsobený konštantou. Ide o tzv. náhodnú prechádzku s posunom. Skúste si nasimulovať niekoľko krát nasledujúce procesy s 500 pozorovaniami:  $y_t = y_{t-1} + e_t$  a  $y_t = 0.08 + y_{t-1} + e_t$ ,  $e_t \sim N(0, 1)$ .

vhodný počet oneskorení z intervalu od  $k = 1, 2, \dots, k_{\max}$ , kde  $k_{\max}$  sa zvolí podľa vzťahu  $k_{\max} = \text{int}[12(T/100)^{0.25}]$  (Schwert, 1989).

Počet oneskorení  $k_{\text{opt}}$  sa potom vyberie pri takej špecifikácii modelu, pri ktorom je najmenšia hodnota upraveného Akaike informačného kritéria (tzv. MAIC, pozri Ng – Perron, 2001)<sup>29</sup>. Inou možnosťou ako odhadnúť správnu špecifikáciu modelu, je odhadnúť tento model s počtom oneskorení  $k_{\max}$ . Ak je na hladine významnosti  $\alpha = 0.1$  koeficient pri poslednom oneskorení významný, potom  $k_{\text{opt}} = k_{\max}$ . Ak nie je významný, potom sa odhadne model s  $k_{\max} - 1$  počtom oneskorení a postup sa opakuje dotedy, pokiaľ nebude koeficient pri poslednom oneskorení významný, resp. pokiaľ nebude počet oneskorení = 1. Tento postup bol navrhnutý v práci Ng – Perron (1995). Pre tieto dva postupy výberu  $k_{\text{opt}}$  môžeme použiť kritické hodnoty od Cook – Manning (2004), ktorý pomocou simulácií zostavili približné kritické hodnoty pre rôzne veľkosti vzoriek  $T = \{50, 100, 250, 2500\}$  a príslušný spôsob výberu počtu oneskorení  $k_{\text{opt}}$ . Ako sme už spomenuli, počet oneskorení sa pri ADF testoch používa z dôvodu autokorelácie rezíduí. Preto uvádzame ešte jeden postup, kde  $k_{\text{opt}}$  sa vyberá tak, že sa na začiatku  $k$  stanoví ako 0 a overuje sa hypotéza o nulovej autokorelácií rezíduí z takto špecifikovaného modelu. K tomu slúžil napríklad štandardný Ljung – Box test, kde počet oneskorení (používaných v teste) je spravidla od 1 do  $\text{int}\{0.05T\}$ . Ak nulovú hypotézu o neprítomnosti autokorelácie zamietneme, pokračujeme ďalej odhadom modelu s  $k + 1$  počtom oneskorení. Tento postup môžeme opakovať dotedy, pokiaľ použitím Ljung – Box testu hypotézu o neprítomnosti autokorelácie nevieme zamietnuť. Tento postup je blízky názoru prezentovaného v Cheung – Lai (1995), ktorí tvrdili, že počet oneskorení sa volí tak, aby v rezíduách nebola prítomná autokorelácia. Pri overovaní hypotézy, že časový rad je  $I(1)$ , potom môžeme použiť ich kritické hodnoty, ktoré sa vypočítajú zo vzťahov pre výpočet kritických hodnôt pre  $\alpha = 0.05, 0.1$  (pozri Cheung – Lai, 1995; Tabuľka 1). V nasledujúcej kapitole ešte spomenieme dva výbery  $k_{\text{opt}}$  pomocou modifikovaného informačného kritéria Akaike.

### 5.3.2 ADF-GLS test

V súčasnosti jedným z najpoužívanejších testov je zrejme ADF-GLS test od Elliott et al. (1996). Ak pozorovania časového radu označíme ako  $y_t$ , potom overujeme nasledujúcu

---

<sup>29</sup> V súčasnosti sa už táto forma výberu počtu oneskorení neodporúča, pozri Perron – Qu (2007).

nulovú hypotézu  $\phi = 0$ , oproti alternatívnej hypotéze  $\phi < 0$  z modelu (5.103). Ide o test, ktorého nulová hypotéza hovorí o nestacionarite vo forme jednotkového koreňa(ov), kým alternatívna hypotéza hovorí o neprítomnosti jednotkového koreňa:

$$\Delta\tilde{y}_t = \phi\tilde{y}_{t-1} + \sum_{i=1}^k \delta_i \Delta\tilde{y}_{t-i} + \varepsilon_t \quad (5.103)$$

kde  $k > 0$  je počet oneskorení diferencií detrendovaného časového radu  $y_t$ , ktorý sa môže zvoliť pomocou jednej z metód uvedených vyššie. Tento detrendovaný časový rad sa vytvoril nasledujúcim spôsobom:

$$\tilde{y}_t = y_t - \hat{\beta}_0 - \hat{\beta}_1 t \quad (5.104)$$

Z výrazu (5.104) vyplýva, že od pôvodného časového radu sa odpočítava deterministická časť. Parametre  $\hat{\beta}_0$  a  $\hat{\beta}_1$  sa získali pomocou metódy najmenších štvorcov z jednoduchej lineárnej regresie, kde závislá premenná  $\bar{y}$  a vektor nezávislých premenných  $\bar{z}$  boli definované nasledovne:

$$\bar{y} = [y_1, (1 - \bar{\alpha}L)y_2, \dots, (1 - \bar{\alpha}L)y_T] \quad (5.105)$$

$$\bar{z} = [z_1, (1 - \bar{\alpha}L)z_2, \dots, (1 - \bar{\alpha}L)z_T] \quad (5.106)$$

pričom  $z_t = (1, t)'$  pre model s lineárnym trendom a  $z_t = (1)'$  pre model s konštantou. Ďalej  $\bar{\alpha} = 1 + \bar{c}/T$ , kde  $\bar{c}$  je konštanta, ktorá pre model s lineárnym trendom a konštantou nadobúda hodnotu  $-13.5$  a pre model s konštantou  $-7$ . Počet pozorovaní časového radu  $y_t$  označujeme ako  $T$ ,  $t = 1, 2, \dots, T$ . Testovaciu štatistiku týkajúcu sa koeficientu  $\phi = 0$  si môžeme označiť ako  $\tau_\mu^{\text{GLS}}$  alebo  $\tau_\tau^{\text{GLS}}$ , a to v závislosti od toho, či pôjde o model s konštantou alebo o model s konštantou a trendom. Kritické hodnoty potom nájdeme napr. v Cook – Manning (2004) alebo Cheung – Lai (1995).

V Prílohe 1 je kód (funkcia `ADF_GLS()`), ktorý slúži na výpočet ADF-GLS testu, pričom ponúka testovacie štatistiky pre štyri metódy voľby počtu oneskorení. Hodnoty testovacej štatistiky pre rôzne oneskorenia použitím funkcie v Prílohe 1 a použitím funkcie `ur.ers()` sú prakticky rovnaké. Aby sme ilustrovali citlivosť výsledkov na počet oneskorení, uskutočnili sme jednoduchý experiment. Vybrali sme veľkosť vzorky  $T = 500$ . Pre túto veľkosť vzorky sme simulovali časový rad  $x_t = \mu + \rho x_{t-1} + e_t$ ,  $e_t \sim N(0, 1)$  a  $\mu \sim U(-2, 2)$ , ktorý bol teda stacionárny. Pre  $\rho = 0$  je vhodný počet oneskorení pritom zjavne 0 a vhodná špecifikácia je model s konštantou. Na tomto časovom rade sme uskutočnili nasledujúce špecifikácie ADF-GLS testu: 1) počet oneskorení sme zvolili pomocou Ng – Perron (1995)

procedúry a kritické hodnoty vybrali na základe Cook – Manning (2004), 2) počet oneskorení sme vybral pomocou modifikovaného informačného kritéria Akaike (MAIC), Ng – Perron (2001) a kritické hodnoty na základe Cook – Manning (2004), 3) počet oneskorení sme vybrali pomocou MAIC s korekciou tak, ako bola uvedená v Perron – Qu (2007), pričom sme použili asymptotické kritické hodnoty, a na záver 4) sa počet oneskorení zvolili pomocou testovania neprítomnosti autokorelácie v rezíduách, kde kritické hodnoty sa zvolili pomocou Cheung – Lai (1995). Vykonali sme 1000 iterácií a uložili si počet chybných nezamietnutí nulovej hypotézy o stacionarite. Následne sme tento pokus vykonali pre hodnoty  $\rho$  od 0.00 po 0.99 (s krokmi po 0.01). Na záver sme simulovali aj procesy s  $\rho = 1$ , aby sme zistili tzv. veľkosť testu, teda počet nesprávne nezamietnutých nulových hypotéz. Pre túto analýzu sú výsledky v obrázku (pozri Obrázok 21) zaznačené červenou farbou. Modrou je čiara znázorňujúca hladinu významnosti, na ktorej sme testy vykonali, t. j.  $\alpha = 0.05$ .

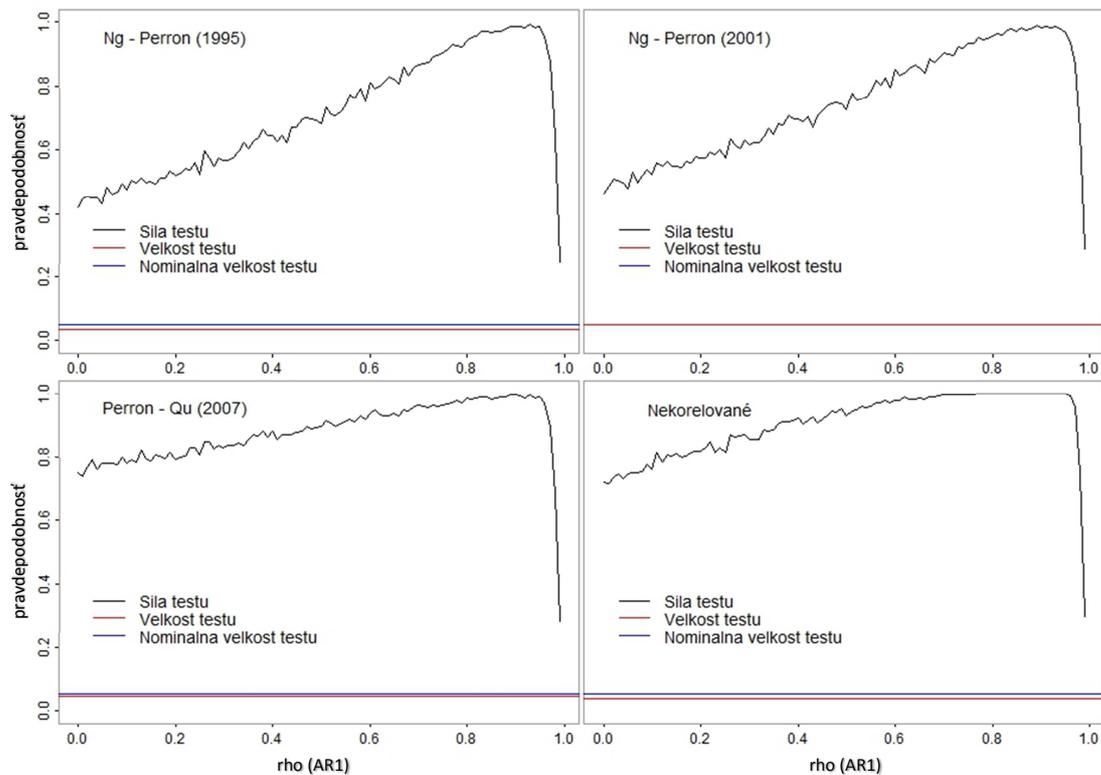
```

> TT = 500; I = 1000; rhos = 100; metody = 4;
> results <- matrix(nrow = rhos, ncol = metody)
> rho <- seq(from = 0, to = 0.99, by = 0.01)
> for (j in 1:length(rho)) {
+   r <- rho[j]
+   res <- matrix(ncol = metody, nrow = I)
+   for (i in 1:I) {
+     srs <- arlsim(TT, rho = r, shock = 0.0, size = 0, errors =
"different")[[1]] + runif(1, min = -2, max = 2)
+     a <- ADF_GLS(srs, model = "mu", kmax = NULL, mLBo = NULL)
+     res[i,1] <- c(a[1,1] < a[1,4])
+     res[i,2] <- c(a[2,1] < a[2,4])
+     res[i,3] <- c(a[3,1] < a[3,4])
+     res[i,4] <- c(a[4,1] < a[4,4])
+   }
+   results[j,1] <- sum(res[,1])/I
+   results[j,2] <- sum(res[,2])/I
+   results[j,3] <- sum(res[,3])/I
+   results[j,4] <- sum(res[,4])/I
+ }

```

Na obrázku nižšie (pozri Obrázok 21) je čiernou čiarou zaznačená sila testu. To znamená pravdepodobnosť správneho zamietnutia nulovej hypotézy o prítomnosti jednotkového koreňa. Na osi  $x$ -ovej je pritom veľkosť parametra  $\rho$ . Problematické správanie sa testu je vidno pri procedúrach Ng – Perron (1995), ako aj Ng – Perron (2001). V oboch prípadoch je sila testu vo väčšine prípadov nižšia ako pri Perron – Qu (2007) procedúre a pri voľbe, kde sa uprednostňujú nekorelované rezíduá. Dokonca sa javí, ako by sa pri nekorelovaných rezíduách vyššia úroveň sily testu dosiahla skôr (v závislosti od parametra  $\rho$ ). Veľkosť testu je v každom z týchto procedúr adekvátna, t. j. ak sme spustili test na časovom

rade vygenerovanom cez  $\rho = 1$ , tak nesprávne nezamietnutie nulovej hypotézy nastalo v približne 5 % prípadov (čo je vidno pri porovnaní modrej a červenej čiary). Táto analýza nezahŕňa ďalšie typické charakteristiky časových radov v ekonómii. Ide najmä o ARMA ako aj GARCH procesy a samozrejme meniacu sa veľkosť vzorky. V našom prípade sme uvažovali o pomerne veľkých vzorkách.



Obrázok 21: Veľkosť a neupravená sila ADF-GLS testu pre rôzne metódy voľby počtu oneskorení a perzistentnosti procesu

*Zdroj: výstup zo softvéru R*

```
> dev.off()
null device
      1
> rgb <- hcl(c(0), c = c(0), l = c(50), alpha = 0.25)
> par(mfrow = c(2, 2))
> par(cex = 0.6)
> par(mar = c(1.8, 0.2, 0.2, 0.2), oma = c(2.8, 2.0, 0.5, 0.5))
> par(tcl = -0.25)
> par(mgp = c(2, 0.6, 0))
>
> plot(rho, results[,1], ylim = c(0, 1), main = NA, cex.axis =
  1.2, yaxt = "n", type = "l", lty = 1, lwd = 1.5)
> text(x = 0.20, y = 0.95, labels = "Ng - Perron (1995)", lty =
  1, bty = "n", cex = 1.5)
> axis(side = 2, cex.axis = 1.25)
> box(col = "grey60")
```

```

> abline(h = 0.05, col = "blue")
> abline(h = 0.036, col = "red")
> legend(x = 0.0, y = 0.4, legend = c("Sila testu", "Velkost
  testu", "Nominalna velkost testu"), col = c("black", "red",
  "blue"), lty = 1, bty = "n", cex = 1.5)
-----
> plot(rho, results[,2], ylim = c(0, 1), main = NA, cex.axis =
  1.2, yaxt = "n", type = "l", lty = 1, lwd = 1.5)
> text(x = 0.20, y = 0.95, labels = "Ng - Perron (2001)", lty =
  1, bty = "n", cex = 1.5)
> box(col = "grey60")
> abline(h = 0.05, col = "blue")
> abline(h = 0.051, col = "red")
> legend(x = 0.0, y = 0.4, legend = c("Sila testu", "Velkost
  testu", "Nominalna velkost testu"), col = c("black", "red",
  "blue"), lty = 1, bty = "n", cex = 1.5)
-----
> plot(rho, results[,3], ylim = c(0, 1), main = NA, cex.axis =
  1.2, yaxt = "n", type = "l", lty = 1, lwd = 1.5)
> text(x = 0.20, y = 0.95, labels = "Perron - Qu (2007)", lty =
  1, bty = "n", cex = 1.5)
> axis(side = 2, cex.axis = 1.25)
> box(col = "grey60")
> abline(h = 0.05, col = "blue")
> abline(h = 0.042, col = "red")
> legend(x = 0.0, y = 0.4, legend = c("Sila testu", "Velkost
  testu", "Nominalna velkost testu"), col = c("black", "red",
  "blue"), lty = 1, bty = "n", cex = 1.5)
-----
> plot(rho, results[,4], ylim = c(0, 1), main = NA, cex.axis =
  1.2, yaxt = "n", type = "l", lty = 1, lwd = 1.5)
> text(x = 0.20, y = 0.95, labels = "Nekorelované", lty = 1, bty
  = "n", cex = 1.5)
> box(col = "grey60")
> abline(h = 0.05, col = "blue")
> abline(h = 0.036, col = "red")
> legend(x = 0.0, y = 0.4, legend = c("Sila testu", "Velkost
  testu", "Nominalna velkost testu"), col = c("black", "red",
  "blue"), lty = 1, bty = "n", cex = 1.5)

```

### 5.3.3 KPSS test

Nasledujúci postup testovania overuje hypotézu, že časový rad nemá jednotkový koreň. Ide o tzv. KPSS test od Kwiatkowski et al. (1992). Tomuto testu sa hovorí aj test stacionarity. V tejto publikácii uvádzame verziu testu odporúčanú v štúdií Hobijn et al. (2004). Na rozdiel od väčšiny štatistických testov je pri KPSS nulová hypotéza postavená tak, že časový rad je „stacionárny“ (čo však znamená len to, že nie je  $I(1)$ ). Uvažujme o nasledujúcom modeli:

$$y_t = \alpha + \beta t + d \sum_{i=1}^t u_i + \varepsilon_t \quad (5.107)$$

kde  $u_i$  a  $\varepsilon_i$  sú kovariančne stacionárne časové rady so strednou hodnotou 0 a  $d = \{0, 1\}$ . V prípade, ak  $d = 1$  ide o nestacionárny časový rad s jednotkovým koreňom. Ak  $d = 0$  ide o stacionárny časový rad. Testovacia štatistika má potom nasledujúci tvar:

$$w = \frac{\sum_{t=1}^T \left( \sum_{i=1}^t e_i \right)^2}{T^2 \hat{\sigma}^2} \quad (5.108)$$

kde  $e_i$  predstavujú odhadované rezíduá z modelu (5.107), pričom v závislosti od typu nulovej hypotézy môžeme v modeli (5.107) uvažovať o  $\alpha, \beta \neq 0$  (model s konštantou a trendom) alebo o  $\alpha \neq 0$  a zároveň  $\beta = 0$  (model s konštantou). V prvom prípade si budeme testovaciu štatistiku označovať ako  $w_\tau$ , v druhom ako  $w_\mu$ . Okrem toho platí, že model (5.107) sa odhaduje pod nulovou hypotézou, t. j.  $d = 0$ . V menovateli vo vzťahu (5.108) používame odhad dlhodobého rozptylu  $\sigma^2$  rezíduí  $e_i$ . Hobijn et al. (2004) uvažovali o neparametrickom odhade dlhodobého rozptylu  $e_i$ , ktorý je možné použiť aj v situáciách, keď je v  $e_t$  prítomná autokorelácia a heteroskedasticita, čo je častý prípad práve časových radov, s ktorými sa stretávame v ekonómii. Test, ktorý zoberie do úvahy tieto dva aspekty časového radu, predstavuje značnú výhodu oproti ADF-GLS testu, kde sa autokorelácia rieši voľbou vhodnej špecifikácie pomocnej regresie. Ako sme mohli vidieť, od voľby vhodnej špecifikácie sa do značnej miery odvíjajú aj štatistické vlastnosti testu. Alternatívou k neparametrickému odhadu dlhodobého rozptylu, je použiť tzv. parametrický odhad.<sup>30</sup> V našej práci sa budeme držať postupu od Hobijn et al. (2004). Na odhad dlhodobého rozptylu použijeme nasledujúci vzťah:

$$\hat{\sigma}^2 = \frac{1}{T} \sum_{t=1}^T e_t^2 + 2 \sum_{j=1}^{T-1} k\left(\frac{j}{M}\right) \frac{1}{T} \sum_{t=1}^T e_t e_{t-j} \quad (5.109)$$

pričom  $k(\cdot)$  je kernel funkcia. Pre funkcie  $k(\cdot)$  platí nasledovné:

$$k(\cdot): R \rightarrow [-1, 1], \quad k(0) = 1, \quad k(z) = k(-z) \forall z \in R, \quad \int_{-\infty}^{\infty} |k(z)| dz < \infty \quad (5.110)$$

pričom  $z = j/M$  a funkcia  $k(\cdot)$  je spojitá v bode 0 a vo všetkých bodoch až na konečný počet bodov (pozri napr. Andrews, 1991; Cushing – McGarvey, 1999). Parameter  $M$  budeme nazývať bandwidth. V slovenskej literatúre sa môžeme stretnúť aj s pojmi ako šírka okna, prípadne vyhladávací parameter. V mnohých prípadoch a pre mnohé kernel funkcie môžeme

<sup>30</sup> Ďalšou neparametrickou alternatívou je korekcia, ktorú navrhovali Sul et al. (2005).

bandwidth interpretovať ako počet vážených oneskorení. V ekonometrii medzi často používané kernel funkcie patrí: Bartlett ( $B$ ), Quadratic Spectral ( $QS$ ), Parzen ( $P$ ), Daniell ( $D$ ):

$$B(j, M) = \begin{cases} 1 - |j/M|, & |j/M| \leq 1 \\ 0 & , \text{inak} \end{cases} \quad (5.111)$$

$$P(j, M) = \begin{cases} 1 - 6(j/M)^2 + 6|j/M|^3, & 0 \leq |j/M| \leq 1/2 \\ 2(1 - |j/M|)^3 & , 1/2 \leq |j/M| \leq 1 \\ 0 & , \text{inak} \end{cases} \quad (5.112)$$

$$D(j, M) = \sin(\pi(j/M)) / \pi(j/M) \quad (5.113)$$

$$QS(j, M) = \frac{25}{12\pi^2(j/M)^2} \left( \frac{\sin(6\pi(j/M)/5)}{6\pi(j/M)/5} - \cos(6\pi(j/M)/5) \right) \quad (5.114)$$

Odhad dlhodobého rozptylu je spravidla viac náchylný na voľbu parametra  $M$ , ako na špecifikáciu kernel funkcie. Hobijn et al. (2004, Tabuľka 3) pre potreby výberu hodnoty parametra  $M$  použili postup prvý krát navrhnutý v Newey – West (1994). Vypočítanú testovaciu štatistiku porovnávali s asymptotickými kritickými hodnotami, ktoré je možné nájsť taktiež v Hobijn et al. (2004, pozri Tabuľku 5). Pre úplnosť uvádzame, že v programe R existuje funkcia `ur.kpss()` z balíka `urca()`, kde sa ako kernel funkcia používa Bartlett. Počet oneskorení je však potrebné zadať manuálne. V Prílohe 2 uvádzame funkciu `KPSS.NEW()`, ktorá vráti testovaciu štatistiku podľa postupu opísaného v tejto kapitole, ako aj testovaciu štatistiku podľa funkcie `ur.kpss()`, kde hodnota parametra bandwidth je rovnaká akú sme získali použitím procedúry od Newey – West (1994).

### 5.3.4 Lee – Strazicich (2003, 2004) test

Vo svojej práci Lee – Strazicich (2003, 2004) navrhli test, ktorý umožňuje zohľadniť prípadný výskyt rôznych foriem štrukturálnych zlomov v deterministických zložkách časového radu, tak ako sme to naznačili v Kapitole 5.3.1. V tejto podkapitole prezentujeme pomerne všeobecnú špecifikáciu ich modelu, ktorá umožňuje zohľadniť výskyt štrukturálnych zlomov v konštante a v trende. Uvažujme o nasledujúcom procese:

$$y_t = \delta' z_t + x_t, \quad x_t = \beta x_{t-1} + \varepsilon_t \quad (5.115)$$

kde podobne ako v prípade ADF-GLS testu vektor  $z_t$  predstavuje deterministické premenné charakterizujúce konštantu a lineárny trend vrátane štrukturálnych zlomov. Uvažovali sme o modeli s dvoma štrukturálnymi zlomami, kde  $z_t = [1, t, D^1_t, DT^1_t, D^2_t, DT^2_t]^T$ ,

$D_t^1 = 1$ ,  $DT_t^1 = t - T_B^1$  pre  $t \geq T_B^1 + 1$  a 0 pre  $t \leq T_B^1$ ,  $D_t^2 = 1$ ,  $DT_t^2 = t - T_B^2$  pre  $t \geq T_B^2 + 1$  a 0 pre  $t \leq T_B^2$ . Pozorovanie, v ktorom dôjde k štruktúrnemu zlomu označujeme ako  $T_B^i$  ( $i = 1, 2$ ). Ďalej platí, že  $T_B^2 > T_B^1$ . Zlom sa nepripúšťa pre všetky možné pozorovania. Spravidla sa ignoruje určitý podiel prvých a posledných pozorovaní. Tomuto podielu hovoríme „pristrihnutie“ (z angl. *trim*), avšak budeme sa na to odkazovať len ako na *trim*. Hodnotu tohto parametra označíme ako  $h$  a je to spravidla jedno z čísel 0.05, 0.10, 0.15, 0.2 (ide o podiel k  $T$ ). Jeho voľba je do istej miery subjektívna. Určitým vodítkom pri voľbe tohto parametra by mal byť počet pozorovaní. Idea spočíva v tom, že  $p$  štruktúrnych zlomov rozdeľuje časový rad na  $p + 1$  rôznych režimov. Počet pozorovaní pre jeden režim by nemal byť príliš nízky. Parameter  $h$  znamená, že v jednom režime môže byť najmenej približne  $Th$  pozorovaní. Ak si stanovíme, že v jednom režime by nemalo byť menej ako povedzme 30 pozorovaní (napr. z dôvodu použitia Centrálnej limitnej vety), potom je možné odvodiť aj minimálnu hodnotu parametra *trim*. Testovacia štatistika je vypočítaná z nasledujúcej regresnej rovnice:

$$\Delta y_t = \delta' \Delta z_t + \phi \tilde{S}_{t-1} + \sum_{j=1}^k \gamma_j \Delta \tilde{S}_{t-j} + u_t \quad (5.116)$$

kde  $\tilde{S}_t = y_t - \tilde{\psi}_x - \mathbf{z}_t' \tilde{\delta}$  pre  $t = 2, 3, \dots, T$  a prvé pozorovanie je  $= 0$ . Ďalej  $\tilde{\delta}$  je vektor regresných koeficientov z regresie, kde závislou premennou je  $\Delta y_t$  a vektor nezávislých premenných je definovaný ako  $\Delta z_t$  ( $\Delta z_t = [1, \Delta D_t^1, \Delta DT_t^1, \Delta D_t^2, \Delta DT_t^2]$ ). A nakoniec,  $\tilde{\psi}_x$  je definované ako  $y_1 - \mathbf{z}_1' \tilde{\delta}$ . Nulovú hypotézu o prítomnosti jednotkového koreňa overujeme pomocou testovania významnosti koeficientu  $\phi = 0$ . Testovaciu štatistiku si môžeme označiť ako  $\tau^{LS}$ . Počet oneskorení  $k$  môžeme zvoliť pomocou sekvenčného postupu od Ng – Perron (1995), ktorý bol navrhnutý aj pri ADF-GLS teste. Ostáva nám určiť okamih štruktúrného zlomu v časovom rade. Lee – Strazicich (2003, 2004) navrhujú určiť zlom  $T_B^1$  a  $T_B^2$  na takom mieste, kde je hodnota testovacej štatistiky koeficientu  $\phi$  čo najmenšia. Na druhej strane výskum od Perron – Zhu (2005) naznačuje, že odhad štruktúrného zlomu pomocou kritéria minimalizácie sumy štvorcov rezíduí vedie ku konzistentnému odhadu a to bez ohľadu na to, či rezíduá z takéhoto modelu sú alebo nie sú stacionárne. Niektoré štúdie naznačujú, že tento spôsob odhadu štruktúrného zlomu vedie k testom (ne)stacionarity, ktoré majú lepšie štatistické vlastnosti, ako keď sa použije metóda (maxi)minimalizácie testovacej štatistiky príslušného testu (Kim – Perron, 2009).

Keďže kritické hodnoty závisia od polohy štrukturálneho zlomu, je potrebné ich pre každý časový rad vytvoriť pomocou simulácií. Tá sa uskutoční pomocou modelu (5.115), kde  $\beta = 1$ , takže simulácia časového radu prebieha pod „nulovou“ hypotézou, kde pre rezíduá platí  $\varepsilon_t \sim N(0, 1)$ . Hľadanie štrukturálneho zlomu a odhad kritických hodnôt pomocou simulácií je aj pre výpočtovú techniku časovo náročný proces. V programe R neexistuje funkcia, ktorá by tento test aplikovala. V Prílohe 3 uvádzame našu pracovnú verziu kódu.

### 5.3.5 Carrion-i-Silvestre – Sansó (2007) test

Test je obdobný KPSS testu s tým rozdielom, že umožňuje štrukturálny zlom v konštante, prípadne lineárnom trende pomocnej regresie. Budeme uvažovať o modeloch s konštantou a trendom, pričom v oboch prípadoch dovoľíme výskyt dvoch štrukturálnych zlomov. Ide o test, ktorý vychádza z Kwiatkowski et al. (1992), ale tiež Lee – Strazicich (2001), Presno – López (2003), no najmä Carrion-i-Silvestre – Sansó (2007), ktorí na rozdiel od predchádzajúcich autorov predstavili verziu s dvoma štrukturálnymi zlomami. Tie sa môžu odhadnúť pomocou minimalizácie sumy štvorcov rezíduí z nasledujúceho modelu (ide o model CC z Tabuľky 1 v Carrion-i-Silvestre – Sansó, 2007):

$$y_t = \alpha + \beta t + \sum_{i=1}^2 \gamma_i D_t^i + \sum_{i=1}^2 \delta_i DT_t^i + d \sum_{i=1}^t u_i + \varepsilon_t \quad (5.117)$$

kde podobne ako v predošlom Lee – Strazicich (2003, 2004) teste,  $D_t^1 = 1$ ,  $DT_t^1 = t - T_B^1$  pre  $t \geq T_B^1 + 1$  a 0 pre  $t \leq T_B^1$ ,  $D_t^2 = 1$ ,  $DT_t^2 = t - T_B^2$  pre  $t \geq T_B^2 + 1$  a 0 pre  $t \leq T_B^2$ . Pozorovanie, v ktorom dôjde k štrukturálnemu zlomu označujeme ako  $T_B^i$  ( $i = 1, 2$ ). Ďalej platí, že  $T_B^2 > T_B^1$ . Ďalej  $u_i$  a  $\varepsilon_i$  sú kovariančne stacionárne časové rady so strednou hodnotou 0 a  $d = \{0, 1\}$ . Testovacia štatistika je rovnaká ako pri KPSS teste. Môžeme si ju označiť ako ako  $w_{break}$ . Rozdiel oproti KPSS testu spočíva v tom, že rezíduá sa odhadnú z rovnice (5.117), kde  $d = 0$ . Odhad dlhodobého rozptylu je znovu uskutočnený cez (5.109). Carrion-i-Silvestre – Sansó (2007) pomocou simulácií dospeli ku vzťahom, pomocou ktorých je možné vypočítať asymptotické kritické hodnoty. Tie zároveň závisia od polohy štrukturálnych zlomov. Keďže podobne ako v predošlom prípade neexistuje funkcia v programe R, ktorá by tieto testy vykonávala, túto funkciu sme navrhli a jej pracovnú verziu je možné nájsť v Prílohe 4. Kritické hodnoty sme vypočítali pomocou vzťahov z Tabuľky 2 od Carrion-i-Silvestre – Sansó (2007).

### 5.3.6 Ostatné testy na stacionaritu

V softvéri R máme v knižnici `urca` k dispozícii aj ďalšie testy:

- PP test (Phillips – Perron, 1988): spolu s ADF, ADF-GLS a KPSS testom patrí medzi najčastejšie používané testy. Hlavnou výhodou tohto testu oproti ADF testu je, že PP test je robustný voči heteroskedasticite v chybových členoch a taktiež nie je nutné stanoviť explicitne počet oneskorení v testovacej regresii. Rovnako ako pri KPSS teste, vo funkcii sa nachádza argument `lags`, pomocou ktorého môžeme použiť klasické Schwertovo kritérium (`lags = "long"`) alebo upravené (`lags = "short"`), kde počet oneskorení je,  $\text{int}\{4(T/100)^{0.4}\}$ . V knižnici `urca` je výpočet tohto testu dostupný cez funkciu `ur.pp()`.
- SP test (Schmidt – Phillips, 1992): ide o menej bežný test, v ktorom v nulovej aj alternatívnej hypotéze sú zahrnuté aj deterministické zložky testu. Prostredníctvom funkcie `ur.sp()` máme k dispozícii dva typy testov – rho test a tau test (oba počítané na základe LM štatistiky).
- ZA test (Zivot – Andrews, 1992): je špecifický oproti ostatným testom z knižnice `urca` v tom, že umožňuje testovať stacionaritu pri výskyte štrukturálneho zlomu (bez exogénne určenej polohy) v konštante (`model = "intercept"`), v trende (`model = "trend"`), alebo aj v konštante a aj v trende (`model = "both"`). Výpočet tohto testu je v danej knižnici možný pomocou funkcie `ur.za()`.

Prehľad niektorých ďalších testov na prítomnosť jednotkového koreňa (angl. *unit root tests*, v ktorých je nulová hypotéza spravidla formulovaná tak, že časový rad obsahuje jednotkový koreň), a taktiež aj rôznych testov na overovanie stacionarity (angl. *stationarity tests*, nulová hypotéza je spravidla formulovaná tak, že časový rad je stacionárny), je uvedený v nasledujúcich tabuľkách. Nevenovali sme sa tzv. panelovým testom, ktoré sú v čase publikovania týchto materiálov pomerne dosť populárne, ale ich rámec presahuje nami stanovené ciele pre túto publikáciu. Tieto testy sa hodia najmä v prípadoch, ak je počet pozorovaní pre jeden časový rad obmedzený.

Tabuľka 11: Prehľad testov jednotkového koreňa a testov na stacionaritu

<b>Testy jednotkového koreňa s <math>H_0</math>: Časový rad obsahuje jednotkový koreň, <math>H_1</math>: Časový rad je stacionárny</b>
Dickey – Fuller test (1979); <i>DF test</i>
Dickey – Fuller test (1979, 1981); <i>ADF test</i>
Phillips – Perron (1988) - <i>PP test</i>
Elliott et al. (1996); <i>DF-GLS test</i>
Point optimal test; <i>P test</i>
Ng – Perron (2001); $MZ^d_{\alpha}$ test
Ng – Perron (2001); $MZ^d_t$ test
Ng – Perron (2001); $MSB^d$ test
Ng – Perron (2001); $MP^d_T$ test
<b>Testy stacionarity s <math>H_0</math>: Časový rad je stacionárny, <math>H_1</math>: Časový rad obsahuje jednotkový koreň</b>
Kwiatkowski et al. (1992); <i>KPSS test</i>
Leybourne – McCabe (1994)
<b>Testy jednotkového koreňa so štruktúrnymi zlomami v <math>H_0</math> a aj v <math>H_1</math></b>
Perron (1989) <sup>a</sup>
Perron – Vogelsang (1992) <sup>b</sup>
Clemente et al. (1998) <sup>c</sup>
Lee – Strazicich (2003) <sup>d</sup>
Lee – Strazicich (2004) <sup>e</sup>
Kim – Perron (2009) <sup>f</sup>
<b>Testy jednotkového koreňa so štruktúrnymi zlomami v <math>H_1</math></b>
Zivot – Andrews (1992) <sup>g</sup> ; <i>ZA test</i>
Lumsdaine – Papell (1997) <sup>h</sup> ; <i>LP test</i>
Perron (1997) <sup>i</sup>
<b>Testy na stacionaritu so štruktúrnymi zlomami</b>
Lee – Strazicich (2001) <sup>j</sup>
Carrion-i-Silvestre – Sansó (2007) <sup>k</sup>
<b>Špeciálne testy jednotkového koreňa</b>
Kapetanios et al. (2003) <sup>l</sup> ; <i>KSS test</i>
Elliott – Jansson (2003) <sup>m</sup> ; <i>EJ test</i>

Zdroj: Lyócsa et al. (2011)

Pozn.: a – Zlomky sú špecifikované exogénne.  $H_0$ : Časový rad obsahuje jednotkový koreň so štruktúrnym zlomom,  $H_1$ : Časový rad je stacionárny so štruktúrnym zlomom. Navrhnuté sú tri modely so štruktúrnymi zlomami: A) zlom v strednej hodnote časového radu, B) zlom v sklone lineárneho trendu a C) súčasná zmena (zlom) v strednej hodnote a aj v trende. Pokiaľ štruktúrný zlom vznikol náhle, používa sa tzv. AO model (z angl. additive outlier model), pokiaľ zlom vzniká postupne, používa sa IO model (z angl. innovation outlier model). Tieto dva modely slúžia na popísanie mechanizmu presunu štruktúrných zlomov. Jednoduchý príklad AO modelu s dvomi zlomami:  $yt = \alpha + \delta 1DU_{t,1} + \delta 2DU_{t,2} + \epsilon_t$ , kde  $DU_t$  je umelá premenná, pričom  $DU_t = 1$  pre  $t > T_b + 1$ , 0 v opačnom prípade a  $T_b$  je dátum výskytu zlomu. Tento model predpokladá dva zlomy v strednej hodnote časového radu. Jednoduchý príklad na IO model:  $yt = \alpha + \omega 1DT_{t,1} + \omega 2DT_{t,2} + \delta 1DU_{t,1} + \delta 2DU_{t,2} + \rho y_{t-1} + \epsilon_t$ , kde  $DT_t$  je umelá premenná, pričom  $DT_t = 1$  ak  $t = T_b + 1$  a 0 v opačnom prípade. Zároveň musí platiť  $|\rho| < 1$ . b – Zlomky sú determinované endogénne. c – Clemente et al. (1998) uvažovali len s netrendujúcimi časovými radmi (AO a IO modely).  $H_0$ : Časový rad obsahuje jednotkový koreň s dvomi štruktúrnymi zlomami,  $H_1$ : Časový rad je stacionárny s dvomi štruktúrnymi zlomami. d – Lee – Strazicich (2001, 2003, 2004) uvažovali len s modelmi so zlomom v strednej hodnote alebo so zlomom v strednej hodnote a súčasne aj v trende. V Lee – Strazicich (2003, 2004) ich metodológia implikuje, že zamietnutím nulovej hypotézy je časový rad stacionárny so štruktúrnymi zlomami, bez ohľadu na to, či sa zlomy vyskytovali v nulovej hypotéze o existencii jednotkového koreňa. Z toho dôvodu sme ich zaradili do tejto kategórie. e –  $H_0$ : Časový rad obsahuje jednotkový koreň,  $H_1$ : Časový rad je stacionárny s dvomi štruktúrnymi zlomami. f –  $H_0$ : Časový rad obsahuje jednotkový koreň,  $H_1$ : Časový rad je stacionárny s jedným štruktúrnym zlomom. g –  $H_0$ : Časový rad obsahuje jednotkový koreň so štruktúrnym zlomom,  $H_1$ : Časový rad je stacionárny so štruktúrnym zlomom. h –  $H_0$ : Časový rad obsahuje jednotkový koreň,  $H_1$ : Časový rad je stacionárny s jedným štruktúrnym zlomom. i –  $H_0$ : Časový rad obsahuje jednotkový koreň,  $H_1$ : Časový rad je stacionárny s dvomi štruktúrnymi zlomami v trende. j –  $H_0$ : Časový rad obsahuje jednotkový koreň,  $H_1$ : Časový rad je stacionárny s jedným štruktúrnym zlomom. k –  $H_0$ : Časový rad je stacionárny so štruktúrnym zlomom,  $H_1$ : Časový rad obsahuje jednotkový koreň. l –  $H_0$ : Časový rad je stacionárny s dvomi štruktúrnymi zlomami,  $H_1$ : Časový rad obsahuje jednotkový koreň. m –  $H_0$ : Časový rad obsahuje jednotkový koreň,  $H_1$ : Časový rad je globálne stacionárny ESTAR proces (z angl. Exponential Smooth Transition AutoRegressive). n – test používa stacionárne kovariáty (z angl. covariates), bližšie napr. v Amara – Papell (2006),  $H_0$ : Časový rad obsahuje jednotkový koreň,  $H_1$ : Časový rad je stacionárny.

Tabuľka 12: Prehľad panelových testov jednotkového koreňa a testov na stacionaritu

<b>Testy jednotkového koreňa s <math>H_0</math>: Všetky časové rady obsahujú jednotkový koreň, <math>H_1</math>: Všetky časové rady sú stacionárne</b>
Levin – Lin (1993); <i>LL test</i>
Levin – Lin – Chu (2002); <i>LLC test</i>
Harris – Tzavalis (1999)
Breitung – Das (2005) <sup>*a</sup>
<b>Testy jednotkového koreňa s <math>H_0</math>: Všetky časové rady obsahujú jednotkový koreň, <math>H_1</math>: Niektoré časové rady sú stacionárne</b>
Im – Pesaran – Shin (2003); <i>IPS test</i>
Maddala – Wu (1999)
Chang (2002) <sup>*</sup>
Pesaran (2007) <sup>*</sup>
Phillips – Sul (2003) <sup>*</sup>
Moon – Perron (2004) <sup>*</sup>
Choi (2006) <sup>*</sup>
Pesaran – Smith – Yamagata (2009) <sup>*</sup>
Taylor – Sarno (1998) <sup>*b</sup> ; <i>MADF test</i>
<b>Testy stacionarity s <math>H_0</math>: Všetky časové rady sú stacionárne, <math>H_1</math>: Niektoré časové rady obsahujú jednotkový koreň</b>
Choi (2001)
Hadri (2000)
Harris – Leybourne – McCabe (2004) <sup>*</sup>
Demetrescu – Hassler – Tarcolea (2010) <sup>*</sup>
<b>Testy jednotkového koreňa so štruktúrnymi zlomami</b>
Lee – Im – Tieslau (2005) <sup>*c</sup> ; <i>LIT test</i>
Hadri – Rao (2008) <sup>*d</sup>
Im – Lee – Tieslau (2010) <sup>*e</sup> ; <i>ILT test</i>
<b>Špeciálne panelové testy jednotkového koreňa</b>
Taylor – Sarno (1998) <sup>*f</sup> ; <i>JLR test</i>
Breuer – McNown – Wallace (2001) <sup>*g</sup>

Zdroj: Lyócsa et al. (2011)

Pozn.: Symbol \* označuje tie testy, ktoré berú do úvahy prierezovú závislosť časových radov (v angl. *cross-sectional dependence*) v rámci panela, tzv. testy druhej generácie. a – Používa sa model SUR (z angl. *Seemingly Unrelated Regression*). b –  $H_0$ : všetky časové rady v paneli obsahujú jednotkový koreň,  $H_1$ : niektoré časové rady v systéme sú stacionárne. c – Test povoľuje zmenu v strednej hodnote časového radu.  $H_0$ : všetky časové rady v paneli obsahujú jednotkový koreň,  $H_1$ : niektoré časové rady v systéme sú stacionárne so štruktúrnym zlomom. d – Test umožňuje zlom v strednej hodnote a v trende časových radov.  $H_0$ : všetky časové rady sú stacionárne so štruktúrnym zlomom (použitie rôznych modelov obsahujúcich štruktúrne zlomy,  $H_1$ : niektoré časové rady v paneli obsahujú jednotkový koreň. e – Test umožňuje zlom v strednej hodnote a v trende časových radov.  $H_0$ : všetky časové rady v paneli obsahujú jednotkový koreň,  $H_1$ : niektoré časové rady v paneli sú stacionárne so štruktúrnymi zlomami. f – Podľa Taylor – Sarno (1998): „Engle – Granger (1987) ukázali, že v systéme  $N I(1)$  časových radov môže byť najviac  $N-1$  kointegračných vektorov. Preto ak zamietneme nulovú hypotézu, že v systéme je menej ako  $N$  kointegračných vektorov medzi  $N$  časovými radmi, tak to je ekvivalentné k zamietnutiu hypotézy o nestacionárnosti všetkých časových radov“. Implikáciou je, že ak sa v paneli vyskytuje  $N$  kointegračných vektorov, tak všetky časové rady v danom paneli sú stacionárne. Túto myšlienku autori využili pri rozšírení MADF testu. g – V tomto teste je možné rozlíšiť, ktoré časové rady v paneli sú stacionárne, a ktoré nie.

## 5.4 Problém falošnej regresie

(Ne)stacionarita časových radov je dôležitým konceptom najmä pri regresných modeloch, do ktorých keď zahrnieme nestacionárne premenné, môže dôjsť k tzv. falošnej

regresii.<sup>31</sup> Granger – Newbold (1974) pomocou Monte Carlo simulácii ukázali, že v regresných modeloch s nestacionárnymi premennými (presnejšie  $I(1)$ ), dochádza k vysokej chybe I. druhu<sup>32</sup>, a to napriek tomu, že tieto dve premenné sú v skutočnosti nezávislé. Medzi inými formulovali praktické pravidlo, pomocou ktorého je možné odhaliť prítomnosť falošnej regresie – ak je koeficient determinácie vyšší ako Durbin – Watsonova štatistika, tak zrejme ide o falošnú regresiu.

Ukážeme si to na jednoduchom príklade. Uvažujme o nasledujúcom časovom rade vygenerovaného z jednoduchého autoregresného procesu prvého rádu. Pre dve rôzne hodnoty autoregresného parametra si vygenerujeme dva rôzne nestacionárne časové rady s 500 pozorovaniami:

$$\text{AR\_1: } y_t = 1.05y_{t-1} + e_t \quad (5.118)$$

$$\text{AR\_2: } x_t = 1.1x_{t-1} + w_t \quad (5.119)$$

Chybové členy  $e_t$  a  $w_t$  boli generované nezávisle z  $N(0, 1)$ . Keďže časové rady sú generované náhodne, nemal by medzi nimi existovať žiadny vzťah. Pre naše potreby budeme vychádzať z jednoduchého regresného modelu, v ktorom jeden časový rad budeme vysvetľovať pomocou druhého:

$$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t \quad (5.120)$$

kde  $\varepsilon_t$  je poruchový člen. Odhad regresného parametra  $\beta_1$  by nemal byť významný a koeficient determinácie by mal byť nízky<sup>33</sup>.

```
> library (lmtest)
> x <- rnorm(500,0,1)
> AR_1 <- filter(x, filter = c(1.05), method = "recursive")
> AR_2 <- filter(x, filter = c(1.1), method = "recursive")
-----
> model <- lm(AR_1 ~ AR_2); summary (model)

Call:
lm(formula = AR_1 ~ AR_2)

Residuals:
    Min       1Q   Median       3Q      Max
-7.702e+10 -3.885e+09 -3.884e+09 -3.420e+09  4.795e+10

Coefficients:
```

<sup>31</sup> To isté samozrejme platí aj pri výpočte korelácií medzi dvoma nestacionárnymi časovými radmi. V takom prípade môžeme hovoriť o falošnej korelácií.

<sup>32</sup> Teda že je zamietnutá hypotéza o neprítomnosti vzťahu medzi takými dvoma premennými.

<sup>33</sup> Keďže pôjde o náhodne generované časové rady, výsledky budú samozrejme odlišné po zopakovaní postupu čitateľom.

```

              Estimate Std. Error t value Pr(>|t|)
(Intercept) 3.885e+09  5.690e+08   6.828 2.52e-11 ***
AR_2        2.336e-10  3.559e-12  65.625 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.245e+10 on 498 degrees of freedom
Multiple R-squared:  0.8963,    Adjusted R-squared:  0.8961
F-statistic:  4307 on 1 and 498 DF,  p-value: < 2.2e-16
-----
> dwtest(model, alternative = "two.sided")

              Durbin-Watson test

data:  model
DW = 0.0186, p-value < 2.2e-16
alternative hypothesis: true autocorrelation is not 0

```

Napriek tomu, že ide o náhodne vygenerované časové rady, tak odhad koeficientu  $\beta_1$  je vysoko štatisticky významný a koeficient determinácie je vysoký (0.8961). Na základe výsledkov z takého regresného modelu by sme dospeli k záveru, že medzi premennými je štatisticky významná závislosť. Vieme, že dané časové rady sú nestacionárne (poznáme dáta generujúci proces) a keď sa pozrieme na Durbin – Watson štatistiku, tak môžeme vidieť, že je veľmi blízka nule. Aj keby sme teda nevedeli, že ide o nestacionárne premenné, na základe pravidla o vysokom koeficiente determinácie a nižšej Durbin – Watson štatistike by sme mali mať podozrenie, že ide o falošnú regresiu.

K rovnako nesprávnemu výsledku by sme dospeli v prípade, keď by nás zaujímala korelácia medzi týmito dvoma premennými. Výsledkom by bol vysoký a štatisticky významný koeficient korelácie. Odmocninou z koeficientu determinácie vieme jednoducho získať Pearsonov korelačný koeficient (0.9467), prípadne vieme na ukážku falošnej korelácie použiť funkciu `cor.test()`.

```

> cor.test(AR_1, AR_2, alternative = "two.sided", conf.level =
  0.95)

              Pearson's product-moment correlation

data:  AR_1 and AR_2
t = 65.6247, df = 498, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.9368451 0.9551497
sample estimates:
      cor
0.9467572

```

Na základe uvedeného príkladu sme chceli demonštrovať významnosť konceptu stacionarity pri skúmaní vzájomných vzťahov medzi časovými radmi. Zrejme je čitateľovi v tomto bode jasné, že nie je vždy vhodné do výpočtu korelácie alebo regresných modelov zahrnúť časové rady, ktoré nie sú stacionárne. Jednou z významných „výnimiek“ je situácia, ak sú časové rady kointegrované. Tejto problematike sa však nebudeme bližšie venovať. Iný prípad uvádza Hamilton (1994, s. 561). Ak sa do modelu pridajú oneskorené hodnoty závislej, ako aj nezávislej premennej tak, aby rezíduá boli  $I(0)$ , potom OLS odhad modelu vedie ku konzistentným odhadom regresných koeficientov a je možné vykonať aspoň  $t$ -testy jednotlivých koeficientov regresného modelu.

Pozrime sa ešte na spomínané praktické pravidlo, pomocou ktorého môžeme na prvý pohľad z výsledkov regresie predpokladať, že ide o falošnú regresiu. Pracovať budeme s takými istými náhodne vygenerovanými časovými radmi a pomocou jednoduchej simulácie sa pokúsime zistiť, či budú vychádzať štatisticky významné výsledky medzi náhodne vygenerovanými nestacionárnymi časovými radmi (opäť využijeme jednoduchý lineárny regresný model). Taktiež sa pozrieme na to, aký je vzťah medzi koeficientom determinácie a Durbin – Watson štatistikou. V zásade teda pôjde o taký istý príklad ako v predchádzajúcom texte, avšak teraz celý postup zopakujeme 500 krát.<sup>34</sup>

```

> rm(list = ls())
> library(lmtest)
-----
> r2s <- c()
> dwt_stats <- c()
> dwt_pvals <- c()
> tstats <- c()
> pvs <- c()
-----
> for (c in 1:500) {
+ x <- filter(rnorm(100,0,1), filter = c(1.05), method =
+ "recursive")
+ y <- filter(rnorm(100,0,1), filter = c(1.1), method =
+ "recursive")

+   model = lm(y~x)
+   dwt = dwtest(model)
+   sumr <- summary(model)
+   n <- length(x)
+   es <- resid(model)
+   s <- sqrt(sum(es^2)/(n-2))
+   b1 <- model$coefficients[2]

```

<sup>34</sup> O niečo zložitejšie simulácie z danej oblasti sú uvedené v práci Baumöhl – Lyócsa (2009).

```

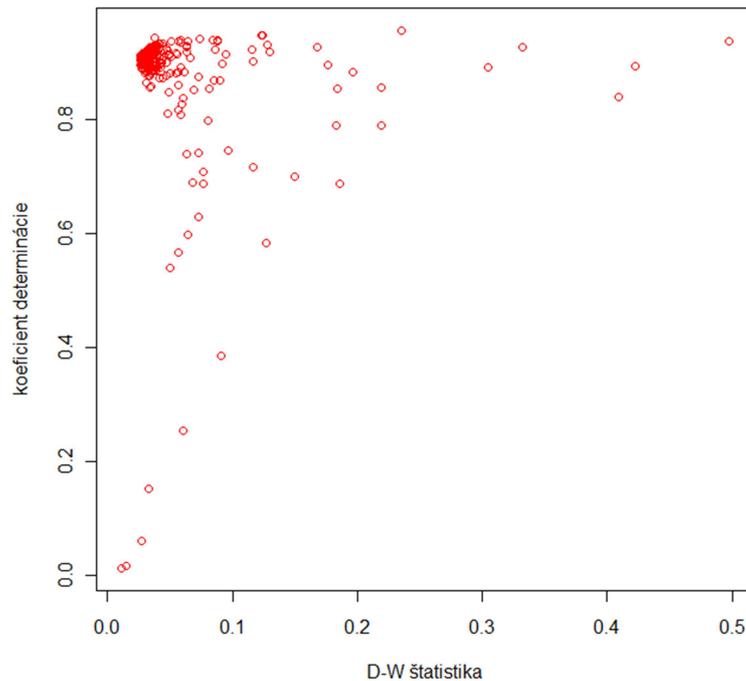
+   SE <- s/sqrt(sum((x-mean(x))^2))
+   t <- ((b1-(0))/SE)
+   pv <- 2*pt(abs(t), n-2, lower.tail = FALSE)
+   r2s <- c(r2s, sumr$r.squared)
+   dwt_stats <- c(dwt_stats, dwt$statistic)
+   dwt_pvals <- c(dwt_pvals, dwt$p.value)
+   tstats <- c(tstats, t)
+   pvs <- c(pvs, pv)
+ }
-----
> df <- data.frame(r2s, dwt_stats, dwt_pvals, tstats, pvs)
> summary(df)

      r2s                dwt_stats                dwt_pvals
Min.   :0.002877      Min.   :0.01049      Min.   :1.908e-24
1st Qu.:0.900486      1st Qu.:0.02876      1st Qu.:4.069e-24
Median :0.908619      Median :0.03043      Median :4.433e-24
Mean   :0.889159      Mean   :0.04191      Mean   :3.248e-19
3rd Qu.:0.912703      3rd Qu.:0.03694      3rd Qu.:6.188e-24
Max.   :0.965143      Max.   :0.39125      Max.   :9.480e-17

      tstats                pvs
Min.   : -52.091      Min.   :3.055e-73
1st Qu.: -31.284      1st Qu.:1.081e-53
Median : -17.615      Median :1.018e-52
Mean   :  -1.128      Mean   :1.193e-03
3rd Qu.:  31.090      3rd Qu.:6.666e-51
Max.   :  42.438      Max.   :5.961e-01

```

Pri 500 iteráciách môžeme vidieť, že vo väčšine prípadov boli koeficienty  $\beta_i$  štatisticky významné (v 499 prípadoch na hladine významnosti 1 %). Keďže zamietame nulové hypotézy (o existencii vzťahu medzi náhodne generovanými premennými), ktoré by sme zamietat' nemali, chyba I. druhu je do značnej miery vysoká. Taktiež pravidlo o vysokom koeficiente determinácie a nízkej Durbin – Watson štatistike sa javí ako opodstatnené pri identifikácii falošnej regresie. Pre lepšiu prehľadnosť tieto dve kritériá uvedieme aj v  $x$ - $y$  grafe (kde Durbin – Watson štatistika je označená ako D-W štatistika).



Obrázok 22: Falošná regresia – vzťah medzi koeficientom determinácie a D-W štatistikou

*Zdroj: výstup zo softvéru R*

Z uvedeného grafu je zrejmé, ako vzniklo toto praktické odporúčanie na odhalenie falošnej regresie. Najpodstatnejšie je samozrejme exaktné testovanie časových radov na stacionaritu. Pokiaľ sú premenné v regresnom modeli (prípadne pri počítaní korelácie) stacionárne, tak výskyt falošnej regresie by mal byť podstatne menším problémom.

## 5.5 Transformácia časových radov

Jedným zo spôsobov, ako sa prípadne vyhnúť problému falošnej regresie, je transformácia nestacionárnych časových radov na stacionárne. V závislosti od toho, aký typ transformácie musíme uplatniť na získanie stacionárnych časových radov, môžeme hovoriť o stacionárnych premenných získaných (A) diferencovaním (v angl. *difference stationary*), (B) odstránením trendu (v angl. *trend stationary*) alebo (C) zahrnutím štrukturálnych zlomov v trende (v angl. *broken trend stationary*).

Najbežnejšie sa stretávame s diferencovaním. V mnohých prípadoch ide o prirodzený krok. V ekonómii nás totiž často zaujíma miera zmeny nejakej premennej, ako napríklad pri inflácii (zmena cenovej hladiny). Pripomenieme, že pokiaľ je nutné časový rad diferencovať  $d$  krát, aby sa z neho stal  $I(0)$  časový rad, tak hovoríme, že časový rad je integrovaný  $d$ -tého

rádu, čo značíme ako  $I(d)$ . Časový rad, ktorý nemá jednotkový koreň môžeme potom označiť ako  $I(0)$ . Najčastejšie sa stretávame s prípadom, že časový rad je  $I(1)$ .<sup>35</sup>

Diferencie prvého stupňa ( $\Delta x_t = x_t - x_{t-1}$ ) obvykle postačujú na získanie  $I(0)$  časového radu. Môžeme sa stretnúť aj s diferenciami druhého stupňa  $\Delta^2 x_t = \Delta x_t - \Delta x_{t-1}$ , ktoré sú vlastne lineárne zmeny získané z lineárnych zmien (po prvých diferenciách).<sup>36</sup> S vyšším stupňom diferencií sa už v ekonometrii veľmi nestretávame, keďže jednak pri každej diferencii strácame pozorovania, ale čo je zrejme podstatnejšie, stráca sa ekonomická interpretácia takto získaných časových radov.

Pred samotným diferencovaním sa odporúča najprv hodnoty logaritmovať. Dôvodom je hneď niekoľko:

- Môžu sa tým odstrániť možné „nelineárnosti“. Najmä ekonomické premenné zvyknú obsahovať exponenciálny trend. Napríklad HDP krajiny, ktoré rastie každý rok približne o 3 %, bude obsahovať exponenciálny trend. Po logaritmovaní tento trend odstránime a získame časový rad, ktorý rastie lineárne. Vzhľadom na to, že väčšinou používame lineárne štatistické modely, ide o vhodnú transformáciu údajov. S touto transformáciou sa môžeme stretnúť vo väčšine empirických prác.
- Ekonomické časové rady sú často pravostranne zošikmené, pričom táto transformácia zníži mieru zošikmenia, čo v závislosti od použitých štatistických metód môže pomôcť pri odhadoch parametrov modelu.<sup>37</sup>
- V ekonómii má množstvo časových radov heteroskedastický charakter, čo sa môže prejavovať tak, že so zmenou úrovne premennej sa mení aj rozptyl hodnôt. Spravidla sa rozptyl hodnôt zvyšuje s rastom absolútnej hodnoty premennej. Za určitých okolností, logaritmická transformácia tento efekt zmierni (pozri Box – Cox, 1964). Taktiež spravidla zmierni vplyv extrémnych hodnôt.

---

<sup>35</sup> Definícia predpokladá, že  $d$  je celé číslo. Môžeme sa stretnúť aj s tzv. *fractional integration*, kde tento predpoklad nie je nutný. Kvôli zameraniu tejto publikácie je bližšia charakteristika zbytočná, čitateľ však môže danú problematiku nájsť napr. v práci Mills (1999).

<sup>36</sup> Diferencie prvého stupňa sa obvykle pre zjednodušenie nazývajú „prvé diferencie“ a diferencie druhého stupňa „druhé diferencie“.

<sup>37</sup> Len ako príklad uvedieme štandardný  $t$ -test, ktorého testovacia štatistika sa (za predpokladu platnosti nulovej hypotézy) riadi Studentovým  $t$  rozdelením pravdepodobnosti. Jedným z predpokladov je, že realizácie majú pochádzať z normálneho rozdelenia pravdepodobnosti. Je však známe, že ak tomu tak nie je, pri dostatočne veľkej vzorke je možné rozdelenie testovacej štatistiky možné aproximovať Studentovým  $t$ -rozdelením. Okrem veľkosti vzorky je však dôležité aj rozdelenie pravdepodobnosti. Čím je toto rozdelenie symetrickejšie, o to menšia vzorka je potrebná na to, aby bolo možné rozdelenie testovacej štatistiky aproximovať Studentovým  $t$  rozdelením. Logaritmická transformácia pri jednostranne zošikmených realizáciách tomu napomáha.

- Ďalej vieme, že ak platí  $\Delta x_t \ll 1$ , potom  $\ln(1 + \Delta x_t) \approx \Delta x_t$  (do  $\Delta x_t = 0.1$  je chyba menej ako 5 %, pri  $\Delta x_t = 0.15$  je chyba menej ako 7 %). Uvedená vlastnosť sa využíva pri interpretácii regresných koeficientov. Ak je nezávislá premenná transformovaná pomocou prirodzeného logaritmu, potom regresný koeficient v jednoduchom lineárnom modeli získava interpretáciu elasticity (ak aj závislá premenná je transformovaná). Zvýšenie nezávislej premennej o určité (nízke) percento vyvolá vždy rovnakú zmenu závislej premennej.
- Pre nás však považujeme zrejme za najdôležitejšiu nasledujúcu vlastnosť. Ak zistíme, že časový rad má jednotkový koreň, potrebujeme uskutočniť jeho diferenciu aby sme dostali časový rad bez jednotkového koreňa. Taktiež nás často bude zaujímať stupeň integrácie časového radu. Z vlastností vyššie vyplýva, že ak by sme pracovali s pôvodnými (nie logaritmovanými) hodnotami, rovnako veľké diferencie nemusia znamenať rovnako veľké zmeny. Interpretácie by tak mohli byť do značnej miery skreslené. Ak by sme namiesto toho vypočítali percentuálne zmeny, ktoré by sa ukázali byť bez jednotkového koreňa, nejde už o diferencie pôvodného časového radu, a preto nám takýto postup nemôže povedať nič o stupni integrácie pôvodného časového radu. Ak však použijeme logaritmické diferencie, pre malé zmeny budú výsledky porovnateľné s percentuálnymi zmenami a zároveň budeme schopní zistiť stupeň integrácie pôvodného logaritmovaného časového radu.

Ekonomické premenné zvyknú rásť exponenciálne, čo znamená, že pre premennú  $X$  máme rovnicu rastu definovanú ako  $X_t = (1 + g_t) X_{t-1}$  v diskretnom prípade, alebo  $X_t = X_{t-1}e^{g_t}$  v spojitom prípade, kde  $g_t$  je miera rastu medzi dvoma po sebe nasledujúcimi obdobiami. Logaritmickými diferenciami dostávame  $\ln(X_t) - \ln(X_{t-1}) = \ln(1 + g_t) \approx g_t$  v diskretnom prípade a  $\ln(X_t) - \ln(X_{t-1}) = g_t$  v spojitom prípade.<sup>38</sup>

Keď premenné obsahujú zjavný trend (čo je v ekonómii bežné), tak je tento trend neraz nutné zohľadniť. Časové rady obsahujúce trend totiž nemôžu byť stacionárne, keďže ich stredná hodnota nie je konštantná. V jednoduchom lineárnom regresnom modeli by sme mohli prítomnosť deterministického trendu zohľadniť vhodnou špecifikáciou modelu, t. j. pridaním trendovej zložky. Odstránenie trendu sa najjednoduchšie realizuje cez regresný model s konštantou a s trendom (či už lineárnym alebo vyšším polynomickým stupňom).

<sup>38</sup> Kočenda – Černý (2007).

Rezíduá z takéhoto regresného modelu tak predstavujú časový rad po odstránení trendu, ktorý už môže byť stacionárny. Aby sme vedeli, ktorú z metód vybrať, je vhodné najprv časové rady testovať pomocou vyššie spomínaných testov.

Na ukážku si vygenerujeme jeden časový rad typu náhodnej prechádzky.<sup>39</sup> Vieme, že takto vygenerovaný časový rad stacionárny nebude, čo je zrejme aj z jeho grafickej vizualizácie. Pod vygenerovaným procesom typu náhodnej prechádzky si však môžeme predstaviť nejaký ekonomický časový rad (napr. vývoj cien finančných aktív na trhu). Z danej grafickej podoby musíme pripustiť, že veľa ekonomických premenných môže mať podobný vývoj. Keďže však pri týchto premenných nepoznáme dáta generujúci proces, tak exaktné testovanie stacionarity je vždy vhodné.

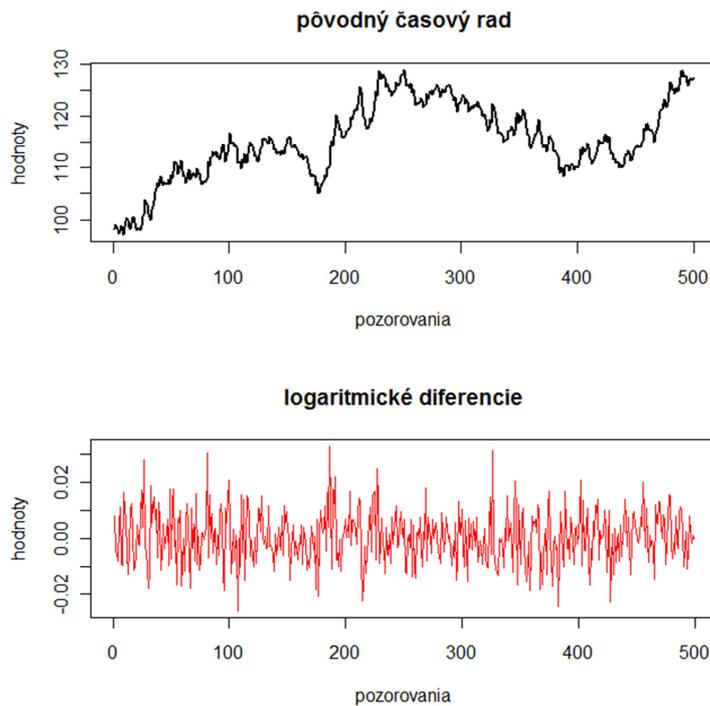
```
> par(mfrow = c(2,1))
> w <- rnorm(500,0,1)
> x <- cumsum(w)
> wd <- 0.1*(1:length(w))+x+100
> diff <- diff(log(wd))
> plot(wd, ylab = "hodnoty", xlab = "pozorovania", main =
  "pôvodný časový rad", type = "l",lwd = 2)
> plot(diff, ylab = "hodnoty", xlab = "pozorovania", main =
  "logaritmické diferencie", type = "l",lwd = 1, col = "red")
```

Na druhej strane ak sa pozrieme na logaritmické diferencie, tak tie sa už javia ako stacionárne. Aj po otestovaní na prítomnosť jednotkového koreňa pomocou ADF testu môžeme vidieť, že pôvodný časový rad (náhodná prechádzka) stacionárny nie je, avšak jeho logaritmické diferencie už za stacionárne môžeme považovať (nulovú hypotézu o existencii jednotkového koreňa zamietame pri  $p$ -hodnote  $6.411689e-32$ ).

```
> library(urca)
> p_ADF_wd <- punitroot(ur.df(wd, type = "trend", lags =
  1)@teststat[1], N = length(wd), trend = "ct", statistic =
  "t"); p_ADF_wd
[1] 0.4270368
> p_ADF_diff <- punitroot(ur.df(diff, type = "trend", lags =
  1)@teststat[1], N = length(diff), trend = "ct", statistic =
  "t"); p_ADF_diff
[1] 6.411689e-32
```

---

<sup>39</sup> Ku vygenerovanému časovému radu sme ešte prirátali hodnotu 100 (čitateľ môže pripočítať  $0.01 + \max(x_t)$ ), aby sme zabezpečili, že hodnoty budú kladné, keďže sme ich chceli následne logaritmovať.



Obrázok 23: Časový rad typu random walk a jeho logaritmicke diferencie

*Zdroj: výstup zo softvéru R*

Z reálnych ekonomických časových radov si na ukážku zoberieme uzatváracie ceny akciových indexov. V práci Baumöhl – Lyócsa (2009) sa pracovalo s dennými uzatváracími cenami indexov krajín V4 (Maďarsko – BUX, Poľsko – WIG, Česká republika – PX, Slovensko – SAX) za obdobie od 1. septembra 1999 do 1. septembra 2009. Stacionarita sa testovala s využitím 4 testov, pričom výsledky boli dosť jednoznačné. Uzatváracie ceny ani jedného z indexov nebolo možné považovať za stacionárne, avšak ich logaritmicke diferencie (tzv. spojité výnosy) už áno. Výsledky sú zobrazené v nasledujúcej tabuľke.

Tabuľka 13: Akciové index krajín V4 – testovanie stacionarity

	Uzatváracie ceny		Logaritmické diferencie		Uzatváracie ceny		Logaritmické diferencie	
	<b>ADF test</b>				<b>ADF-GLS test</b>			
Index	c	ct	c	ct	c	ct	c	ct
BUX	NR	NR	R	R	NR	NR	R	R
WIG	NR	NR	R	R	NR	NR	R	R
PX	NR	NR	R	R	NR	NR	R	R
SAX	NR	NR	R	R	NR	NR	R	R
	<b>ZA test</b>				<b>PP test</b>			
Index	c	ct	c	ct	c	ct	c	ct
BUX	NR	NR	R	R	NR	NR	R	R
WIG	NR	NR	R	R	NR	NR	R	R
PX	NR	NR	R	R	NR	NR	R	R
SAX	NR	NR	R	R	NR	NR	R	R

Zdroj: Baumöhl – Lyócsa (2009)

Pozn.: „c“ označuje zahrnutie konštanty, „ct“ označuje zahrnutie konštanty aj trendu. „R“ znamená zamietnutie nulovej hypotézy (časový rad je nestacionárny) a „NR“ znamená nezamietnutie nulovej hypotézy.

Pre úplnosť ešte uvedieme výsledky z jednoduchého regresného modelu, kde ako vysvetľovaná premenná vystupuje akciový index a ako vysvetľujúca nejaký iný index. Tu opäť vidno, aké sú rozdielne výsledky v prípade výskytu falošnej regresie. V ľavej časti tabuľky sú do regresných modelov zahrnuté indexy v uzatváracích cenách, ktoré sú nestacionárne. Môžeme vidieť, že všetky koeficienty sú štatisticky významné a koeficienty determinácie sú dosť vysoké. Záverom takejto analýzy by bolo, že akciové index krajín V4 sú vo veľkej miere integrované. Tieto závery by však boli samozrejme nesprávne. V pravej časti tabuľky sú do regresných modelov zahrnuté spojité výnosy (logaritmické diferencie) týchto indexov a tu už sú koeficienty determinácie do značnej miery nižšie. Taktiež pri indexe SAX sú vzťahy nevýznamné, čo vzhľadom na nízku efektívnosť slovenského kapitálového trhu môžeme považovať za očakávaný výsledok.

Tabuľka 14: Akciové index krajín V4 – výsledky z regresných modelov

		Vysvetľovaná premenná v regresnom modeli							
		Uzatváracie ceny				Logaritmické diferencie			
		BUX	WIG	PX	SAX	BUX	WIG	PX	SAX
BUX	<i>t</i> -test		0.0000	0.0000	0.0000		0.0000	0.0000	0.2867
	R <sup>2</sup>	-	0.9222	0.9835	0.8286	-	0.3124	0.3411	0.0013
	DW		0.0126	0.0643	0.0117		2.0063	2.0578	2.0584
WIG	<i>t</i> -test	0.0000		0.0000	0.0000	0.0000		0.0000	0.7280
	R <sup>2</sup>	0.9222	-	0.9293	0.6771	0.3124	-	0.3446	0.0001
	DW	0.0131		0.0115	0.0043	2.0026		1.9879	2.0640
PX	<i>t</i> -test	0.0000	0.0000		0.0000	0.0000	0.0000		0.7537
	R <sup>2</sup>	0.9835	0.9293	-	0.8293	0.3411	0.3446	-	0.0001
	DW	0.0646	0.0114		0.0095	2.0158	1.9495		2.0646
SAX	<i>t</i> -test	0.0000	0.0000	0.0000		0.2927	0.7257	0.7577	
	R <sup>2</sup>	0.8286	0.6771	0.8293	-	0.0013	0.0001	0.0001	-
	DW	0.0127	0.0049	0.0102		1.9226	1.932	1.9709	

Zdroj: Baumöhl – Lyócsa (2009)

Pozn.: *t*-test označuje *t*-štatistiku pri vysvetľujúcej premennej v regresnom modeli; R<sup>2</sup> je koeficient determinácie a DW označuje Durbin – Watson testovaciu štatistiku. Regresné modely sú kvantifikované pomocou matíc konzistentných na prítomnosť autokorelácie a heteroskedasticity HAC.

## 5.6 Kauzalita v Grangerovom zmysle

Overovať kauzalitu v ekonómii je prakticky nemožné. Väčšinou sme sa pri analýze vzťahu medzi ekonomickými premennými obmedzili len na testovanie korelácie medzi nimi. Pri časových radoch je však do istej miery možné skúmať aj kauzalitu, konkrétne však ide len o tzv. kauzalitu v Grangerovom zmysle. V ekonómii sa pojem kauzalita veľmi často viaže práve s tzv. Grangerovou kauzalitou. Granger (1980) publikoval prehľadovú štúdiu, v ktorej bolo jeho cieľom zdefinovať pojem kauzalita tak, aby jeho definíciu mohli zmysluplne používať ekonómovia v empirickom výskume. Pri definovaní Grangerovej kauzality tak budeme v nasledujúcich odsekoch vychádzať najmä z práce Granger (1980).

Logika Grangerovho modelu kauzality je priamočiara. V prípade jednoduchého lineárneho regresného modelu, závislosť jednej premennej od druhej nemusí indikovať kauzálny vzťah v pravom slova zmysle (pri konštrukcii modelu vychádzame z apriórnej informácie). Ak disponujeme časovým radom a využijeme dynamický regresný model, situácia už môže byť odlišná. Uvažujme dve udalosti A a B, pričom ak udalosť A sa stane pred udalosťou B je možné, že A zapríčiňuje výskyt B. V žiadnom prípade však nie naopak. Inak povedané, udalosti v minulosti môžu zapríčiniť udalosti v súčasnosti, ale budúce udalosti

nemôžu zapríčiniť súčasne. Upozorňujeme, že to samozrejme neznamená, že udalosť A zapríčinila udalosť B.

V prípade Grangerovho modelu sa zvykne často uvádzať, že jedna premenná ovplyvňuje, resp. zapríčiňuje druhú. Presnejšie je však tvrdenie, že súčasné a historické hodnoty jednej premennej nám poskytujú informácie pre vysvetlenie a predikciu druhej premennej. Ak vieme vysvetliť premennú  $x_t$  lepšie pomocou jej minulých hodnôt, ako s využitím aj minulých hodnôt premennej  $y_t$  hovoríme, že  $y_t$  neovplyvňuje  $x_t$  v Grangerovom zmysle kauzality. Nadefinujme si tieto pojmy presnejšie.

Uvažujme o situácii, kde premenné, ktoré sú predmetom nášho záujmu sú merané k určitému okamihu  $t$ ,  $t = 1, 2, \dots, T$ . Množinu všetkých možných premenných, ktorých realizácie (súčasné aj minulé) sú známe v čase  $n$ , si označme ako  $\Omega_n$ . Množinu  $\Omega_n$  budeme označovať ako tzv. informačnú množinu, pričom tá zahŕňa všetky informácie v čase  $t \leq n$ . To znamená, že nezahŕňa žiadne informácie z budúcnosti, avšak ako Granger (1980) poznamenáva, môže zahŕňať očakávania ohľadom budúcnosti. Všetky tieto očakávania však budú determinované množinou  $\Omega_n$ . Ak z tejto množiny vylúčime premennú  $Y_n \in \Omega_n$ , dostaneme množinu, v ktorej sa nenachádzajú informácie sprostredkované výlučne premennou  $Y_n$ . Túto množinu si označíme ako  $\Omega_n - Y_n$ . Ďalej uvažujme o dvoch axiómoch: (i) minulosť a prítomnosť môže ovplyvniť budúcnosť, pričom budúcnosť nemôže ovplyvniť minulosť, (ii) množina  $\Omega_n$  nezahŕňa žiadne nadbytočné premenné.<sup>40</sup> Množina  $\Omega_n$  zahŕňa väčší počet premenných a rovnako aj  $Y_n$  môže predstavovať vektor premenných. Všetky tieto premenné považujeme za náhodné premenné.

Ak nás zaujíma, či premenná  $Y$  ovplyvňuje premennú  $X$ , potom nás zaujíma, či nám v čase  $n$  pomáha premenná  $Y_n$  zlepšiť odhad premennej  $X_{n+1}$ .<sup>41</sup> Keďže aj  $X_{n+1}$  budeme považovať za náhodnú premennú, zaujíma nás, či premenná  $Y_n$  ovplyvňuje pravdepodobnosť, že náhodná premenná  $X_{n+1}$  nadobudne určitú hodnotu z množiny javov  $A$ . Hovoríme, že  $Y_n$  ovplyvňuje  $X_{n+1}$ , ak:

$$P(X_{n+1} \in A | \Omega_n) \neq P(X_{n+1} \in A | \Omega_n - Y_n) \quad (5.121)$$

<sup>40</sup> Potreba tejto axiomy je zrejmá napríklad v situácii, kde medzi dvoma premennými patriacimi do informačnej množiny  $\Omega_n$  existuje deterministický vzťah. Ak budeme chcieť z informačnej množiny  $\Omega_n$  vylúčiť jednu z týchto dvoch premenných, to ešte nebude znamenať, že nová informačná množina nebude obsahovať informácie o vylúčenej premennej. Uvedené je možné zovšeobecniť aj na viac ako dve premenné.

<sup>41</sup> Nemusí ísť nutne o odhad v čase  $n+1$ . Možné sú aj odhady  $n+p$ ,  $p = 1, 2, \dots$ . Budeme ďalej predpokladať  $n+1$ .

Toto je všeobecná definícia Grangerovej kauzality, z ktorej vyplýva, že premenná  $Y_n$  by mala obsahovať určitú informáciu, ktorá ovplyvní  $P(X_{n+1} \in A)$ . Granger ďalej zaviedol tretiu axiómu: (iii) smer všetkých kauzálnych vzťahov ostáva nemenný. Táto axióma pripúšťa zmenu sily kauzality, ako aj zmenu oneskorenia, s akou dochádza k ovplyvňovaniu medzi premennými. Na druhej strane neumožňuje, aby došlo k zániku kauzality, prípadne k zmene smeru kauzality.

Definujme si informačné množiny  $J_n$  a  $J_n'$ , kde prvá zahŕňa rôzne premenné vrátane  $\mathbf{X}_n$ , avšak nie  $\mathbf{Y}_n$ , a druhá zahŕňa navyše práve  $\mathbf{Y}_n$ . Označme si podmienenú distribučnú funkciu premennej  $\mathbf{X}_{n+1}$  ako  $F(\mathbf{X}_{n+1}|J_n)$ , pričom uvedené rozdelenie má strednú hodnotu  $E[\mathbf{X}_{n+1}|J_n]$ . Potom hovoríme, že vzhľadom na informačnú množinu  $J_n'$ ,  $\mathbf{Y}_n$  neovplyvňuje  $\mathbf{X}_{n+1}$ , ak:

$$F(\mathbf{X}_{n+1}|J_n) = F(\mathbf{X}_{n+1}|J_n') \quad (5.122)$$

Vzhľadom na informačnú množinu  $J_n'$ ,  $\mathbf{Y}_n$  zrejme ovplyvňuje  $\mathbf{X}_{n+1}$ , ak:

$$F(\mathbf{X}_{n+1}|J_n) \neq F(\mathbf{X}_{n+1}|J_n') \quad (5.123)$$

Granger (1980) pri vzťahu (5.123) schválne zvolil pomerne diplomatické tvrdenie: „ $\mathbf{Y}_n$  zrejme ovplyvňuje  $\mathbf{X}_{n+1}$ “, keďže sa týka iba určitej množiny informácií, ktoré máme v danom okamihu  $n$  k dispozícii (nie všetky možné informácie).

Granger (1980) pripúšťa, že s uvedenými definíciami môže vzniknúť niekoľko problémov. Napríklad hneď prvá axióma nepripúšťa tzv. okamžitú kauzalitu v stredných hodnotách medzi  $\mathbf{X}_{n+1}$  a  $\mathbf{Y}_{n+1}$ . Ide o nasledujúcu situáciu:

$$E(\mathbf{X}_{n+1}|J_n, \mathbf{Y}_{n+1}) \neq E(\mathbf{X}_{n+1}|J_n) \quad (5.124)$$

V empirickom výskume sa môžeme s okamžitou kauzalitou stretnúť pomerne často. Interpretácia je však často do značnej miery komplikovaná. Bez dodatočných deduktívnych dôvodov často nevieme rozhodnúť, ktorá premenná vlastne ovplyvňuje ktorú. Stále môže platiť, že reagujú rovnako na spoločný faktor.

Predošlá definícia hovorí o kauzalite v stredných hodnotách. Je možné nadefinovať aj kauzalitu pre ostatné charakteristiky. Mierne odlišným prípadom je nasledovná definícia. Prítomnosť kauzality sa môže posudzovať pomocou toho, či premenná pomáhala znížiť odchýlky predikovaných hodnôt od skutočných. Týmto postupom zodpovedá nasledujúca definícia Grangerovej kauzality. Vzhľadom na informačnú množinu  $J_n'$ ,  $\mathbf{Y}$  zrejme ovplyvňuje  $\mathbf{X}$ , ak:

$$\sigma^2(\mathbf{X}|J_n') < \sigma^2(\mathbf{X}|J_n) \quad (5.125)$$

kde  $\sigma^2(\mathbf{X}|J_n')$  je rozptyl predikovaných hodnôt  $\mathbf{X}_{n+1}$  (predikcia na jedno budúce pozorovanie).

S definíciami, ktoré sú podobné tým, ktoré boli prezentované vyššie, sa môžeme stretnúť aj v iných publikáciách. Napríklad Green (2003, s. 382) uvádza nasledovnú rovnosť podmienených funkcií hustôt:

$$f(\mathbf{X}_t|\mathbf{X}_{t-1}, \mathbf{Y}_{t-1}) = f(\mathbf{X}_t|\mathbf{X}_{t-1}) \quad (5.126)$$

kde  $\mathbf{X}_t, \mathbf{Y}_t$  sú matice premenných, pričom výraz (5.126) hovorí, že oneskorené hodnoty premennej  $\mathbf{X}_t$  a  $\mathbf{Y}_t$ , t. j.  $\mathbf{X}_{t-1}$  a  $\mathbf{Y}_{t-1}$  nepomáhajú vysvetliť hodnoty  $\mathbf{X}_t$  lepšie ako len oneskorené hodnoty  $\mathbf{X}_{t-1}$ . Prípadne Hansen (2012, s. 278) uvádza:

$$E(Y_t|I_{1,t-1}) = E(Y_t|I_{2,t-1}) \quad (5.127)$$

kde  $I_{1t} = (Y_t, Y_{t-1}, Y_{t-2}, \dots)$  a  $I_{2t} = (Y_t, Z_t, Y_{t-1}, Z_{t-1}, Y_{t-2}, Z_{t-2}, \dots)$  sú dve informačné množiny. Prvá tvoria len súčasné a oneskorené hodnoty premennej  $Y_t$  a druhá je tvorená súčasnými a oneskorenými hodnotami premennej  $Y_t$  a  $Z_t$ . Výraz (5.127) tak hovorí, že  $Z_t$  neovplyvňuje v Grangerovom zmysle strednú hodnotu premennej  $Y_t$ .

Medzi dvoma časovými radmi môže existovať spolu osem rôznych stavov. Ak však vylúčime tzv. okamžitú kauzalitu, ostanú nám štyri prípady. Majme dva časové rady  $X_t, Y_t$  a prípadne vektor ďalších časových radov  $Z_t, t = 1, 2, \dots$ , potom budeme vzájomné vzťahy medzi premennými  $X_t$  a  $Y_t$  zapisovať nasledujúcimi všeobecnými zápismi:

- vzhľadom na informačnú množinu  $J(Z, X, Y)$  sa  $X$  a  $Y$  neovplyvňujú;  $X \nleftrightarrow Y$
- vzhľadom na informačnú množinu  $J(Z, X, Y)$   $X$  ovplyvňuje  $Y$ ;  $X \Rightarrow Y$
- vzhľadom na informačnú množinu  $J(Z, X, Y)$   $Y$  ovplyvňuje  $X$ ;  $Y \Rightarrow X$
- vzhľadom na informačnú množinu  $J(Z, X, Y)$   $X$  ovplyvňuje  $Y$  a  $Y$  ovplyvňuje  $X$ ;  $X \Leftrightarrow Y$

Vzhľadom na to, že pri overovaní hypotéz sa často pracuje s hypotézou o neexistencii vzťahu, budeme používať nasledujúce značenia:  $X \nRightarrow Y, Y \nRightarrow X$ . Tieto tak budú znamenať hypotézu, že premenná  $X$  neovplyvňuje v Grangerovom zmysle premennú  $Y$  a naopak.

Existuje pomerne mnoho metód, ako overovať prítomnosť Grangerovej kauzality. Ukážeme si základný parametrický model. V tomto prístupe sa vychádza z toho, že dynamika vzťahov medzi premennými v Grangerovom modeli je zabezpečená časovými posunmi závislej aj nezávislej premennej. Model bol navrhnutý Grangerom (1969) a vychádza z autoregresného modelu s rozloženým oneskorením (angl. *AutoRegressive Distributed Lag* – ARDL rádu  $(p, q)$ ), ktorý dostaneme, ak k modelu s rozloženým oneskorením, ktorý okrem

premennej  $x_t$  obsahuje aj jej  $r$  oneskorených hodnôt (teda  $x_{t-i}$ ), pridáme  $q$  oneskorených hodnôt vysvetľovanej premennej  $y_{t-i}$ :

$$y_t = \alpha_0 + \sum_{i=1}^q \alpha_i y_{t-i} + \sum_{j=0}^r \beta_j x_{t-j} + u_t \quad (5.128)$$

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \dots + \alpha_q y_{t-q} + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \dots + \beta_r x_{t-r} + u_t \quad (5.129)$$

kde  $y_t$  je vysvetľovaná premenná,  $x_t$  je vysvetľujúca premenná,  $\alpha_0, \alpha_i, \beta_j$  pre ( $i = 1, 2, \dots, q$  a  $j = 0, 1, 2, \dots, r$ ) sú regresné koeficienty,  $u_t$  je náhodná chyba a  $q, r$  maximálne posunutia (lagy). Indexy  $i, j$  označujú dané posunutie.

Grangerov model predstavuje odklon od ARDL ( $p, q$ ) v tom, že vysvetľujeme premennú  $y$  ako lineárnu funkciu jej a) minulých hodnôt a b) minulých hodnôt vektora premennej  $x$ . Takto zostavený model predstavuje teoretický základ pre testovanie tzv. Grangerovej kauzality. Grangerov model má tvar:

$$y_t = \alpha_0 + \sum_{i=1}^m \alpha_i y_{t-i} + \sum_{i=1}^m \beta_i x_{t-i} + u_t \quad (5.130)$$

Oproti vzťahu (5.128) a (5.129) nám v Grangerovom modeli chýba člen vysvetľujúcej premennej  $x$  s hodnotou v čase  $t$ . Prirodzene musí platiť  $m+1 \leq t, m \in N$  vzhľadom na počet oneskorení, avšak tiež musí platiť  $2m+1 \leq t$  podľa predpokladu P4. Všimnime si, že takto postavený model vlastne zisťuje, či okrem premennej  $y$  nám minulé hodnoty premennej  $x$  pomáhajú predikovať vývoj v hodnotách  $y$ . Pri overovaní jednosmerného vzťahu potom testujeme dve pomocné regresie:

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \dots + \alpha_m y_{t-m} + \beta_1 x_{t-1} + \dots + \beta_m x_{t-m} + u_t \quad (5.131)$$

$$x_t = \gamma_0 + \gamma_1 x_{t-1} + \dots + \gamma_m x_{t-m} + \delta_1 y_{t-1} + \dots + \delta_m y_{t-m} + u'_t \quad (5.132)$$

Keďže meníme funkčný tvar, dostávame rozdielne odhady parametrov a náhodné chyby. Z tohto dôvodu je značenie rozdielne. Nulová hypotéza v prvom prípade je, že  $x$  neovplyvňuje v Grangerovom zmysle  $y$ , teda neplatí  $x \Rightarrow y$ . V druhej rovnici potom analogicky testujeme hypotézu, že neplatí kauzalita  $y \Rightarrow x$ . Ak nevieme zamietnuť nulové hypotézy na stanovenej hladine významnosti v oboch rovniciach, potom sú premenné v zmysle Grangerovej kauzality nezávislé.<sup>42</sup>

<sup>42</sup> Bližšie k danej problematike napr. Mills (1999), Hatrák (2007).

Na overenie hypotéz vo vyššie uvedenom modeli testujeme združenú nulovú hypotézu oproti alternatívnej hypotéze na základe Waldovej  $F$ -štatistiky<sup>43</sup>:

$$H_0: \beta_1 = \beta_2 = \dots = \beta_m = 0$$

$$H_1: \exists i \in \{1, 2, \dots, m\}: \beta_i \neq 0$$

Aplikácia Grangerovho modelu v softvéri R je možná viacerými spôsobmi. Môžeme odhadnúť autoregresné modely s rozloženým oneskorením (ARDL), v ktorých raz budú vystupovať oneskorené hodnoty vysvetľujúcej premennej a raz nie. V takom prípade skúmame, či hodnoty jednej premennej prispievajú k predikcii druhej premennej, o čom môžeme rozhodnúť na základe modelu ANOVA.

Druhý spôsob je zrejme jednoduchší. Keďže v softvéri R sa nachádzajú rôzne funkcie na výpočet Grangerovho modelu, môžeme požadované výpočty realizovať cez jednu z nich. Napríklad v knižnici MSBVAR sa nachádza funkcia `granger.test()`, pomocou ktorej vieme vypočítať Grangerovu kauzalitu medzi všetkými premennými, ktoré máme v databáze.

Na ukážku použijeme údaje o HDP a akciových indexoch z krajín V4 (Česká republika, Maďarsko, Poľsko a Slovensko). K dispozícii máme kvartálne údaje od Q1:1996 do Q4:2009. Tieto údaje sú už po transformácii na logaritmické diferencie kvôli stacionarite premenných.

Podľa teórie, by akciové trhy mali predbiehať reálny výstup z ekonomiky o 3 až 6 mesiacov. Uvedenú skutočnosť nie je ťažké dokázať na základe dynamických regresných modelov časových radov (kauzalita v Grangerovom zmysle). Dôvodom je záujem o akciové trhy zo strany investorov, ktorí sa pokúšajú predvídať rast HDP. Tým sa spätne ovplyvňuje aj rast reálneho produktu, pretože spolu s rastom akciových trhov rastie aj bohatstvo akcionárov (zvyšuje sa agregátny dopyt). Najmä rýchlo rastúci HDP indikuje expandujúcu ekonomiku s rozsiahlymi možnosťami pre podniky a rastu ich výnosov.

Keďže funkcia `granger.test()`<sup>44</sup> vypočíta všetky možné varianty, vyberieme len tie, ktoré sa týkajú vzťahov medzi HDP a akciovými indexmi danej krajiny. Pre zjednodušenie uvedieme tiež len prvé oneskorenia (argument vo funkcii `granger.test()` nastavíme na `p = 1`), t. j. bude skúmať, či HDP krajiny z predchádzajúceho kvartálu vysvetľuje vývoj akciových indexov, resp. budeme overovať

---

<sup>43</sup> Využitý môže byť aj štandardný  $F$ -test, v tomto koncepte sa však zvykne častejšie používať Wald. Waldova štatistika je asymptoticky ekvivalentná k  $F$ -testu (Davidson – MacKinnon, 2003).

<sup>44</sup> Naše značenie pre výraz „neovplyvňuje v Grangerovom zmysle kauzalitu“ symbolom “ $\neq$ ” je ekvivalentné symbolu “ $\rightarrow$ ” vo funkcii `granger.test()`.

hypotézu (v zmysle uvedeného výroku, že akciové trhy predbiehajú HDP o 3 až 6 mesiacov), či akciové indexy v krajinách V4 predbiehajú ich HDP o jeden kvartál, teda o 3 mesiace. Ak by sme nastavili argument  $p = 2$ , bolo by možné vyjadriť sa aj k predbiehaniu akciových indexov o 6 mesiacov (túto možnosť necháme na čitateľa).

```

> data <- read.csv(file = "...cesta
  k súboru...\granger_gdp.csv", sep = ";", dec = ".", header =
  T)
-----
> library(MSBVAR)
> granger.test(data, p = 1)

                F-statistic    p-value
px      -> hdp_cz    1.935497e+01    5.544050e-05
hdp_cz  -> px      2.773891e-04    9.867768e-01
bux     -> hdp_hu    8.705816e+00    4.781141e-03
hdp_hu  -> bux     1.361087e-01    7.137078e-01
wig     -> hdp_pl    4.291958e+00    4.336889e-02
hdp_pl  -> wig     5.725715e-01    4.527204e-01
sax     -> hdp_sk    4.671527e+00    3.538461e-02
hdp_sk  -> sax     5.180993e-02    8.208533e-01

```

V prípade českého indexu PX môžeme na hladine významnosti 1 % zamietnuť nulovú hypotézu (pri  $p$ -hodnote 0.0000554), že tento index neovplyvňuje HDP v Českej republike v Grangerovom zmysle kauzality. Opačný smer Grangerovej kauzality sa nepreukázal, keďže nulovú hypotézu „HDP v Českej republike neovplyvňuje index PX“ nevieme zamietnuť ( $p$ -hodnota 0.9867768). Obdobné výsledky dostávame aj pri ostatných krajinách V4.

Z uvedených výsledkov teda vyplýva, že všetky akciové indexy v krajinách V4 predbiehajú HDP v danej krajine o 3 mesiace, avšak ide len o jednosmernú závislosť. Ani v jednom prípade sa nepotvrdilo, že HDP skúmaných krajín ovplyvňujú ich akciové indexy.

## 6 Príklady

Táto kapitola obsahuje vybrané príklady, ktoré tematicky súvisia s problematikou, ktorej sme sa venovali v predchádzajúcich kapitolách. Väčšinou predstavujú rozsiahlejšie analýzy, ktorých cieľom je demonštrovať odhad konkrétnych modelov, ako aj testovanie predpokladov modelov v reálnych podmienkach. Z dôvodu ich väčšieho rozsahu sme ich nezaradovali priamo do textu, ale venovali sme im samostatný priestor. Keďže predstavujú spravidla komplexné riešenie, využívajú často poznatky hneď z niekoľkých kapitol. Pred snahou o reprodukciu ich riešení preto odporúčame oboznámiť sa najprv s relevantnými teoretickými výsledkami prezentovanými hlavne v kapitolách 3 až 5.

### 6.1 Lineárne modely a ich diagnostika

#### Príklad 6.1

V makroekonómii je vzťah medzi infláciou a nezamestnanosťou známy ako Phillipsova krivka (aj keď v pôvodnej verzii Phillips skúmal nezamestnanosť a mieru zmeny nominálnych miezd). Na základe práce Samuelson – Solow (1960) sa v základnej teórii predpokladá, že medzi nezamestnanosťou a infláciou je nepriamo úmerný vzťah. V databáze `unemployment` (knížnica `lmtest`) máme k dispozícii ročné údaje z USA o nezamestnanosti (premenná `UN`) a inflácii (premenná `p`) za obdobie 1890 – 1979. Overte, či platí nepriamo úmerný vzťah medzi týmito dvoma premennými.

---

#### Príklad 6.1 – Riešenie

---

V tomto príklade máme overiť platnosť Phillipsovej krivky, teda ideme skúmať vzťah medzi infláciou a nezamestnanosťou (údaje z databázy `unemployment`, z knižnice `lmtest`). Keďže pracujeme s časovými radmi, je nutné ich pred samotnou analýzou otestovať na stacionaritu. Výsledky v prípade týchto dvoch premenných neboli úplne jednoznačné. Budeme preto skúmať jednak pôvodné časové rady, ale taktiež nás bude zaujímať vzťah medzi zmenou rastu cenovej hladiny a zmenou nezamestnanosti. Tieto časové rady už zrejme za stacionárne môžeme považovať (bližšie pozri Príklad 6.14).

Vzhľadom k tomu, že skúmame vzťah medzi dvoma premennými, môžeme použiť korelačnú alebo regresnú analýzu. Pristúpme najprv ku korelačnej analýze, pričom okrem

Pearsonovho korelačného koeficientu vypočítame aj Spearmanov, ktorý je vhodný na zachytenie inej ako lineárnej závislosti (avšak len monotónnej).

```
> UN <- unemployment[2:90,1]; p <- diff(unemployment[,3]);
> cor.test(UN, p, method = "pearson", alternative = "two.sided",
  exact = FALSE)
-----
Pearson's product-moment correlation

data: UN and p
t = -2.8701, df = 87, p-value = 0.005151
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.47336246 -0.09144211
sample estimates:
      cor
-0.2940982
-----
> cor.test(UN, p, method = "spearman", alternative =
  "two.sided", exact = FALSE)

Spearman's rank correlation rho

data: UN and p
S = 151827.5, p-value = 0.005432
alternative hypothesis: true rho is not equal to 0
sample estimates:
      rho
-0.2923691
```

Oba vypočítané korelačné koeficienty poukazujú na slabší nepriamy vzťah. V prípade oboch koeficientov sme pomocou štandardného testu vyhodnotili koeficienty za štatisticky významné na hladine významnosti 1 %. Na tomto mieste je potrebné si uvedomiť, že ak v oboch časových radoch je prítomná autokorelácia a heteroskedasticita, môže sa ľahko stať, že náš štandardný *t*-test porovnávania zhody korelačného koeficientu s konštantou (s nulou) môže ponúkať zavádzajúce výsledky. V rámci regresnej analýzy sa však s týmito problémami vieme lepšie vysporiadať.

```
> model = lm(p ~ UN); summary(model)

Call:
lm(formula = p ~ UN)

Residuals:
    Min       1Q   Median       3Q      Max
-0.084697 -0.016725 -0.007764  0.007188  0.109131

Coefficients:
```

```

              Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.0297807  0.0055983    5.32 7.98e-07 ***
UN           -0.0020558  0.0007163   -2.87  0.00515 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.02926 on 87 degrees of freedom
Multiple R-squared:  0.08649, Adjusted R-squared:  0.07599
F-statistic: 8.237 on 1 and 87 DF,  p-value: 0.005151

```

Regresný koeficient pri nezamestnanosti je samozrejme taktiež záporný. Hodnotenie významnosti však nateraz odložíme. Najprv skontrolujeme dva kľúčové predpoklady modelu. Pripomenieme, že ešte okrem týchto dvoch predpokladov (heteroskedasticita a autokorelácia rezíduí), stále platí že ak sú obe premenné nestacionárne (procesy s jednotkovým koreňom), tak potom stále ešte riskujeme aj výskyt falošnej regresie.

```

> dwtest(model, alternative = "two.sided")

Durbin-Watson test

data:  model
DW = 0.5115, p-value < 2.2e-16
alternative hypothesis: true autocorrelation is not 0
-----
> bptest(model)

studentized Breusch-Pagan test

data:  model
BP = 0.0092, df = 1, p-value = 0.9235

```

Na základe výsledkov z Durbin – Watson testu môžeme konštatovať, že autokorelácia v modeli je prítomná. Pomocou Breusch – Paganovho testu sme však nulovú hypotézu o homoskedasticite rezíduí zamietnuť nevedeli. Pri odhade štandardnej chyby regresných koeficientov použijeme variančno-kovariančnú maticu konzistentnú aj v prípade výskytu heteroskedasticity a autokorelácie (HAC). Pri tejto analýze je stále regresný koeficient významný na hladine menšej ako 1 %.

```

> coeftest(model, vcov = vcovHAC(model))

t test of coefficients:

              Estimate Std. Error t value  Pr(>|t|)
(Intercept)  0.02978074  0.00784786  3.7948 0.0002726 ***
UN           -0.00205582  0.00042692 -4.8154 6.149e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Tieto výsledky hovoria stále v prospech existencie štatisticky významného záporného vzťahu medzi nezamestnanosťou a infláciou. Ak sme pozorovali väčšiu mieru nezamestnanosti, tak zároveň sme pozorovali aj menšiu mieru inflácie. Koeficient determinácie však bol relatívne nízky a preto zrejme existujú aj iné, možno dôležitejšie faktory, ktoré determinujú zmenu cenovej hladiny (napr. zmena množstva peňazí v ekonomike).

Pristúpime teraz aj ku kvantifikácii regresného modelu na zmenách inflácie a nezamestnanosti. Takto špecifikovaný model nám hovorí o krátkodobom vzťahu medzi premennými. Kvantifikujeme, či so zmenou jednej premennej došlo k zmenám druhej premennej.

```

> dUN = diff(UN)
> dp = diff(p)
-----
> model_2 = lm(dp ~ dUN); summary(model_2)

Call:
lm(formula = dp ~ dUN)

Residuals:
    Min       1Q   Median       3Q      Max
-0.120027 -0.008874  0.000591  0.008628  0.055498

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.0014785  0.0022219   0.665  0.50755
dUN          -0.0029660  0.0008097  -3.663  0.00043 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.02084 on 86 degrees of freedom
Multiple R-squared:  0.135,    Adjusted R-squared:  0.1249
F-statistic: 13.42 on 1 and 86 DF,  p-value: 0.0004297

```

Výsledky regresného modelu kvantifikovaného na diferenciách sú do značnej miery podobné, čo do významnosti regresného koeficientu, tak aj koeficientu determinácie.

```

> dwtest(model_2, alternative = "two.sided")

Durbin-Watson test

data:  model_2
DW = 2.3375, p-value = 0.1157
alternative hypothesis: true autocorrelation is not 0
-----
> bptest(model_2)

```

```
studentized Breusch-Pagan test
```

```
data: model_2  
BP = 5.2007, df = 1, p-value = 0.02258
```

Autokoreláciu sme v tomto prípade nenašli, avšak hypotézu o homoskedasticite sme použitím Breusch – Paganovho testu zamietli. Z tohto dôvodu sme sa rozhodli použiť odhad štandardných chýb regresných koeficientov pomocou HAC matíc.

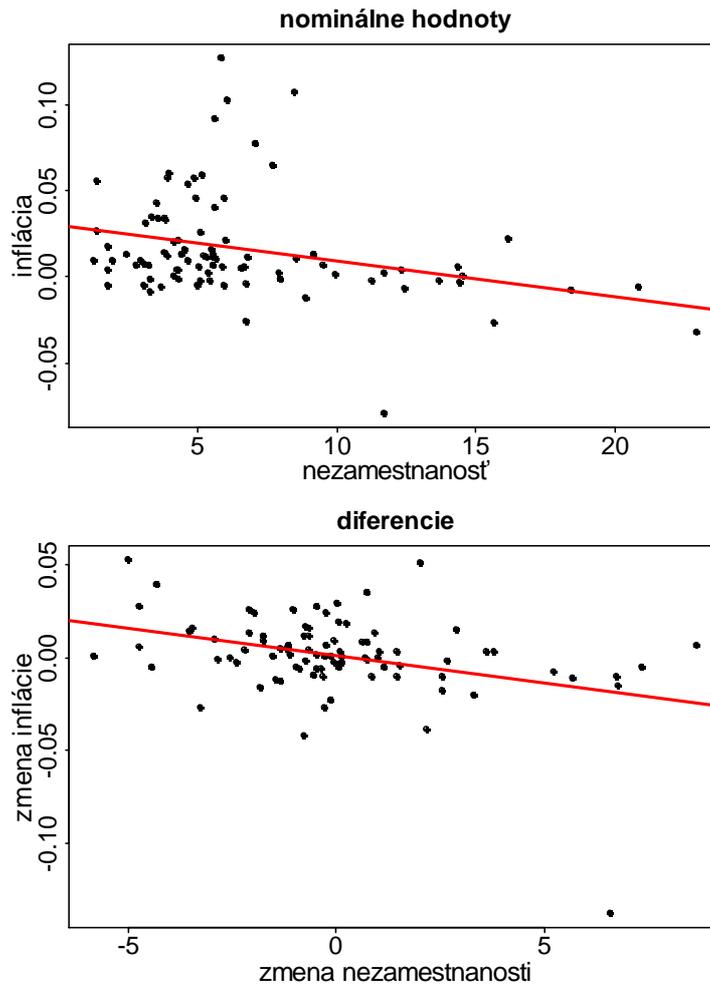
```
> coeftest(model_2, vcov = vcovHAC(model_2))  
  
t test of coefficients:  
  
              Estimate Std. Error t value Pr(>|t|)  
(Intercept)  0.0014785  0.0015084  0.9802  0.32973  
dUN          -0.0029660  0.0016433 -1.8050  0.07458 .  
---  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Môžeme vidieť, že koeficient pri nezamestnanosti je stále záporný, avšak pri diferenciách je významný len na hladine významnosti  $\alpha = 0.10$ . Zo všetkých uvedených výsledkov by sme mohli prijať tvrdenie, že existuje záporný vzťah medzi nezamestnanosťou a infláciou, avšak táto závislosť nie je veľmi silná. Túto skutočnosť môžeme vidieť aj pri grafickej vizualizácii vzájomného vzťahu týchto dvoch premenných<sup>45</sup>.

```
> library(car)  
> par(mfrow = c(2, 1))  
> par(cex = 0.6)  
> par(mar = c(3.4, 3.2, 3.2, 2.1), oma = c(2, 4.0, 0.5, 0.5))  
> par(tcl = -0.25)  
> par(mgp = c(2, 0.6, 0))  
> plot(p ~ UN, main = "nominálne hodnoty", col = "black", xlab =  
"nezamestnanosť", ylab = "inflácia", pch = 19, cex.lab = 1.4,  
cex.axis = 1.2)  
> abline(model, col = "red", lwd = 2)  
> plot(dp ~ dUN, main = "diferencie", col = "black", xlab =  
"zmena nezamestnanosti", ylab = "zmena inflácie", pch = 19,  
cex.lab = 1.4, cex.axis = 1.2)  
> abline(model_2, col = "red", lwd = 2)
```

---

<sup>45</sup> Čitateľ si môže vyskúšať odhadnúť nelineárny model (nelineárny v premenných, nie v parametroch):  
 $p_t = \beta_0 + \beta_1 UN_t + \beta_2 UN_t^2 + \varepsilon_t$ . Koeficient determinácie bude síce vyšší, avšak ani jeden z regresných koeficientov (okrem konštanty) nie je štatisticky významný ani na hladine 10 %.



Obrázok 24: Vzťah medzi nezamestnanosťou a infláciou

*Zdroj: výstup zo softvéru R*

### Príklad 6.2

V tomto príklade budeme pracovať s databázou *Anscombe* (knížnica *car*), z ktorej nás budú zaujímať údaje o výdavkoch na vzdelávanie na jednu osobu (*education*) a príjem na jednu osobu (*income*). Údaje sú z roku 1970 za 51 štátov z USA. Zistite, či medzi týmito dvoma premennými existuje lineárna závislosť a overte predpoklad homoskedasticity a normality rezíduí.

---

### Príklad 6.2 – Riešenie

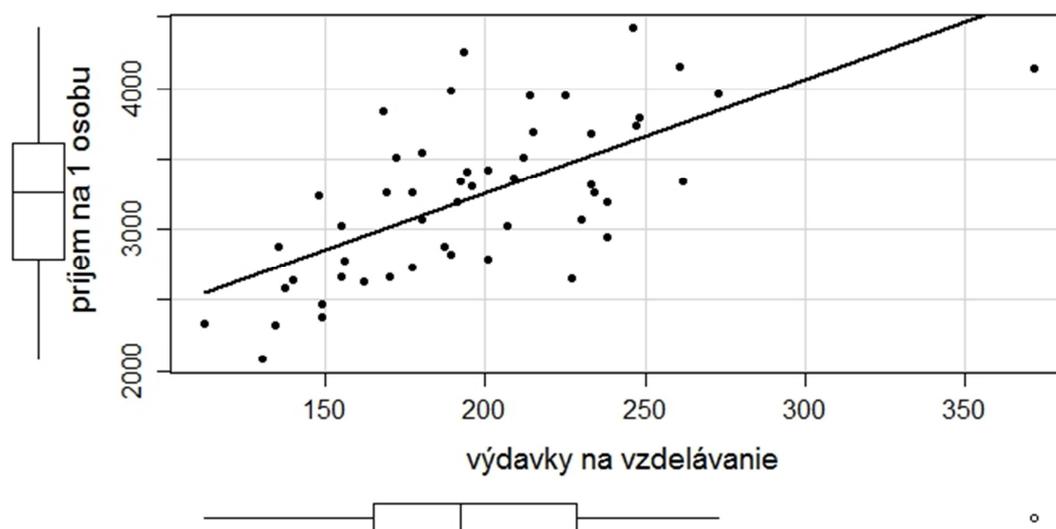
---

Z údajov dostupných v databáze *Anscombe* (knížnica *car*) máme zistiť, či existuje lineárna závislosť medzi príjmom (premenná *income*) a výdavkami na vzdelávanie (premenná *education*). Pozrime sa najprv na grafickú vizualizáciu ich vzájomného vzťahu vo forme *x-y* grafu.

```

> library(car)
> attach(Anscombe)
-----
> scatterplot(education ~ income, smooth = F, col = "black",
  xlab = " príjem na 1 osobu", ylab = " výdavky na vzdelávanie",
  lwd = 2, lty = 1, pch = 19, cex.lab = 1.4, cex.axis = 1.2)

```



Obrázok 25: Vzťah medzi výdavkami na vzdelávanie a príjmom

*Zdroj: výstup zo softvéru R*

Z uvedeného obrázku môžeme vidieť, že medzi týmito dvoma skúmanými premennými je dosť zjavná priama lineárna závislosť. Otázkou ostáva, aká silná je táto závislosť a za týmto účelom kvantifikujeme jednoduchý regresný model.

```

> model = lm(income ~ education); summary(model)

Call:
lm(formula = income ~ education)

Residuals:
    Min       1Q   Median       3Q      Max
-821.25 -341.15    8.47  196.23 1057.37

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 1645.383   258.539   6.364 6.41e-08 ***
education     8.048     1.282   6.276 8.76e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 421.2 on 49 degrees of freedom
Multiple R-squared:  0.4457,    Adjusted R-squared:  0.4343

```

```
F-statistic: 39.39 on 1 and 49 DF, p-value: 8.762e-08 08
```

Týmto modelom sme vysvetlili 43.43 % variability (upravený koeficient determinácie). Ak sa teda v jednotlivých štátoch U.S. zvýšili výdavky na vzdelávanie (na osobu) o jeden dolár, bol pozorovaný nárast ročného príjmu obyvateľstva o niečo vyššie 8 dolárov. Keďže v tomto príklade pracujeme s prierezovými údajmi, je nutné overiť predpoklad o homoskedasticite rezíduí. Na testovanie tohto predpokladu využijeme Breusch – Paganov test.

```
> library(lmtest)
> bptest(model)

studentized Breusch-Pagan test

data: model
BP = 0.4507, df = 1, p-value = 0.502
```

Na hladine významnosti  $\alpha = 0.05$ , nulovú hypotézu o konštantnom rozptyle rezíduí nevieme zamietnuť. Vzhľadom na to, že sme nemali k dispozícii veľký počet pozorovaní, je vhodné overiť aj predpoklad normality.

Na testovanie normality využijeme v tomto príklade tri testy, a to Anderson – Darlingov (funkcia `ad.test()` z knižnice `nortest`), Shapiro – Wilkov (funkcia `shapiro.test()` z knižnice `stats`) a Jarque – Berov (funkcia `rjb.test()` z knižnice `lawstat`). Pri Jarque – Berovom test budeme vychádzať z empirických (simulovaných) kritických hodnôt a použijeme obe jeho verzie, t. j. klasickú (`option = c("JB")`) aj modifikovanú (`option = c("RJB")`).

```
> ad.test(resids)

Anderson-Darling normality test

data: resids
A = 0.5873, p-value = 0.1193
-----
> shapiro.test(resids)

Shapiro-Wilk normality test

data: resids
W = 0.9685, p-value = 0.1911
-----
> rjb.test(resids, option = c("JB"), crit.values =
c("empirical"), N = 1000)
```

```

Jarque Bera Test

data:  resid
X-squared = 1.9845, df = 2, p-value = 0.2091
-----
> rjb.test(resid, option = c("RJB"), crit.values =
  c("empirical"), N = 1000)

Robust Jarque Bera Test

data:  resid
X-squared = 1.7968, df = 2, p-value = 0.2657

```

Pri všetkých týchto testoch sa overuje nulová hypotéza o normalite. Túto hypotézu sme nevedeli zamietnuť pri každom z týchto testov.

### Príklad 6.3

Zo stránky [pages.stern.nyu.edu/~adamodar/](http://pages.stern.nyu.edu/~adamodar/) sme vytvorili databázu Eurocompfirm, ktorá obsahuje len tie údaje za podniky, ktoré boli kompletne (žiadne z požadovaných údajov neboli chýbajúce). Databáza obsahuje údaje len za európske podniky, pričom ide o nasledujúce premenné:

- *Ticker* – ide identifikačnú skratku podniku, pod ktorou sa obchoduje na burze.
- *Sub.Group* – táto premenná rozdeľuje obsiahnuté podniky na tri úrovne: UK, EU a Eastern Europe & Russia.
- *Beta* – je štandardný ukazovateľ z modelu CAPM (z angl. Capital Asset Pricing Model), ktorý charakterizuje volatilitu akcií daného podniku. V tomto príklade ho budeme využívať ako aproximáciu rizikovosti akcií.
- *PB* – je pomerový ukazovateľ trhovej hodnoty podniku. Dáva do pomeru trhovú cenu akcie a účtovnú hodnotu akcie.
- *Payout* – je výplatný pomer, ktorý sa počíta ako pomer dividendy na akciu a zisku pripadajúceho na jednu akciu. Hovorí teda o tom, aký pomer sa vypláca akcionárom v podobe dividend z vytvoreného zisku.
- *Growth* – je historická miera rastu tržieb za posledné tri roky.
- *ROE* – je rentabilita vlastného kapitálu.

Kvantifikujte nasledujúci oceňovací model:

$PB_i = \beta_0 + \beta_1 Payout_i + \beta_2 Growth_i + \beta_3 Beta_i + \beta_4 ROE_i + \varepsilon_i$  (kde  $i = 1, 2, \dots, n$  a  $n$  je počet podnikov/pozorovaní), pomocou ktorého je možné určiť nesprávne ocenené tituly. Vychádzame pritom z jednoduchšej myšlienky, že akcie nachádzajúce sa na regresnej

nadvrhne sú správne ocenené (vnútorná hodnota = trhova cena), akcie nachadzajuce sa nad ňou su nadhodnotene (vnútorna hodnota < trhova cena) a akcie nachadzajuce sa pod ňou su podhodnotene (vnútorna hodnota > trhova cena)<sup>46</sup>. Model kvantifikujte na celej vzorke, ako aj na podvzorkach rozdelenych podla premennej Sub.Group.

### Prıklad 6.3 – Rieenie

V tomto prıklade je naouuou zostaviť jednoduchy oceovaci model na vzorke europskych podnikov (Eurocompfirm), na zaklade ktoreho by sme mali vedieť zistiť, ktore akcie su podhodnotene alebo nadhodnotene. V modeli vysvetlujeme ukazovate P/B prostrednictvom dalsichutyroch fundamentalnych ukazovateov. Najskôr odhadneme model na celej vzorke.

```
> data = read.csv(file = "...cesta
  k suboru...\Eurocompfirm.csv", sep = ";", dec = ".", header =
  T)
> attach(data)
-----
> model = lm(PB ~ Payout + Growth + Beta + ROE); summary(model)

Call:
lm(formula = PB ~ Payout + Growth + Beta + ROE)

Residuals:
    Min       1Q   Median       3Q      Max
-14.740  -0.988  -0.575   0.292  51.047

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.852197    0.094161  19.670 < 2e-16 ***
Payout       0.002964    0.019377   0.153  0.87844
Growth       0.061270    0.082324   0.744  0.45679
Beta        -0.199664    0.071604  -2.788  0.00534 **
ROE          0.462695    0.049895   9.273 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.499 on 2547 degrees of freedom
Multiple R-squared:  0.03587,    Adjusted R-squared:  0.03436
F-statistic: 23.69 on 4 and 2547 DF,  p-value: < 2.2e-16
```

Dva ukazovatele sutatisticky vyznamne, a to proxy ukazovate rizika (Beta) a rentabilita vlastneho kapitalu (ROE). Koeficient pri riziku je zaporny, zoho vyplyva, že ak sa zvrysi riziko, tak ukazovate P/B sa znizi. Vztah medzi tym, koľko platime za jednotku

<sup>46</sup> Blišie k problematike vyuitia regresnych modelov pri oceovani akcii pozri Baumohl et al. 2011.

vlastného kapitálu (P/B) a rentabilitou vlastného kapitálu (ROE) je prirodzene s kladným znamienkom. Model ako celok je síce štatisticky významný, ale týmto modelom sme popísali len veľmi nízke percento celkovej variability.

Keďže v modeli vystupuje viac vysvetľujúcich premenných, v tomto príklade už má zmysel overiť aj výskyt multikolinearity (na rozdiel od predchádzajúcich príkladov). V prvom rade je vždy vhodné pred samotnou kvantifikáciou modelu zostaviť korelačnú maticu a pozrieť sa na vzájomný vzťah medzi jednotlivými vysvetľujúcimi premennými. Za týmto účelom môžeme napríklad využiť funkciu `rcor.test()` z knižnice `ltm`. Túto funkciu je možné použiť priamo na skupinu dátových vektorov (objekt získaný cez funkciu `data.frame()`), takže najprv si z databázy oddelíme tie premenné, ktorých korelácie chceme skúmať. Len pre úplnosť zahrnieme aj vysvetľovanú premennú P/B.

```
> dataFrame = data.frame(PB, Payout, Growth, Beta, ROE)
> library(ltm)
> rcor.test(dataFrame, method = "pearson")
```

	PB	Payout	Growth	Beta	ROE
PB	*****	-0.002	0.016	-0.056	0.181
Payout	0.934	*****	-0.018	0.008	-0.022
Growth	0.425	0.356	*****	-0.029	-0.001
Beta	0.005	0.693	0.143	*****	-0.007
ROE	<0.001	0.272	0.956	0.739	*****

upper diagonal part contains correlation coefficient estimates  
lower diagonal part contains corresponding p-values

Výstupom z tejto funkcie je matica, ktorá nad hlavnou diagonálou obsahuje korelačné koeficienty (zvolili sme Pearsonov korelačný koeficient) a pod hlavnou diagonálou sú  $p$ -hodnoty k prislúchajúcim obojstranným  $t$ -testom. Z uvedenej matice môžeme vidieť, že korelácie nie sú vysoké (okrem iného ani štatisticky významné), a teda multikolinearita by v našom regresnom modeli nemala predstavovať problém<sup>47</sup>. Na overenie výskytu multikolinearity môžeme vypočítať inflačný faktor rozptylu (v angl. VIF, Variance Inflation Factor) napríklad prostredníctvom funkcie `vif()` z knižnice `HH`.

```
> library(HH)
> vif(model)
Payout    Growth      Beta      ROE
```

<sup>47</sup> Aj keď korelácie vysvetľujúcich premenných s vysvetľovanou premennou sú nízke, ich lineárna kombinácia môže zapríčiniť, že koeficienty v regresnom modeli budú významné a ich vzájomný vplyv bude väčší, ako by sa mohlo javiť len z izolovaného skúmania v rámci korelačnej analýzy. Preto aj pri nízkych koreláciách má zmysel pristúpiť k použitiu regresnej analýzy.

```
1.000861 1.001170 1.000935 1.000518
```

Všetky hodnoty VIF sú pomerne nízke a nepribližujú sa ani k hodnote 5 ani k hodnote 10 (čo sa obvykle považujú za hraničné čísla od ktorých vyššie hodnoty naznačujú možný problém s multikolinearitou). Na základe týchto výsledkov môžeme zhodnotiť, že koeficienty pri premenných Growth a Payout v regresnom modeli nie sú nevýznamné z dôvodu multikolinearity. Overme ešte predpoklad o homoskedasticite rezíduí, aby sme významnosti koeficientov mohli veriť.

```
> library(lmtest)
> bptest(model)

                studentized Breusch-Pagan test

data:  model
BP = 34.0193, df = 4, p-value = 7.384e-07
```

Nulovú hypotézu o konštantnom rozptyle rezíduí môžeme zamietnuť na hladine významnosti 1 %. Keďže v kvantifikovanom modeli sa preukázal výskyt heteroskedasticity, využijeme na odhad významnosti koeficientov maticu konzistentnú na prítomnosť heteroskedasticity, konkrétne typ HC0.

```
> library(sandwich)
> coeftest(model, vcov = vcovHC(model, type = "HC0"))

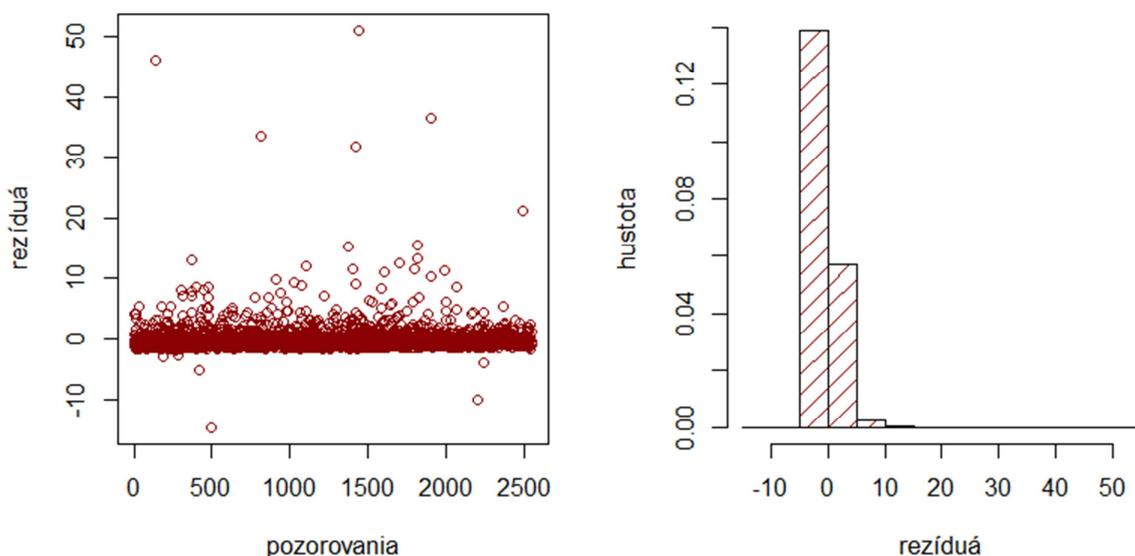
t test of coefficients:

              Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.8521966  0.0901970  20.5350 < 2.2e-16 ***
Payout       0.0029638  0.0123576   0.2398  0.810476
Growth       0.0612699  0.0823738   0.7438  0.457064
Beta        -0.1996637  0.0619125  -3.2249  0.001276 **
ROE          0.4626947  0.2065245   2.2404  0.025152 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

S využitím matice konzistentnej na prítomnosť heteroskedasticity nám významnosť koeficientu pri premennej ROE mierne klesla, stále však koeficient ostáva významný na hladine 5 %. Odhliadnuc od toho, že koeficient determinácie je naozaj nízky, ukážme si, akým spôsobom by sme z daného modelu identifikovali nesprávne ocenené akcie. Vychádzajúc z predpokladu, že nami vytvorená regresná nadroviná predstavuje správne ocenené akcie, za nadhodnotenú akcie môžeme považovať tie, ktoré sa nachádzajú nad regresnou nadrovinou. Ako podhodnotenú potom môžeme vyhodnotiť tie akcie, ktoré sa

nachádzajú pod regresnou nadrovinou. Grafická reprezentácia modelu nám v tomto prípade nepomôže (keďže máme 5 premenných), avšak za účelom zistenia, ktoré akcie sa nachádzajú pod/nad regresnou nadrovinou vieme využiť rezíduá.

```
> par(mfrow = c(1, 2))
> plot(model$residuals, type = "p", ylab = "rezíduá", main = NA,
      cex.lab = 1.1, cex.axis = 1.0, xlab = "pozorovania", col =
      "dark red")
> hist(model$residuals, density = 10, main = NA, cex.lab = 1.1,
      cex.axis = 1.0, freq = FALSE, ylab = "hustota", xlab =
      "rezíduá", col = "dark red", border = "black")
```



Obrázok 26: Rezíduá z regresného modelu ukazovateľa P/B

*Zdroj: výstup zo softvéru R*

Okrem grafickej vizualizácie rezíduí sa vieme tiež pozrieť na ich základnú deskriptívnu štatistiku. V zásade sú pre nás najzaujímavejšie hlavne tie pozorovania, ktoré sú najviac vzdialené od regresnej nadroviny o oboch smeroch. Pozorovania blízke nule by sme mohli vylúčiť ako správne ocenené podniky.

```
> summary(model$residuals)
      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
-14.7400 -0.9883  -0.5753   0.0000  0.2916  51.0500
```

Môžeme vidieť, že priemerná hodnota rezíduí je rovná nule, inak je rozdelenie rezíduí výrazne pravostranne zošikmené. Aj z grafickej vizualizácie je vidieť, že viac podnikov je nadhodnotených ako podhodnotených. Ak by sme chceli zistiť, ktoré podniky sú najviac podhodnotené alebo nadhodnotené (v zmysle finančného investovania vieme vytvoriť pozíciu

k obom prípadom – uzatvorením dlhej pozície pri podhodnotených akciách a uzatvorením krátkej pozície pri nadhodnotených akciách), je vhodné si rezíduá zoradiť od najmenšieho po najväčšie, teda zostaviť variačný rad. Za týmto účelom využijeme funkciu `order()` a následne si zobrazíme 10 podnikov najviac podhodnotených (označíme ako `top10_under`) a 10 podnikov najviac nadhodnotených (označíme ako `top10_over`). Keďže využívame funkcie `head()` a `tail()`, ktoré nám vrátia prvých 10 a posledných 10 podnikov z variačného radu, je nutné si uvedomiť, že poradie pri nadhodnotených podnikoch nie je zoradené od najviac nadhodnoteného po najmenej nadhodnotený podnik, ale naopak.

```

> resid = data.frame(model$residuals, Ticker)
> order_resid = resid[with(resid, order(-model$residuals,
  Ticker, decreasing = TRUE)),]
-----
> top10_under = head(order_resid, n = 10); top10_under
  model.residuals      Ticker
500      -14.739831      LSE:WRN
2206      -10.163555 ENXTAM:ATVAM
419       -5.200815       DB:MXC
2246      -3.812095      ENXTPA:AC
189       -2.947760      NGM:AIK B
286       -2.610962       DB:HET
263       -1.886634      AIM:HCP
245       -1.806668      OB:PLUG
33        -1.764288      ENXTPA:BLEE
2539      -1.756454      BDM:FIA
-----
> top10_over = tail(order_resid, n = 10); top10_over
  model.residuals      Ticker
365         13.00447      AIM:IDG
1821        13.27235      OM:AXIS
1374        15.17558      CATS:ZOT
1816        15.53945      AIM:MUL
2495        21.18255      ENXTPA:PRC
1424        31.71512      HLSE:TLA1V
810         33.59534      LSE:NXT
1907        36.58852      NYSE:CLB
141         46.11601      LSE:DTY
1444        51.04712      LSE:RMV

```

Zaujímavé je, že tak najviac nadhodnotený podnik (RMV), ako aj najviac podhodnotený podnik (WRN), oba sú obchodované na Londýnskej burze (LSE, z angl. London Stock Exchange) a oba podniky pracujú s nehnuteľnosťami (WRN ako developer a RMV ako online inzerent).

Ukázali sme si, ako je možné využiť jednoduchý lineárny regresný model pri oceňovaní akcií. Samozrejme pre investičné rozhodnutie by bola nutná hlbšia analýza, pričom výsledky z kvantifikovaného modelu sú spochybniteľné aj kvôli veľmi nízkemu koeficientu determinácie. Pozrime sa na kvantifikáciu tohto modelu po jednotlivých regionálnych skupinách (rozdelíme vzorku podľa premennej Sub.Group).

```
> model_EU = lm(PB ~ Payout + Growth + Beta + ROE, data =
  subset(data, Sub.Group=="EU")); summary(model_EU)

Call:
lm(formula = PB ~ Payout + Growth + Beta + ROE, data =
  subset(data, Sub.Group == "EU"))

Residuals:
    Min       1Q   Median       3Q      Max
-13.269  -0.817  -0.443   0.280  36.348

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.52889     0.08782  17.410  <2e-16 ***
Payout       0.01049     0.01482   0.708   0.4790
Growth      0.01719     0.06332   0.271   0.7861
Beta       -0.14931     0.06475  -2.306   0.0212 *
ROE        1.16289     0.11189  10.393  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.889 on 1933 degrees of freedom
Multiple R-squared:  0.0559,    Adjusted R-squared:  0.05395
F-statistic: 28.62 on 4 and 1933 DF,  p-value: < 2.2e-16
```

Model kvantifikovaný na vzorke podnikov z podkategórie EU sa nejaví nijako zvlášť odlišne v porovnaní s modelom kvantifikovaným na celej vzorke. Významné sú opäť len dva koeficienty (pri ukazovateli Beta a ROE) a znamienka pri týchto koeficientoch sú taktiež rovnaké. Čo sa týka popísanej variability, tak koeficient determinácie tiež nadobúda len veľmi nízke hodnoty. Opäť sa môžeme pozrieť, či multikolarita na danej podvzorke nezapríčiňuje nevýznamnosť niektorých regresných koeficientov<sup>48</sup>.

```
> library(HH)
> vif(model_EU)
  Payout   Growth   Beta   ROE
1.001952 1.002027 1.001896 1.002159
```

<sup>48</sup> Zrejme to nebude ten prípad, keďže korelácie medzi vysvetľujúcimi premennými v rámci celej vzorky boli veľmi nízke, ale napriek tomu je vždy vhodné overiť výskyt všetkých možných problémov súvisiacich s kvantifikáciou regresného modelu.

Všetky VIF sú opäť veľmi nízke, čiže multikolinearita ani v tomto modeli nepredstavuje problém. Skontrolujeme ešte predpoklad o homoskedasticite rezíduí.

```
> bptest(model_EU)

              studentized Breusch-Pagan test

data:  model_EU
BP = 48.064, df = 4, p-value = 9.152e-10
```

Môžeme vidieť, že aj tento druhý kvantifikovaný model je zaťažený výskytom heteroskedasticity, avšak ani po použití matice konzistentnej na jej prítomnosť sa významnosti koeficientov výrazne nezmenili.

```
> library(sandwich)
> coeftest(model_EU, vcov = vcovHC(model, type = "HC0"))

t test of coefficients:

              Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.528887   0.090197 16.9505 < 2.2e-16 ***
Payout       0.010495   0.012358  0.8492  0.39585
Growth       0.017186   0.082374  0.2086  0.83475
Beta        -0.149314   0.061912 -2.4117  0.01597 *
ROE         1.162888   0.206524  5.6308 2.057e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Týmto modelom sme si teda nijak zvlášť nepomohli, ale minimálne máme možnosť porovnať výsledky na podvzorke. Pozrime sa na vzorku podnikov z UK.

```
> model_UK = lm(PB ~ Payout + Growth + Beta + ROE, data =
  subset(data, Sub.Group=="UK")); summary(model_UK)

Call:
lm(formula = PB ~ Payout + Growth + Beta + ROE, data =
  subset(data, Sub.Group == "UK"))

Residuals:
    Min       1Q   Median       3Q      Max
-11.959  -1.398  -0.844   0.316  50.789

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.31520    0.27972   8.277 8.19e-16 ***
Payout      -0.08014    0.20044  -0.400  0.689
Growth       0.82470    0.69460   1.187  0.236
Beta        -0.22091    0.20156  -1.096  0.274
ROE         0.35027    0.08111   4.319 1.84e-05 ***
---
```

```

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.795 on 601 degrees of freedom
Multiple R-squared:  0.03371,    Adjusted R-squared:  0.02728 
F-statistic: 5.242 on 4 and 601 DF,  p-value: 0.0003722

```

Pri modeli kvantifikovanom na vzorke podnikov z UK máme síce stále významný koeficient pri premennej ROE, avšak všetky ostatné premenné sa javia ako nevýznamné. Koeficient determinácie je opäť veľmi nízky. Výskyt multikolinearity sa ani v tomto regresnom modeli nepotvrdil, čo je však zaujímavé, že ani výskyt heteroskedasticity. Získané výsledky tak môžeme považovať za konečné.

```

> library(HH)
> vif(model_UK)
  Payout   Growth   Beta   ROE
1.018426 1.017323 1.006007 1.006208
-----
> bptest(model_UK)

              studentized Breusch-Pagan test

data:  model_UK
BP = 6.5678, df = 4, p-value = 0.1606

```

Poslednou podvzorkou sú podniky z východnej Európy a Ruska, teda z trhov, ktoré by sme mohli považovať za rozvíjajúce sa (v angl. emerging markets).

```

> model_ER = lm(PB ~ Payout + Growth + Beta + ROE, data =
  subset(data, Sub.Group=="Eastern Europe & Russia"));
summary(model_ER)

Call:
lm(formula = PB ~ Payout + Growth + Beta + ROE, data =
  subset(data,
    Sub.Group == "Eastern Europe & Russia"))

Residuals:
    303     761     970    1032    1295    1636
 1910    2350
 0.108556 -0.179954 -0.021932 -0.066690  0.492535 -0.417128 -
 0.005532  0.090144

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.9122     1.0572   0.863   0.452
Payout        -0.0617     0.4464  -0.138   0.899
Growth        -3.4332     2.1963  -1.563   0.216
Beta          -0.6764     0.4082  -1.657   0.196
ROE           7.6300     5.0672   1.506   0.229

```

```
Residual standard error: 0.3974 on 3 degrees of freedom
Multiple R-squared: 0.8317, Adjusted R-squared: 0.6073
F-statistic: 3.706 on 4 and 3 DF, p-value: 0.1552
```

```
-----
> library(HH)
> vif(model_ER)
  Payout   Growth      Beta      ROE
3.113564 1.766338 2.062694 1.469194
```

```
-----
> bptest(model_ER)
```

```
studentized Breusch-Pagan test
```

```
data: model_ER
BP = 0.7101, df = 4, p-value = 0.9501
```

Pri vzorke podnikov z východnej Európy a Ruska sme dosiahli výrazne vyšší koeficient determinácie, avšak ani jeden z koeficientov nie je štatisticky významný (rovnako tiež ani model ako celok). Inflačné faktory rozptylu sú síce pri tomto modeli vyššie ako v predchádzajúcich prípadoch, stále však neindikujú problém s multikolinearitou. Výskyt heteroskedasticity sa v tomto modeli nepotvrdil.

V prípade tohto posledného modelu je zrejším problémom nízky počet pozorovaní – v našej databáze máme len 8 podnikov priradených k týmto trhom, pričom všetky sú z maďarskej burzy (v databáze skratka BUSE, z angl. Budapest Stock Exchange).

```
> data_ER = matrix(subset(Ticker, Sub.Group == "Eastern Europe &
  Russia"))
> colnames(data_ER) = c("Podniky"); data_ER
  Podniky
[1,] "BUSE:ZWACK"
[2,] "BUSE:MTELEKOM"
[3,] "BUSE:EGIS"
[4,] "BUSE:ANY"
[5,] "BUSE:RICHTER"
[6,] "BUSE:FHB"
[7,] "BUSE:MOL"
[8,] "BUSE:OTP"
```

Problém s nízkym počtom pozorovaní je možné rýchlo identifikovať aj priamo vo výstupe z regresného modelu, v ktorom nám softvér R nedáva deskriptívnu štatistiku rezíduí, ale priamo vypíše všetky príslušné rezíduá.

Tento model sa teda určite vzhľadom na počet pozorovaní nedá považovať za vhodný, avšak ani predošlé kvantifikované modely nemôžeme vyhodnotiť ako veľmi vhodné, práve kvôli nízkemu koeficientu determinácie. Evidentne sú použité údaje veľmi rozptýlené, a preto

nie je možné popísať vysoké percento variability. Riešením by mohla byť kvantifikácia modelu po jednotlivých odvetviach. Minimálne by sme však radi poukázali na skutočnosť, že pred samotnou analýzou je vždy vhodné pozrieť sa na údaje, s ktorými pracujeme (čo bolo najviac markantné pri poslednom kvantifikovanom modeli). Či už na základnú deskriptívnu štatistiku alebo na ich grafickú vizualizáciu. Hneď na začiatku by sme tak mali možnosť posúdiť, či má zmysel v analýze pokračovať, resp. či nie je nutné údaje očistiť o extrémne hodnoty.

```
> summary(data.frame(PB, Payout, Growth, Beta, ROE, Sub.Group))
```

PB		Payout		Growth	
Min.	: 0.03239	Min.	: 0.0000	Min.	:-0.91700
1st Qu.:	0.70417	1st Qu.:	0.0000	1st Qu.:	-0.01862
Median :	1.13211	Median :	0.2434	Median :	0.03790
Mean :	1.74982	Mean :	0.4950	Mean :	0.07440
3rd Qu.:	2.04468	3rd Qu.:	0.4954	3rd Qu.:	0.10100
Max. :	53.59673	Max. :	83.2836	Max. :	28.25800

Beta		ROE	
Min.	:-0.668	Min.	: 0.000016
1st Qu.:	0.627	1st Qu.:	0.072084
Median :	1.011	Median :	0.131125
Mean :	1.092	Mean :	0.236951
3rd Qu.:	1.465	3rd Qu.:	0.226425
Max. :	10.155	Max. :	29.181818

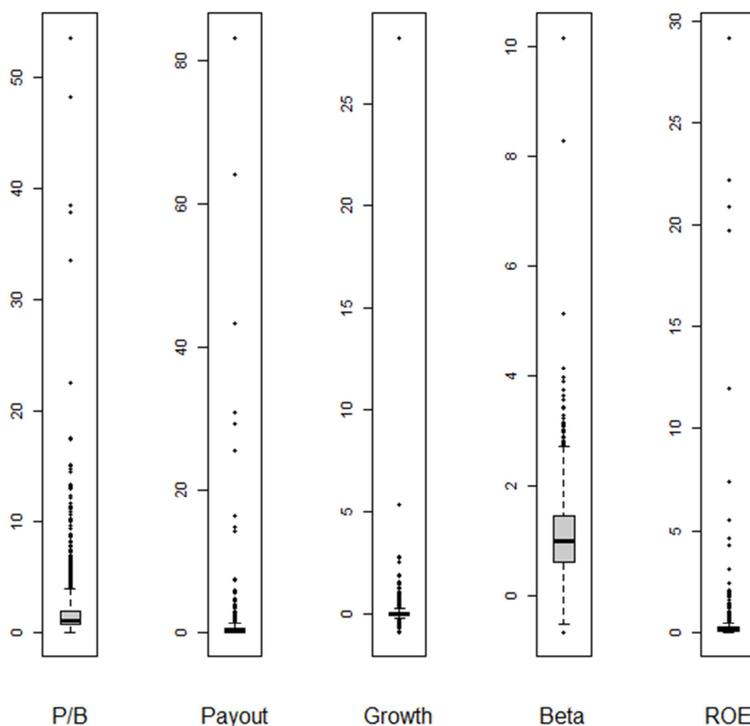
  

Sub.Group	
Eastern Europe & Russia:	8
EU	:1938
UK	: 606

Z deskriptívnej štatistiky môžeme vidieť, že výskyt extrémnych hodnôt je výrazný pri všetkých skúmaných premenných. Zrejme najmarkantnejšie je to v prípade výplatného pomeru (premenná Payout) a miery rastu (premenná Growth), pri ktorých tretí kvartil je číslo veľmi blízke nule, ale maximálne hodnoty dosahujú niekoľko násobne vyššie hodnoty. Výskyt extrémnych hodnôt je dobre vidieť aj na box – plotoch.

```
> par(mfrow = c(1, 5))
> boxplot(PB, xlab = "P/B", col = gray(0.8), pch = 19, cex.axis
= 1, cex.lab = 1.5)
> boxplot(Payout, xlab = "Payout", col = gray(0.8), pch = 19,
cex.axis = 1, cex.lab = 1.5)
> boxplot(Growth, xlab = "Growth", col = gray(0.8), pch = 19,
cex.axis = 1, cex.lab = 1.5)
> boxplot(Beta, xlab = "Beta", col = gray(0.8), pch = 19,
cex.axis = 1, cex.lab = 1.5)
```

```
> boxplot(ROE, xlab = "ROE", col = gray(0.8), pch = 19, cex.axis = 1, cex.lab = 1.5)
```



Obrázok 27: Box – plot skúmaných premenných

*Zdroj: výstup zo softvéru R*

### Príklad 6.4

Svetová banka realizovala v roku 2009 prieskum podnikateľských subjektov v rôznych krajinách sveta (bližšie informácie na [www.enterprisesurveys.org](http://www.enterprisesurveys.org)). Nás v tomto príklade budú zaujímať krajiny východnej Európy a strednej Ázie. Celkovo máme k dispozícii pozorovania za 11668 podnikateľských subjektov z 30 krajín daného regiónu. Z množstva údajov z tohto prieskumu sme pre naše potreby vybrali nasledujúce premenné:

- `ownership_W` – v tejto premennej je odpoveď na otázku, či je jedným z vlastníkov podniku žena. Premenná nadobúda tri úrovne: áno, nie, neviem.
- `management_W` – táto premenná obsahuje odpoveď na otázku, či je vo vedení podniku žena. Premenná nadobúda tri úrovne: áno, nie, neviem.
- `country` – v ktorej krajine má podnikateľský subjekt sídlo.
- `industry` – zaradenie podniku do odvetvia podľa Svetovej banky.
- `sales` – celkové tržby za posledný fiškálny rok vyjadrené v domácej mene.

- `employees` – počet zamestnancov (na plný úväzok) ku koncu posledného fiškálneho roka.
- `labor_costs` – kompletne mzdové náklady za posledný fiškálny rok.
- `sales_3y` – celkové tržby dosiahnuté za fiškálny rok tri roky dozadu.

Kvantifikujte regresný model, v ktorom ako vysvetľovaná premenná vystupujú tržby (`sales`) a vysvetľujúce premenné sú `ownership_W`, `management_W`, `sales_3y`, `employees` a `labor_costs`. Pri kvantifikácii použite celú vzorku a potom vzorku podnikov zo Slovenskej republiky (premenná `country` = "Slovak Republic").

---

### Príklad 6.4 – Riešenie

---

Za úlohu máme kvantifikovať regresný model s použitím údajov o podnikateľských subjektoch z krajín východnej Európy a strednej Ázie (prieskum Svetovej banky. Pristúpme najprv ku kvantifikácii modelu na celej vzorke pozostávajúcej z 11668 podnikateľských subjektov. Z premenných `ownership_W` a `management_W` si musíme v prvom rade vytvoriť tzv. umelé premenné, ktoré budú nadobúdať hodnoty 1 (ak je vlastníkom/manažérom žena) a 0 (ak nie je vlastníkom/manažérom žena). Keďže obe tieto premenné nadobúdajú tri úrovne, spojíme triedy a hodnotu „neviem“ nahradíme hodnotou „nie“.

```
> data = read.csv(file = "...cesta
  k súboru...\\enterprise_survey.csv", sep = ";", dec = ".",
  header = T)
> attach(data)
-----
> bin <- data.frame(ownership_W, management_W)
> NEVIEM <- bin == "neviem"
> bin_ok <- replace(bin, NEVIEM, "nie")
```

Po získaní dvoch úrovní v oboch premenných môžeme pristúpiť k vytvoreniu umelých premenných (v angl. dummy variable). Všetky hodnoty „ano“ nahradíme číslom 1 a všetky hodnoty „nie“ číslom 0. Za týmto účelom využijeme jednoduchú funkciu `recode()` z knižnice `epicalc`.

```
> library(epicalc)
> recode_1 = recode(vars = c(ownership_W, management_W),
  old.value = "ano", new.value = "1", dataFrame = bin_ok)
> recode_0 = recode(vars = c(ownership_W, management_W),
  old.value = "nie", new.value = "0", dataFrame = bin_ok)
> dummy_own = bin_ok$ownership_W
> dummy_mng = bin_ok$management_W
```

Získali sme tak dve premenné, ktoré už je možné použiť vo funkcii `lm()` na kvantifikáciu požadovaného regresného modelu. Napriek tomu, že nám vo viacerých premenných chýbajú pozorovania, nemusíme tento problém explicitne riešiť, keďže funkcia `lm()` chýbajúce pozorovania automaticky odstráni. Vo výstupe z modelu môžeme vidieť, že bolo odstránených až 3944 pozorovaní.

```
> model = lm(sales ~ sales_3y + employees + labor_costs +
  dummy_own + dummy_mng); summary(model)

Call:
lm(formula = sales ~ sales_3y + employees + labor_costs +
  dummy_own +
  dummy_mng)

Residuals:
      Min       1Q   Median       3Q      Max
-1.463e+11 -4.478e+10 -2.213e+09 -1.935e+09  9.995e+13

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  7.298e+09  3.055e+10  0.239   0.811
sales_3y      1.046e+00  3.056e+00  0.342   0.732
employees     1.483e+06  2.119e+07  0.070   0.944
labor_costs   4.625e+00  2.641e+01  0.175   0.861
dummy_own0   -4.314e+10  2.866e+10 -1.505   0.132
dummy_mng0    3.775e+10  3.631e+10  1.040   0.298

Residual standard error: 1.138e+12 on 7718 degrees of freedom
(3944 observations deleted due to missingness)
Multiple R-squared:  0.0004028, Adjusted R-squared: -0.0002447
F-statistic: 0.6221 on 5 and 7718 DF, p-value: 0.683
```

Uvedený regresný model je na prvý pohľad prakticky nepoužiteľný. Žiadny z koeficientov nie je štatisticky významný, koeficient determinácie je veľmi blízky nule (upravený koeficient determinácie je dokonca záporný) a model ako celok taktiež nie je štatisticky významný. Evidentne je použitie celej vzorky nie veľmi vhodné, keďže údaje sú až príliš heterogénne. Z tohto dôvodu sa týmto modelom ďalej zaoberať nebudeme a prejdeme rovno na kvantifikáciu modelu na podvzorke podnikateľských subjektov zo Slovenskej republiky. Pre istotu celý postup zopakujeme po vymazaní všetkých objektov v programe R (príkaz `rm(list=ls(all=TRUE))`) a nahráme údaje len za požadovanú podvzorku.

```
> rm(list=ls(all=TRUE))
> SR <- read.table(file = "...cesta
  k súboru...\\enterprise_survey.csv", sep = ";", dec = ".",
  header = T)
```

```

> SR <- subset(SR, subset = country == "Slovak Republic", select
= c(sales, sales_3y, employees, labor_costs, ownership_W,
management_W))
> attach(SR)

```

Celý postup so spájaním tried a prekódovania v rámci premenných `ownership_W` a `management_W` zopakujeme za účelom vytvorenia umelých premenných.

```

> bin <- data.frame(ownership_W, management_W)
> NEVIEM <- bin == "neviem"
> bin_ok <- replace(bin, NEVIEM, "nie")
-----
> library(epicalc)
> recode_1 = recode(vars = c(ownership_W, management_W),
old.value = "ano", new.value = "1", dataFrame = bin_ok)
> recode_0 = recode(vars = c(ownership_W, management_W),
old.value = "nie", new.value = "0", dataFrame = bin_ok)
> dummy_own = bin_ok$ownership_W
> dummy_mng = bin_ok$management_W
-----
> table(ownership_W)
ownership_W
  ano neviem  nie
  84      9  182
> table(management_W)
management_W
  ano neviem  nie
  69      1  205

```

Môžeme si všimnúť, že tieto dve premenné nenadobúdali veľa krát hodnotu „neviem“ (spolu 10 krát), čiže nemalo by ísť o zásadný zásah do použitých dát. V ďalšom kroku môžeme pristúpiť ku kvantifikácii modelu.

```

> model = lm(sales ~ sales_3y + employees + labor_costs +
dummy_own + dummy_mng); summary(model)

Call:
lm(formula = sales ~ sales_3y + employees + labor_costs +
dummy_own + dummy_mng)

Residuals:
    Min       1Q   Median       3Q      Max
-363954006 -23135469 -15455488  17265004  772401175

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.295e+07  1.846e+07  1.243  0.21581
sales_3y      9.633e-01  2.702e-02 35.652 < 2e-16 ***
employees    -1.192e+05  1.108e+05 -1.076  0.28385
labor_costs   9.935e-01  3.454e-01  2.876  0.00464 **
dummy_own0    3.538e+07  1.875e+07  1.887  0.06116 .

```

```

dummy_mng0  -3.810e+07  2.083e+07  -1.829  0.06945  .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 99990000 on 143 degrees of freedom
(126 observations deleted due to missingness)
Multiple R-squared:  0.9772,    Adjusted R-squared:  0.9764
F-statistic: 1225 on 5 and 143 DF,  p-value: < 2.2e-16

```

Model kvantifikovaný na podvzorke podnikov zo Slovenskej republiky vyzerá do značnej miery odlišne (v porovnaní s predchádzajúcim modelom). Štatisticky významné máme koeficienty pri premenných `sales_3y` a `labor_costs`. Evidente je súvis medzi tržbami za posledný fiškálny rok a tržbami dosiahnutými pred 3 rokmi. Vytvorené umelé premenné sú významné na hladine 10 %, ale keď sa pozrieme na odhadnuté koeficienty pri týchto premenných, tak si môžeme všimnúť, že sa ich vplyv na vysvetľovanú premennú prakticky kompenzuje. Zrejme žiadna z týchto premenných dosadená do modelu samostatne by nebola štatisticky významná (čitateľ si to môže vyskúšať). Za zmienku určite stojí koeficient determinácie, pretože ten je naozaj neobvykle vysoký. Príliš vysoký koeficient determinácie je samozrejme podozrivý. Pozorný čitateľ si akiste všimol, že premenná `sales_3y` v sebe zahrňuje aj závislú premennú, keďže predstavuje celkové tržby za posledné tri roky, kým závislá premenná celkové tržby za posledný rok. Necháme na čitateľovi aby analýzu zopakoval pri upravenej premennej `sales_2y = sales_3y - sales`.

Zaujímavý je ešte koeficient pri počte zamestnancov (`employees`). Ten síce nevyšiel štatisticky významný, ale zrejme je to z toho dôvodu, že vplyv tejto premennej je sprostredkovaný v mzdových nákladoch (premenná `labor_costs`). Tieto dve premenné hovoria v zásade o tom istom, čo môžeme vidieť aj na vysokej korelácii medzi nimi.

```

> cor.test(employees, labor_costs, method = "pearson",
  alternative = c("two.sided"))

Pearson's product-moment correlation

data:  employees and labor_costs
t = 30.2832, df = 206, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.8751943 0.9258669
sample estimates:
      cor
0.9036452

```

Z dôvodu vysokej korelácie medzi vysvetľujúcimi premennými dochádza k problému s multikolinearitou. Inflačný faktor rozptylu (v angl. VIF, Variance Inflation Factor) si môžeme vypočítať prostredníctvom funkcie `vif()` z knižnice `HH`.

```
> library(HH)
> vif(model)
  sales_3y  employees labor_costs  dummy_own0  dummy_mng0
  4.199279  12.051638  11.759485   1.179084   1.185098
```

Skutočne teda môžeme vidieť, že problém multikolinearity sa v našom regresnom modeli vyskytuje, čo zapríčinilo nevýznamnosť koeficientu pri premennej `employees`. Túto premennú z modelu preto odstránime a až potom sa pozrieme na výskyt heteroskedasticity.

```
> model_2 = lm(sales ~ sales_3y + labor_costs + dummy_own +
  dummy_mng); summary(model_2)

Call:
lm(formula = sales ~ sales_3y + labor_costs + dummy_own +
  dummy_mng)

Residuals:
    Min       1Q   Median       3Q      Max
-354934696 -21117071 -14717760  19186385  788708090

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.909e+07  1.806e+07   1.057  0.292449
sales_3y      9.550e-01  2.578e-02  37.042 < 2e-16 ***
labor_costs   6.909e-01  1.965e-01   3.516  0.000584 ***
dummy_own0    3.487e+07  1.848e+07   1.887  0.061108 .
dummy_mng0   -3.598e+07  2.038e+07  -1.765  0.079610 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 99460000 on 146 degrees of freedom
(124 observations deleted due to missingness)
Multiple R-squared:  0.977,    Adjusted R-squared:  0.9764
F-statistic: 1550 on 4 and 146 DF,  p-value: < 2.2e-16
```

Máme teda kvantifikovaný model, ktorý na prvý pohľad vyzerá veľmi dobre – vzhľadom na významné koeficienty a vysoký koeficient determinácie. Aby však celá analýza bola korektná, musíme ešte otestovať predpoklad o homoskedasticite rezíduí.

```
> library(lmtest)
> bptest(model_2)

studentized Breusch-Pagan test
```

```
data: model_2
BP = 14.7838, df = 4, p-value = 0.005171
```

Keďže rezíduá z daného modelu nie sú homoskedastické (zamietli sme nulovú hypotézu na hladine významnosti 1 %), musíme pri odhade modelu použiť maticu konzistentnú na prítomnosť heteroskedasticity.

```
> library(sandwich)
> coeftest(model_2, vcov = vcovHC(model, type = "HC0"))

t test of coefficients:

              Estimate  Std. Error t value Pr(>|t|)
(Intercept) 1.9085e+07  1.9854e+07  0.9613  0.3380
sales_3y     9.5498e-01  2.6707e-02 35.7572 <2e-16 ***
labor_costs  6.9089e-01  7.1809e-01  0.9621  0.3376
dummy_own0   3.4872e+07  2.4916e+07  1.3996  0.1638
dummy_mng0  -3.5983e+07  3.4250e+07 -1.0506  0.2952
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Po zohľadnení výskytu heteroskedasticity rezíduí nám ostal významný už iba jeden koeficient, a to pri dosiahnutých tržbách pred 3 rokmi (*sales\_3y*). Na tomto príklade sme tak mali možnosť vidieť, že pri nedodržaní predpokladov regresného modelu môžeme dospieť k zavádzajúcim výsledkom.

### Príklad 6.5

#### Úlohy

1. Vygenerujte si akýmkoľvek spôsobom dve premenné (*x* a *y*) o rozsahu 500 pozorovaní a kvantifikujte lineárny regresný model popisujúci ich závislosť. Premenné skúste vygenerovať tak, aby rezíduá z regresného modelu vykazovali prítomnosť heteroskedasticity.
2. Zostrojte *x-y* graf a skúste posúdiť, či rezíduá z tohto modelu budú homoskedastické.
3. Následne zostrojte graf rezíduí a tiež skúste vizuálne posúdiť, či sú homoskedastické.
4. Na záver otestujte prítomnosť heteroskedasticity s využitím Breusch – Paganovho testu a autokorelácie s využitím Durbin – Watson testu. Ak sa preukáže nesplnenie týchto predpokladov regresného modelu, kvantifikujte nové modely, ktoré budú zohľadňovať prípadný výskyt heteroskedasticity a autokorelácie.

---

### Príklad 6.5 – Riešenie

---

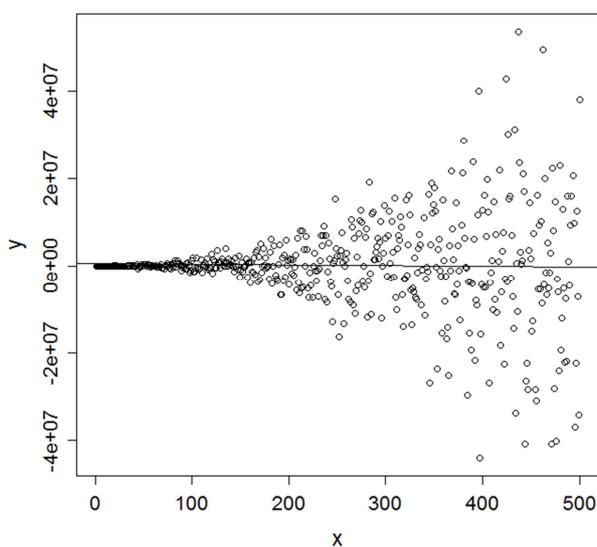
Na tomto jednoduchom príklade zameranom na testovanie heteroskedasticity a autokorelácie ukážeme jednoduchý spôsob, ktorým vieme ošetriť prípadný výskyt týchto problémov v regresných modeloch. Vygenerujeme dve premenné ( $x$  a  $y$ ) a kvantifikujeme regresný model, ktorý bude popisovať závislosť premennej  $y$  od  $x$ .

```
> x = 1:500
> y = rnorm(500,0,100)*x^2
> plot(x,y, cex.axis = 1.3, cex.lab = 1.5)
> Reg.Model = lm(y~x)
> abline(Reg.Model)
```

Ak sa pozrieme na vzťah týchto dvoch premenných (Obrázok 29) je zrejmé, že po kvantifikácii lineárneho regresného modelu nebudú rezíduá homoskedastické (pri ich vytváraní sme toto cielene chceli dosiahnuť). Upozorňujeme, že heteroskedasticita rezíduí môže byť neraz sama o sebe zaujímavým javom. Jej dôsledky sa týkajú induktívnej štatistiky, ktorá nasleduje po odhade modelu.

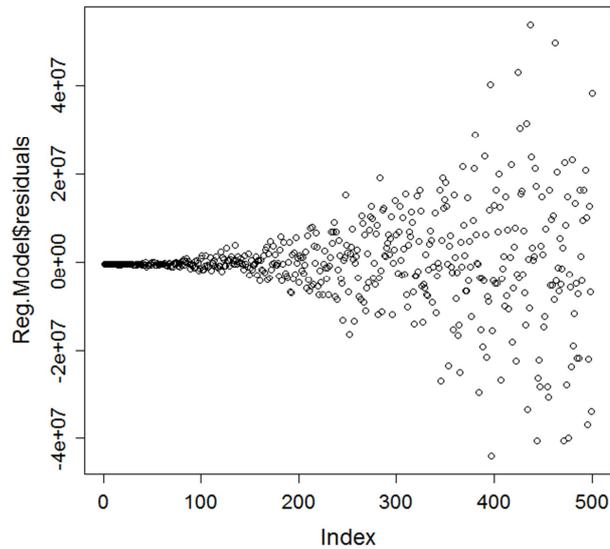
Rezíduá získame po odhade modelu a zostrojíme si okrem  $x$ - $y$  grafu aj graf rezíduí (Obrázok 29):

```
> plot(Reg.Model$residuals, cex.axis = 1.3, cex.lab = 1.5)
```



Obrázok 28: Závislosť vygenerovaných premenných

*Zdroj: výstup zo softvéru R*



Obrázok 29: Rezíduá z regresného modelu

*Zdroj: výstup zo softvéru R*

Výstup z kvantifikovaného modelu poukazuje na štatisticky významnú lineárnu závislosť medzi vygenerovanými premennými ( $p$ -hodnota  $< 0.01$ ). Koeficient determinácie je však pomerne nízky. Vzhľadom na to, že vieme ako tieto údaje vznikli, tieto výsledky nie sú prekvapivé. Štvorec lineárne sa zvyšujúcej nezávislej premennej, v kombinácii s náhodne generovanými inováciami spôsobuje, že výsledky budú v empirických modeloch ovplyvňované extrémnymi hodnotami. Napríklad v našom prípade je koeficient kladný a teda zrejme prevažujú kladné extrémny nad zápornými. Samozrejme, vieme že z dlhodobého hľadiska by medzi premennými nemal existovať žiaden lineárny vzťah.

```
> summary (Reg.Model)

Call:
lm(formula = y ~ x)

Residuals:
    Min       1Q   Median       3Q      Max
-68996008 -3582805  783635  3036416 45467068

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -1447988     966790  -1.498  0.13484
x              8976         3344   2.684  0.00752 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 10790000 on 498 degrees of freedom
Multiple R-squared:  0.01426,    Adjusted R-squared:  0.01228
F-statistic: 7.204 on 1 and 498 DF,  p-value: 0.007516
```

V ďalšom kroku pristúpime k testovaniu autokorelácie pomocou Durbin – Watson testu a heteroskedasticity pomocou Breusch – Pagan testu.

```
> library(lmtest)
> bptest(Reg.Model)

studentized Breusch-Pagan test

data:  Reg.Model
BP = 73.0006, df = 1, p-value < 2.2e-16

> dwtest (Reg.Model)

Durbin-Watson test

data:  Reg.Model
DW = 1.8855, p-value = 0.09205
alternative hypothesis: true autocorrelation is greater than 0
```

Výsledky poukazujú na výskyt heteroskedasticity, keďže nulovú hypotézu o konštantnom rozptyle rezíduí v Breusch – Paganovom teste zamietame. Taktiež pri Durbin – Watson teste je možné zamietnuť nulovú hypotézu na hladine 10 %, teda rezíduá môžeme považovať za autokorelované. Tieto výsledky znemožňujú priamočiare používanie odhadu štandardných chýb regresných koeficientov, pozri Kapitolu 3.4.3.

Jednou z alternatív je pri odhade variančno-kovariačnej matice rezíduí použiť také metódy, ktoré budú viesť ku konzistentným odhadom variančno-kovariančnej matice a to aj v prípade, ak sú rezíduá zaťažené heteroskedasticitou a autokoreláciou. V programe R môžeme za týmto účelom využiť funkcie z knižnice `sandwich`. Pri odhade matice konzistentných aj v prípade prítomnosti heteroskedasticity môžeme použiť funkciu `vcovHC()`, v prípade prítomnosti autokorelácie aj heteroskedasticity, funkciu `vcovHAC()`, prípadne iné.<sup>49</sup>

```
> library(sandwich)
> coeftest(Reg.Model, vcov = vcovHC(Reg.Model))
t test of coefficients:

              Estimate Std. Error t value Pr(>|t|)
(Intercept) -1447987.7   584863.6  -2.4758  0.01363 *
x              8975.6     3940.1   2.2780  0.02315 *
```

---

<sup>49</sup> Funkcia `vcovHAC()` vypočítava variančno-kovariančnú maticu, kde jednotlivé pozorovania sú vážené použitím váh z kernej funkcie `Quadratic – Spectral`, vyhladávací parameter sa volí podľa Andrews (1991). S týmito základnými nastaveniami budeme túto funkciu používať aj v ďalšom texte. Upozorňujeme ale, že možností je podstatne viac, napr. Kiefer – Vogelsang (2005).

```

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
-----
> library(sandwich)
> coeftest(Reg.Model, vcov = vcovHAC(Reg.Model))
t test of coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -1447987.7   594764.4 -2.4346  0.01526 *
x              8975.6     3976.9  2.2570  0.02444 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
-----
> library(sandwich)
> coeftest(Reg.Model, vcov = NeweyWest(Reg.Model))
t test of coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -1447987.7   584298.1 -2.4782  0.01354 *
x              8975.6     3864.0  2.3229  0.02059 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

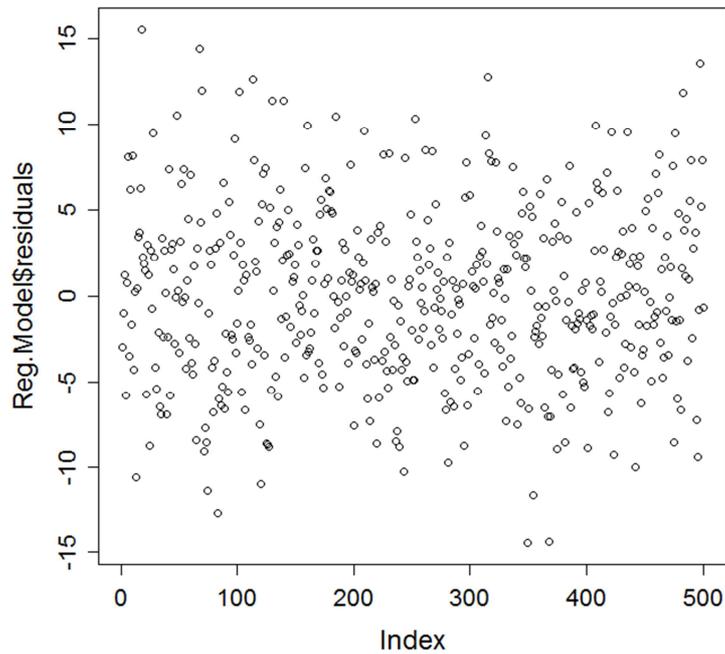
Ako môžeme vidieť, pri tomto spôsobe zohľadnenia autokorelácie a/alebo heteroskedasticity sa menia štandardné chyby (pri ktorých výpočte sa použijú zvolené matice) a teda aj  $t$ -štatistika. Vo všetkých troch prípadoch sa nám znížila významnosť regresného koeficientu pri vysvetľovanej premennej  $x$ , pričom predtým bol tento koeficient významný na hladine 1 %. Keďže stále ostáva významný na hladine 5 %, v reálnom príklade by nám to interpretáciu výsledkov nezmenilo.

Keďže jednou z úloh v tomto príklade bolo preskúmať graf rezíduí a na jeho základe sa pokúsiť o identifikáciu nesplnených predpokladov, na záver ešte uvedieme niekoľko vzorových príkladov toho, ako je možné na základe vizuálnej inšpekcie rezíduí zistiť, že s kvantifikovaným modelom sa spájajú určité problémy spojené s heteroskedasticitou alebo autokoreláciou.

```

> x = 1:500
> y = rnorm(500,0,5)
> Reg.Model = lm(y ~ x)
> plot(Reg.Model$residuals, cex.axis = 1.3, cex.lab = 1.5)

```

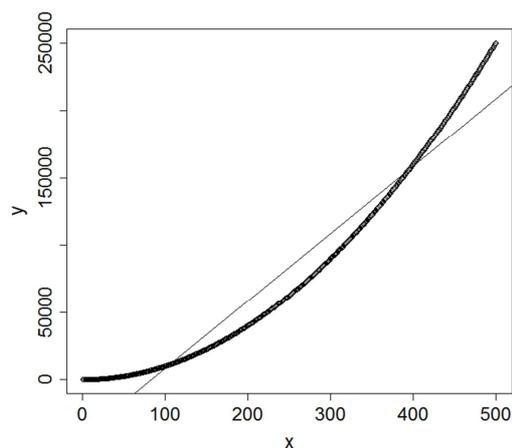


Obrázok 30: Rezíduá z regresného modelu – bez zjavného vzoru

*Zdroj: výstup zo softvéru R*

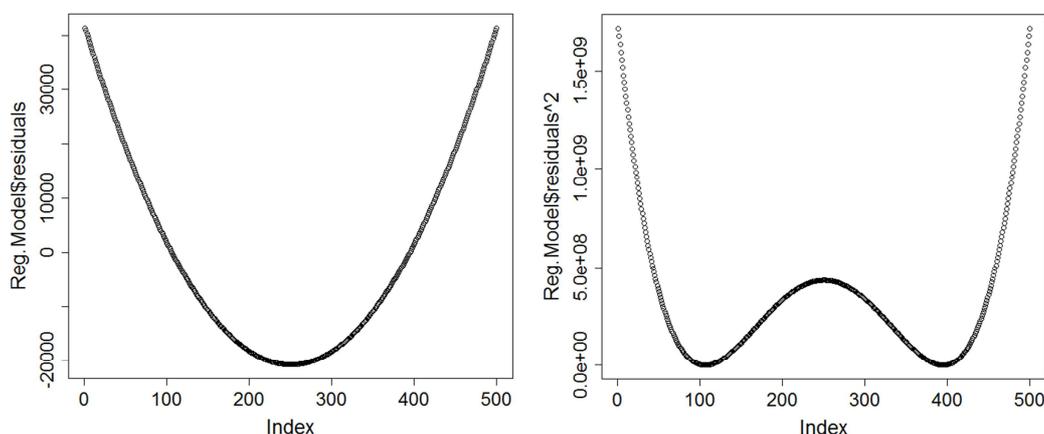
Na základe vizuálnej inšpekcie môže byť na prvý pohľad zrejmé, že údaje sú zaťažené problémom heteroskedasticity, alebo že špecifikácia modelu nie je správna. Napríklad, že miesto lineárneho vzťahu by bolo vhodné uvažovať o nejakom inom funkčnom vzťahu (model môže ostať lineárny v parametroch, ale nie v použitých premenných). Na ilustráciu uvádzame nasledujúci príklad:

```
> x = 1:500
> y = x^2
> Reg.Model = lm(y~x)
> plot(y~x, cex.axis = 1.6, cex.lab = 1.7)
> abline(Reg.Model)
> plot(Reg.Model$residuals, cex.axis = 1.6, cex.lab = 1.7)
> plot(Reg.Model$residuals^2, cex.axis = 1.6, cex.lab = 1.7)
```



Obrázok 31: Kvadratická závislosť dvoch premenných

*Zdroj: výstup zo softvéru R*



Obrázok 32: Rezíduá z regresného modelu – so zjavným vzorom

*Zdroj: výstup zo softvéru R*

### Príklad 6.6

V tomto príklade budeme pracovať s údajmi z databázy RegDat Štatistického úradu Slovenskej republiky (pozri aj Lyócsa et al., 2013), konkrétne s reálnou mzdou a nezamestnanosťou po okresoch v rámci SR za rok 2010. Jednotlivé okresy sú priradené ku krajom prostredníctvom premennej kraj (Bratislavský [BA] = 1, Trnavský [TT] = 2, Trenčiansky [TN] = 3, Nitriansky [NR] = 4, Žilinský [ZA] = 5, Banskobystrický [BB] = 6, Prešovský [PO] = 7, Košický [KE] = 8). Zrejme môžeme očakávať, že reálna mzda v Bratislavskom kraji je vyššia ako priemerná reálna mzda v SR. Taktiež nezamestnanosť v Banskobystrickom a Prešovskom kraji bude zrejme vyššia ako je celoslovenský priemer. Pokúste sa zistiť, ktoré úrovne nezamestnanosti a reálnej mzdy môžeme v okresoch považovať za odľahlé hodnoty – zaujímajú nás extrémne veľké hodnoty. Uvažujeme

o hladine významnosti 5 %. Kvantifikujte regresný model, v ktorom reálnu mzdu vysvetľujete pomocou nezamestnanosti. Otestujte tiež prítomnosť extrémnych hodnôt v skúmaných premenných.

```
okres <- c("OkresBratislavaI", "OkresBratislavaII",
  "OkresBratislavaIII", "OkresBratislavaIV", "OkresBratislavaV",
  "OkresMalacky", "OkresPezinok", "OkresSenec",
  "OkresDunajskáStreda", "OkresGalanta", "OkresHlohovec",
  "OkresPiešťany", "OkresSenica", "OkresSkalica", "OkresTrnava",
  "OkresBánovcenadBebravou", "OkresIlava", "OkresMyjava",
  "OkresNovéMestonadVáhom", "OkresPartizánske",
  "OkresPovažskáBystrica", "OkresPrievidza", "OkresPúchov",
  "OkresTrenčín", "OkresKomárno", "OkresLevice", "OkresNitra",
  "OkresNovéZámky", "OkresŠaľa", "OkresTopoľčany",
  "OkresZlatéMoravce", "OkresTvrdošín", "OkresŽilina",
  "OkresBytča", "OkresČadca", "OkresDolnýKubín",
  "OkresKysuckéNovéMesto", "OkresLiptovskýMikuláš",
  "OkresMartin", "OkresNámestovo", "OkresRužomberok",
  "OkresTurčianskeTeplice", "OkresVeľkýKrtíš", "OkresZvolen",
  "OkresŽarnovica", "OkresŽiarnadHronom", "OkresBanskáBystrica",
  "OkresBanskáŠtiavnica", "OkresBrezno", "OkresDetva",
  "OkresKrupina", "OkresLučenec", "OkresPoltár", "OkresRevúca",
  "OkresRimavskáSobota", "OkresStaráLubovňa", "OkresStropkov",
  "OkresSvidník", "OkresVranovnadTopľou", "OkresBardejov",
  "OkresHumenné", "OkresKežmarok", "OkresLevoča",
  "OkresMedzilaborce", "OkresPoprad", "OkresPrešov",
  "OkresSabinov", "OkresSnina", "OkresSpišskáNováVes",
  "OkresTrebišov", "OkresGelnica", "OkresKošiceI",
  "OkresKošiceII", "OkresKošiceIII", "OkresKošiceIV",
  "OkresKošice-okolie", "OkresMichalovce", "OkresRožňava",
  "OkresSobrance")
kraj <- c(1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 3, 3, 3,
  3, 3, 3, 3, 3, 4, 4, 4, 4, 4, 4, 4, 5, 5, 5, 5, 5, 5, 5, 5,
  5, 5, 5, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 7, 7, 7, 7, 7,
  7, 7, 7, 7, 7, 7, 7, 7, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8)
real_mzda <- c(1136.262365, 1182.604046, 1045.361376,
  945.5485251, 956.2427591, 803.8499242, 611.3537118,
  691.560467, 597.0947331, 647.0011585, 726.3167276,
  654.1306479, 616.7008288, 726.3167276, 800.2851796,
  569.4679619, 657.6953926, 566.7944034, 691.560467,
  571.2503342, 616.7008288, 644.3276, 713.8401212, 689.7780946,
  519.5615364, 712.0577489, 684.4309776, 573.0327065,
  663.0425096, 571.2503342, 603.3330363, 560.5561002,
  732.5550307, 619.3743873, 544.5147491, 630.0686213,
  702.254701, 602.4418501, 756.6170573, 531.1469566,
  726.3167276, 556.1001693, 555.2089832, 675.5191159,
  708.4930042, 652.3482756, 696.0163978, 553.4266108,
  581.0533821, 653.2394617, 509.7584885, 539.1676321,
  485.696462, 564.1208448, 517.7791641, 553.4266108,
  499.9554407, 537.3852598, 513.3232332, 438.463595,
  562.3384725, 544.5147491, 598.8771054, 509.7584885,
  657.6953926, 652.3482756, 552.5354247, 482.1317173,
```

```

581.0533821, 548.97068, 517.7791641, 805.6322966, 877.8183763,
622.939132, 786.0262009, 606.0065948, 606.8977809, 612.244898,
626.5038767)
nezamestnanost <- c(3.18, 4.6, 3.8, 3.58, 3.98, 7.46, 5.96,
5.71, 11.01, 6.29, 7.87, 6.98, 10.92, 8.76, 6.15, 9.34, 6.72,
8.76, 7.63, 12.27, 12.49, 12.05, 7.51, 7.25, 16.34, 14.55,
7.52, 12.48, 10.11, 11, 10.31, 13.02, 8.41, 14.87, 11.04,
13.87, 12.23, 11.18, 9.15, 13.58, 11.17, 11.47, 23.71, 9.24,
20.23, 14.39, 8.95, 17.1, 18, 16.16, 19.95, 23.2, 22.06,
28.83, 33.64, 13.63, 17.14, 18.8, 19.68, 19.43, 15.66, 26.18,
18.41, 19.4, 10.65, 16.6, 25.71, 19.38, 16.28, 24.42, 19.14,
8.07, 9.27, 8.59, 7.82, 21.27, 17.21, 26.82, 20.34)

```

### Príklad 6.6 – Riešenie

Pre lepšiu predstavu o použitých údajoch si najskôr premenné zobrazíme v podobe box – plotov. Jednotlivé okresy sú priradené ku krajom prostredníctvom premennej kraj (Bratislavský [BA] = 1, Trnavský [TT] = 2, Trenčiansky [TN] = 3, Nitriansky [NR] = 4, Žilinský [ZA] = 5, Banskobystrický [BB] = 6, Prešovský [PO] = 7, Košický [KE] = 8).

```

> data = read.csv(file = "...cesta k súboru...\\sk_data.csv",
  sep = ";", dec = ".", header = T)
> attach(data)
-----
> BA = subset(real_mzda, subset = kraj == 1)
> TT = subset(real_mzda, subset = kraj == 2)
> TN = subset(real_mzda, subset = kraj == 3)
> NR = subset(real_mzda, subset = kraj == 4)
> ZA = subset(real_mzda, subset = kraj == 5)
> BB = subset(real_mzda, subset = kraj == 6)
> PO = subset(real_mzda, subset = kraj == 7)
> KE = subset(real_mzda, subset = kraj == 8)

```

Keď už máme okresy zotriedené podľa príslušnosti ku kraju, môžeme pristúpiť k tvorbe box – plotov.

```

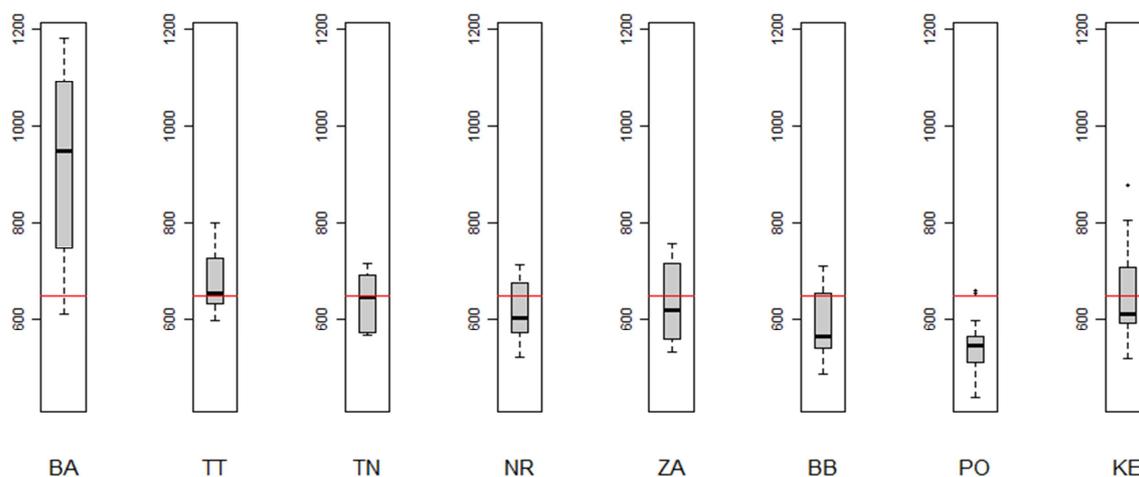
> par(mfrow = c(1, 8))
> minimum = min(real_mzda)
> maximum = max(real_mzda)
-----
> boxplot(BA, xlab = "BA", col = gray(0.8), pch = 19, cex.axis =
  1, cex.lab = 1.5, ylim = c(minimum, maximum))
> abline(h = mean(real_mzda), lwd = 1, lty = 1, col = "red")
> boxplot(TT, xlab = "TT", col = gray(0.8), pch = 19, cex.axis =
  1, cex.lab = 1.5, ylim = c(minimum, maximum))
> abline(h = mean(real_mzda), lwd = 1, lty = 1, col = "red")
> boxplot(TN, xlab = "TN", col = gray(0.8), pch = 19, cex.axis =
  1, cex.lab = 1.5, ylim = c(minimum, maximum))
> abline(h = mean(real_mzda), lwd = 1, lty = 1, col = "red")

```

```

> boxplot(NR, xlab = "NR", col = gray(0.8), pch = 19, cex.axis =
  1, cex.lab = 1.5, ylim = c(minimum, maximum))
> abline(h = mean(real_mzda), lwd = 1, lty = 1, col = "red")
> boxplot(ZA, xlab = "ZA", col = gray(0.8), pch = 19, cex.axis =
  1, cex.lab = 1.5, ylim = c(minimum, maximum))
> abline(h = mean(real_mzda), lwd = 1, lty = 1, col = "red")
> boxplot(BB, xlab = "BB", col = gray(0.8), pch = 19, cex.axis =
  1, cex.lab = 1.5, ylim = c(minimum, maximum))
> abline(h = mean(real_mzda), lwd = 1, lty = 1, col = "red")
> boxplot(PO, xlab = "PO", col = gray(0.8), pch = 19, cex.axis =
  1, cex.lab = 1.5, ylim = c(minimum, maximum))
> abline(h = mean(real_mzda), lwd = 1, lty = 1, col = "red")
> boxplot(KE, xlab = "KE", col = gray(0.8), pch = 19, cex.axis =
  1, cex.lab = 1.5, ylim = c(minimum, maximum))
> abline(h = mean(real_mzda), lwd = 1, lty = 1, col = "red")

```



Obrázok 33: Box – ploty reálnych miezd v krajoch SR

*Zdroj: výstup zo softvéru R*

Z uvedených grafov je zrejmé, že reálna mzda v Bratislavskom kraji je výrazne vyššia ako v ostatných krajoch SR. Okrem Trnavského a Bratislavského kraja, medián reálnej mzdy je vo všetkých ostatných krajoch nižší ako priemer (v grafoch zvýraznený červenou čiarou). Najnižšie reálne mzdy sa dosahujú v okresoch Prešovského kraja. Pri vizualizácii druhej premennej (nezamestnanosť v okresoch SR) budeme postupovať rovnakým spôsobom.

```

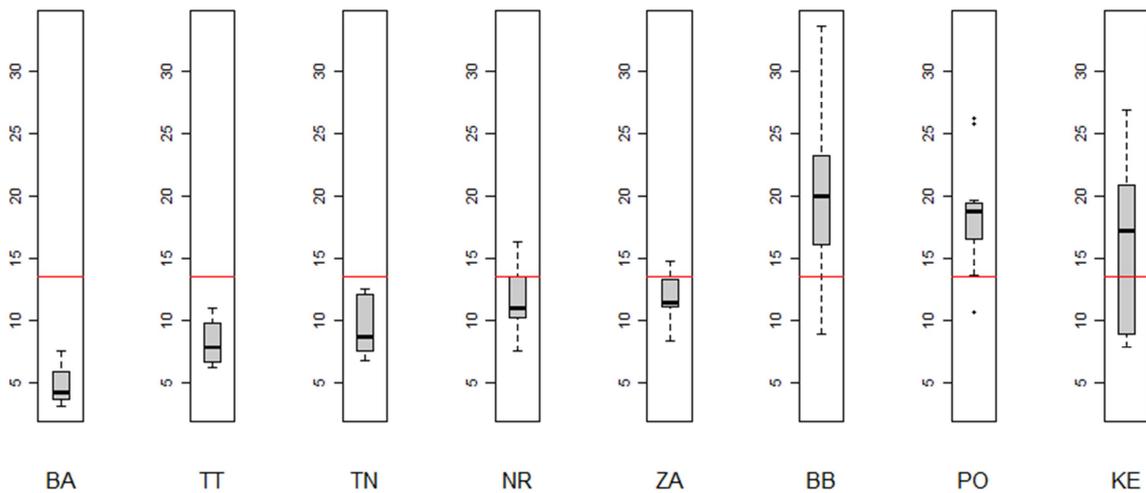
> data = read.csv(file = "...cesta k súboru...\sk_data.csv",
  sep = ";", dec = ".", header = T)
> attach(data)
-----
> BA = subset(nezamestnanost, subset = kraj == 1)
> TT = subset(nezamestnanost, subset = kraj == 2)
> TN = subset(nezamestnanost, subset = kraj == 3)
> NR = subset(nezamestnanost, subset = kraj == 4)
> ZA = subset(nezamestnanost, subset = kraj == 5)
> BB = subset(nezamestnanost, subset = kraj == 6)

```

```
> PO = subset(nezamestnanost, subset = kraj == 7)
> KE = subset(nezamestnanost, subset = kraj == 8)
```

Najprv sme si opäť rozdelili údaje za jednotlivé okresy podľa premennej kraj, teda podľa príslušnosti okresu ku kraju v SR. Následne môžeme prísť k vizualizácii danej premennej formou box – plotov.

```
> par(mfrow = c(1, 8))
> minimum = min(nezamestnanost)
> maximum = max(nezamestnanost)
-----
> boxplot(BA, xlab = "BA", col = gray(0.8), pch = 19, cex.axis =
1, cex.lab = 1.5, ylim = c(minimum, maximum))
> abline(h = mean(nezamestnanost), lwd = 1, lty = 1, col =
"red")
> boxplot(TT, xlab = "TT", col = gray(0.8), pch = 19, cex.axis =
1, cex.lab = 1.5, ylim = c(minimum, maximum))
> abline(h = mean(nezamestnanost), lwd = 1, lty = 1, col =
"red")
> boxplot(TN, xlab = "TN", col = gray(0.8), pch = 19, cex.axis =
1, cex.lab = 1.5, ylim = c(minimum, maximum))
> abline(h = mean(nezamestnanost), lwd = 1, lty = 1, col =
"red")
> boxplot(NR, xlab = "NR", col = gray(0.8), pch = 19, cex.axis =
1, cex.lab = 1.5, ylim = c(minimum, maximum))
> abline(h = mean(nezamestnanost), lwd = 1, lty = 1, col =
"red")
> boxplot(ZA, xlab = "ZA", col = gray(0.8), pch = 19, cex.axis =
1, cex.lab = 1.5, ylim = c(minimum, maximum))
> abline(h = mean(nezamestnanost), lwd = 1, lty = 1, col =
"red")
> boxplot(BB, xlab = "BB", col = gray(0.8), pch = 19, cex.axis =
1, cex.lab = 1.5, ylim = c(minimum, maximum))
> abline(h = mean(nezamestnanost), lwd = 1, lty = 1, col =
"red")
> boxplot(PO, xlab = "PO", col = gray(0.8), pch = 19, cex.axis =
1, cex.lab = 1.5, ylim = c(minimum, maximum))
> abline(h = mean(nezamestnanost), lwd = 1, lty = 1, col =
"red")
> boxplot(KE, xlab = "KE", col = gray(0.8), pch = 19, cex.axis =
1, cex.lab = 1.5, ylim = c(minimum, maximum))
> abline(h = mean(nezamestnanost), lwd = 1, lty = 1, col =
"red")
```



Obrázok 34: Box – ploty nezamestnanosti v krajoch SR

Zdroj: výstup zo softvéru R

Pri premennej nezamestnanosť je situácia v porovnaní s reálnou mzdou mierne odlišná. V Bratislavskom kraji, v ktorom boli dosahované najvyššie reálny mzdy je nezamestnanosť najnižšia. Naproti tomu v krajoch, v ktorých bola reálna mzda najnižšia (Banskobystrický a Prešovský), je vykazovaná najvyššia nezamestnanosť. Box – ploty síce nie sú najvhodnejším typom grafu na odhalenie funkčnej závislosti ( $x$ - $y$  graf zobrazíme nižšie), napriek tomu však určitú mieru vzťahu medzi týmito dvoma premennými môžeme pozorovať. Pristúpme ku kvantifikácii regresného modelu.

```
> model = lm(real_mzda ~ nezamestnanost); summary(model)

Call:
lm(formula = real_mzda ~ nezamestnanost)

Residuals:
    Min       1Q   Median       3Q      Max
-144.86  -81.57  -28.40   49.67  414.70

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   830.116    27.638  30.035 < 2e-16 ***
nezamestnanost  -13.523     1.841  -7.344 1.84e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 106.8 on 77 degrees of freedom
Multiple R-squared:  0.4119,    Adjusted R-squared:  0.4043
F-statistic: 53.94 on 1 and 77 DF, p-value: 1.841e-10
```

Môžeme vidieť, že ak sa nezamestnanosť zvýši o jeden percentuálny bod, tak reálna mzda poklesne o približne 13 EUR. Takýto inverzný vzťah medzi mzdou a nezamestnanosťou sme mohli predpokladať už na základe box – plotov. Tento vzťah teda vyjadruje, že v regiónoch s vyššou nezamestnanosťou sú reálne mzdy nižšie (a vice versa). Tento výsledok je do značnej miery logický. Ak by sme považovali ponuku práce medzi jednotlivými okresmi za konštantnú a mieru nezamestnanosti vnímali ako prejav dopytu po práci, potom zjavne pri väčšej miere nezamestnanosti je menšia ponuka práce a teda aj reálna mzda by mala byť nižšia. Náš regresný model tejto intuícii neodporuje.

Po využití Breusch – Pagan testu však môžeme skonštatovať, že v modeli je prítomná heteroskedasticita.

```
> library(lmtest)
> bptest(model)

                studentized Breusch-Pagan test

data:  model
BP = 5.9289, df = 1, p-value = 0.01489
```

Pre overenie významnosti regresných koeficientov využijeme maticu konzistentnú na prítomnosť heteroskedasticity, konkrétne typ HC4 (ktorý je vhodný pre menšie vzorky, Zeileis; 2004).

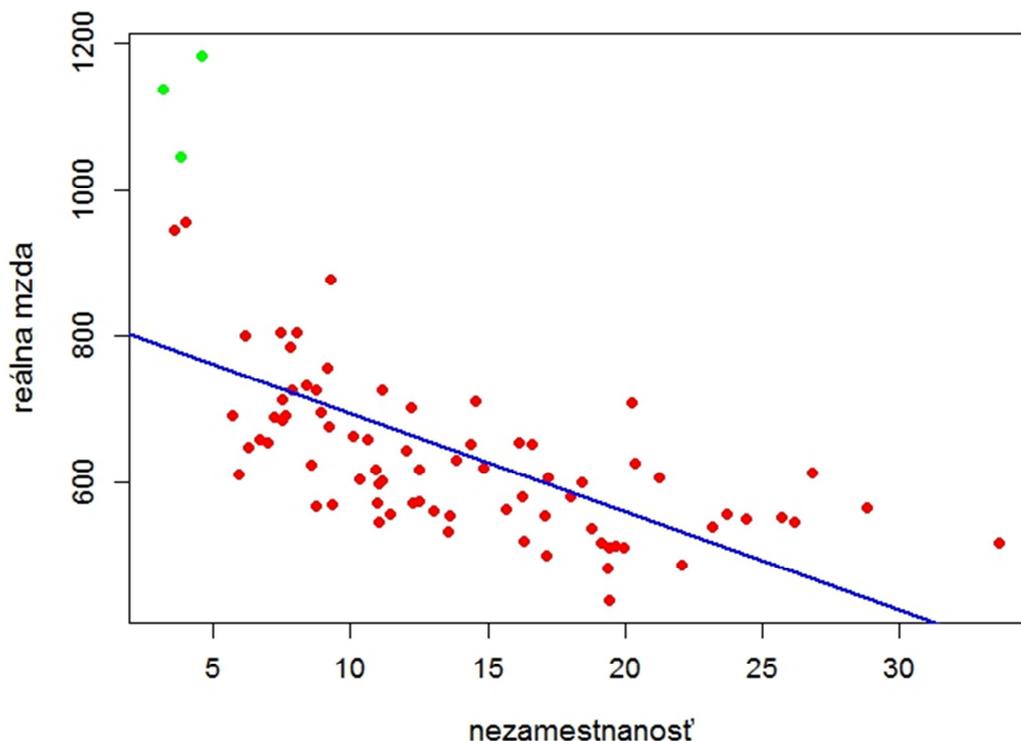
```
> library(sandwich)
> coeftest(model, vcov = vcovHC(model, type = "HC4"))

t test of coefficients:

                Estimate Std. Error t value Pr(>|t|)
(Intercept)    830.1160    41.5112  19.9974 < 2.2e-16 ***
nezamestnanost -13.5234     2.6035  -5.1942 1.633e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Regresný koeficient pri nezamestnanosti ostáva významný aj po ošetrení modelu na prítomnosť heteroskedasticity. Pozrime sa, ako vyzerá funkčný vzťah medzi skúmanými premennými v grafickej podobe.

```
> plot(real_mzda ~ nezamestnanost, xlab = "nezamestnanosť", ylab =
"reálna mzda", col = "red", type = "p", pch = 19, cex.axis =
1.1, cex.lab = 1.2, )
> abline(model, lwd = 2, col = "blue")
> points(x = nezamestnanost[1:5], y = real_mzda[1:5], col =
"green", pch=19)
```



Obrázok 35: Funkčný vzťah medzi reálnou mzdou a nezamestnanosťou v okresoch SR

*Zdroj: výstup zo softvéru R*

V uvedenom grafe sme zelenou farbou zvýraznili okresy, ktoré by sme mohli považovať za extrémne. Na identifikáciu odľahlých hodnôt využijeme funkciu `grubbs.test()` v knižnici `outliers`. Nevýhodou tohto testu je, že v nulovej hypotéze testuje prítomnosť vždy len jednej odľahlej hodnoty. Preto po identifikovaní prvého outliera je nutné túto hodnotu odstrániť a test zopakovať. Extrémne hodnoty sme sa rozhodli identifikovať na rezíduách. Hľadáme teda také pozorovania, pre ktoré bol model najviac nepresný. Držíme sa myšlienky, že také pozorovania zrejme patria do inej populácie, keďže náš model ich nevedel dostatočne presne vysvetliť.

```
> rezidua <- round(residuals(model), 4)
> library(outliers)
> grubbs.test(rezidua)

Grubbs test for one outlier

data:  round(rezidua, 4)
G.2 = 3.9093, U = 0.8016, p-value = 0.001537
alternative hypothesis: highest value 414.6956 is an outlier
-----
> extrem <- c(414.6956)
> rezidua_ciste = rezidua[!(rezidua %in% extrem)]
> grubbs.test(rezidua_ciste)
```

```

Grubbs test for one outlier

data: rezidua_ciste
G.1 = 3.7083, U = 0.8191, p-value = 0.004043
alternative hypothesis: highest value 349.1507 is an outlier
-----
> extrem <- c(extrem, 349.1507)
> rezidua_ciste = rezidua[!(rezidua %in% extrem)]
> grubbs.test(rezidua_ciste)

Grubbs test for one outlier

data: rezidua_ciste
G.3 = 3.1759, U = 0.8655, p-value = 0.03995
alternative hypothesis: highest value 266.6342 is an outlier
-----
> extrem <- c(extrem, 266.6342)
> rezidua_ciste = rezidua[!(rezidua %in% extrem)]
> grubbs.test(rezidua_ciste)

Grubbs test for one outlier

data: rezidua_ciste
G.5 = 2.3729, U = 0.9239, p-value = 0.6035
alternative hypothesis: highest value 179.9498 is an outlier
-----
> okres[rezidua %in% extrem]
[1] "OkresBratislavaI"    "OkresBratislavaII"
    "OkresBratislavaIII"

```

Na hladine významnosti 5 % sme postupne identifikovali tri extrémne hodnoty. Konkrétne ide o okresy Bratislava I až Bratislava III. Naše podozrenie sa teda nepotvrdilo a z piatich označených pozorovaní sme pomocou testu vybrali iba tri.

Následne kvantifikujeme model bez extrémnych hodnôt

```

> real_mzda_cista = real_mzda[!(rezidua %in% extrem)]
> nezamestnanost_cista = nezamestnanost[!(rezidua %in% extrem)]
> model_2 = lm(real_mzda_cista ~ nezamestnanost_cista);
summary(model_2)

Call:
lm(formula = real_mzda_cista ~ nezamestnanost_cista)

Residuals:
    Min       1Q   Median       3Q      Max
-133.874  -59.714   -4.018   47.764  226.760

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)          769.96      21.85  35.241  <2e-16 ***

```

```

nezamestnanost_cista    -10.17      1.43   -7.115    6e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 79.21 on 74 degrees of freedom
Multiple R-squared:  0.4062,    Adjusted R-squared:  0.3982 
F-statistic: 50.62 on 1 and 74 DF,  p-value: 5.996e-10
-----
> bptest(model_2)

studentized Breusch-Pagan test

data:  model_2
BP = 3.265, df = 1, p-value = 0.07078

```

Hodnota odhadovaného regresného koeficientu sa mierne zvýšila. Množstvo variability v reálnej mzde, ktorú sa nám modelom podarilo vysvetliť, sa prakticky nezmenilo a ostalo na  $\approx 40\%$ . Na hladine významnosti  $\alpha = 0.05$  sa heteroskedasticita v tomto modeli už nepreukázala. Na záver sme sa rozhodli ešte overiť, či sa nevytvorili nové extrémne hodnoty.

```

> rezidua <- round(residuals(model_2),4)
> summary(rezidua)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
-133.900 -59.710  -4.018   0.000  47.760  226.800
> grubbs.test(rezidua)

Grubbs test for one outlier

data:  rezidua
G.2 = 2.8820, U = 0.8878, p-value = 0.1176
alternative hypothesis: highest value 226.76 is an outlier

```

Najprv sme si zostrojili prehľad rezíduí, ktorý ukázal, že kandidátom je kladné rezíduum, 226.8. Test však nulovú hypotézu nevedel zamietnuť. Pozornému čitateľovi zrejme neušlo, že pri overovaní prítomnosti extrémnych hodnôt je potrebné skúmať hodnoty tak maximálne ako aj minimálne. My sme sa zámerne venovali iba tým kladným, keďže z povahy Bratislavského kraja sme mali podozrenie, že náš model bude okresy v tomto kraji podhodnocovať.

V tomto príklade sme mali možnosť (okrem iného) vidieť, že extrémne hodnoty často krát ovplyvňujú výšku koeficientu determinácie. Napriek skutočnosti, že podiel vysvetlenej variability klesne po odstránení extrémov, mali by sme vždy k tomuto kroku pristúpiť. Inak sa môže stať, že necháme niekoľko málo hodnôt ovplyvniť celý model.

## 6.2 Stacionarita a kauzalita v Grangerovom zmysle

### Príklad 6.7

Na základe nasledujúceho procesu  $w_t \sim iid(0, 1)$  si vygenerujte dva časové rady s počtom pozorovaní 200, jeden stacionárny  $x_t = 0.7x_{t-1} + w_t$  a jeden nestacionárny  $x_t = 1.2x_{t-1} + w_t$ . Tieto časové rady otestujte pomocou rozšíreného ADF testu s tým, že maximálny počet oneskorení pomocnej regresii zvolíte pomocou Schwertovho pravidla a optimálny počet oneskorení podľa Akaikeho informačného kritéria.

---

### Stacionarita a kauzalita v Grangerovom zmysle

---

#### Príklad 6.7 – Riešenie

---

Na testovanie stacionarity pomocou ADF testu využijeme funkciu `ur.df()` z knižnice `urca` a na výpočet prislúchajúcich  $p$ -hodnôt funkciu `punitroot()`. Môžeme vidieť, že pri testovaní časového radu s počtom pozorovaní 200, nám podľa Schwertovho kritéria vyšlo 14 lagov, teda toľko oneskorení nastavíme ako maximálny počet. Keďže optimálny počet lagov máme zvoliť na základe Akaikeho informačného kritéria, tak nastavíme argument funkcie `selectlags = "AIC"`.

```
> library(urca)
> x = rnorm(200,0,1)
> AR_1 = filter(x, filter = c(0.7), method = "recursive")
> AR_2 = filter(x, filter = c(1.2), method = "recursive")
-----
> k_max = as.integer(12*(length(AR_1)/100)^(1/4)); k_max
> 14
-----
> p_ADF_1n = punitroot(ur.df(AR_1, type = "none", lags = k_max,
  selectlags = "AIC")@teststat [1], N = length(AR_1), trend =
  "nc", statistic = "t")
> p_ADF_1c = punitroot(ur.df(AR_1, type = "drift", lags = k_max,
  selectlags = "AIC")@teststat [1], N = length(AR_1), trend =
  "c", statistic = "t")
> p_ADF_1ct = punitroot(ur.df(AR_1, type = "trend", lags =
  k_max, selectlags = "AIC")@teststat [1], N = length(AR_1),
  trend = "ct", statistic = "t")
-----
> p_ADF_2n = punitroot(ur.df(AR_2, type = "none", lags = k_max,
  selectlags = "AIC")@teststat [1], N = length(AR_2), trend =
  "nc", statistic = "t")
> p_ADF_2c = punitroot(ur.df(AR_2, type = "drift", lags = k_max,
  selectlags = "AIC")@teststat [1], N = length(AR_2), trend =
  "c", statistic = "t")
```

```

> p_ADF_2ct = punitroot(ur.df(AR_2, type = "trend", lags =
  k_max, selectlags = "AIC")@teststat [1], N = length(AR_2),
  trend = "ct", statistic = "t")
-----
> cat("AR_1", "\t", p_ADF_1n, "\t", p_ADF_1c, "\t", p_ADF_1ct,
  "\n", "AR_2", "\t", p_ADF_2n, "\t", p_ADF_2c, "\t", p_ADF_2ct,
  "\n")
AR_1      6.993958e-05      0.001038081      0.006448258
AR_2              1              1              1

```

Pri prvom časovom rade nám vo všetkých možnostiach (model bez konštanty a bez trendu, model s konštantou, model s konštantou aj s trendom) vyšla  $p$ -hodnota menšia ako hladina významnosti 0.01, a teda nulovú hypotézu o existencii jednotkového koreňa môžeme zamietnuť. Pri druhom časovom rade to vyzerá byť jednoznačné, nulovú hypotézu zamietnuť nevieme. V praxi to znamená, že predpokladáme (v skutočnosti viac urobiť nevieme), že časový rad je nestacionárny.

### Príklad 6.8

V tomto príklade si vyskúšame testovanie stacionarity na reálnych údajoch. V databáze `EuStockMarkets` (knižnica `datasets`) máme k dispozícii údaje zo 4 európskych akciových indexov. Ide o denné uzatváracie ceny indexov DAX (Nemecko), SMI (Švajčiarsko), CAC (Francúzsko) a FTSE (Anglicko) za obdobie od roku 1991 do roku 1998. Z daných uzatváracích cien vypočítajte tzv. spojité výnosy podľa vzťahu  $\ln(P_{t+1}/P_t)$ , kde  $P_t$  je uzatváracia cena indexu v čase  $t$ . Následne otestujte pomocou ADF testu, či je možné považovať uzatváracie ceny a/alebo spojité výnosy týchto indexov za stacionárne. Maximálny počet oneskorení v pomocných regresiach zvolte opäť pomocou Schwertovho pravidla a optimálny počet oneskorení podľa Akaikeho informačného kritéria.

---

### Príklad 6.8 – Riešenie

---

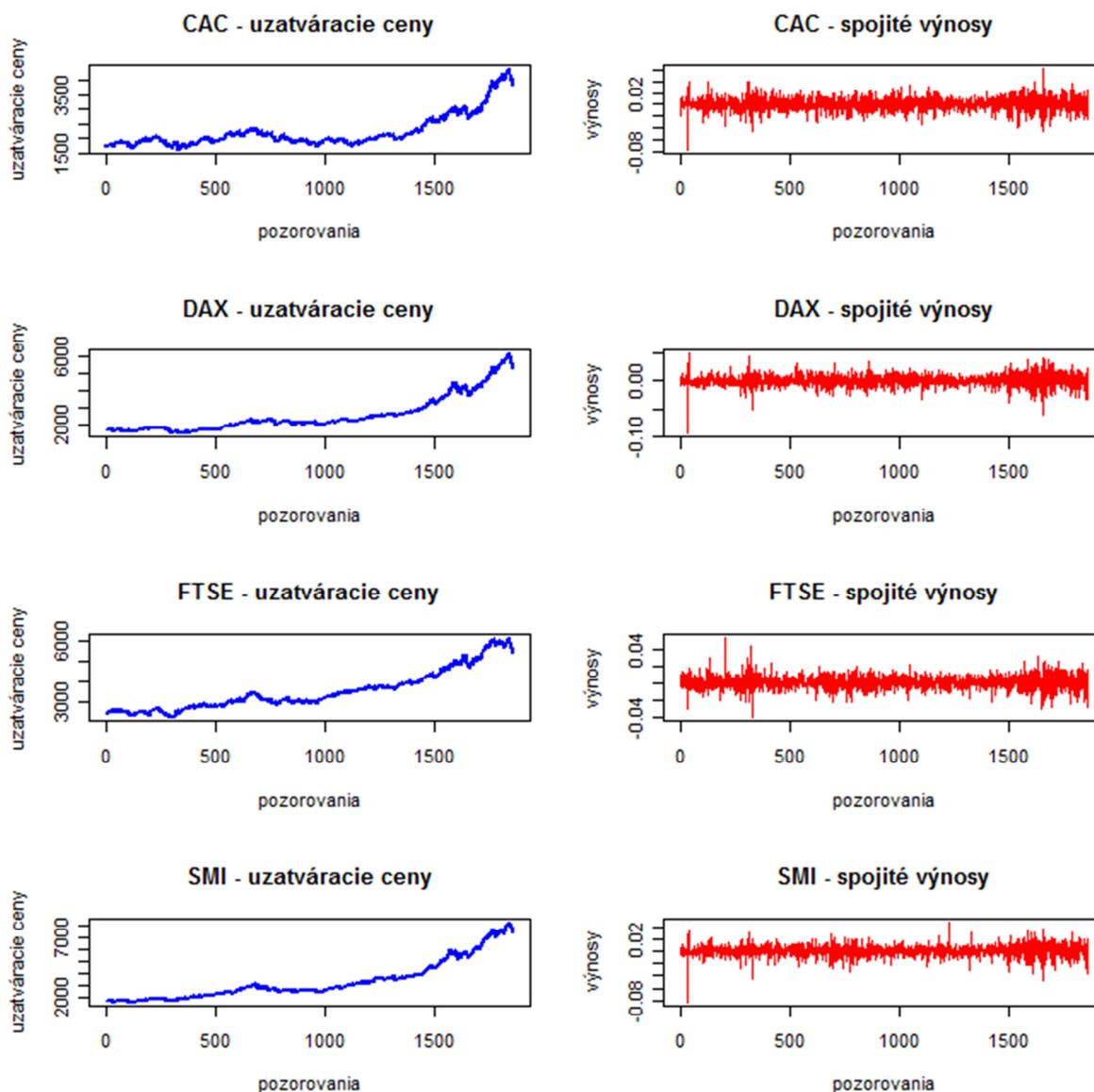
Pre lepšiu predstavu o skúmaných časových radoch (najmä kvôli lepšej predstave o rozdielnosti uzatváracích cien a spojitých výnosov) si ich zobrazíme graficky.

```

> library(datasets)
> attach(data.frame(EuStockMarkets))
> rCAC = diff(log(CAC))
> rDAX = diff(log(DAX))
> rFTSE = diff(log(FTSE))
> rSMI = diff(log(SMI))
-----
> par(mfrow = c(4, 2))
> plot(CAC, type = "l", ylab = "uzatváracie ceny", main = "CAC -
  uzatváracie ceny", xlab = "pozorovania", col = "blue", lwd=2)

```

```
> plot(rCAC, type = "l", ylab = "výnosy", main = "CAC - spojité  
výnosy", xlab = "pozorovania", col = "red", lwd=1)  
> plot(DAX, type = "l", ylab = "uzatváracie ceny", main = "DAX -  
uzatváracie ceny", xlab = "pozorovania", col = "blue", lwd=2)  
> plot(rDAX, type = "l", ylab = "výnosy", main = "DAX - spojité  
výnosy", xlab = "pozorovania", col = "red", lwd=1)  
> plot(FTSE, type = "l", ylab = "uzatváracie ceny", main = "FTSE  
- uzatváracie ceny", xlab = "pozorovania", col = "blue",  
lwd=2)  
> plot(rFTSE, type = "l", ylab = "výnosy", main = "FTSE -  
spojité výnosy", xlab = "pozorovania", col = "red", lwd=1)  
> plot(SMI, type = "l", ylab = "uzatváracie ceny", main = "SMI -  
uzatváracie ceny", xlab = "pozorovania", col = "blue", lwd=2)  
> plot(rSMI, type = "l", ylab = "výnosy", main = "SMI - spojité  
výnosy", xlab = "pozorovania", col = "red", lwd=1)
```



Obrázok 36: Uzatváracie ceny a spojité výnosy skúmaných indexov

*Zdroj: výstup zo softvéru R*

Z uvedených grafov je zjavné, že uzatváracie ceny stacionárne nebudú, zjavne je prítomný rastúci trend. Samozrejme, uzatváracie ceny by stále mohli byť stacionárne v okolí deterministického trendu. Na druhej strane, spojité výnosy by stacionárne byť mohli. Spojité výnosy oscilujú v okolí konštanty a tak minimálne stredná hodnota sa zdá byť konštantná. Samozrejme, že korektné otestovanie je nutné. Keďže pri uzatváracích cenách je trend zrejmý, pri ADF teste budeme uvažovať o modeli s konštantou aj s trendom. Pri spojitých výnosoch nemáme dôvod domnievať sa, že sú rastúce. Preto sme použili iba model s konštantou.

```
> library(urca)
```

```

> library(datasets)
> attach(data.frame(EuStockMarkets))
> rCAC = diff(log(CAC))
> rDAX = diff(log(DAX))
> rFTSE = diff(log(FTSE))
> rSMI = diff(log(SMI))
-----
> k_max = as.integer(12*(length(CAC)/100)^(1/4)); k_max
[1] 24
> p_ADF_CAC = punitroot(ur.df(CAC, type = "trend", lags = k_max,
  selectlags = "AIC")@teststat [1], N = length(CAC), trend =
  "ct", statistic = "t")
> p_ADF_DAX = punitroot(ur.df(DAX, type = "trend", lags = k_max,
  selectlags = "AIC")@teststat [1], N = length(DAX), trend =
  "ct", statistic = "t")
> p_ADF_FTSE = punitroot(ur.df(FTSE, type = "trend", lags =
  k_max, selectlags = "AIC")@teststat [1], N = length(FTSE),
  trend = "ct", statistic = "t")
> p_ADF_SMI = punitroot(ur.df(SMI, type = "trend", lags = k_max,
  selectlags = "AIC")@teststat [1], N = length(SMI), trend =
  "ct", statistic = "t")
-----
> k_max = as.integer(12*(length(rCAC)/100)^(1/4)); k_max
[1] 24
> p_ADF_rCAC = punitroot(ur.df(rCAC, type = "drift", lags =
  k_max, selectlags = "AIC")@teststat [1], N = length(rCAC),
  trend = "ct", statistic = "t")
> p_ADF_rDAX = punitroot(ur.df(rDAX, type = "drift", lags =
  k_max, selectlags = "AIC")@teststat [1], N = length(rDAX),
  trend = "ct", statistic = "t")
> p_ADF_rFTSE = punitroot(ur.df(rFTSE, type = "drift", lags =
  k_max, selectlags = "AIC")@teststat [1], N = length(rFTSE),
  trend = "ct", statistic = "t")
> p_ADF_rSMI = punitroot(ur.df(rSMI, type = "drift", lags =
  k_max, selectlags = "AIC")@teststat [1], N = length(rSMI),
  trend = "ct", statistic = "t")
-----
> results[1,] <- c(p_ADF_CAC, p_ADF_DAX, p_ADF_FTSE, p_ADF_SMI)
> results[2,] <- c(p_ADF_rCAC, p_ADF_rDAX, p_ADF_rFTSE,
  p_ADF_rSMI)
> results

```

	CAC	DAX	FTSE	SMI
ceny	9.924439e-01	9.962897e-01	5.22122e-01	9.937517e-01
vynosy	6.388754e-58	6.388754e-58	1.68316e-23	6.388754e-58

Z uvedených výsledkov vyplýva, že uzatváracie ceny skúmaných akciových indexov nie sú stacionárne, keďže ani v jednom prípade sme nulovú hypotézu o existencii jednotkového koreňa nevedeli zamietnuť. Opačná situácia je v prípade spojitých výnosov, kde je možné nulovú hypotézu zamietnuť pri všetkých indexoch. Tieto časové rady tak môžeme považovať za stacionárne.

## Príklad 6.9

V tomto príklade opäť použijete databázu `EuStockMarkets` (knížnica `datasets`). Teraz však otestujte uzatváracie ceny indexov a ich spojité výnosy s pomocou ďalších často používaných testov z knižnice `urca`, konkrétne **ERS test** (Elliott et al., 1996), **PP test** (Phillips – Perron, 1988) a **KPSS test** (Kwiatkowski et al., 1992). Po porovnaní výsledkov z týchto troch testov ešte na záver otestujte pomocou **ZA testu** (Zivot – Andrews, 1992) tie premenné, ktoré nebudete môcť považovať za stacionárne. Zhodnoťte, či ich môžete považovať za stacionárne so štrukturálnym zlomom.

---

### Príklad 6.9 – Riešenie

---

Začnime najprv funkciou `ur.ers()`, pomocou ktorej vieme vypočítať dva typy ERS testov: P-test a DF-GLS test. Kritické hodnoty pre DF-GLS test s konštantou a trendom sú: – 3.48 (1 %), –2.89 (5 %), –2.57 (10 %). Kritické hodnoty pre P-test s konštantou a trendom sú: 3.96 (1 %), 5.62 (5 %), 6.89 (10 %). K týmto kritickým hodnotám sa vieme ľahko dostať po spustení testu cez funkciu `summary()`, napríklad `summary(ers_DAX)`, alebo v tomto konkrétnom prípade pomocou príkazu `ers_DAX@cval`, ktorý vráti vektor kritických hodnôt.

```
> library(datasets)
> library(urca)
> attach(data.frame(EuStockMarkets))
-----
> k_max = as.integer(12*(length(CAC)/100)^(1/4)); k_max
[1] 24
-----
> ers_CAC = ur.ers(CAC, type = "DF-GLS", model = "trend",
  lag.max = k_max); ers_CAC

#####
# Elliot, Rothenberg and Stock Unit Root / Cointegration Test #
#####

The value of the test statistic is: -0.4678

> ers_DAX = ur.ers(DAX, type = "DF-GLS", model = "trend",
  lag.max = k_max); ers_DAX

#####
# Elliot, Rothenberg and Stock Unit Root / Cointegration Test #
#####

The value of the test statistic is: -0.0377
```

```

> ers_FTSE = ur.ers(FTSE, type = "DF-GLS", model = "trend",
  lag.max = k_max); ers_FTSE

#####
# Elliot, Rothenberg and Stock Unit Root / Cointegration Test #
#####

The value of the test statistic is: -1.3004

> ers_SMI = ur.ers(SMI, type = "DF-GLS", model = "trend",
  lag.max = k_max); ers_SMI

#####
# Elliot, Rothenberg and Stock Unit Root / Cointegration Test #
#####

The value of the test statistic is: -0.1747

```

Všetky testovacie štatistiky sú väčšie ako ľavostranné kritické hodnoty. Z tohto dôvodu sme ani v jednom z týchto prípadov nevedeli zamietnuť nulovú hypotézu o stacionarite premenných. Upozorňujeme, že pri funkcii `ur.ers()`, pomocné regresie sa odhadovali s počtom oneskorení závislej premennej na úrovni *kmax*. Takáto voľba väčšinou vedie k slabšej sile testov. Pri použití nasledujúceho P-testu zrejme nebudú výsledky odlišné.

```

> ers_CAC = ur.ers(CAC, type = "P-test", model = "trend",
  lag.max = k_max); ers_CAC

#####
# Elliot, Rothenberg and Stock Unit Root / Cointegration Test #
#####

The value of the test statistic is: 42.0187

> ers_DAX = ur.ers(DAX, type = "P-test", model = "trend",
  lag.max = k_max); ers_DAX

#####
# Elliot, Rothenberg and Stock Unit Root / Cointegration Test #
#####

The value of the test statistic is: 75.8729

> ers_FTSE = ur.ers(FTSE, type = "P-test", model = "trend",
  lag.max = k_max); ers_FTSE

#####
# Elliot, Rothenberg and Stock Unit Root / Cointegration Test #
#####

The value of the test statistic is: 23.1932

```

```

> ers_SMI = ur.ers(SMI, type = "P-test", model = "trend",
  lag.max = k_max); ers_SMI

#####
# Elliot, Rothenberg and Stock Unit Root / Cointegration Test #
#####

The value of the test statistic is: 72.0248

```

Pri P-teste sú zas všetky testovacie štatistiky mimo oboru kritických hodnôt (ide taktiež o ľavostranný test), takže opäť nevieme zamietnuť nulovú hypotézu o nestacionarite časových radov. Uzatváracie ceny tak stále považujeme za nestacionárne.

V predchádzajúcom príklade nám vyšlo, že po transformácií uzatváracích cien indexov na spojité výnosy už dostávame stacionárne časové rady. Vyskúšajme, či ERS test potvrdí dané zistenie. Kritické hodnoty pre DF-GLS test s konštantou sú: -2.57 (1 %), -1.94 (5 %), -1.62 (10 %).

```

> rCAC = diff(log(CAC))
> rDAX = diff(log(DAX))
> rFTSE = diff(log(FTSE))
> rSMI = diff(log(SMI))

-----
> k_max = as.integer(12*(length(rCAC)/100)^(1/4)); k_max
[1] 24
-----

> ers_rCAC = ur.ers(rCAC, type = "DF-GLS", model = "constant",
  lag.max = k_max); ers_rCAC

#####
# Elliot, Rothenberg and Stock Unit Root / Cointegration Test #
#####

The value of the test statistic is: -1.5089

> ers_rDAX = ur.ers(rDAX, type = "DF-GLS", model = "constant",
  lag.max = k_max); ers_rDAX

#####
# Elliot, Rothenberg and Stock Unit Root / Cointegration Test #
#####

The value of the test statistic is: -1.8191

> ers_rFTSE = ur.ers(rFTSE, type = "DF-GLS", model = "constant",
  lag.max = k_max); ers_rFTSE

#####
# Elliot, Rothenberg and Stock Unit Root / Cointegration Test #

```

```
#####
The value of the test statistic is: -1.7819

> ers_rSMI = ur.ers(rSMI, type = "DF-GLS", model = "constant",
  lag.max = k_max); ers_rSMI

#####
# Elliot, Rothenberg and Stock Unit Root / Cointegration Test #
#####

The value of the test statistic is: -2.7413
```

Pri DF-GLS teste s konštantou sme nulovú hypotézu nevedeli zamietnuť v prípade spojitých výnosov indexu CAC, pri indexoch DAX a FTSE môžeme nulovú hypotézu zamietnuť na hladine významnosti 10 % a pri indexe SMI môžeme spojité výnosy považovať za stacionárne na hladine významnosti 1 %.

Pri P-teste s konštantou sme vo všetkých prípadoch mohli zamietnuť nulovú hypotézu, a teda všetky časové rady spojitých výnosov môžeme považovať za stacionárne na hladine významnosti 1 %. Kritické hodnoty pre P-test s konštantou sú: 1.99 (1 %), 3.26 (5 %), 4.48 (10 %).

```
> ers_rCAC = ur.ers(rCAC, type = "P-test", model = "constant",
  lag.max = k_max); ers_rCAC

#####
# Elliot, Rothenberg and Stock Unit Root / Cointegration Test #
#####

The value of the test statistic is: 0.0796

> ers_rDAX = ur.ers(rDAX, type = "P-test", model = "constant",
  lag.max = k_max); ers_rDAX

#####
# Elliot, Rothenberg and Stock Unit Root / Cointegration Test #
#####

The value of the test statistic is: 0.0798

> ers_rFTSE = ur.ers(rFTSE, type = "P-test", model = "constant",
  lag.max = k_max); ers_rFTSE

#####
# Elliot, Rothenberg and Stock Unit Root / Cointegration Test #
#####

The value of the test statistic is: 0.0423
```

```

> ers_rSMI = ur.ers(rSMI, type = "P-test", model = "constant",
  lag.max = k_max); ers_rSMI

#####
# Elliot, Rothenberg and Stock Unit Root / Cointegration Test #
#####

The value of the test statistic is: 0.0378

```

V tomto príklade sme mohli vidieť, že použitím rôznych testov, výsledky nemusia byť celkovo jednoznačné. Takáto situácia však v praxi samozrejme môže vzniknúť. Preto je vhodné kombinovať vždy viacero testov a pridať prípadne ešte KPSS test, ktorý má nestacionaritu v alternatívnej hypotéze.

Pozrime sa teda na výsledky z ďalšieho (veľmi často používaného) testu, konkrétne Phillips – Perron testu (PP test). Pre zjednodušenie budeme uvádzať len výsledky pre "Z-tau" štatistiky. Počet oneskorení stanovíme podľa klasického Schwertovho kritéria (pomocou argumentu funkcie lags = "long").

```

> library(datasets)
> library(urca)
> attach(data.frame(EuStockMarkets))
-----
> pp_CAC = ur.pp(CAC, type = "Z-tau", model = "trend", lags =
  "long", use.lag = NULL); pp_CAC

#####
# Phillips-Perron Unit Root / Cointegration Test #
#####

The value of the test statistic is: -0.2392
> pp_DAX = ur.pp(DAX, type = "Z-tau", model = "trend", lags =
  "long", use.lag = NULL); pp_DAX

#####
# Phillips-Perron Unit Root / Cointegration Test #
#####

The value of the test statistic is: -0.4263
> pp_FTSE = ur.pp(FTSE, type = "Z-tau", model = "trend", lags =
  "long", use.lag = NULL); pp_FTSE

#####
# Phillips-Perron Unit Root / Cointegration Test #
#####

The value of the test statistic is: -1.8744

```

```

> pp_SMI = ur.pp(SMI, type = "Z-tau", model = "trend", lags =
  "long", use.lag = NULL); pp_SMI

#####
# Phillips-Perron Unit Root / Cointegration Test #
#####

The value of the test statistic is: -0.2679

```

Kritické hodnoty získame pomocou príkazu `pp_CAC@cval` (keďže časové rady sú rovnako dlhé, kritické hodnoty sú rovnaké pre všetky premenné). Dostaneme vektor kritických hodnôt pre významnosti 1 %, 5 % a 10 %.

```

> pp_CAC@cval
                1pct      5pct      10pct
critical values -3.968307 -3.414778 -3.129203

```

Na základe týchto kritických hodnôt môžeme vidieť, že ani jeden z akciových indexov nemôžeme považovať za stacionárny, keďže ani v jednom prípade nevieme zamietnuť nulovú hypotézu. Pozrime sa ešte na výsledky pre spojité výnosy.

```

> pp_rCAC = ur.pp(rCAC, type = "Z-tau", model = "constant", lags
  = "long", use.lag = NULL); pp_rCAC

#####
# Phillips-Perron Unit Root / Cointegration Test #
#####

The value of the test statistic is: -41.8419

> pp_rDAX = ur.pp(rDAX, type = "Z-tau", model = "constant", lags
  = "long", use.lag = NULL); pp_rDAX

#####
# Phillips-Perron Unit Root / Cointegration Test #
#####

The value of the test statistic is: -43.0781

> pp_rFTSE = ur.pp(rFTSE, type = "Z-tau", model = "constant",
  lags = "long", use.lag = NULL); pp_rFTSE

#####
# Phillips-Perron Unit Root / Cointegration Test #
#####

The value of the test statistic is: -39.1084

> pp_rSMI = ur.pp(rSMI, type = "Z-tau", model = "constant", lags
  = "long", use.lag = NULL); pp_rSMI

```

```
#####
# Phillips-Perron Unit Root / Cointegration Test #
#####

The value of the test statistic is: -40.9957
```

Kritické hodnoty pre spojité výnosy sú len mierne odlišné (ide o jedno pozorovanie kratšie časové rady):

```
> pp_rCAC@cval
                1pct      5pct      10pct
critical values -3.436737 -2.863576 -2.567875
```

Pri spojitých výnosoch sa teda opäť potvrdilo, že tieto časové rady už môžeme považovať za stacionárne. Posledný test, ktorý by sme v tomto príklade mali ešte vyskúšať z knižnice `urca` je KPSS test. Tento test je odlišný od predchádzajúcich dvoch v tom, že v nulovej hypotéze sú časové rady považované za stacionárne.

KPSS test je v knižnici `urca` dostupný prostredníctvom funkcie `ur.kpss()`. Aj v tomto prípade budeme pre ceny akcií uvažovať o modeloch s konštantou a s trendom a pri spojitých výnosoch o modeloch s konštantou.

```
> library(datasets)
> library(urca)
> attach(data.frame(EuStockMarkets))
-----
> pp_CAC = ur.kpss(CAC, type = "tau", lags = "long", use.lag =
  NULL); pp_CAC

#####
# KPSS Unit Root / Cointegration Test #
#####

The value of the test statistic is: 1.2285

> pp_DAX = ur.kpss(DAX, type = "tau", lags = "long", use.lag =
  NULL); pp_DAX

#####
# KPSS Unit Root / Cointegration Test #
#####

The value of the test statistic is: 1.3619

> pp_FTSE = ur.kpss(FTSE, type = "tau", lags = "long", use.lag =
  NULL); pp_FTSE

#####
```

```

# KPSS Unit Root / Cointegration Test #
#####

The value of the test statistic is: 1.3513

> pp_SMI = ur.kpss(SMI, type = "tau", lags = "long", use.lag =
  NULL); pp_SMI

#####
# KPSS Unit Root / Cointegration Test #
#####

The value of the test statistic is: 1.4049

```

Kritické hodnoty pre KPSS test vieme získať rovnako ako pri PP teste:

```

> pp_CAC@cval
              10pct  5pct  2.5pct  1pct
critical values 0.119 0.146  0.176 0.216

```

Pri všetkých indexoch môžeme zamietnuť nulovú hypotézu o stacionarite ich uzatváracích cien a prijať alternatívnu, teda že dané časové rady nie sú stacionárne. Keďže už predpokladáme na základe predchádzajúcich testov, že spojité výnosy stacionárne sú, pri tomto teste by sme nulovú hypotézu zamietnuť nemali. Až na prípad spojitých výnosov pri akciovom indexe DAX, sme pre všetky testy nevedeli zamietnuť nulovú hypotézu o stacionarite. Aj v prípade akciového indexu DAX sme zamietli nulovú hypotézu iba na hladine významnosti  $\alpha = 0.10$ .

```

> rCAC = diff(log(CAC))
> rDAX = diff(log(DAX))
> rFTSE = diff(log(FTSE))
> rSMI = diff(log(SMI))
-----
> pp_rCAC@cval
              10pct  5pct  2.5pct  1pct
critical values 0.347 0.463  0.574 0.739
-----
> pp_rCAC = ur.kpss(rCAC, type = "mu", lags = "long", use.lag =
  NULL); pp_rCAC

#####
# KPSS Unit Root / Cointegration Test #
#####

The value of the test statistic is: 0.3231

> pp_rDAX = ur.kpss(rDAX, type = "mu", lags = "long", use.lag =
  NULL); pp_rDAX

```

```
#####
# KPSS Unit Root / Cointegration Test #
#####

The value of the test statistic is: 0.4151

> pp_rFTSE = ur.kpss(rFTSE, type = "mu", lags = "long", use.lag
  = NULL); pp_rFTSE

#####
# KPSS Unit Root / Cointegration Test #
#####

The value of the test statistic is: 0.076

> pp_rSMI = ur.kpss(rSMI, type = "mu", lags = "long", use.lag =
  NULL); pp_rSMI

#####
# KPSS Unit Root / Cointegration Test #
#####

The value of the test statistic is: 0.271
```

Na záver si vyskúšajme ešte jeden špecifický test z knižnice `urca`, ktorý umožňuje pri testovaní jednotkového koreňa (nulová hypotéza) prítomnosť endogénne určeného štrukturálneho zlomu v konštante, v trende alebo konštante aj v trende súčasne. Ide o Zivot – Andrews (1992) test (ZA test), ktorý je v tejto knižnici dostupný prostredníctvom funkcie `ur.za()`. Podľa zadanie máme otestovať už len uzatváracie ceny.

Pre zjednodušenie povolíme zlom v konštante aj v trende (`model = "both"`). Ak by sme chceli zahrnúť zlom len v konštante, argument nastavíme na `model = "intercept"` a ak len v lineárnom trende tak `model = "trend"`).

```
> library(datasets)
> library(urca)
> attach(data.frame(EuStockMarkets))
-----
> k_max = as.integer(12*(length(CAC)/100)^(1/4)); k_max
[1] 24
-----
> za_CAC = ur.za(CAC, model = "both", lag = k_max); za_CAC

#####
# Zivot-Andrews Unit Root / Cointegration Test #
#####

The value of the test statistic is: -3.6032
```

```

> za_DAX = ur.za(DAX, model = "both", lag = k_max); za_DAX

#####
# Zivot-Andrews Unit Root / Cointegration Test #
#####

The value of the test statistic is: -3.9036

> za_FTSE = ur.za(FTSE, model = "both", lag = k_max); za_FTSE

#####
# Zivot-Andrews Unit Root / Cointegration Test #
#####

The value of the test statistic is: -3.7842

> za_SMI = ur.za(SMI, model = "both", lag = k_max); za_SMI

#####
# Zivot-Andrews Unit Root / Cointegration Test #
#####

The value of the test statistic is: -3.9713

```

Kritické hodnoty sú nasledujúce (prvý člen vektora je kritická hodnota pri významnosti 1 %, potom nasleduje významnosť 5 % a posledný člen je kritická hodnota pri významnosti 10 %):

```

> za_CAC@cval
[1] -5.57 -5.08 -4.82

```

Z uvedených výsledkov teda vyplýva, že uzatváracie ceny indexov nemôžeme považovať za stacionárne ani po zahrnutí štrukturálneho zlomu. Pri použití funkcie `summary()` môžeme vidieť kompletne výsledky spolu s identifikovaným zlomom. Keďže však máme 4 premenné, tak pre lepšiu prehľadnosť je zrejme jednoduchšie využiť nasledujúce príkazy:

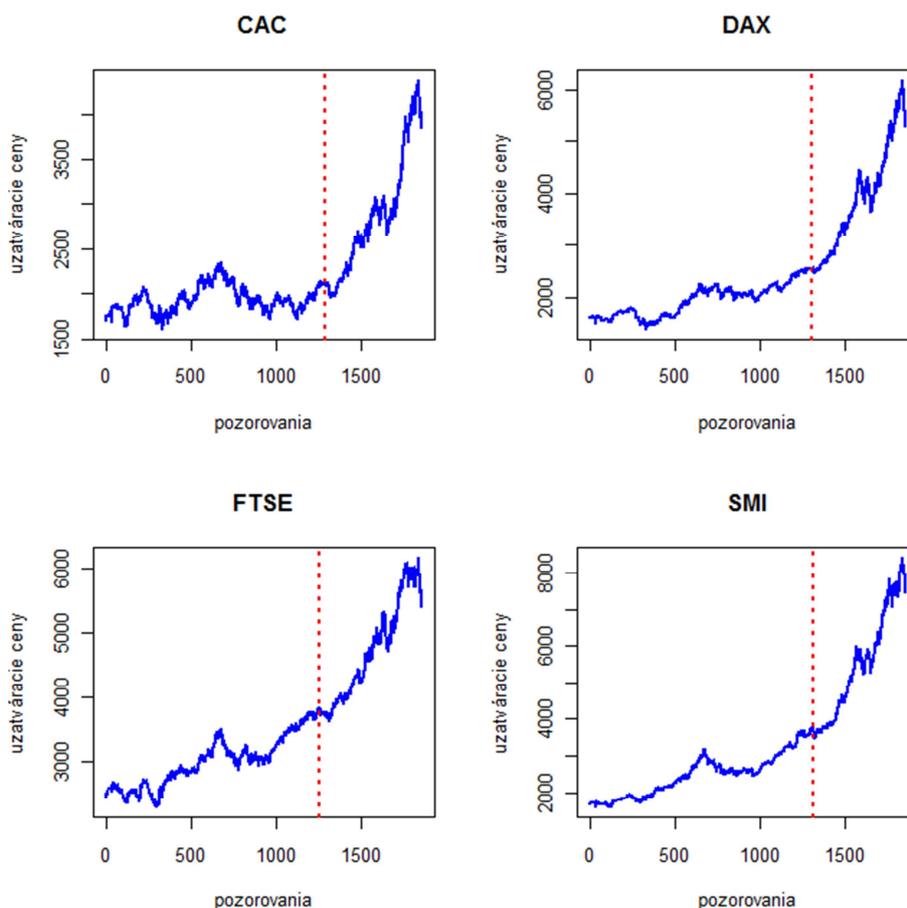
```

> za_CAC@bpoint
[1] 1292
> za_DAX@bpoint
[1] 1309
> za_FTSE@bpoint
[1] 1255
> za_SMI@bpoint
[1] 1314

```

Veľmi jednoducho je potom možné tieto identifikované zlomy naniest' aj do grafu uzatváracích cien:

```
> par(mfrow = c(2,2))
-----
> plot(CAC, type = "l", ylab = "uzatváracie ceny", main = "CAC",
      xlab = "pozorovania", col = "blue", lwd=2)
> abline(v = za_CAC@bpoint, lty=3, lwd=2, col = "red")
-----
> plot(DAX, type = "l", ylab = "uzatváracie ceny", main = "DAX",
      xlab = "pozorovania", col = "blue", lwd=2)
> abline(v = za_DAX@bpoint, lty=3, lwd=2, col = "red")
-----
> plot(FTSE, type = "l", ylab = "uzatváracie ceny", main =
      "FTSE", xlab = "pozorovania", col = "blue", lwd=2)
> abline(v = za_FTSE@bpoint, lty=3, lwd=2, col = "red")
-----
> plot(SMI, type = "l", ylab = "uzatváracie ceny", main = "SMI",
      xlab = "pozorovania", col = "blue", lwd=2)
> abline(v = za_SMI@bpoint, lty=3, lwd=2, col = "red")
```



Obrázok 37: Štrukturálne zlomy identifikované pomocou ZA testu

*Zdroj: výstup zo softvéru R*

Je na škodu, že použité údaje neobsahujú presné dátumy. Inak by sme vedeli povedať úplne presne, kedy došlo k daným štrukturálnym zlomom. Vidíme však, že ich výskyt je do značnej miery podobný pri všetkých skúmaných indexoch. Približne ide o rok 1997, kedy akciové trhy začali prudšie rásť. Uvedené obdobie sa spája s tzv. dot.com bublinou, kedy sa akcie technologických firiem do značnej miery nadhodnocovali. Išlo prevažne o podniky z USA avšak akciové trhy vyspelých krajín EÚ boli už v tom čase z veľkej časti integrované s ostatnými vyspelými trhami. Ak by sme mali k dispozícii dlhšie časové obdobie, mohli by sme pozorovať pokles všetkých týchto vyspelých trhov od roku 2000, kedy špekulatívna bublina praskla.

### Príklad 6.10

V tomto príklade budeme pracovať s údajmi z prebiehajúcej dlhovej krízy v EÚ. Z voľne dostupných zdrojov je možné stiahnuť údaje za ukazovateľ D/HDP, teda pomer štátneho dlhu a hrubého domáceho produktu krajiny (v angl. *debt-to-GDP ratio*). K dispozícii máme údaje z databázy Eurostat, za obdobie od 4. kvartálu 2000 do 3. kvartálu 2010, čo je 40

pozorovaní pre každú krajinu. Pre konkrétne údaje pozri Príklad 7.6 v Lyócsa et al. (2013). Zvolili sme krajiny, o ktorých sa predpokladá výskyt problémov s výškou štátneho dlhu. Tieto krajiny sú známe pod mierne degradujúcim akronymom „PIIGGS“ – Portugal (Portugalsko), Ireland (Írsko), Italy (Taliansko), Greece (Grécko), Great Britain (Veľká Británia), Spain (Španielsko).

V predchádzajúcom príklade sme si ukázali, že je možné využiť ZA test na identifikáciu štrukturálnych zlomov. Otestujte, či je možné považovať tieto premenné za stacionárne so zlomom (aj keď máme menší počet pozorovaní) a zistite, kde sa nachádzajú štrukturálne zlomy.

---

### Príklad 6.10 – Riešenie

---

Predtým ako pristúpime k testovaniu stacionarity časových radov, je vždy vhodné pozrieť sa na ich grafickú vizualizáciu. Pri týchto údajoch je zaujímavé sledovať zmenu v trende. Ak by sme vypočítali jednoduchú lineárnu regresiu, mohli by sme vidieť, že koeficient pri časovom trende sa do značnej miery zmenil od roku 2008. Inými slovami, rast verejného dlhu v pomere k HDP sa v danom období (počas dlhovej krízy v EÚ) výrazne zvyšoval. Lineárnu regresiu v tvare  $D/HDP_t = \beta_0 + \beta_1 t + \varepsilon_{i,t}$  vypočítame pomocou funkcie `lm()` pre každú krajinu vo vzorke a trend nanesieme do bodového grafu (presnejšie tzv. vyrovnané hodnoty, v angl. *fitted values*).

```

> data = read.csv(file = "...cesta k súboru...\\debt_gdp.csv",
  sep = ";", dec = ".", header = T)
> attach(data)
> library(zoo)
> par(mfrow = c(3,2))
-----
> plot(x = time[1:40], y = Portugal[1:40], type = "p", main =
  "Portugal", col = "black", pch = 19, xaxt = "n", ylab =
  "D/HDP", xlab = "čas", xlim = c(1, max(time)), ylim =
  c(min(Portugal), max(Portugal)))
> axis(side = 1, at = time, labels = obs)
> trend_1 = lm(Portugal[1:29] ~ time[1:29])$fitted
> trend_2 = lm(Portugal[30:40] ~ time[30:40])$fitted
> line_trend_1 = zoo(trend_1,time[1:29])
> line_trend_2 = zoo(trend_2,time[30:40])
> lines(line_trend_1,type="l",col=3,lty=2,lwd=2)
> lines(line_trend_2,type="l",col=2,lty=2,lwd=2)
-----
> plot(x = time[1:40], y = Ireland[1:40], type = "p", main =
  "Ireland", col = "black", pch = 19, xaxt = "n", ylab =
  "D/HDP", xlab = "čas", xlim = c(1, max(time)), ylim =
  c(min(Ireland), max(Ireland)))

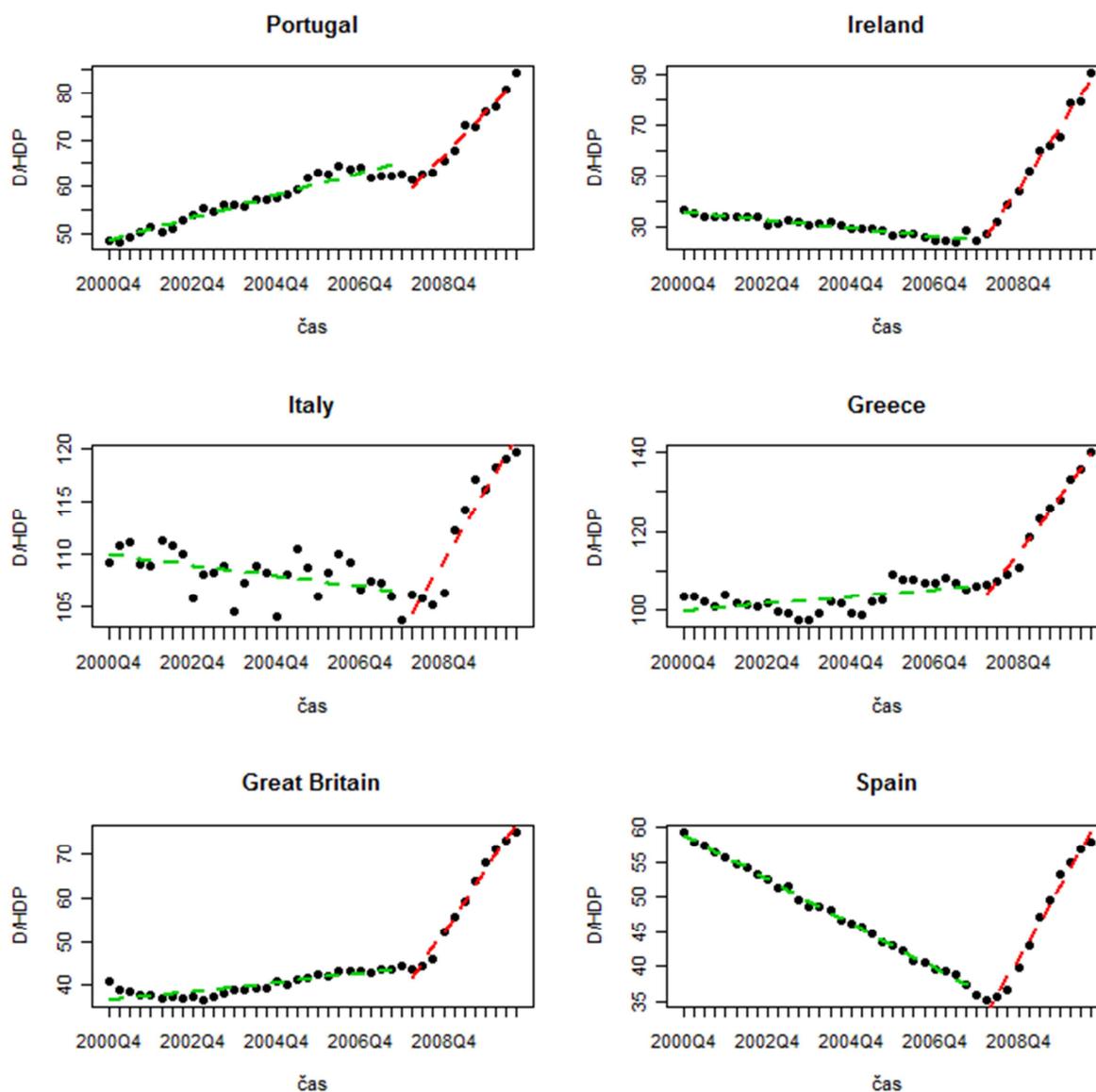
```

```

> axis(side = 1, at = time, labels = obs)
> trend_1 = lm(Ireland[1:29] ~ time[1:29])$fitted
> trend_2 = lm(Ireland[30:40] ~ time[30:40])$fitted
> line_trend_1 = zoo(trend_1,time[1:29])
> line_trend_2 = zoo(trend_2,time[30:40])
> lines(line_trend_1,type="l",col=3,lty=2,lwd=2)
> lines(line_trend_2,type="l",col=2,lty=2,lwd=2)
-----
> plot(x = time[1:40], y = Italy[1:40], type = "p", main =
  "Italy", col = "black", pch = 19, xaxt = "n", ylab = "D/HDP",
  xlab = "čas", xlim = c(1, max(time)), ylim = c(min(Italy),
  max(Italy)))
> axis(side = 1, at = time, labels = obs)
> trend_1 = lm(Italy[1:29] ~ time[1:29])$fitted
> trend_2 = lm(Italy[30:40] ~ time[30:40])$fitted
> line_trend_1 = zoo(trend_1,time[1:29])
> line_trend_2 = zoo(trend_2,time[30:40])
> lines(line_trend_1,type="l",col=3,lty=2,lwd=2)
> lines(line_trend_2,type="l",col=2,lty=2,lwd=2)
-----
> plot(x = time[1:40], y = Greece[1:40], type = "p", main =
  "Greece", col = "black", pch = 19, xaxt = "n", ylab = "D/HDP",
  xlab = "čas", xlim = c(1, max(time)), ylim = c(min(Greece),
  max(Greece)))
> axis(side = 1, at = time, labels = obs)
> trend_1 = lm(Greece[1:29] ~ time[1:29])$fitted
> trend_2 = lm(Greece[30:40] ~ time[30:40])$fitted
> line_trend_1 = zoo(trend_1,time[1:29])
> line_trend_2 = zoo(trend_2,time[30:40])
> lines(line_trend_1,type="l",col=3,lty=2,lwd=2)
> lines(line_trend_2,type="l",col=2,lty=2,lwd=2)
-----
> plot(x = time[1:40], y = Great.Britain[1:40], type = "p", main =
  "Great Britain", col = "black", pch = 19, xaxt = "n", ylab =
  "D/HDP", xlab = "čas", xlim = c(1, max(time)), ylim =
  c(min(Great.Britain), max(Great.Britain)))
> axis(side = 1, at = time, labels = obs)
> trend_1 = lm(Great.Britain[1:29] ~ time[1:29])$fitted
> trend_2 = lm(Great.Britain[30:40] ~ time[30:40])$fitted
> line_trend_1 = zoo(trend_1,time[1:29])
> line_trend_2 = zoo(trend_2,time[30:40])
> lines(line_trend_1,type="l",col=3,lty=2,lwd=2)
> lines(line_trend_2,type="l",col=2,lty=2,lwd=2)
-----
> plot(x = time[1:40], y = Spain[1:40], type = "p", main =
  "Spain", col = "black", pch = 19, xaxt = "n", ylab = "D/HDP",
  xlab = "čas", xlim = c(1, max(time)), ylim = c(min(Spain),
  max(Spain)))
> axis(side = 1, at = time, labels = obs)
> trend_1 = lm(Spain[1:29] ~ time[1:29])$fitted
> trend_2 = lm(Spain[30:40] ~ time[30:40])$fitted
> line_trend_1 = zoo(trend_1,time[1:29])
> line_trend_2 = zoo(trend_2,time[30:40])
> lines(line_trend_1,type="l",col=3,lty=2,lwd=2)

```

```
> lines(line_trend_2,type="l",col=2,lty=2,lwd=2)
```



Obrázok 38: Vývoj ukazovateľov D/HDP v krajinách PIIGGS

Zdroj: výstup zo softvéru R

Prítomnosť štrukturálneho zlomu je z uvedených obrázkov zrejmalá, konkrétne ide o zmenu v trende. Z toho dôvodu nastavíme potrebný argument vo funkcii `ur.za()` na `model = "trend"`.

```
> data = read.csv(file = "...cesta k súboru...\\debt_gdp.csv",
  sep = ";", dec = ".", header = T)
> attach(data)
> library(urca)
-----
> k_max = as.integer(12*(length(time)/100)^(1/4)); k_max
[1] 9
```

```

-----
> za_Portugal = ur.za(Portugal, model = "trend", lag = k_max);
  za_Portugal

#####
# Zivot-Andrews Unit Root / Cointegration Test #
#####

The value of the test statistic is: -2.0554

> za_Ireland = ur.za(Ireland, model = "trend", lag = k_max);
  za_Ireland

#####
# Zivot-Andrews Unit Root / Cointegration Test #
#####

The value of the test statistic is: -2.3039

> za_Italy = ur.za(Italy, model = "trend", lag = k_max);
  za_Italy

#####
# Zivot-Andrews Unit Root / Cointegration Test #
#####

The value of the test statistic is: -4.8542

> za_Greece = ur.za(Greece, model = "trend", lag = k_max);
  za_Greece

#####
# Zivot-Andrews Unit Root / Cointegration Test #
#####

The value of the test statistic is: -3.2637

> za_Great.Britain = ur.za(Great.Britain, model = "trend", lag =
  k_max); za_Great.Britain

#####
# Zivot-Andrews Unit Root / Cointegration Test #
#####

The value of the test statistic is: -6.9259

> za_Spain = ur.za(Spain , model = "trend", lag = k_max);
  za_Spain

#####
# Zivot-Andrews Unit Root / Cointegration Test #
#####

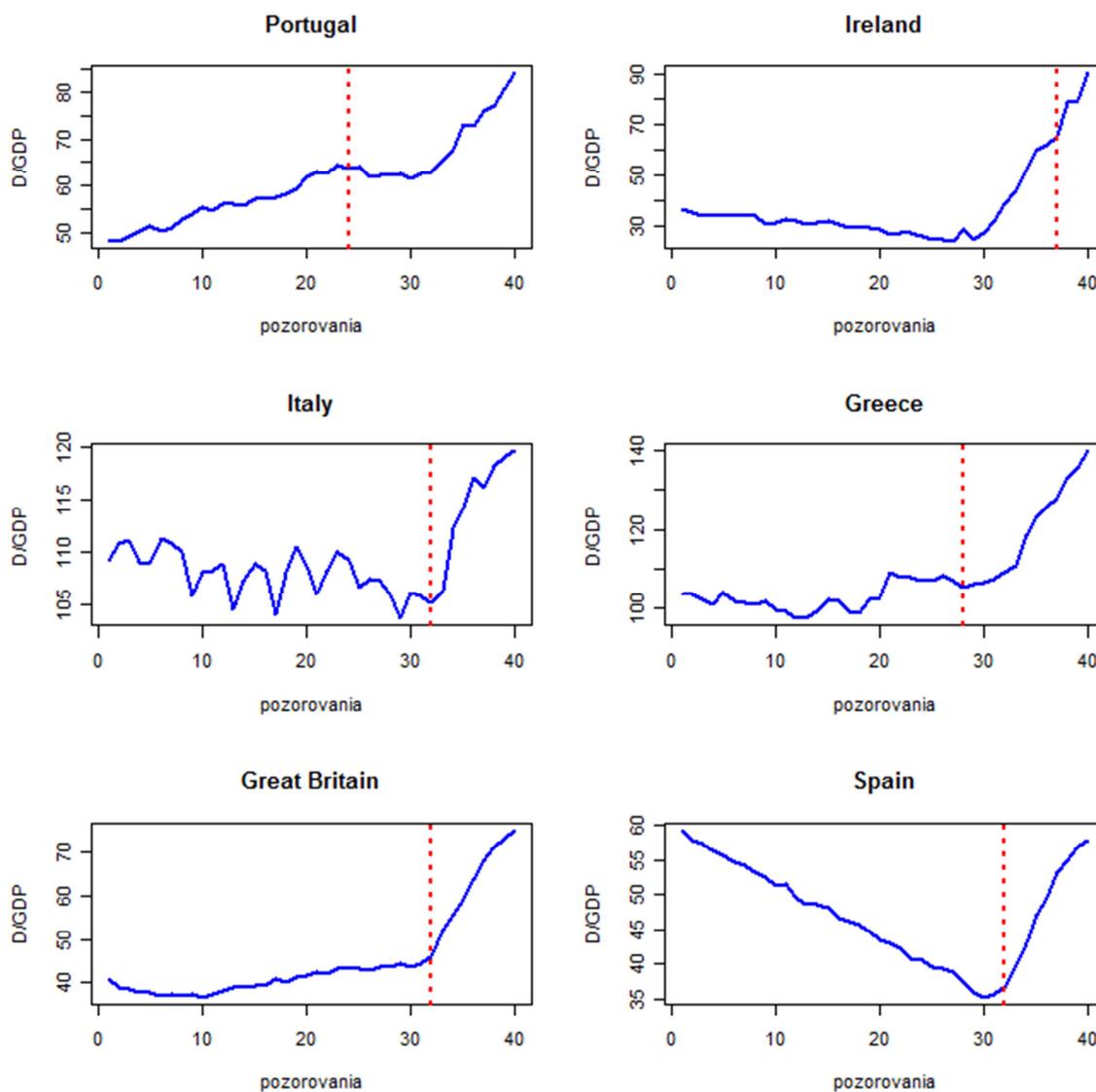
The value of the test statistic is: -3.2929

```

```
-----  
> za_Portugal@cval  
[1] -4.93 -4.42 -4.11
```

Bez transformácie časových radov (na logaritmickej diferencii) môžeme vidieť, že v prípade Veľkej Británie a Talianska môžeme po zahrnutí štrukturálneho zlomu hovoriť o stacionárnych ukazovateľoch D/HDP. Aj vzhľadom na nízky počet pozorovaní je tento výsledok mierne prekvapujúci. Pri takýchto údajoch by bolo samozrejme vhodnejšie použiť panelový test. Nás v tomto bode bude zaujímať, či je možné použiť ZA test na identifikáciu štrukturálnych zlomov, resp. ako presne ZA test daný zlom určí. Zámerne sme totiž vybrali údaje, kde je zlom v trende úplne zrejмый.

```
> par(mfrow = c(3, 2))  
> plot(Portugal, type = "l", ylab = "D/GDP", main = "Portugal",  
      xlab = "pozorovania", col = "blue", lwd=2)  
> abline(v = za_Portugal@bpoint, lty=3, lwd=2, col = "red")  
> plot(Ireland, type = "l", ylab = "D/GDP", main = "Ireland",  
      xlab = "pozorovania", col = "blue", lwd=2)  
> abline(v = za_Ireland@bpoint, lty=3, lwd=2, col = "red")  
> plot(Italy, type = "l", ylab = "D/GDP", main = "Italy", xlab =  
      "pozorovania", col = "blue", lwd=2)  
> abline(v = za_Italy@bpoint, lty=3, lwd=2, col = "red")  
> plot(Greece, type = "l", ylab = "D/GDP", main = "Greece", xlab =  
      "pozorovania", col = "blue", lwd=2)  
> abline(v = za_Greece@bpoint, lty=3, lwd=2, col = "red")  
> plot(Great.Britain, type = "l", ylab = "D/GDP", main = "Great  
      Britain", xlab = "pozorovania", col = "blue", lwd=2)  
> abline(v = za_Great.Britain@bpoint, lty=3, lwd=2, col = "red")  
> plot(Spain, type = "l", ylab = "D/GDP", main = "Spain", xlab =  
      "pozorovania", col = "blue", lwd=2)  
> abline(v = za_Spain@bpoint, lty=3, lwd=2, col = "red")
```



Obrázok 39: Štrukturálne zlomy identifikované pomocou ZA testu

Zdroj: výstup zo softvéru R

Pozn.: pri Portugalsku sme do grafu naniesli zlom, ktorý bol identifikovaný v strednej hodnote (príslušný argument funkcie `ur.za()` bol zvolený na `model = "intercept"`).

Pri Taliansku, Grécku, Veľkej Británii a Španielsku nám ZA test určil pozíciu štrukturálneho zlomu v trende celkom presne. V prípade Írska je odhadnutý zlom až ku koncu sledovaného obdobia a rovnaká situácia by bola aj v prípade Portugalska (keby sme sa do grafu nerozhodli naniesť zlom v strednej hodnote, keďže ten sa javil ako presnejší). ZA test teda je možné použiť na endogénnu identifikáciu zlomov, ale ako môžeme vidieť, nie vždy sú výsledky úplne presné.

### Príklad 6.11

V databáze *Commodities* z knižnice *FitAR* sú dostupné uzatváracie (angl. *close*), najvyššie (angl. *high*) a najnižšie (angl. *low*) denné ceny rôznych komodít. Otestujte, či uzatváracie ceny zlata sú stacionárne, prípadne, či sú stacionárne po transformácii na logaritmické diferencie.

### Príklad 6.12

V tomto príklade využijeme údaje zo 4 európskych akciových indexov, ktoré sú dostupné v databáze *EuStockMarkets* (knižnica *datasets*). Ide o denné uzatváracie ceny indexov DAX (Nemecko), SMI (Švajčiarsko), CAC (Francúzsko) a FTSE (Anglicko) za obdobie od roku 1991 do roku 1998. Najprv z údajov vypočítajte tzv. spojité výnosy podľa vzťahu  $\ln(P_{t+1}/P_t)$ , kde  $P_t$  je uzatváracia cena indexu v čase  $t$ . Následne zistite, či priemerný denný výnos týchto indexov je menší ako 0.1 %.

### Príklad 6.13

Vygenerujte dva časové rady (s počtom pozorovaní 500), konkrétne náhodné prechádzky s posunom (angl. *random walk with drift*)  $x_t = \mu + x_{t-1} + w_t$ . Posun zvolte na úrovni 0.1. Na základe jednoduchšej simulácie s 500 iteráciami regresných modelov vytvorených z týchto dvoch premenných sa pokúste zistiť, v koľkých prípadoch získate štatisticky významný regresný koeficient  $\beta_1$  (z regresie  $y_t = \beta_0 + \beta_1 x_t + \varepsilon_t$ ) a overte praktické pravidlo identifikácie falošnej regresie na základe koeficientu determinácie a Durbin – Watson štatistiky.

---

### Príklad 6.13 – Riešenie

---

Okrem toho, že máme zistiť v koľkých prípadoch dostaneme významný regresný koeficient z regresie náhodne vygenerovaných časových radov, pozrieme sa aj na vzťah medzi koeficientom determinácie a Durbin – Watson štatistikou.

```
> rm(list = ls())
> library(lmtest)
-----
> r2s = c()
> dwt_stats = c()
> dwt_pvals = c()
> tstats = c()
> pvs = c()
-----
> for (c in 1:500) {
+   x = 0.1*(1:500)+cumsum(rnorm(500,0,1))
+   y = 0.1*(1:500)+cumsum(rnorm(500,0,1))
```

```

+   model = lm(y~x)
+   dwt = dwtest(model)
+   sumr = summary(model)
+   n = length(x)
+   es = resid(model)
+   s = sqrt(sum(es^2)/(n-2))
+   b1 = model$coefficients[2]
+   SE = s/sqrt(sum((x-mean(x))^2))
+   t = ((b1-(0))/SE)
+   pv = 2*pt(abs(t), n-2, lower.tail = FALSE)
+   r2s = c(r2s, sumr$r.squared)
+   dwt_stats = c(dwt_stats, dwt$statistic)
+   dwt_pvals = c(dwt_pvals, dwt$p.value)
+   tstats = c(tstats, t)
+   pvs = c(pvs, pv)
+ }
> df = data.frame(r2s, dwt_stats, dwt_pvals, tstats, pvs)
> summary(df)
      r2s                dwt_stats                dwt_pvals
Min.   : 0.001965      Min.   : 0.002058      Min.   : 1.067e-111
1st Qu.: 0.480917      1st Qu.: 0.021409      1st Qu.: 1.348e-109
Median : 0.725055      Median : 0.031680      Median : 1.735e-108
Mean   : 0.630203      Mean   : 0.037378      Mean   : 9.950e-98
3rd Qu.: 0.832126      3rd Qu.: 0.048547      3rd Qu.: 1.128e-106
Max.   : 0.967595      Max.   : 0.160588      Max.   : 4.719e-95

      tstats                pvs
Min.   : -27.75      Min.   : 0.000e+00
1st Qu.: 20.82       1st Qu.: 3.537e-195
Median : 36.29       Median : 6.087e-142
Mean   : 37.01       Mean   : 1.001e-02
3rd Qu.: 49.74       3rd Qu.: 5.315e-72
Max.   : 122.07      Max.   : 8.842e-01

```

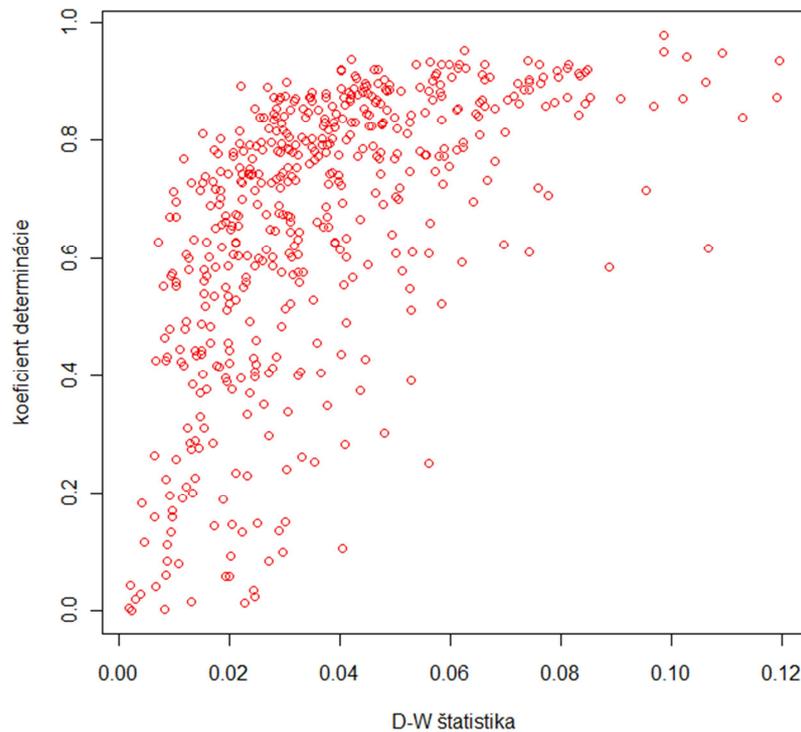
Z uvedenej deskriptívnej štatistiky vyplýva, že koeficient  $\beta_1$  je vo väčšine prípadov štatisticky významný (v 464 prípadoch je  $p$ -hodnota koeficientu  $\beta_1$  (premenná pvs) menšia ako hladina významnosti 10 %, v 459 prípadoch je menšia ako 5 % a v 445 prípadoch je menšia ako 1 % hladina významnosti).

Taktiež môžeme vidieť, že pri 500 iteráciách je v priemere koeficient determinácie 0.63 a Durbin – Watson štatistika je v priemere 0.037. Čiže naozaj vo väčšine prípadov sa potvrdilo pravidlo o identifikácii falošnej regresie na základe týchto dvoch kritérií. Pre lepšiu prehľadnosť zobrazíme tieto dve kritéria aj v podobe  $x$ - $y$  grafu.

```

> plot(r2s ~ dwt_stats, ylab = "koeficient determinácie", xlab =
      "D-W štatistika", col = "red")

```



Obrázok 40: Vzťah medzi koeficientom determinácie a D-W štatistikou

*Zdroj: výstup zo softvéru R*

### **Príklad 6.14**

V knižnici `lmtest` sa nachádza databáza `unemployment`, ktorá obsahuje 5 makroekonomických veličín. Precvičte si využívanie testov (hlavne interpretácie) z knižnice `urca` a otestujte všetky premenné a ich logaritmické diferencie na prítomnosť jednotkového koreňa, resp. na stacionárnosť.

### **Príklad 6.15**

V predchádzajúcom príklade (Príklad 6.14) sme už testovali na stacionaritu makroekonomické veličiny z databázy `unemployment` (knižnica `lmtest`). Ak by sme chceli zistiť, o koľko sa zmení inflácia (v danej databáze charakterizovaná pomocou deflátoru  $-p$ ), ak vzrastie množstvo peňazí v obehu (premenná peňažná zásoba  $-m$ ), mohli by sme využiť jednoduchý regresný model. Môžeme predpokladať, že čím viac peňazí v obehu, tým vyššia bude inflácia. Inými slovami, regresný koeficient pri peňažnej zásobe by mal byť kladný. Otázkou ostáva, aká bude hodnota tohto koeficientu a koľko variability inflácie dokážeme popísať prostredníctvom peňažnej zásoby.

---

### Príklad 6.15 – Riešenie

---

V tomto príklade máme kvantifikovať vzťah medzi infláciou a množstvom peňazí v obehu prostredníctvom jednoduchého lineárneho regresného modelu. My už však v tomto bode vieme, že ak do regresného modelu zahrnieme nestacionárne premenné, výsledky môžu byť zavádzajúce. Keďže stacionaritu týchto premenných sme testovali v jednom z predchádzajúcich príkladov (Príklad 6.14), taktiež vieme, že ani infláciu a ani peňažnú zásobu nemôžeme považovať za stacionárne. Ich logaritmické diferencie už za stacionárne považovať môžeme. Pozrime sa, aké rozdielne budú výsledky z regresného modelu, v ktorom budú vystupovať nestacionárne časové rady (označíme ho ako `model_1`), oproti modelu, do ktorého zahrnieme transformované premenné (`model_2`).

```
> library(lmtest)
> attach(data.frame(unemployment))
-----
> model_1 = lm(p ~ m)
-----
> summary(model_1)

Call:
lm(formula = p ~ m)

Residuals:
    Min       1Q   Median       3Q      Max
-0.09871 -0.05815 -0.00939  0.04689  0.14869

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 2.238e-01  8.299e-03  26.96  <2e-16 ***
m           1.625e-03  3.295e-05  49.31  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.06371 on 88 degrees of freedom
Multiple R-squared:  0.9651,    Adjusted R-squared:  0.9647
F-statistic: 2432 on 1 and 88 DF,  p-value: < 2.2e-16
-----
> dwtest(model_1, alternative = "two.sided")$statistic
      DW
0.09760078
```

V prvom modeli môžeme vidieť, že vzťah medzi skúmanými premennými je vysoko štatisticky významný a koeficient determinácie dosahuje vysokú hodnotu. Durbin – Watson štatistika je zas na druhej strane veľmi nízka. Na prvý pohľad z týchto výsledkov by malo byť zrejmé, že môže ísť o falošnú regresiu (keďže my vieme, že premenné sú nestacionárne, tak

vieme, že naozaj ide o falošnú regresiu). Pozrime sa, ako sa zmenia výsledky po transformácii premenných.

```
> dm = diff(log(m))
> dp = diff(log(p))
-----
> model_2 = lm(dp ~ dm)
-----
> summary(model_2)

Call:
lm(formula = dp ~ dm)

Residuals:
    Min       1Q   Median       3Q      Max
-0.1611109 -0.0263650  0.0005227  0.0242959  0.1288100

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.009474   0.006906  -1.372   0.174
dm           0.571207   0.080876   7.063 3.8e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.04539 on 87 degrees of freedom
Multiple R-squared:  0.3644,    Adjusted R-squared:  0.3571
F-statistic: 49.88 on 1 and 87 DF,  p-value: 3.801e-10
-----
> dwtest(model_2, alternative = "two.sided")$statistic
      DW
1.422353
```

Pri druhom modeli je regresný koeficient pri peňažnej zásobe stále významný, avšak koeficient determinácie výrazne poklesol a už neplatí, že by bol vyšší ako Durbin – Watson štatistika. Výskyt autokorelácie sa však transformáciou premenných neodstránil. Pre overenie významnosti koeficientu môžeme pri odhade modelu použiť maticu konzistentnú na prítomnosť autokorelácie a heteroskedasticity (HAC) – pomocou funkcie `coefTest()` z knižnice `sandwich()`.

```
> library(sandwich)
> coefTest(model_2, vcov = vcovHAC(model_2))

t test of coefficients:

            Estimate Std. Error t value  Pr(>|t|)
(Intercept) -0.0094739  0.0101132  -0.9368 0.3514645
dm           0.5712068  0.1623972   3.5173 0.0006952 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

$p$ -hodnota pri regresnom koeficiente sa síce znížila, avšak koeficient je stále vysoko štatisticky významný. Koeficient je kladný a štatisticky významný, čo naznačuje že pri zvyšovaní množstva peňazí v obehu došlo aj k zvyšovaniu cenovej hladiny.

### Príklad 6.16

Vráťme sa k databáze `EuStockMarkets` (knižnica `datasets`), v ktorej máme k dispozícii údaje zo 4 európskych akciových indexov. Ide o denné uzatváracie ceny indexov DAX (Nemecko), SMI (Švajčiarsko), CAC (Francúzsko) a FTSE (Anglicko) za obdobie od roku 1991 do roku 1998. V predchádzajúcom príklade sme testovali stacionárnosť týchto uzatváracích cien a dospeli sme k záveru, že ich za stacionárne považovať nemôžeme. Údaje môžeme považovať za stacionárne až po transformácii uzatváracích cien na logaritmické diferencie (tzv. spojité výnosy podľa vzťahu  $\ln(P_{t+1}/P_t)$ , kde  $P_t$  je uzatváracia cena indexu v čase  $t$ ).

V tomto príklade zistíte, či existujú jednosmerné alebo obojsmerné závislosti medzi týmito 4 indexmi, resp. medzi ich spojitými výnosmi (do Grangerovho modelu musia vstupovať stacionárne premenné).

---

### Príklad 6.16 – Riešenie

---

Existenciu jednosmerných alebo obojsmerných závislostí medzi skúmanými akciovými indexmi budeme overovať prostredníctvom konceptu Grangerovej kauzality. Na výber máme (minimálne) dve možnosti. Po prvé, môžeme odhadnúť autoregresné modely s rozloženým oneskorením (ARDL), v ktorých raz budú vystupovať oneskorené hodnoty druhého indexu a raz nie. O tom, či hodnoty druhého indexu prispievajú k predikcii prvého indexu, môžeme rozhodnúť na základe modelu ANOVA.

Druhý spôsob je zrejme jednoduchší. Keďže v softvéri R sa nachádzajú rôzne funkcie na výpočet Grangerovho modelu, môžeme požadované výpočty realizovať cez jednu z nich. Napríklad v knižnici `MSBVAR` sa nachádza funkcia `granger.test()`, pomocou ktorej vieme vypočítať Grangerovu kauzalitu medzi všetkými premennými v datase.

Ukážme si najprv prvý spôsob na indexoch CAC a DAX. Oneskorenia do regresných modelov vieme pridať viacerými spôsobmi, ale zrejme najjednoduchšie je využiť knižnicu `dynlm`, ktorá pri časových radoch ponecháva aj niektoré ich základné charakteristiky (dĺžku, začiatok a koniec obdobia, a iné). V rámci odhadu lineárneho regresného modelu cez funkciu `dynlm()` je možné veľmi jednoducho vytvoriť diferencie – pomocou funkcie `d()` alebo oneskorenia – pomocou funkcie `L()`. K výpočtu modelu ANOVA súži funkcia `anova()`,

ale nás v tomto bode nebude zaujímať celý výstup z danej funkcie, ale len vypočítaná  $p$ -hodnota z  $F$ -štatistiky. Výsledky potom chceme porovnať s druhým spôsobom výpočtu Grangerovho modelu.

```
> library(datasets)
> library(dynlm)
-----
> model1 = dynlm(d(log(CAC)) ~ L(d(log(CAC)), 1) +
  L(d(log(DAX)), 1), data = EuStockMarkets)
> model2 = dynlm(d(log(CAC)) ~ L(d(log(CAC)), 1), data =
  EuStockMarkets)
> model3 = dynlm(d(log(DAX)) ~ L(d(log(DAX)), 1) +
  L(d(log(CAC)), 1), data = EuStockMarkets)
> model4 = dynlm(d(log(DAX)) ~ L(d(log(DAX)), 1), data =
  EuStockMarkets)
-----
> format(anova(model1,model2)[["Pr(>F)"]][2],digits=10)
[1] "0.11915"
> format(anova(model3,model4)[["Pr(>F)"]][2],digits=10)
[1] "0.25722"
```

V rámci prvých dvoch modelov (model1 a model2) testujeme nulovú hypotézu, že nemecký akciový index DAX neovplyvňuje v Grangerovom zmysle kauzality francúzsky index CAC (značíme  $DAX \rightarrow CAC$ ). Inými slovami, hodnoty indexu DAX neprispievajú k predikcii hodnôt indexu CAC. V druhých dvoch modeloch (model3 a model4) naopak, že CAC neovplyvňuje v Grangerovom zmysle DAX ( $CAC \rightarrow DAX$ ).

Z uvedených výsledkov vyplýva, že medzi týmito dvoma akciovými indexmi neexistujú ani jednosmerné a ani obojsmerné závislosti. Ani v jednom prípade sme totiž nulovú hypotézu nevedeli zamietnuť.

Výsledky pre ďalšie indexy už nebudeme uvádzať týmto prvým spôsobom, keďže daný postup je zbytočne prácny. Zameriame sa radšej na výsledky z Grangerovho modelu medzi všetkými premennými s využitím funkcie `granger.test()` a môžeme porovnať, či pri indexoch DAX a CAC dostávame rovnaké výsledky.

```
> attach(data.frame(EuStockMarkets))
> rCAC = diff(log(CAC))
> rDAX = diff(log(DAX))
> rFTSE = diff(log(FTSE))
> rSMI = diff(log(SMI))
> data = data.frame(rCAC, rDAX, rFTSE, rSMI)
-----
> library(MSBVAR)
> granger.test(data, p = 1)
              F-statistic      p-value
```

```

rDAX -> rCAC      2.4307580      1.191459e-01
rFTSE -> rCAC     0.8821802      3.477275e-01
rSMI  -> rCAC     8.4895148      3.614753e-03
rCAC  -> rDAX     1.2844742      2.572153e-01
rFTSE -> rDAX     1.0414589      3.076155e-01
rSMI  -> rDAX     4.2900430      3.847417e-02
rCAC  -> rFTSE    3.1867806      7.439942e-02
rDAX  -> rFTSE    5.9821860      1.454343e-02
rSMI  -> rFTSE   15.5647175      8.269723e-05
rCAC  -> rSMI     5.2469308      2.209724e-02
rDAX  -> rSMI     1.7492293      1.861367e-01
rFTSE -> rSMI     6.8904559      8.736883e-03
-----
> format(1.191459e-01, digits = 5, format = "f")
[1] "0.11915"
> format(2.572153e-01, digits = 5, format = "f")
[1] "0.25722"

```

Výsledky pre indexy DAX a CAC sú rovnaké v porovnaní s predchádzajúcim postupom, jediný rozdiel spočíva v tom, že funkcia `granger.test()` uvádza *p*-hodnoty v tzv. vedeckej notácii. Cez funkciu `format()` však čísla vieme pre ľahšie porovnanie upraviť.

Z kompletných výsledkov Grangerovho modelu na daných údajoch vyplýva, že švajčiarsky index SMI ovplyvňuje na hladine významnosti 5 % indexy CAC (*p*-hodnota 0.00361), DAX (*p*-hodnota 0.03847) a FTSE (*p*-hodnota 0.00008). Ďalej index DAX ovplyvňuje FTSE (*p*-hodnota 0.01454), index CAC ovplyvňuje SMI (*p*-hodnota 0.0221) a taktiež index FTSE ovplyvňuje SMI (*p*-hodnota 0.00874). O obojsmerných závislostiach teda môžeme hovoriť vo vzťahu indexov SMI-CAC a SMI-FTSE. Jednosmerné závislosti sa preukázali medzi indexmi SMI-DAX a DAX-FTSE.

Grangerov model sa pri akciových indexoch využíva kvôli zisteniu stupňa prepojenosti medzi rôznymi trhmi, teda stupňa integrácie akciových trhov. Treba zdôrazniť, že použité údaje sú za obdobie od roku 1991 do roku 1998. V súčasnej dobe je integrácie trhov na vyššej úrovni a pri použití neskorších dát by výsledky Grangerovho modelu poukázali na výskyt obojsmerných závislostí medzi všetkými indexmi vyspelých trhov z EÚ.

### Príklad 6.17

V jednom z predchádzajúcich príkladov (Príklad 6.15) sme kvantifikovali vzťah medzi infláciou a množstvom peňazí v obehu. Pracovali sme s databázou `unemployment` (knižnica `lmtest`). Na základe regresného modelu sme dospeli k záveru, že ak sa zvýši peňažná zásoba (premenná *m*), tak sa zvýši aj inflácia (premenná *p*). Implicitne sme

vychádzali z takého funkčného tvaru modelu, v ktorom infláciu vysvetľujeme pomocou peňažnej zásoby. V tomto príklade sa pokúste zistiť, či jedna z týchto premenných ovplyvňuje druhú v Grangerovom zmysle kauzality.

---

### Príklad 6.17 – Riešenie

---

V tomto príklade máme zistiť, či existuje jednosmerná závislosť medzi infláciou (premenná  $p$ ) a peňažnou zásobou (premenná  $m$ ). Pre zjednodušenie budeme na určenie Grangerovej kauzality využívať už len funkciu `granger.test()`. Overovať teda ideme nulové hypotézy  $m \nrightarrow p$  (peňažná zásoba neovplyvňuje infláciu v Grangerovom zmysle kauzality) a  $p \nrightarrow m$  (inflácia neovplyvňuje peňažnú zásobu v Grangerovom zmysle kauzality). Pripomíname, že symbol “ $\nrightarrow$ ” je ekvivalentný symbolu “ $->$ ” vo funkcii `granger.test()`.

```
> library(lmtest)
> library(MSBVAR)
> attach(data.frame(unemployment))
-----
> dm = diff(log(m))
> dp = diff(log(p))
> data = data.frame(dm, dp)
-----
> granger.test(data, p = 1)

              F-statistic      p-value
dp -> dm  0.001981622      0.96459798
dm -> dp  4.948227071      0.02876739
```

Výsledky sú dosť jednoznačné. Nulovú hypotézu  $p \nrightarrow m$  nevieme zamietnuť, a teda nemôžeme tvrdiť, že inflácia ovplyvňuje peňažnú zásobu v Grangerovom zmysle kauzality. Na druhej strane hypotézu  $m \nrightarrow p$  vieme zamietnuť na hladine významnosti 5 %. Môžeme teda prijať alternatívnu hypotézu, teda že peňažná zásoba ovplyvňuje infláciu v Grangerovom zmysle kauzality.

### Príklad 6.18

V tomto príklade opäť využijeme údaje o akciových indexoch a HDP krajín V4 (Česká republika, Maďarsko, Poľsko a Slovensko) za obdobie od Q1:1996 do Q4:2009. Tieto údaje sú už po transformácii na logaritmické diferencie kvôli stacionarite premenných. Na týchto údajoch sme už zistili, že akciové index v krajinách V4 predbiehajú ich HDP o 3 mesiace. Teraz však overte, či akciové indexy predbiehajú HDP v daných krajinách o 6 mesiacov.

---

### Príklad 6.18 – Riešenie

---

Keďže teraz máme zistiť, či akciové indexy krajín V4 predbiehajú ich HDP o 6 mesiacov, musíme vo funkcii `granger.test()` nastaviť argument `p = 2` (keďže pracujeme s kvartálnymi údajmi). Funkcia vráti výsledky Grangerovho modelu pre všetky premenné v datasete, preto si z nich vyberieme len tie, ktoré nás zaujímajú.

```
> data = read.csv(file = "...cesta
k súboru...\\granger_gdp.csv", sep = ";", dec = ".", header =
T)
-----
> library(MSBVAR)
> granger.test(data, p = 2)
```

		F-statistic	p-value
px	-> hdp_cz	11.28135738	9.637912e-05
hdp_cz	-> px	0.29005467	7.495248e-01
bux	-> hdp_hu	5.33989588	8.054974e-03
hdp_hu	-> bux	1.14969256	3.253011e-01
wig	-> hdp_pl	2.47714901	9.465870e-02
hdp_pl	-> wig	0.44504344	6.434137e-01
sax	-> hdp_sk	2.24984352	1.164203e-01
hdp_sk	-> sax	0.53875082	5.869634e-01

Z týchto výsledkov vyplýva, že v prípade slovenského akciového indexu SAX sa pri druhom oneskorení v Grangerovom modeli už nepreukázal výskyt jednosmernej závislosti smerom k slovenskému HDP ( $p$ -hodnota 0.11642). Pri poľskom indexe WIG môžeme nulovú hypotézu  $WIG \nRightarrow HDP\_PL$  (akciový index WIG neovplyvňuje HDP v Poľsku v Grangerovom zmysle kauzality) zamietnuť na hladine významnosti 10 % (pri  $p$ -hodnote 0.094659). V prípade českého a maďarského akciového indexu ostávajú (v porovnaní s lagom 1) vzťahy významné na hladine 1 %, s  $p$ -hodnotou 0.000096 v prípade indexu PX a 0.008055 pri indexe BUX. Pri oboch teda môžeme zamietnuť nulové hypotézy a prijať tvrdenie, že český PX a maďarský BUX ovplyvňujú HDP v daných krajinách aj pri druhom oneskorení, teda akciový trh v týchto dvoch krajinách predbieha ich HDP aj o 6 mesiacov.

## Príloha 1

```
> ADF_GLS <- function(x, model = c("mu", "tau"), kmax = NULL,
  mLBo = NULL) {
+   TT <- length(x)
+   if(length(kmax) == 0) kmax = floor(12*(TT/100)^0.25);
+   if(length(mLBo) == 0) mLBo = floor(0.05*TT);
+
+   ## GLS detrending
+   glsd <- function(x, model = c("mu", "tau")) {
+     TT <- length(x)
+     alpha <- switch(model, mu = 1+-7/TT, tau = 1+-13.5/TT)
+     newx <- x[1]
+     for (i in 2:TT) newx[i] <- x[i]-x[i-1]*alpha
+     z <- switch(model, tau = cbind(rep(1,TT),1:TT), mu =
+ cbind(rep(1,TT)))
+     za <- z[1,]
+     for (i in 2:TT) za <- rbind(za, z[i,]-z[i-1,]*alpha)
+     psi <- lm(newx ~ -1 + za)$coefficients
+     y_tilda <- x - z*%cbind(psi)
+     return(y_tilda)
+   }
+
+   ## Tvorba datasetu pre pomocne regresie
+   m_data <- function(x, k = 1) {
+     TT <- length(x)
+     dataset <- matrix(ncol = 2+k, nrow = TT-k-1)
+     if (k == 0) colnames(dataset) <- c("yt", "delta_yt")
+     if (k > 0) colnames(dataset) <- c("yt", "delta_yt",
+ paste("lag",c(1:k)))
+     dataset[,1] <- x[(k+1):(TT-1)]
+     difx <- diff(x)
+     dataset[,2] <- difx[(k+1):(TT-1)]
+     if (k > 0) for (i in 1:k) dataset[,2+i] <- difx[(k+1-
+ i):(TT-1-i)]
+     return(dataset)
+   }
+
+   ## Hladanie specifikacie s nekorelovanymi reziduami
+   UC <- function(kmax = kmax, y, mLBo = mLBo) {
+     for (i in 0:kmax) {
+       dataset <- m_data(y, k = i)
+       if (i == 0) te <- lm(dataset[,2]~dataset[,1]-1)
+       if (i > 0) te <- lm(dataset[,2]~dataset[,1]-
+ 1+dataset[,3:(2+i)])
+       res <- na.omit(residuals(te))
+       pval <- c()
+       for (j in 1:mLBo) {
+         pval[j] <- Box.test(res, lag = j, type = "Ljung-
+ Box")$p.value
+       }
+       mpval <- min(pval)
+     }
+   }
+ }
```

```

+     if (mpval > 0.05) {
+       LBp <- mpval;
+       break
+     }
+     if (mpval <= 0.05) LBp <- NULL
+   }
+   if (length(LBp) == 0) {
+     warning("Suitable Auxiliary Regression not found")
+   }
+   results <- list()
+   results[["statistics"]] <- summary(te)$coefficients[1,3];
+   results[["lag"]] <- i;
+   return(results)
+ }
+
+ ## Hladanie specifikacie podla Ng - Perron (1995)
+ SEQ <- function(kmax = kmax, y) {
+   for (i in kmax:1) {
+     dataset <- m_data(y, k = i)
+     te <- lm(dataset[,2]~dataset[,1]-
1+dataset[,3:dim(dataset)[2]])
+     coef_val <- summary(te)$coefficients[i+1,4]
+     if (coef_val < 0.1) {
+       break
+     }
+   }
+   results <- list()
+   results[["statistics"]] <- summary(te)$coefficients[1,3];
+   results[["lag"]] <- i;
+   return(results)
+ }
+
+ ## Hladanie specifikacie podla MAIC
+ maic <- function(kmax = kmax, y) {
+   TT <- length(y); maic <- c(); coefs <- c()
+   for (i in kmax:1) {
+     dataset <- m_data(y, k = i)
+     te <- lm(dataset[,2]~dataset[,1]-1+dataset[,3:(2+i)])
+     coefs[kmax - i + 1] <- summary(te)$coefficients[1,3]
+     res <- residuals(te)
+     sig_squared <- ((TT-kmax)^-1)*sum(res[(kmax-
i+1):length(res)]^2)
+     b0 <- summary(te)$coefficients[1,1]
+     tauT <- (sig_squared^-1)*(b0^2)*sum(dataset[(kmax-
i+1):length(res),1]^2)
+     maic[kmax - i + 1] <- log(sig_squared)+2*(tauT+i)/(TT-
kmax)
+   }
+   a <- which(maic==min(maic));
+   tstat <- coefs[a];
+   results <- list()
+   results[["statistics"]] <- tstat;
+   results[["lag"]] <- kmax + 1 - a;
+   return(results)

```

```

+   }
+
+   ## OLS detrending
+   olsd <- function(x, model = c("mu", "tau")) {
+     TT <- length(x)
+     z <- switch(model, tau = cbind(rep(1,TT),1:TT), mu =
+ cbind(rep(1,TT)))
+     y_hat <- lm(x ~ -1 + z)$residuals
+     return(y_hat)
+   }
+
+   ## Hladanie specifikacie podla MAIC ale Perron - QU (2007)
+   maic_new_kopt <- function(kmax = kmax, y_hat) {
+     TT <- length(y_hat); maic <- c(); coefs <- c()
+     for (i in kmax:1) {
+       dataset <- m_data(y_hat, k = i)
+       te <- lm(dataset[,2]~dataset[,1]-1+dataset[,3:(2+i)])
+       coefs[kmax - i + 1] <- summary(te)$coefficients[1,3]
+       res <- residuals(te)
+       sig_squared <- ((TT-kmax)^-1)*sum(res[(kmax-
+ i+1):length(res)]^2)
+       b0 <- summary(te)$coefficients[1,1]
+       tauT <- (sig_squared^-1)*(b0^2)*sum(dataset[(kmax-
+ i+1):length(res),1]^2)
+       maic[kmax - i + 1] <- log(sig_squared)+2*(tauT+i)/(TT-
+ kmax)
+     }
+     a <- which(maic==min(maic));
+     lag <- kmax + 1 - a;
+     return(lag)
+   }
+
+   ## ADF-GLS statistika pre Perron - Qu (2007)
+   maic_new_stats <- function(kmax = kmax, y, kopt = kopt) {
+     TT <- length(y);
+     dataset <- m_data(y, k = kopt)
+     if (kopt > 0) te <- lm(dataset[,2]~dataset[,1]-
+ 1+dataset[,3:(2+kopt)])
+     if (kopt == 0) te <- lm(dataset[,2]~dataset[,1]-1)
+     tstat <- summary(te)$coefficients[1,3]
+     return(tstat)
+   }
+
+   # GLS detrending na datach
+   y <- glsd(x,model=model)
+
+   results <- matrix(ncol = 5, nrow = 4);
+   rownames(results) <- c("tau_seq", "tau_maic",
+ "tau_maic_new", "tau_uc");
+   colnames(results) <- c("statistics", "lag", "1%", "5%",
+ "10%")
+
+   temp <- abs(TT - c(50, 100, 250, 2500)); temp <-
+ which(temp==min(temp));

```

```

+   seq_mu <- c(-3.19, -2.48, -2.13, -2.90, -2.22, -1.88, -2.73,
-2.04, -1.72, -2.58, -1.95, -1.62)
+   maic_mu <- c(-2.74, -2.14, -1.86, -2.62, -2.05, -1.77, -
2.57, -1.98, -1.69, -2.57, -1.94, -1.64)
+   seq_tau <- c(-4.15, -3.52, -3.19, -3.82, -3.22, -2.90, -
3.63, -3.02, -2.71, -3.43, -2.87, -2.58)
+   maic_tau <- c(-3.59, -3.00, -2.72, -3.42, -2.87, -2.61, -
3.42, -2.85, -2.58, -3.40, -2.85, -2.56)
+
+   CM_mu <- matrix()
+   a <- maic(kmax = kmax, y); b <- SEQ(kmax = kmax, y); c <-
UC(kmax = kmax, y = y, mLBo = mLBo);
+   y_hat <- olsd(x, model = model); kopt <- maic_new_kopt(kmax
= kmax, y_hat = y_hat); stat <- maic_new_stats(kmax = kmax, y
= y, kopt = kopt);
+   results[1,1] <- a$statistics; results[1,2] <- a$lag;
+   results[2,1] <- b$statistics; results[2,2] <- b$lag;
+   results[3,1] <- stat; results[3,2] <- kopt;
+   results[4,1] <- c$statistics; results[4,2] <- c$lag;
+   if (model == "mu") {
+     results[1,3] <- seq_mu[(temp-1)*3+1]; results[1,4] <-
seq_mu[(temp-1)*3+2]; results[1,5] <- seq_mu[(temp-1)*3+3];
+     results[2,3] <- maic_mu[(temp-1)*3+1]; results[2,4] <-
maic_mu[(temp-1)*3+2]; results[2,5] <- maic_mu[(temp-1)*3+3];
+     results[3,3] <- -2.58; results[3,4] <- -1.98; results[3,5]
<- -1.62;
+     cv5mu <- -1.948-
17.839*(1/TT)^1+104.086*(1/TT)^2+0.802*(c$lag/TT)^1+5.558*(c$la
g/TT)^2-18.332*(c$lag/TT)^3
+     cv10mu <- -1.624-
19.888*(1/TT)^1+155.231*(1/TT)^2+0.709*(c$lag/TT)^1+5.48*(c$la
g/TT)^2-16.055*(c$lag/TT)^3
+     results[4,4] <- cv5mu; results[4,5] <- cv10mu;
+   }
+   if (model == "tau") {
+     results[1,3] <- seq_tau[(temp-1)*3+1]; results[1,4] <-
seq_tau[(temp-1)*3+2]; results[1,5] <- seq_tau[(temp-1)*3+3];
+     results[2,3] <- maic_tau[(temp-1)*3+1]; results[2,4] <-
maic_tau[(temp-1)*3+2]; results[2,5] <- maic_tau[(temp-
1)*3+3];
+     results[3,3] <- -3.42; results[3,4] <- -2.91; results[3,5]
<- -2.62;
+     cv5tau <- -2.838-
20.328*(1/TT)^1+124.191*(1/TT)^2+1.267*(c$lag/TT)^1+10.530*(c$
lag/TT)^2-24.6*(c$lag/TT)^3
+     cv10tau <- -2.550-
20.166*(1/TT)^1+155.215*(1/TT)^2+1.133*(c$lag/TT)^1+9.808*(c$la
g/TT)^2-20.313*(c$lag/TT)^3
+     results[4,4] <- cv5tau; results[4,5] <- cv10tau;
+   }
+   return(round(results, 3))
+ }

```

## Príloha 2

```
KPSS.NEW <- function(x, model = c("mu", "tau"), kern = c("QS",
"Bartlett")) {
  library(sandwich); library(urca)
  TT <- length(x)

  # Hľadanie Bandwidth - Newey-West (1994), Hobijn et al. (2004)
  ABS <- function(x) {
    TT <- length(x);
    n <- round(4*(TT/100)^(2/25));
    Gamma <- function(x,j) {
      sum(x[(1+j):TT]*x[((1+j):TT)-j])/TT
    }
    temp <- c()
    for (i in 1:n) temp[i] <- Gamma(x,i)
    s_0 <- Gamma(x,0) + 2 * sum(temp)
    ## STEP 3: s_2
    temp <- c()
    for (i in 1:n) temp[i] <- Gamma(x,i)*i^2
    s_2 <- 2 * sum(temp)
    ## STEP 4: GH
    GH <- 1.3221*((s_2/s_0)^2)^(1/5)
    ## STEP 5: Bandwidth - I decided to have minimum Bandwidth
of 1 - computational issues.
    M <- round(max(min(TT, GH*TT^(1/5)),1))
    return(M);
  }

  # Vyber rezidui podľa špecifikácie modelu pod nulovou H
  res <- switch(model, mu = residuals(lm(x~1)), tau =
residuals(lm(x~1+I(1:TT))))
  M <- ABS(res)

  # Long Run Variance - QS method
  if (kern == "QS") LRVc <- kernHAC(lm(res~1), prewhite = 0, bw
= M, kernel = "Quadratic Spectral", sandwich = T)*TT
  if (kern == "Bartlett") LRVc <- kernHAC(lm(res~1), prewhite =
0, bw = M, kernel = "Bartlett", sandwich = T)*TT

  # Testovacie statistiky - aj s porovnaním
  KPSS1 <- sum(cumsum(res)^2)/(TT^2)
  KPSS_H <- round(KPSS1/LRVc, 4)
  KPSS_UR <- round(ur.kpss(x, type = model, use.lag =
M)@teststat, 4)
  urca_crit <- ur.kpss(x, type = model, use.lag = M)@cval
  hobijn_crit <- switch(model, mu = c(0.348,0.460,0.580,0.754),
tau = c(0.119,0.148,0.178,0.219))
  rs <- matrix(ncol = 5, nrow = 2)
  first <- c(KPSS_H, M, hobijn_crit[4], hobijn_crit[2],
hobijn_crit[1])
}
```

```
second <- c(KPSS_UR, M, urca_crit[4], urca_crit[2],  
urca_crit[1])  
rs[1,] <- first; rs[2,] <- second;  
colnames(rs) <- c("statistics", "bandwidth", "1%", "5%",  
"10%")  
rownames(rs) <- c("hobijn", "urca-Bartlett")  
return(rs)  
}
```

## Príloha 3

```
> LS <- function(x, h = 0.10, SIMUL = 50, kmax = NULL, breaks =
  c(1, 2), model = c("coupled", "decoupled"), dating = c("RSS",
  "endogenous")) {
+   TT <- length(x)
+   if(length(kmax) == 0) kmax = floor(12*(TT/100)^0.25);
+   if(kmax > h*TT) kmax = floor(h*TT)-3
+   library(strucchange)
+
+   #Tvorba datasetu s S_wave
+   S_wave <- function(zt, x) {
+     TT <- length(x); dx <- diff(x); dzt <- diff(zt);
+     delta_wave <- cbind(na.omit(lm(dx ~ dzt -
  1)$coefficient));
+     S_w <- c(0, x[2:TT]-(x[1]-zt[1,2:dim(zt)[2]]**delta_wave)
  - zt[2:TT,2:dim(zt)[2]]**delta_wave)
+     return(S_w)
+   }
+
+   ## Tvorba datasetu pre pomocne regresie
+   #Tvorba zakladneho datasetu
+   datasets <- function(x, zt, S_w, k = 0, breaks = breaks) {
+     TT <- length(x);
+     NoV <- (3+2*breaks+k); aTT <- TT-1-k;
+     dat <- c(rep(0, aTT*NoV)); dim(dat) <- c(aTT,NoV);
+     dat[,1] <- diff(x)[(1+k):(TT-1)];
+     dat[,2] <- S_w[(1+k):(TT-1)]
+     dat[,3:(NoV-k)] <-
  apply(zt[(1+k):TT,2:dim(zt)[2]],2,diff);
+     if (k > 0) {
+       r = 1
+       for (i in (NoV-k+1):(NoV)) {
+         dat[,i] <- diff(S_w)[(1+k-r):(TT-1-r)]
+         r = r + 1
+       }
+     }
+     return(dat)
+   }
+
+   # Pocet oneskoreni pomocou Ng - Perron
+   SEQ <- function(kmax = kmax, x, S_w = S_w, zt = zt, breaks =
  breaks) {
+     for (i in kmax:1) {
+       dataset <- datasets(x = x, zt = zt, S_w = S_w, k = i,
  breaks = breaks)
+       te <- lm(dataset[,1]~dataset[,2:dim(dataset)[2]]-1)
+       last_pval <- summary(te)$coefficients[(dim(zt)[2]+i),4]
+       if (last_pval < 0.1) {
+         break
+       }
+     }
+   }
+ }
```

```

+   results <- list()
+   results[["statistics"]] <- summary(te)$coefficients[1,3];
+   results[["lag"]] <- i;
+   return(results)
+ }
+
+ # RSS kod pre situaciu konstanta trend po vlastnom brejku.
+ Neoptimalizovany postup - hruba sila.
+ RSS <- function(x, h = h) {
+   TT <- length(x);
+   start <- floor(h*TT)
+   space <- TT-2*start;
+   zts <- matrix(ncol = 4, nrow = TT)
+   zts[, 1] <- 1
+   zts[, 2] <- 1:TT
+   pb <- winProgressBar(title = "Dating Breaks", min = 0, max
+ = (space+1)^2, width = 300)
+   RSS <- c(); k = 1; omega <- c();
+   for (i in start:(TT-start)) {
+     zts[, 3] <- 0
+     zts[(i+1):TT, 3] <- 1
+     for (j in start:(TT-start)) {
+       zts[, 4] <- 0
+       zts[(j+1):TT, 4] <- 1:(TT-j)
+       RSS[k] <- sum(residuals(lm(x ~ zts - 1))^2)
+       if (RSS[k]==min(RSS)) {
+         zt <- zts
+         omega[1] <- i
+         omega[2] <- j
+       }
+     }
+     setWinProgressBar(pb, k, title =
+ paste(round(k/((space+1)^2)*100, 0), "% done", ": RSS break
+ search"))
+     k = k + 1
+   }
+   close(pb)
+   results <- list()
+   results[["omega"]] <- omega;
+   results[["zt"]] <- zt;
+   return(results)
+ }
+
+ # Hladanie zlomu - "coupled RSS"
+ UB <- function(x, h = h, breaks = breaks) {
+   TT <- length(x);
+   zt <- c(rep(0, TT*(2+2*breaks))); dim(zt) <-
+ c(TT,(2+2*breaks));
+   Lam <- c();
+   a <- breakpoints(x~1+I(1:TT), h = h, breaks = breaks)
+   Lam <- a$breakpoints
+   zt[,1] <- 1; zt[,2] <- 1:TT;
+   zt[,3] <- c(rep(0,Lam[1]),rep(1,TT-Lam[1]))
+   zt[,4] <- c(rep(0,Lam[1]),1:(TT-Lam[1]))

```

```

+   if (breaks == 2) {
+     zt[,5] <- c(rep(0,Lam[2]),rep(1,TT-Lam[2]))
+     zt[,6] <- c(rep(0,Lam[2]),1:(TT-Lam[2]))
+   }
+   results <- list()
+   results[["omega"]] <- Lam;
+   results[["zt"]] <- zt;
+   return(results)
+ }
+
+ # Procedura
+ # Hladanie zlomu - "endogenne" iba pre model s 1 zlomom
+ if (dating == "endogenous") {
+   TT <- length(x);
+   start <- floor(h*TT)
+   space <- TT-2*start;
+   zts <- matrix(ncol = 4, nrow = TT)
+   zts[, 1] <- 1
+   zts[, 2] <- 1:TT
+   stats <- c(); k = 1; omega <- c(); lag <- c(); statistics
  <- c()
+   if (model == "coupled") pb <- winProgressBar(title =
  "Dating Breaks", min = 0, max = (space+2), width = 300)
+   if (model == "decoupled") pb <- winProgressBar(title =
  "Dating Breaks", min = 0, max = (space+1)^2, width = 300)
+   tra <- c()
+   for (i in start:(TT-start)) {
+     zts[, 3] <- 0
+     zts[(i+1):TT, 3] <- 1
+     if (model == "coupled") {
+       zts[, 4] <- 0
+       zts[(i+1):TT, 4] <- 1:(TT-i)
+       S_w <- S_wave(zt = zts, x = x); tra[k] <- mean(S_w);
+       b <- SEQ(kmax = kmax, x = x, S_w = S_w, zt = zts,
  breaks = breaks)
+       rm(S_w)
+       stats[k] <- b$statistics
+       if (stats[k]==min(stats)) {
+         zt <- zts
+         omega[1] <- i
+         omega[2] <- i
+         lag <- b$lag
+         statistics <- stats[k]
+       }
+       setWinProgressBar(pb, k, title =
  paste(round(k/(space+2)*100, 0), "% done", ": Endogenous break
  search"))
+       k = k + 1
+     }
+   }
+   if (model == "decoupled") {
+     for (j in start:(TT-start)) {
+       zts[, 4] <- 0
+       zts[(j+1):TT, 4] <- 1:(TT-j)
+       S_w <- S_wave(zt = zts, x = x)

```

```

+         b <- SEQ(kmax = kmax, x = x, S_w = S_w, zt = zts,
breaks = breaks)
+         stats[k] <- b$statistics
+         if (stats[k]==min(stats)) {
+             zt <- zts
+             omega[1] <- i
+             omega[2] <- j
+             lag <- b$lag
+             statistics <- stats[k]
+         }
+         setWinProgressBar(pb, k, title =
paste(round(k/((space+1)^2)*100, 0), "% done", ": Endogenous
break search"))
+         k = k + 1
+     }
+ }
+ }
+ close(pb)
+ }
+
+ # RSS
+ if (dating == "RSS") {
+     if(model == "coupled") a <- UB(x = x, h = h, breaks =
breaks);
+     if(model == "decoupled") a <- RSS(x = x, h = h)
+     omega <- a$omega;
+     zt <- a$zt
+     S_w <- S_wave(zt = zt, x = x)
+     b <- SEQ(kmax = kmax, x = x, S_w = S_w, zt = zt, breaks =
breaks)
+     statistics <- b$statistics
+     lag <- b$lag
+ }
+
+ # Simulacia kritickych hodnot
+ pb <- winProgressBar(title = "Estimating critical values",
min = 0, max = SIMUL, width = 300)
+ sim <- c();
+ for (w in 1:SIMUL) {
+     temp <- predict(lm(x~zt-1))+cumsum(rnorm(TT))
+     S_w <- S_wave(zt = zt, x = temp)
+     sim[w] <- SEQ(kmax = kmax, x = temp, S_w = S_w, zt = zt,
breaks = breaks)$statistics
+     setWinProgressBar(pb, w, title =
paste(round(w/(SIMUL)*100, 0), "% done", ": Estimating
critical values"))
+ }
+ close(pb);
+
+ # Tvorba vysledkov
+ results <- matrix(ncol = 7, nrow = 1)
+ colnames(results) <- c("break 1", "break 2", "statistics",
"lag", "1%", "5%", "10%");

```

```

+   rownames(results) <- c("tau_sequential");
+   if (breaks == 1) {
+     if (model == "coupled") {
+       if (dating == "RSS") results[1,] <- c(omega, omega,
+ statistics, lag, quantile(sim, p = c(0.01, 0.05, 0.1)))
+       if (dating == "endogenous") results[1,] <- c(omega,
+ statistics, lag, quantile(sim, p = c(0.01, 0.05, 0.1)))
+     }
+     if (model == "decoupled") results[1,] <- c(omega[1],
+ omega[2], statistics, lag, quantile(sim, p = c(0.01, 0.05,
+ 0.1)))
+   }
+   if (breaks == 2) results[1,] <- c(omega, statistics, lag,
+ quantile(sim, p = c(0.01, 0.05, 0.1)))
+   return(results)
+ }

```

## Príloha 4

```
LSKPSS <- function(x, h = 0.10, kern = c("QS", "Bartlett"),
  breaks = c(1, 2), model = c("coupled", "decoupled")) {
  library(sandwich); library(urca); library(strucchange);
  TT <- length(x)
  # Automatic Bandwidth Selection - Newey-West (1994) a Hobijn
  et al. (2004)
  ABS <- function(x) {
    TT <- length(x);
    n <- round(4*(TT/100)^(2/25));
    Gamma <- function(x, j) {
      sum(x[(1+j):TT]*x[((1+j):TT)-j])/TT
    }
    temp <- c()
    for (i in 1:n) temp[i] <- Gamma(x,i)
    s_0 <- Gamma(x,0) + 2 * sum(temp)
    ## STEP 3: s_2
    temp <- c()
    for (i in 1:n) temp[i] <- Gamma(x,i)*i^2
    s_2 <- 2 * sum(temp)
    ## STEP 4: GH
    GH <- 1.3221*((s_2/s_0)^2)^(1/5)
    ## STEP 5: Bandwidth - I decided to have minimum Bandwidth
    of 1 - computational issues.
    M <- round(max(min(TT, GH*TT^(1/5)),1))
    return(M);
  }

  # RSS kod pre situaciu konstanta trend po vlastnom zlome.
  Neoptimalizovany - hruba sila.
  RSS <- function(x, h = h) {
    TT <- length(x);
    start <- floor(h*TT)
    space <- TT-2*start;
    zts <- matrix(ncol = 4, nrow = TT)
    zts[, 1] <- 1
    zts[, 2] <- 1:TT
    pb <- winProgressBar(title = "Dating Breaks", min = 0, max =
    (space+1)^2, width = 300)
    RSS <- c(); k = 1; omega <- c();
    for (i in start:(TT-start)) {
      zts[, 3] <- 0
      zts[(i+1):TT, 3] <- 1
      for (j in start:(TT-start)) {
        zts[, 4] <- 0
        zts[(j+1):TT, 4] <- 1:(TT-j)
        RSS[k] <- sum(residuals(lm(x ~ zts - 1))^2)
        if (RSS[k]==min(RSS)) {
          zt <- zts
          omega[1] <- i
          omega[2] <- j
        }
      }
    }
  }
}
```

```

    }
    setWinProgressBar(pb, k, title =
paste(round(k/((space+1)^2)*100, 0), "% done", ": RSS break
search"))
    k = k + 1
  }
}
close(pb)
results <- list()
results[["omega"]] <- omega;
results[["zt"]] <- zt;
return(results)
}

# Hladanie zlomu - "coupled RSS"
UB <- function(x, h = h, breaks = breaks) {
  TT <- length(x);
  zt <- c(rep(0, TT*(2+2*breaks))); dim(zt) <-
c(TT,(2+2*breaks));
  Lam <- c();
  a <- breakpoints(x~1+I(1:TT), h = h, breaks = breaks)
  Lam <- a$breakpoints
  zt[,1] <- 1; zt[,2] <- 1:TT;
  zt[,3] <- c(rep(0,Lam[1]),rep(1,TT-Lam[1]))
  zt[,4] <- c(rep(0,Lam[1]),1:(TT-Lam[1]))
  if (breaks == 2) {
    zt[,5] <- c(rep(0,Lam[2]),rep(1,TT-Lam[2]))
    zt[,6] <- c(rep(0,Lam[2]),1:(TT-Lam[2]))
  }
  results <- list()
  results[["omega"]] <- Lam;
  results[["zt"]] <- zt;
  return(results)
}

#Pocitanie statistiky
STS <- function(x = x, zt = zt, kern = kern) {
  res <- residuals(lm(x~zt-1))
  M <- ABS(res)
  LRVQS <- switch(kern, QS = kernHAC(lm(res~1), prewhite = 0,
bw = M, kernel = "Quadratic Spectral", sandwich = T)*TT,
Bartlett = kernHAC(lm(res~1), prewhite = 0, bw = M, kernel =
"Bartlett", sandwich = T)*TT)
  statistics <- sum(cumsum(res)^2)/(LRVQS*TT^2)
  results <- list()
  results[["statistics"]] <- statistics;
  results[["BW"]] <- M
  return(results)
}

if (model == "coupled") {
  a <- UB(x = x, h = h, breaks = breaks)
  omega <- a$omega;
  zt <- a$zt
}

```

```

    b <- STS(x = x, zt = zt, kern = kern)
    statistics <- b$statistics
    BW <- b$BW
  }

  if (model == "decoupled") {
    a <- RSS(x, h = h)
    omega <- a$omega;
    zt <- a$zt
    b <- STS(x = x, zt = zt, kern = kern)
    statistics <- b$statistics
    BW <- b$BW
  }

# Kriticke hodnoty
if (model == "coupled" & breaks == 1) {
  results <- matrix(ncol = 7, nrow = 2)
  colnames(results) <- c("level break", "trend break",
"statistics", "BW", "1%", "5%", "10%");
  rownames(results) <- c("Presno Lopez CV", "Lee Strazicich
CV");
  results[1,1:4] <- c(omega, omega, statistics, BW)
  results[2,1:4] <- c(omega, omega, statistics, BW)
  #LS 2001 asymptotic cv
  L <- omega/TT
  if (L > 0.5) L = 1 - L
  temp <- which(abs(L - c(0.1, 0.2, 0.3, 0.4, 0.5)) ==
min(abs(L - c(0.1, 0.2, 0.3, 0.4, 0.5))))
  #asy
  LS_asy <- matrix(c(0.0981, 0.1223, 0.1467, 0.1788, 0.0794,
0.0984, 0.1171, 0.1429, 0.0654, 0.0792, 0.0933, 0.1134,
0.0556, 0.0656, 0.0761, 0.0904, 0.0528,
0.0615, 0.0701, 0.0814), nrow = 4)
  results[2,5:7] <- LS_asy[c(4,2,1),5]
  # Presno - Lopez 2003
  b0 <- c(0.1188, 0.1473, 0.2160);
  b1 <- c(-0.2060, -0.2579, -0.3827);
  b2 <- c(0, 0, 0);
  b3 <- c(0.2913, 0.3411, 0.4568);
  b4 <- c(0, 0, 0);
  b5 <- c(0.1619, 0.1312, 0);
  b6 <- c(0, 0, -1.3537);
  cv10 <-
b0[1]+b1[1]*L+b2[1]*L^2+b3[1]*L^3+b4[1]*L^4+b5[1]/TT+b6[1]/(TT
^2)
  cv5 <-
b0[2]+b1[2]*L+b2[2]*L^2+b3[2]*L^3+b4[2]*L^4+b5[2]/TT+b6[2]/(TT
^2)
  cv1 <-
b0[3]+b1[3]*L+b2[3]*L^2+b3[3]*L^3+b4[3]*L^4+b5[3]/TT+b6[3]/(TT
^2)
  results[1,5:7] <- c(cv1,cv5,cv10)
  return(results)
}

```

```

if (model == "coupled" & breaks == 2) {
  results <- matrix(ncol = 7, nrow = 1)
  colnames(results) <- c("level break", "trend break",
"statistics", "BW", "1%", "5%", "10%");
  rownames(results) <- c("Carrion-i-Silvestre - Sanso ACV");
  results[1,1:4] <- c(omega[1], omega[2], statistics, BW)
  L1 <- omega[1]/TT; L2 <- omega[2]/TT;
  cv10 <- 0.0552+0.0041*abs(0.5-L1)+0.0072*abs(0.5-L2)-
0.1384*abs(L1-L2)+0.1324*abs(0.5-L1)^2+0.1240*abs(0.5-
L2)^2+0.1372*abs(L1-L2)^2
  cv5 <- 0.0665+0.0106*abs(0.5-L1)+0.0124*abs(0.5-L2)-
0.1830*abs(L1-L2)+0.1616*abs(0.5-L1)^2+0.1568*abs(0.5-
L2)^2+0.1819*abs(L1-L2)^2
  cv1 <- 0.0936+0.0305*abs(0.5-L1)+0.0272*abs(0.5-L2)-
0.2967*abs(L1-L2)+0.2282*abs(0.5-L1)^2+0.2365*abs(0.5-
L2)^2+0.2958*abs(L1-L2)^2
  results[1,5:7] <- c(cv1, cv5, cv10)
  return(results)
}
if (model == "decoupled") {
  results <- matrix(ncol = 7, nrow = 1)
  colnames(results) <- c("level break", "trend break",
"statistics", "BW", "1%", "5%", "10%");
  rownames(results) <- c("Carrion-i-Silvestre - Sanso ACV");
  results[1,1:4] <- c(omega[1], omega[2], statistics, BW)
  L1 <- omega[1]/TT; L2 <- omega[2]/TT;
  cv10 <- 0.0349+0.1508*abs(0.5-L1)+0.0210*abs(0.5-
L2)+(0.1194-0.7194*abs(0.5-L1))*abs(L1-L2)+1.1225*(abs(L1-
L2)^2)*(abs(0.5-L1)^2)
  cv5 <- 0.0371+0.2137*abs(0.5-L1)+0.0219*abs(0.5-L2)+(0.1748-
1.0449*abs(0.5-L1))*abs(L1-L2)+1.6409*(abs(L1-L2)^2)*(abs(0.5-
L1)^2)
  cv1 <- 0.0418+0.3687*abs(0.5-L1)+0.0245*abs(0.5-L2)+(0.3052-
1.8302*abs(0.5-L1))*abs(L1-L2)+2.9115*(abs(L1-L2)^2)*(abs(0.5-
L1)^2)
  results[1,5:7] <- c(cv1, cv5, cv10)
  return(results)
}
}

```

## Zoznam literatúry

- [1] A Free Software Project. 2013. [22. 06. 2013]. Dostupné na: <[http://cran.r-project.org/doc/html/interface98-paper/paper\\_2.html](http://cran.r-project.org/doc/html/interface98-paper/paper_2.html)>.
- [2] AMARA, J. – PAPELL, D. H. 2006. Testing for Purchasing Power Parity using stationary covariates. In: *Applied Financial Economics*, 2006, roč. 16, č. 1-2, s. 29 – 39. ISSN 0960-3107.
- [3] ANDĚL, J. 1985. *Matematická statistika*. Praha : SNTL/ALFA, 1985, 346 s.
- [4] ANDREWS, D. W. K. 1991. Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation. In *Econometrica*, 1991, roč. 59, č. 3, s. 817 – 858, ISSN 0012-9682.
- [5] BAUMÖHL, E. – LYÓCSA, Š. – VÝROST, T. 2011. *Fundamentálna analýza akciových trhov*. Košice : Elfa, 2011. 323 s. ISBN 978-80-8086-191-6.
- [6] BAUMÖHL, E. – LYÓCSA, Š. 2009. Stationarity of Time Series and the Problem of Spurious Regression. In: *SEMAFOR 2009*. Zborník príspevkov z medzinárodnej vedeckej konferencie. Košice : EU PHF Košice, 2009. s. 250 – 65. ISBN 978-80-225-2841-2. Dostupné na: <<http://ideas.repec.org/p/pramprapa/27926.html>>.
- [7] BAUMÖHL, E. – VÝROST, T. – LYÓCSA, Š. 2011. Are we able to capture the EU debt crisis? Evidence from PIIGS countries in panel unit-root framework. In: *TPAVF*. Zborník príspevkov z medzinárodnej vedeckej konferencie. Praha : VŠE, 2011. ISBN 978-80-245-1763-6. Dostupné na: <<http://ideas.repec.org/p/pramprapa/30334.html>>.
- [8] BAUMÖHL, E. – VÝROST, T. 2010. Stock Market Integration: Granger Causality Testing with respect to Nonsynchronous Trading Effects. In: *Finance a úvěr - Czech Journal of Economics and Finance*, 2010, roč. 60, č. 5, s. 414 – 425. ISSN 0015-1920.
- [9] BAUMÖHL, E. 2009. Analýza vzájomného vzťahu akciových trhov a HDP – Grangerov test kauzality. In: *Národohospodársky obzor*, 2009, roč. IX, č. 1, s. 5 – 20. ISSN 1213-2446.
- [10] BOLLERSLEV, T. 1986. Generalized autoregressive conditional heteroskedasticity. In: *Journal of Econometrics*, 1986, roč. 31, č. 3, s. 307 – 327, ISSN 0301-4076.
- [11] BOX, G. E. P. – COX, D. R. 1964. An Analysis of Transformations. In *Journal of the Royal Statistical Society. Series B*, 1964, roč. 26, č. 2, s. 211 – 252, ISSN 1369-7412.
- [12] BREITUNG, J. – DAS, S. 2005. Panel unit root tests under cross-sectional dependence. In: *Statistica Neerlandica*, 2005, roč. 59, č. 4, s. 414 – 33. ISSN 0039-0402.

- [13] BREUER, J. B. – McNOWN, R. – WALLACE, M. 2002. Series-specific Unit Root Tests with Panel Data. In: *Oxford Bulletin of Economics and Statistics*, 2002, roč. 64, č. 5, s. 527 – 46. ISSN 1468-0084.
- [14] BREUSCH, T. S. – PAGAN, A. R. 1979. A Simple Test for Heteroscedasticity and Random Coefficient Variation. In: *Econometrica*, 1979, roč. 47, č. 5, s. 1287 – 1294, ISSN 0012 – 9682.
- [15] BURKE, S. P. – HUNTER, J. 2005. *Modelling Non-Stationary Time Series: A Multivariate Approach*. Houndmills : Palgrave MacMillan, 2005, 253 s. ISBN 1-4039-0202-X.
- [16] CARRION-i-SILVESTRE, J. L. – SANSÓ, A. 2007. The KPSS test with two structural breaks. In: *Spanish Economic review*, 2007, roč. 9, č. 2, s. 105 – 27. ISSN 1435-5469.
- [17] CLEMENTE, J. – MONTANES, A. – REYES, M. 1998. Testing for a unit root in variables with a double change in the mean. In: *Economics Letters*, 1998, roč. 59, č. 2, s. 175 – 82. ISSN 0165-1765.
- [18] COOK, S. – MANNING, N. 2004. Lag optimisation and finite-sample size distortion of unit root tests. In *Economics Letters*, 2004, roč. 84, č. 2, s. 267 – 274, ISSN 0165-1765.
- [19] CUSHING, M. J. – MCGARVEY, M. G. 1999. *Covariance Matrix Estimation*. In Generalized Method of Moments Estimation, editor : MATYAS, L. Cambridge : Cambridge University Press, 1999. 316 s., ISBN 0521669677.
- [20] DAVIDSON, R. – MACKINNON, J. 2003. *Econometric Theory and Methods*. New York - Oxford University Press, 2003. ISBN 0-19512-372-7.
- [21] DEMETRESCU, M. – HASSLER, U. TARCOLEA, A. I. 2010. Testing for Stationarity in Large Panels with Cross-Dependence, and US Evidence on Unit Labor Cost. In: *Journal of Applied Statistics*, 2010, roč. 37, č. 8, s. 1381 – 97. ISSN 0266-4763.
- [22] DICKEY, D. A. – FULLER, W. A. 1979. Distribution of the Estimators for Autoregressive Time Series with a Unit Root. In: *Journal of the American Statistical Association*, 1979, roč. 74, č. 366, s. 427 – 31. ISSN 0162-1459.
- [23] DICKEY, D. A. – FULLER, W. A. 1981. Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root. In: *Econometrica*, 1981, roč. 49, č. 4, s. 1057 – 72. ISSN 0012-9682.
- [24] DUFOUR, J.-M. – FARHAT, A. – GARDIOL, L. – KHALAF, L. 1998. Simulation-based finite sample normality tests in linear regression. In: *Econometrics Journal*, 1998, roč. 1, s. 154 – 173, ISSN 1368-4221.

- [25] ELDER, J. – KENNEDY, P. E. 2001. Testing for Unit Roots: What Should Students Be Taught? In: *The Journal of Economic Education*, 2001, roč. 32, č. 2, s. 137 – 146, ISSN 0022-0485.
- [26] ELLIOTT, G. – JANSSON, M. 2003. Testing for unit roots with stationary covariates. In: *Journal of Econometrics*, 2003, roč. 115, č. 1, s. 75 – 89. ISSN 0304-4076.
- [27] ELLIOTT, G. – ROTHENBERG, T. J. – STOCK, J. H. 1996. Efficient tests for an autoregressive unit root. In: *Econometrica*, 1996, roč. 64, č. 4, s. 813 – 36. ISSN 0012-9682.
- [28] ENGLE, R. F. – GRANGER, C. W. 1987. Co-Integration and Error Correction: Representation, Estimation, Testing. In: *Econometrica*, 1987, roč. 55, č. 2, s. 251 – 76. ISSN 0012-9682.
- [29] ENGLE, R. F. 1982. Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. In: *Econometrica*, 1982, roč. 50, č. 4, s. 987 – 07. ISSN 0012-9682.
- [30] GODFREY, L. G. – TREMAYNE, A. R. 2005. The wild bootstrap and heteroskedasticity-robust tests for serial correlation in dynamic regression models, In: *Computational Statistics & Data Analysis*, roč. 49, č. 2, s. 377 – 395, ISSN 0167-9473.
- [31] GODFREY, L. G. 1997. Hausman tests for autocorrelation in the presence of lagged dependent variables: Some further results. In: *Journal of Econometrics*, roč. 82, č. 2, s. 197 – 207, ISSN 0304 – 4076.
- [32] GODFREY, L. G. 2009. *Bootstrap Tests for Regression*. Basingstoke : PALGRAVE MACMILLAN. 344 s. ISBN 978-0-230-20231-3.
- [33] GRANGER, C. W. – NEWBOLD, P. 1974. Spurious Regressions in Econometrics. In: *Journal of Econometrics*, 1974, roč. 2, č. 2, s. 111 – 20. ISSN 0304-4076.
- [34] GRANGER, C. W. 1969. Investigating Causal Relations by Econometric Models and Cross - Spectral Methods. In: *Econometrica*, 1969, roč. 37, č. 3, s. 424 – 38. ISSN 0012-9682.
- [35] GRANGER, C. W. J. 1980. Testing for Causality: A Personal Viewpoint. In *Journal of Economic Dynamics and Control*, 1980, roč. 2, s. 329-352, ISSN 0165-1889.
- [36] GREENE, W. H. 2012. *Econometric analysis*. 7 vyd. Upper Saddle River : Prentice Hall, 2012. 1189 s. ISBN 0-13-139538-6.
- [37] HADRI, K. – RAO, Y. 2008. Panel Stationarity Test with Structural Break. In: *Oxford Bulletin of Economics and Statistics*, 2008, roč. 70, č. 2, s. 245 – 69. ISSN 1468-0084.

- [38] HADRI, K. 2000. Testing for stationarity in heterogeneous panels. In: *Econometrics Journal*, 2000, roč. 3, č. 2, s. 148 – 61. ISSN 1368-4221.
- [39] HAMILTON, J. D. 1994. *Time Series Analysis*. Princeton : Princeton University Press, 1994, 799 s. ISBN 0-691-04289-6.
- [40] HANSEN, B. E. 2012. *Econometrics*. [online]. Wisconsin : University of Wisconsin, 2012. 334 s. [cit. 28.10.2012.] Dostupné na internete: < <http://www.ssc.wisc.edu/~bhansen/econometrics/Econometrics.pdf> >.
- [41] HARRIS, D. – LEYBOURNE, S. – McCABE, B. 2005. Panel Stationarity Tests for Purchasing Power Parity with Cross-Sectional Dependence. In: *Journal of Business & Economic Statistics*, 2005, roč. 23, č. 4, s. 395 – 409. ISSN 0735-0015.
- [42] HARRIS, R. D. F. – TZAVVALIS, E. 1999. Inference for unit roots in dynamic panels where the time dimension is fixed. In: *Journal of Econometrics*, 1999, roč. 91, č. 2, s. 201 – 26. ISSN 0304-4076.
- [43] HATRÁK, M. 2007. *Ekonometria*. Bratislava : IURA EDITION, 2007. ISBN 978-80-8078-150-7.
- [44] HOBIJN, B. – FRANSES, P. H. – OOMS, M. 2004. Generalizations of the KPSS-test for stationarity. In *Statistica Neerlandica*, 2004, roč. 58, č. 4, s. 483 – 502, ISSN 0039-0402.
- [45] HODOSHIMA, J. – ANDO, M. 2007. The Finite-Sample Performance of White's Test for Heteroskedasticity Under Stochastic Regressors. In: *Communications in Statistics – Simulation and Computation*, 2007, roč. 36, č. 6, s. 1201 – 1215, ISSN 0361 – 0918.
- [46] CHANG, Y. 2002. Nonlinear IV Unit Root Tests in Panels with Cross-Sectional Dependency. In: *Journal of Econometrics*, 2002, roč. 110, č. 2, s. 261 – 92. ISSN 0304-4076.
- [47] CHEUNG, Y. W. – LAI, K. S. 1995. Lag Order and Critical Values of a Modified Dickey-Fuller Test. In *Oxford Bulletin of Economics and Statistics*, 1995, roč. 57, č. 3, s. 411 – 419, ISSN 0305-9049.
- [48] CHOI, I. 2001. Unit Root Tests for Panel Data. In: *Journal of International Money and Finance*, 2001, roč. 20, č. 2, s. 249 – 72. ISSN 0261-5606.
- [49] CHOI, I. 2006. Combination Unit Root Tests for Cross-sectionally Correlated Panels. In: CORBAE, D. – DURLAUF, S. – HANSEN, B. (Eds.), *Econometric Theory and Practice: Frontiers of Analysis and Applied Research*, essays in honor of Peter C. B. Phillips. Cambridge: Cambridge University Press.

- [50] IM, K. – PESARAN, M. – SHIN, Y. 2003. Testing for Unit Roots in Heterogeneous Panels. In: *Journal of Econometrics*, 2003, roč. 115, č. 1, s. 53 – 74. ISSN 0304-4076.
- [51] IM, S. K. – LEE, J. – TIESLAU, M. 2010. Stationarity of Inflation: Evidence from Panel Unit Root Tests with Trend Shifts. 20<sup>th</sup> Annual Meetings of the Midwest Econometrics Group, Oct 1-2, 2010.
- [52] KAPETANIOS, G. – SHIN, Y. – SNELL, A. 2003. Testing for a unit root in the nonlinear STAR framework. In: *Journal of Econometrics*, 2003, roč. 112, č. 2, s. 359 – 79. ISSN 0304-4076.
- [53] KIEFER, N. M. – VOGELSANG, T. 2005. A New Asymptotic Theory for Heteroskedasticity-Autocorrelation Robust Tests. In: *Econometric Theory*, 2005, roč. 21, č. 6, s. 1130 – 1164. ISSN 0266-4666.
- [54] KIM, D. – PERRON, P. 2009. Unit root tests allowing for a break in the trend function at an unknown time under both the null and alternative hypotheses. In: *Journal of Econometrics*, 2009, roč. 148, č. 1, s. 1 – 13. ISSN 0304-4076.
- [55] KOČENDA, E. – ČERNÝ, K. 2007. *Elements of Time Series Econometrics: An Applied Approach*. Praha : Karolinum Press, 2007, 228 s., ISBN 8024613700.
- [56] KOENKER, R. 1981. A note on studentizing a test for heteroscedasticity. In: *Journal of Econometrics*, 1981, roč. 17, č. 1, s. 107 – 112, ISSN 0304 – 4076.
- [57] KWIATKOWSKI, D. – PHILLIPS, P. C. B. – SCHMIDT, P. – SHIN, Y. 1992. Testing the null hypothesis of stationarity against the alternative of a unit root. In: *Journal of Econometrics*, 1992, roč. 54, č. 1-3, s. 159 – 78. ISSN 0304-4076.
- [58] LEE, J. – IM, K. S. – TIESLAU, M. 2005. Panel LM unit root tests with level shifts. In: *Oxford Bulletin of Economics and Statistics*, 2005, roč. 67, č. 3, s. 393 – 419. ISSN 1468-0084.
- [59] LEE, J. – STRAZICICH, M. 2001. Testing the null of stationarity in the presence of a structural break. In: *Applied Economics Letters*, 2001, roč. 8, č. 6, s. 377 – 82. ISSN 1350-4851.
- [60] LEE, J. – STRAZICICH, M. 2003. Minimum lagrange multiplier unit root test with two structural breaks. In: *The Review of Economics and Statistics*, 2003, roč. 85, č. 4, s. 1082 – 89. ISSN 0034-6535.
- [61] LEE, J. – STRAZICICH, M. 2004. Minimum LM Unit Root Test with One Structural Break, Department of Economics, Appalachian State University, 04–16.

- [62] LEVIN, A. – LIN, C.F. – CHU, C.S. 2002. Unit Root Tests in Panel Data: Asymptotic and Finite Sample Properties. In: *Journal of Econometrics*, 2002, roč. 108, č. 1, s. 1 – 24. ISSN 0304-4076.
- [63] LEVIN, A. – LIN, C.F. 1993. Unit Root Test in Panel Data: New Results, Discussion Paper, 93-56, Department of Economics, University of California at San Diego.
- [64] LEYBOURNE, S. J. – McCABE, B. P. M. 1994. A Consistent Test for a Unit Root. In: *Journal of Business & Economic Statistics*, 1994, roč. 12, č. 2, s. 157 – 66. ISSN 0735-0015.
- [65] LUMSDAINE, R. L. – PAPELL, D. H. 1997. Multiple Trend Breaks and the Unit-Root Hypothesis. In: *The Review of Economics and Statistics*, 1997, roč. 79, č. 2, s. 212 – 18. ISSN 0034-6535.
- [66] LYÓCSA, Š. – BAUMÖHL, E. – VÝROST, T. 2013. *Kvantitatívne metódy v ekonómii II*. Košice : Elfa, 2013, 460 s. ISBN 978-80-8086-210-7.
- [67] LYÓCSA, Š. – VÝROST, T. – BAUMÖHL, E. 2011. Unit-Root and Stationarity Testing with Empirical Application On Industrial Production of CEE-4 Countries [online]. *SSRN Working Paper Series*. Dostupné na: <[http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1785223](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1785223)>.
- [68] LYÓCSA, Š. 2013. Ekonomická aktivita a akciový trh v krajinách V4. Habilitačná práca.
- [69] MACKINNON, J. G. 1996. Numerical Distribution Functions for Unit Root and Cointegration Tests. In: *Journal of Applied Econometrics*, 1996, roč. 11, č. 6, s. 601 – 18. ISSN 0883-7252.
- [70] MADDALA, G. – WU, S. 1999. A Comparative Study of Unit Root Tests and a New Simple Test. In: *Oxford Bulletin of Economics and Statistics*, 1999, roč. 61, č. 0 (special issue), s. 631 – 52. ISSN 0305-9049.
- [71] MILLS, T. C. 1999. *The Econometric Modelling of Financial Time Series*. Cambridge : Cambridge University Press, 1999. ISBN 0-521-62413-4.
- [72] MOON, H. – PERRON, B. 2004. Testing for a Unit Root in Panels with Dynamic Factors. In: *Journal of Econometrics*, 2004, roč. 122, č. 1, s. 81 – 126. ISSN 0304-4076.
- [73] NELSON, CH. R. – PLOSSER, CH. R. 1982. Trends and random walks in macroeconomic time series: Some evidence and implications. In: *Journal of Monetary Economics*, 1982, roč. 10, č. 2, s. 139 – 162. ISSN 0304-3932.

- [74] NEUBRUNN T. – RIEČAN B. 1992. *Teória miery*. Bratislava : Veda, 1992, 403 s. ISBN 80-224-0368-7.
- [75] NEWAY, W. K. – WEST, K. D. 1994. Automatic Lag Selection in Covariance Matrix Estimation. In *The Review of Economic Studies*, 1994, roč. 61, č. 4, s. 631 – 653, ISSN 0034-6527.
- [76] NG, S. – PERRON, P. 1995. Unit Root Tests in ARMA Models with Data-Dependent Methods for the Selection of the Truncation Lag. In *Journal of the American Statistical Association*, 1995, roč. 90, č. 429, s. 268 – 281, ISSN 0162-1459.
- [77] NG, S. – PERRON, P. 2001. Lag length selection and the construction of unit root tests with good size and power. In: *Econometrica*, 2001, roč. 69, č. 6, s. 1519 – 54. ISSN 0012-9682.
- [78] PERRON, P. – QU, Z. 2007. A simple modification to improve the finite sample properties of Ng and Perron`s unit root tests. In *Economics Letters*, 2007, roč. 94, č. 1, s. 12 – 19, ISSN 0165-1765.
- [79] PERRON, P. – VOGELSANG, T. J. 1992. Nonstationarity and level shifts with an application to purchasing power parity. In: *Journal of Business & Economic Statistics*, 1992, roč. 10, č. 3, s. 301 – 20. ISSN 0735-0015.
- [80] PERRON, P. – ZHU, X. 2005. Structural breaks with deterministic and stochastic trends. In *Journal of Econometrics*, 2005, roč. 129, č. 1-2, s. 65 – 119, ISSN 0304-4076.
- [81] PERRON, P. 1989. The great crash, the oil price shock, and the unit root hypothesis. In: *Econometrica*, 1989, roč. 57, č. 6, s. 1361 – 401. ISSN 0012-9682.
- [82] PERRON, P. 1997. Further evidence on breaking trend functions in macroeconomic variables. In: *Journal of Econometrics*, 1997, roč. 80, č. 2, s. 355 – 85. ISSN 0304-4076.
- [83] PESARAN, M.H. – SMITH, L.V. – YAMAGATA, T. 2009. A Panel Unit Root Test in the Presence of a Multifactor Error Structure. Working paper, Cambridge University.
- [84] PESARAN, M.H. 2007. A Simple Panel Unit Root Test In The Presence Of Cross Section Dependence. In: *Journal of Applied Econometrics*, 2007, roč. 22, č. 2, s. 265 – 312. ISSN 1099-1255.
- [85] PHILLIPS, P. C. B. – PERRON, P. 1988. Testing for a unit root in time series regression. In: *Biometrika*, 1988, roč. 75, č. 2, s. 335 – 46. ISSN 1464-3510.

- [86] PHILLIPS, P.C.B. – SUL, D. 2003. Dynamic Panel Estimation and Homogeneity Testing Under Cross Section Dependence. In: *Econometrics Journal*, 2003, roč. 6, č. 1, s. 217 – 59. ISSN 1368-4221.
- [87] POO, J. R. 2003. *Computer-Aided Introduction to Econometrics*. Springer, 2003. 350 s. ISBN 978-3-540-44114-4.
- [88] PRESNO, J. – LÓPEZ, A. J. 2003. Response surface estimates of stationarity tests with a structural break. In: *Economics Letters*, roč. 78, č. 3, s. 395 – 399, ISSN 0165-1765.
- [89] RIEČAN B. – LAMOŠ F. – LENÁRT C. 1984. *Pravdepodobnosť a matematická štatistika*. Bratislava : Alfa, 1984, 317 s.
- [90] ROMÃO, X. – DELGADO, R. – COSTA, A. 2010. An empirical power comparison of univariate goodness-of-fit tests for normality. In: *Journal of Statistical Computation and Simulation*, 2010, roč. 80, č. 5, s. 545 – 591, ISSN 0094 – 9655.
- [91] SAMUELSON, P.A. – Solow, R.M. 1960. Analytical Aspects of Anti-Inflation Policy. In: *American Economic Review*, 1960, roč. 50, č. 2, s. 177 – 194, ISSN 0002 – 8282.
- [92] SCHMIDT, P. – PHILLIPS, P.C.B. 1992. LM Test for a Unit Root in the Presence of Deterministic Trends. In: *Oxford Bulletin of Economics and Statistics*, 1992, roč. 54, č. 3, s. 257 – 87. ISSN 1468-0084.
- [93] SCHWERT, G. W. 1989. Tests for Unit Roots: A Monte Carlo Investigation. In: *Journal of Business & Economic Statistics*, 1989, roč. 7, č. 2. s. 147 – 59. ISSN 0735-0015.
- [94] STELLMACH, T. 2013. [22. 06. 2013]. Dostupné online: <[http://commons.wikimedia.org/wiki/File:Dice\\_Distribution\\_\(bar\).svg](http://commons.wikimedia.org/wiki/File:Dice_Distribution_(bar).svg)>.
- [95] STEVENS, S. S. 1946. On the theory of scales of measurement. In: *Science*, roč. 103, č. 2684, s. 677 – 680. ISSN 0036 – 8075.
- [96] SUL, D. – PHILLIPS, P. C. B. – CHOI, CH. Y. 2005. Prewhitening Bias in HAC Estimation. In *Oxford Bulletin of Economics and Statistics*, 2005, roč. 67, č. 4, s. 517 – 546, ISSN 0305-9049.
- [97] TAYLOR, M. P. – SARNO, L. 1998. The behavior of real exchange rates during the post-Bretton Woods period. In: *Journal of International Economics*, 1998, roč. 46, č. 2, s. 281 – 312. ISSN 0022-1996.
- [98] WANG, P. 2009. *Financial econometrics*, 2nd edition. Routledge : New York, 2009, 320 s., ISBN 0-203-89287-9.

- [99] WHITE, H. 1980. A Heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity, In: *Econometrics*, 1980, roč. 48, č. 4, s. 817 – 838, ISSN 0012 – 9682.
- [100] YAZICI, B. – YOLACAN, S. 2007. A comparison of various tests of normality. In: *Journal of Statistical Computation and Simulation*, 2007, roč. 77, č. 2, s. 173 – 185. ISSN 0094 – 9655.
- [101] ZAMAN, A. 2000. The Inconsistency of the Breusch-Pagan Test, In: *Journal of Economic and Social Research*, 2000, roč. 2, č. 1, s. 1 – 11, ISSN 1302 – 1060.
- [102] ZIVOT, E – ANDREWS, D. W. K. 1992. Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis. In: *Journal of Business & Economic Statistics*, 1992, roč. 10, č. 3, s. 251 – 70. ISSN 0735-0015.
- [103] ZLATOŠ, P. 2011. *Lineárna ALGEBRA a geometria*. Bratislava : Marenčin PT, 2011, 744s. ISBN 978-80-8114-111-9.

## Zoznam obrázkov

Obrázok 1: Transponovaná matica.....	19
Obrázok 2: Súčin matíc .....	21
Obrázok 3: Transponovanie súčinu matíc .....	26
Obrázok 4: Ortogonálna projekcia .....	105
Obrázok 5: Možné výsledky dvoch hodov kockou a ich pravdepodobnosti.....	119
Obrázok 6: Príklad lineárneho a log-log modelu .....	140
Obrázok 7: Odhad pomocou priemeru (vľavo) a pomocou regresnej priamky (vpravo).....	149
Obrázok 8 Vizualizácia štvorcov rezíduí .....	151
Obrázok 9: $x$ - $y$ graf vzťahu ceny a rozlohy bytov .....	165
Obrázok 10: Box-plot priemerných miezd mužov (1) a žien (0) .....	168
Obrázok 11: Lineárny model ako ortogonálna projekcia.....	172
Obrázok 12: $x$ - $y$ graf vzťahu spotreby a dôchodku .....	191
Obrázok 13: Rezíduá z regresného modelu.....	193
Obrázok 14: $x$ - $y$ graf vzťahu cien bytov a ich plochy v mestskej časti Košice - Sever .....	218
Obrázok 15: Vizualizácia normality rezíduí .....	220
Obrázok 16: Rozdelenia regresorov použitých v experimente .....	221
Obrázok 17: Rozdelenia rezíduí použitých v experimente .....	223
Obrázok 18: $x$ - $y$ graf štvorcov rezíduí a plochy bytov .....	231
Obrázok 19: Perzistentnosť šoku .....	249
Obrázok 20: Rezíduá a autokorelačné funkcie z ADF testu .....	259
Obrázok 21: Veľkosť a neupravená sila ADF-GLS testu pre rôzne metódy voľby počtu oneskorení a perzistentnosti procesu.....	264
Obrázok 22: Falošná regresia – vzťah medzi koeficientom determinácie a D-W štatistikou	277
Obrázok 23: Časový rad typu random walk a jeho logaritmicke diferencie .....	281
Obrázok 24: Vzťah medzi nezamestnanosťou a infláciou .....	295
Obrázok 25: Vzťah medzi výdavkami na vzdelávanie a príjmom.....	296
Obrázok 26: Rezíduá z regresného modelu ukazovateľa P/B .....	302
Obrázok 27: Box – plot skúmaných premenných .....	309
Obrázok 28: Závislosť vygenerovaných premenných.....	316
Obrázok 29: Rezíduá z regresného modelu.....	317
Obrázok 30: Rezíduá z regresného modelu – bez zjavného vzoru .....	320

Obrázok 31: Kvadratická závislosť dvoch premenných .....	321
Obrázok 32: Rezíduá z regresného modelu – so zjavným vzorom .....	321
Obrázok 33: Box – ploty reálnych miezd v krajoch SR.....	324
Obrázok 34: Box – ploty nezamestnanosti v krajoch SR.....	326
Obrázok 35: Funkčný vzťah medzi reálnou mzdou a nezamestnanosťou v okresoch SR.....	328
Obrázok 36: Uzatváracie ceny a spojité výnosy skúmaných indexov .....	334
Obrázok 37: Štrukturálne zlomy identifikované pomocou ZA testu .....	347
Obrázok 38: Vývoj ukazovateľov D/HDP v krajinách PIIGGS .....	350
Obrázok 39: Štrukturálne zlomy identifikované pomocou ZA testu .....	353
Obrázok 40: Vzťah medzi koeficientom determinácie a D-W štatistikou .....	356

## Zoznam tabuliek

Tabuľka 1: Odpovede respondentov .....	11
Tabuľka 2: Odpovede respondentov (pokračovanie).....	12
Tabuľka 3: Vzťah medzi hodnotami náhodných premenných a ich kovarianciou .....	120
Tabuľka 4: Predpoklady lineárneho modelu .....	160
Tabuľka 5: Cena a rozloha vybraných bytov v Košiciach .....	164
Tabuľka 6: HDP a spotreba za obdobie 1995 – 2010 v SR.....	190
Tabuľka 7: Prístupy k testovaniu heteroskedasticity .....	200
Tabuľka 8: Inflačné faktory rozptylu $VIF_i$ .....	212
Tabuľka 9: Rozdelenia rezíduí .....	222
Tabuľka 10: Empirická chyba I. druhu a sila Monte – Carlo Jarque – Berovho testu.....	224
Tabuľka 11: Prehľad testov jednotkového koreňa a testov na stacionaritu .....	271
Tabuľka 12: Prehľad panelových testov jednotkového koreňa a testov na stacionaritu .....	272
Tabuľka 13: Akciové index krajín V4 – testovanie stacionarity.....	282
Tabuľka 14: Akciové index krajín V4 – výsledky z regresných modelov.....	283

## Zoznam programových knižníc

- [1] strucchange 1.4-7
- [2] MASS 7.3-19
- [3] UsingR 0.1-18
- [4] ggplot2 0.9.1
- [5] lattice 0.20
- [6] triangle 0.5
- [7] mvtnorm 0.9-9992
- [8] car 2.0-12
- [9] lmtest 0.9-30
- [10] psych 1.2.4
- [11] zoo
- [12] datasets 2.14.0
- [13] moments\_0.13
- [14] TeachingDemos\_2.8
- [15] Matching\_4.8-0
- [16] nortest\_1.0-2
- [17] lawstat\_2.3
- [18] outliers\_0.14
- [19] tseries\_0.10-29
- [20] vcd\_1.2-13
- [21] ellipse\_0.3-7
- [22] HSAUR\_1.3-2
- [23] alr\_2.0.0
- [24] stats\_2.15.2
- [25] ltm\_0.9-9
- [26] modeest\_2.1
- [27] diptest\_0.75-4

Názov: Kvantitatívne metódy v ekonómii III.  
Autori: Tomáš Výrost, Eduard Baumöhl, Štefan Lyócsa  
Vydavateľstvo: elfa, s.r.o., Park Komenského 7, 040 01 Košice  
Vydanie: prvé  
Tlač: elfa, s.r.o., Park Komenského 7, 040 01 Košice

ISBN 978-80-8086-211-4

**ISBN 978-80-8086-211-4**