



**MATEMATICKO-FYZIKÁLNÍ
FAKULTA**
Univerzita Karlova

DIPLOMOVÁ PRÁCE

Jan Fučík

**Identifikace modelů finančních
časových řad**

Katedra pravděpodobnosti a matematické statistiky

Vedoucí diplomové práce: RNDr. Jitka Zichová, Dr.

Studijní program: Matematika

Studijní obor: Finanční a pojistná matematika

Praha 2016

Prohlašuji, že jsem tuto diplomovou práci vypracoval samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona v platném znění, zejména skutečnost, že Univerzita Karlova v Praze má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle §60 odst. 1 autorského zákona.

V Praze dne 20. července 2016

Jan Fučík

Název práce: Identifikace modelů finančních časových řad

Autor: Jan Fučík

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí diplomové práce: RNDr. Jitka Zichová, Dr., Katedra pravděpodobnosti a matematické statistiky

Abstrakt: Práce se zabývá identifikací modelů finančních časových řad. Čtenář se seznámí s jednorozměrnými a mnohorozměrnými modely ARMA a způsoby jejich identifikace. Jsou představeny postupy, které využívají korelační strukturu časové řady, a také informační kritéria. Fungování jednotlivých postupů je ověřeno na simulovaných časových řadách AR, MA a ARMA. Kritéria jsou následně porovnávána z hlediska spolehlivosti a jednoduchosti použití. Na závěr jsou uvedeny ukázky identifikace jednorozměrného a mnohorozměrného modelu ARMA pro reálné časové řady z finanční praxe. Použitá data a zdrojové kódy v programu R jsou k dispozici na přiloženém CD.

Klíčová slova: časová řada, identifikační kritéria, ARMA model

Title: Financial time series model identification

Author: Jan Fučík

Department: Department of Probability and Mathematical Statistics

Supervisor: RNDr. Jitka Zichová, Dr., Department of Probability and Mathematical Statistics

Abstract: This thesis deals with the financial time series model identification. The univariate and multivariate ARMA models and their identification criteria are described. The procedures using the correlation structure of the time series and some information criteria are presented. The functioning of the criteria is verified on simulated time series AR, MA and ARMA. Afterwards, the criteria are compared in terms of reliability and simplicity of use. Finally, there are two examples of univariate and multivariate ARMA model identification for the real financial time series. The data and the R programme source code are enclosed on a CD.

Keywords: time series, identification criteria, ARMA model

Na tomto místě bych rád poděkoval vedoucí své diplomové práce RNDr. Jitce Zichové, Dr. za cenné rady, připomínky a ochotu, se kterou se mi věnovala.

Obsah

Úvod	3
1 Lineární modely pro jednorozměrné časové řady	4
1.1 Základní pojmy	4
1.2 Lineární proces	6
1.3 Model klouzavých součtů	8
1.4 Autoregresní model	9
1.5 Smíšený model ARMA	10
1.6 Výstavba modelu	11
2 Identifikace modelů jednorozměrných časových řad	12
2.1 Boxova-Coxova transformace	12
2.2 Transformace diferencováním	13
2.2.1 Dickeyův-Fullerův test	13
2.2.2 Rozšířený Dickeyův-Fullerův test	15
2.3 Identifikace pomocí tvaru ACF a PACF	17
2.4 Identifikace pomocí speciálních funkcí	18
2.4.1 Inverzní autokorelační funkce	18
2.4.2 Rozšířená výběrová autokorelační funkce	20
2.5 Identifikace pomocí informačních kritérií	23
2.5.1 Akaikeho informační kritérium	23
2.5.2 Akaikeho Bayesovské informační kritérium	24
2.5.3 Další informační kritéria	25
2.6 Postup identifikace modelu	26
3 Simulace jednorozměrných časových řad	27
3.1 Simulace procesu AR(1)	27
3.2 Simulace procesu MA(1)	29
3.3 Simulace procesu ARMA(1,1)	31
3.4 Porovnání identifikačních kritérií	33
3.5 Dickeyův-Fullerův test	35
4 Softwarové zpracování jednorozměrné časové řady	37
5 Lineární modely pro mnohorozměrné časové řady	40
5.1 Základní pojmy	40
5.2 Vektorový model klouzavých součtů	44
5.3 Vektorový autoregresní model	45
5.4 Vektorový smíšený model VARMA	46

6	Identifikace modelů mnohorozměrných časových řad	48
6.1	Identifikace pomocí maticové autokorelační funkce	48
6.2	Identifikace pomocí parciální korelační maticové funkce	49
6.3	Identifikace pomocí informačních kritérií	50
7	Simulace mnohorozměrných časových řad	51
7.1	Simulace procesu VAR(1)	51
7.2	Simulace procesu VMA(1)	54
7.3	Simulace procesu VARMA(1,1)	56
7.4	Porovnání identifikačních kritérií	59
8	Softwarové zpracování mnohorozměrné časové řady	61
	Závěr	64
	Seznam použité literatury	65
	Seznam obrázků	67
	Seznam tabulek	68

Úvod

Mnoho dat z oblasti ekonomie a financí se uvádí v podobě tzv. časových řad. Jde o uspořádanou posloupnost pozorování v čase, která máme k dispozici s určitou frekvencí, např. týdenní, denní, hodinovou i několikaminutovou. Může se jednat o ceny akcií, úrokové sazby, cenové indexy, měnové kurzy a další. Za účelem popisu chování takových řad a konstrukce předpovědí jejich budoucího vývoje se používají různé matematické modely. Při jejich aplikaci je klíčová volba vhodného modelu pro konkrétní data, tzv. identifikace modelu. Cílem této práce je představit vybrané matematické modely vhodné pro analýzu finančních časových řad a popsat jejich identifikační kritéria. Na rozdíl od četné literatury, např. [3], jsou ve vzájemných souvislostech uvedeny identifikační procedury jak pro jednorozměrné, tak i mnohorozměrné řady. Fungování jednotlivých kritérií bude ověřeno pomocí softwaru na simulovaných časových řadách a následně budou aplikována na reálná data z finanční praxe.

Práce se zabývá lineárními modely ARMA, které jsou pro svou jednoduchost, snadnou interpretovatelnost a implementaci v softwarových produktech populární a v praxi používané i přes to, že pro analýzu finančních dat byly navrženy různé obecnější a složitější modely. Práce je rozdělena do dvou tematických bloků. Kapitoly 1 – 4 pojednávají o modelech jednorozměrných časových řad a kapitoly 5 – 8 se věnují modelům mnohorozměrných řad. Oba bloky jsou členěny stejným způsobem: nejprve jsou shrnuty základy teorie náhodných procesů, poté jsou představeny jednotlivé modely AR, MA a ARMA a jejich vlastnosti. V dalších kapitolech se čtenář dozví o jednotlivých možnostech identifikace, která spočívá ve volbě vhodného modelu a jeho řádu. Jedná se o postupy založené na korelační struktuře časové řady a informační kritéria. Kromě klasických postupů jsou popsány i některé méně známé, například rozšířená výběrová autokorelační funkce u jednorozměrných řad a některé typy maticových autokorelačních funkcí u mnohorozměrných řad. Dále je ověřeno fungování popsaných kritérií na simulovaných časových řadách. Jednotlivá kritéria jsou porovnána z hlediska jejich spolehlivosti a jednoduchosti použití. V závěrečné kapitole každého bloku pak uvádíme ukázkou identifikace modelu pro konkrétní časové řady z finanční praxe.

Teoretické podklady byly čerpány zejména z knih [20] a [4]. Na příslušných místech v textu jsou citovány i další použité zdroje. Pro zpracování simulovaných i reálných časových řad jsme použili software R. Na příloženém CD jsou umístěny zdrojové kódy a zpracovaná data.

Kapitola 1

Lineární modely pro jednorozměrné časové řady

1.1 Základní pojmy

Nejprve definujeme základní pojmy týkající se problematiky časových řad.

Definice 1. *Nechť (Ω, \mathcal{F}, P) je pravděpodobnostní prostor, nechť $T \subset \mathbb{R}$. Množinu reálných náhodných veličin $\{Z(\omega, t), \omega \in \Omega, t \in T\}$ definovaných na (Ω, \mathcal{F}, P) nazveme náhodný proces.*

Hodnoty $t \in T$ interpretujeme jako čas. Pro pevné $t \in T$ je $Z(\omega, t)$ náhodná veličina. Naopak pro dané $\omega \in \Omega$ odpovídá $Z(\omega, t)$ jako deterministická funkce proměnné t trajektorii náhodného procesu, která je pozorovatelná v praxi.

Definice 2. *Náhodný proces $\{Z(\omega, t), \omega \in \Omega, t \in T\}$, kde $T \subset \mathbb{Z}$, nazveme časovou řadou.*

Dále budeme místo $Z(\omega, t)$ psát pouze Z_t .

Definice 3. *Časovou řadu $\{Z_t, t \in T\}$ nazveme silně stacionární, pokud je pravděpodobnostní rozdělení daného náhodného procesu invariantní vůči posunům v čase, tedy pro libovolné $n \in \mathbb{N}$, pro všechna $t_1, \dots, t_n \in T$ a pro libovolné $k \in \mathbb{Z}$ takové, že $t_i + k \in T$ pro $1 \leq i \leq n$, platí*

$$\mathcal{L}(Z_{t_1}, \dots, Z_{t_n}) = \mathcal{L}(Z_{t_1+k}, \dots, Z_{t_n+k}).$$

Definice 4. *Časovou řadu $\{Z_t, t \in T\}$ s konečnými druhými momenty nazveme slabě stacionární, pokud pro všechna $s, t \in T$ a libovolné $k \in \mathbb{Z}$ takové, že $s+k, t+k \in T$, platí*

$$\begin{aligned} E(Z_t) &= \mu = \text{konst}, \\ \text{cov}(Z_s, Z_t) &= \text{cov}(Z_{s+k}, Z_{t+k}). \end{aligned}$$

Stacionarita znamená určitou stochastickou ustálenost chování časové řady. Striktní stacionarita požaduje invarianci pravděpodobnostního rozdělení daného náhodného procesu vůči posunům v čase; pro slabou stacionaritu stačí, aby byly vůči posunům v čase invariantní pouze první dva momenty. Z definice slabě stacionární časové řady také vyplývá, že $\text{var}(Z_t) = \sigma_Z^2 = \text{konst}$. Slabou stacionaritu budeme dále nazývat pouze stacionaritou.

V praxi je obvyklá silná korelovanost hodnot časové řady. Popisuje se pomocí autokovarianční a autokorelační funkce, které nyní definujeme pro stacionární časové řady.

Definice 5. Autokovarianční funkci pro zpoždění k definujeme jako

$$\gamma_k = \text{cov}(Z_t, Z_{t+k}) = \mathbf{E}(Z_t - \mu)(Z_{t+k} - \mu), \quad k = \dots, -1, 0, 1, \dots$$

Definice 6. Autokorelační funkci (ACF) pro zpoždění k definujeme jako

$$\rho_k = \frac{\text{cov}(Z_t, Z_{t+k})}{\sqrt{\text{var}Z_t}\sqrt{\text{var}Z_{t+k}}} = \frac{\gamma_k}{\gamma_0} = \frac{\gamma_k}{\sigma_Z^2}, \quad k = \dots, -1, 0, 1, \dots$$

Autokovarianční a autokorelační funkce stacionárního procesu mají následující vlastnosti:

- $\gamma_0 = \text{var} Z_t, \rho_0 = 1,$
- $|\gamma_k| \leq \gamma_0, |\rho_k| \leq 1,$
- $\gamma_k = \gamma_{-k}, \rho_k = \rho_{-k},$ tedy γ_k a ρ_k jsou sudé funkce.

Vzhledem k poslední vlastnosti se graficky zobrazují hodnoty autokorelační funkce pouze pro nezáporná zpoždění k . Jejich graf se nazývá *korelogram*.

Korelace mezi náhodnými veličinami Z_t a Z_{t+k} pro $k > 1$ je často ovlivněna veličinami $Z_{t+1}, Z_{t+2}, \dots, Z_{t+k-1}$. Proto nás obvykle zajímá také parciální korelace veličin Z_t, Z_{t+k} očištěná o vliv veličin ležících mezi nimi.

Definice 7. Parciální autokorelační funkci (PACF) pro zpoždění k definujeme jako parciální korelační koeficient ϕ_{kk} mezi Z_t a Z_{t+k} při pevných hodnotách $Z_{t+1}, \dots, Z_{t+k-1}$.

Poznámka. Zřejmě platí $\phi_{00} = 1$ a $\phi_{11} = \rho_1$.

Dále zavedeme vytvořující funkci autokovarianční a autokorelační funkce dle [20], str. 25.

Definice 8. Pro danou autokovarianční funkci $\gamma_k, k = 0, \pm 1, \pm 2, \dots$ definujeme vytvořující funkci autokovarianční funkce jako

$$\gamma(B) = \sum_{k=-\infty}^{\infty} \gamma_k B^k. \quad (1.1)$$

Definice 9. Pro danou autokorelační funkci $\rho_k, k = 0, \pm 1, \pm 2, \dots$ definujeme vytvořující funkci autokorelační funkce jako

$$\rho(B) = \sum_{k=-\infty}^{\infty} \rho_k B^k = \frac{\gamma(B)}{\gamma_0}. \quad (1.2)$$

V praxi máme k dispozici jednu pozorovanou trajektorii časové řady Z_1, \dots, Z_n . Z ní lze odhadnout střední hodnotu, autokovarianční a autokorelační funkci následujícím způsobem.

Definice 10. Pro stacionární časovou řadu definujeme odhad střední hodnoty jako

$$\hat{\mu} = \bar{Z} = \frac{1}{n} \sum_{t=1}^n Z_t.$$

Definice 11. Pro stacionární časovou řadu definujeme odhad autokovarianční funkce jako

$$\hat{\gamma}_k = c_k = \frac{1}{n} \sum_{t=1}^{n-k} (Z_t - \bar{Z}) (Z_{t+k} - \bar{Z}), \quad k = 0, 1, \dots, n-1.$$

Definice 12. Odhad autokorelační funkce definujeme jako

$$\hat{\rho}_k = r_k = \frac{c_k}{c_0} = \frac{\sum_{t=1}^{n-k} (Z_t - \bar{Z}) (Z_{t+k} - \bar{Z})}{\sum_{t=1}^n (Z_t - \bar{Z})^2}, \quad k = 0, 1, \dots, n-1.$$

Odhad parciální autokorelace ϕ_{kk} lze podle [4], str. 331 získat jako odhad parametru φ_{kk} v regresním modelu

$$Z_{t+k} = \delta + \varphi_{k1} Z_{t+k-1} + \varphi_{k2} Z_{t+k-2} + \dots + \varphi_{kk} Z_t + \varepsilon_{t+k},$$

nicméně v praxi se obvykle používá odhad dle následující definice.

Definice 13. Odhad parciální autokorelační funkce definujeme rekurentně jako

$$\begin{aligned} \hat{\phi}_{11} &= r_{11} = r_1, \\ \hat{\phi}_{kk} &= r_{kk} = \frac{r_k - \sum_{j=1}^{k-1} r_{k-1,j} \cdot r_{k-j}}{1 - \sum_{j=1}^{k-1} r_{k-1,j} \cdot r_j}, \quad k > 1, \end{aligned}$$

kde

$$r_{kj} = r_{k-1,j} - r_{kk} \cdot r_{k-1,k-j}, \quad j = 1, \dots, k-1.$$

1.2 Lineární proces

Definice 14. Časovou řadu $\{\varepsilon_t\}$ nazveme bílý šum, pokud $\{\varepsilon_t\}$ je posloupnost náhodných veličin splňující $E(\varepsilon_t) = 0$, $\text{var}(\varepsilon_t) = \sigma_\varepsilon^2 > 0$ a $\gamma_k = \text{cov}(\varepsilon_t, \varepsilon_{t+k}) = 0$ pro všechna $k \neq 0$.

Snadno nahlédneme, že bílý šum má autokorelační funkci

$$\rho_k = \begin{cases} 1 & \text{pro } k = 0, \\ 0 & \text{pro } k \neq 0 \end{cases}$$

a parciální autokorelační funkci

$$\phi_{kk} = \begin{cases} 1 & \text{pro } k = 0, \\ 0 & \text{pro } k \neq 0. \end{cases}$$

Jelikož ale $\rho_0 = \phi_{00} = 1$ platí pro libovolnou časovou řadu, zajímá nás jejich chování pouze pro $k \neq 0$. Proto obvykle říkáme, že autokorelační a parciální autokorelační funkce bílého šumu jsou identicky rovny nule.

Definice 15. *Lineární proces definujeme jako nekonečnou časovou řadu*

$$Z_t = \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \dots = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}, \quad (1.3)$$

kde ε_t je bílý šum a $\psi_j \in \mathbb{R}$ jsou parametry.

Lineární proces tedy představuje lineární kombinaci členů posloupnosti bílého šumu, ve které $\psi_0 = 1$. Jiné vyjádření je možné při zavedení tzv. *operátoru časového posunu* B , který zpozdí veličinu o jedno časové období, tedy

$$BZ_t = Z_{t-1},$$

a jeho j -tá mocnina zpozdí veličinu o j časových období:

$$B^j Z_t = B^{j-1} (BZ_t) = B^{j-1} Z_{t-1} = \dots = Z_{t-j}.$$

Lineární proces pak lze psát jako

$$Z_t = (1 + \psi_1 B + \psi_2 B^2 + \dots) \varepsilon_t = \Psi(B) \varepsilon_t,$$

kde $\Psi(B) = \sum_{j=0}^{\infty} \psi_j B^j$, $\psi_0 = 1$.

V mocninné řadě $\Psi(B)$ se operátor B chová jako proměnná z v klasické mocninné řadě $\Psi(z)$. Podmínka

$$\Psi(z) \text{ konverguje pro } |z| \leq 1,$$

tedy na jednotkovém kruhu v komplexní rovině, je postačující pro existenci a stacionaritu lineárního procesu. Potom má proces střední hodnotu $E(Z_t) = 0$ a rozptyl $\text{var}(Z_t) = \sigma_\varepsilon^2 \sum_{j=0}^{\infty} \psi_j^2$. Jeho autokovarianční funkce je

$$\gamma_k = E(Z_t Z_{t+k}) = E\left(\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \psi_i \psi_j \varepsilon_{t-i} \varepsilon_{t+k-j}\right) = \sigma_\varepsilon^2 \sum_{i=0}^{\infty} \psi_i \psi_{i+k},$$

kde jsme v poslední rovnosti využili, že náhodné veličiny ε_t a ε_{t+k} jsou nekorelované pro $k \neq 0$. Autokorelační funkce má tvar

$$\rho_k = \frac{\sum_{i=0}^{\infty} \psi_i \psi_{i+k}}{\sum_{i=0}^{\infty} \psi_i^2}.$$

Dle (1.1) lze psát vytvořující funkci autokovarianční funkce jako

$$\begin{aligned} \gamma(B) &= \sigma_\varepsilon^2 \sum_{k=-\infty}^{\infty} \sum_{i=0}^{\infty} \psi_i \psi_{i+k} B^k \stackrel{j=i+k}{=} \sigma_\varepsilon^2 \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \psi_i \psi_j B^{j-i} = \\ &= \sigma_\varepsilon^2 \sum_{j=0}^{\infty} \psi_j B^j \sum_{i=0}^{\infty} \psi_i B^{-i} = \sigma_\varepsilon^2 \psi(B) \psi(B^{-1}), \end{aligned} \quad (1.4)$$

kde jsme ve druhé rovnosti využili, že $\psi_j = 0$ pro $j < 0$.

V případě, že je lineární proces invertibilní, lze ho zapsat ve tvaru

$$Z_t = \pi_1 Z_{t-1} + \pi_2 Z_{t-2} + \dots + \varepsilon_t, \quad (1.5)$$

$$\text{tj. } \varepsilon_t = Z_t - \pi_1 Z_{t-1} - \dots = \Pi(B) Z_t,$$

kde $\Pi(B) = 1 - \sum_{j=1}^{\infty} \pi_j B^j$. Toto vyjádření je možné při splnění postačující podmínky

$$\Pi(z) \text{ konverguje pro } |z| \leq 1,$$

tedy na jednotkovém kruhu v komplexní rovině.

1.3 Model klouzavých součtů

Definice 16. Řekneme, že časová řada $\{Z_t\}$ se řídí modelem klouzavých součtů řádu q , pokud platí

$$Z_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} = \Theta(B) \varepsilon_t,$$

kde ε_t je bílý šum, $\theta_1, \dots, \theta_q \in \mathbb{R}$ jsou parametry ($\theta_q \neq 0$) a

$$\Theta(B) = 1 + \theta_1 B + \dots + \theta_q B^q$$

je operátor klouzavých součtů.

Model klouzavých součtů řádu q značíme $MA(q)$ a dostaneme ho z lineárního procesu (1.3) vynulováním jeho parametrů kromě prvních q z nich. Model $MA(q)$ je vždy stacionární s nulovou střední hodnotou a za podmínky, že všechny kořeny polynomu

$$\Theta(z) = 1 + \theta_1 z + \dots + \theta_q z^q$$

leží vně jednotkového kruhu v komplexní rovině, také invertibilní. Jeho rozptyl je

$$\gamma_0 = \sigma_\varepsilon^2 \sum_{j=0}^q \theta_j^2,$$

kde $\theta_0 = 1$.

Autokovarianční funkce má tvar

$$\gamma_k = \begin{cases} (\theta_k + \theta_1 \theta_{k+1} + \dots + \theta_{q-k} \theta_q) \sigma_\varepsilon^2 & \text{pro } k = 1, \dots, q, \\ 0 & \text{pro } k > q \end{cases}$$

a snadno získáme i autokorelační funkci

$$\rho_k = \begin{cases} \frac{\theta_k + \theta_1 \theta_{k+1} + \dots + \theta_{q-k} \theta_q}{1 + \theta_1^2 + \dots + \theta_q^2} & \text{pro } k = 1, \dots, q, \\ 0 & \text{pro } k > q. \end{cases} \quad (1.6)$$

Říkáme, že autokorelační funkce modelu klouzavých součtů má bod useknutí k_0 rovný řádu q . Jak uvádí [4] na str. 334, parciální autokorelační funkce modelu $MA(q)$ je omezena lineární kombinací geometricky klesajících posloupností a sinusoid s geometricky klesajícími amplitudami a nemá bod useknutí.

1.4 Autoregresní model

Definice 17. Řekneme, že časová řada $\{Z_t\}$ se řídí autoregresním modelem řádu p , pokud platí

$$Z_t = \varphi_1 Z_{t-1} + \dots + \varphi_p Z_{t-p} + \varepsilon_t, \quad (1.7)$$

$$\text{tj. } Z_t - \varphi_1 Z_{t-1} - \dots - \varphi_p Z_{t-p} = \Phi(B) Z_t = \varepsilon_t, \quad (1.8)$$

kde ε_t je bílý šum, $\varphi_1, \dots, \varphi_p \in \mathbb{R}$ jsou parametry ($\varphi_p \neq 0$) a

$$\Phi(B) = 1 - \varphi_1 B - \dots - \varphi_p B^p$$

je autoregresní operátor.

Autoregresní model řádu p značíme $\text{AR}(p)$ a dostaneme ho z invertovaného zápisu lineárního procesu (1.5) vynulováním jeho parametrů kromě prvních p z nich. Model $\text{AR}(p)$ je tedy vždy invertibilní a za postačující podmínky, že všechny kořeny polynomu

$$\Phi(z) = 1 - \varphi_1 z - \dots - \varphi_p z^p$$

leží vně jednotkového kruhu v komplexní rovině, také stacionární s nulovou střední hodnotou a rozptylem

$$\gamma_0 = \frac{\sigma_\varepsilon^2}{1 - \varphi_1 \rho_1 - \dots - \varphi_p \rho_p}.$$

Vynásobíme-li rovnici (1.7) hodnotou Z_{t-k} pro $k > 0$, můžeme napsat její střední hodnotu jako

$$\mathbf{E}(Z_t \cdot Z_{t-k}) = \varphi_1 \mathbf{E}(Z_{t-1} \cdot Z_{t-k}) + \dots + \varphi_p \mathbf{E}(Z_{t-p} \cdot Z_{t-k}) + \underbrace{\mathbf{E}(\varepsilon_t \cdot Z_{t-k})}_0,$$

kde poslední střední hodnota je nulová, jelikož z (1.8) je $Z_{t-k} = \Phi^{-1}(B) \varepsilon_{t-k}$ lineární proces vytvořený z hodnot $\varepsilon_{t-k}, \varepsilon_{t-k-1}, \dots$ a náhodné veličiny ε_t a Z_{t-k} jsou tedy pro $k > 0$ nekorelované. Dále lze psát

$$\gamma_k = \varphi_1 \gamma_{k-1} + \dots + \varphi_p \gamma_{k-p},$$

odkud po vydělení rozptylem γ_0 dostaneme diferenční rovnici pro autokorelační funkci modelu $\text{AR}(p)$

$$\rho_k = \varphi_1 \rho_{k-1} + \dots + \varphi_p \rho_{k-p}. \quad (1.9)$$

Zapsáním rovnice (1.9) pro $k = 1, \dots, p$ dostaneme soustavu tzv. *Yule-Walkerových rovnic*, pomocí které lze vypočítat parametry $\varphi_1, \dots, \varphi_p$:

$$\begin{aligned} \rho_1 &= \varphi_1 & + \varphi_2 \rho_1 & + \dots + \varphi_p \rho_{p-1}, \\ \rho_2 &= \varphi_1 \rho_1 & + \varphi_2 & + \dots + \varphi_p \rho_{p-2}, \\ &\vdots \\ \rho_p &= \varphi_1 \rho_{p-1} + \varphi_2 \rho_{p-2} + \dots + \varphi_p. \end{aligned}$$

Autokorelační funkce modelu $\text{AR}(p)$ je lineární kombinace klesajících geometrických posloupností a sinusoid různých frekvencí s geometricky klesajícími amplitudami, viz např. [4], str. 335. Parciální autokorelační funkce má bod useknutí k_0 roven řádu modelu p .

1.5 Smíšený model ARMA

V praxi se stává, že použití AR nebo MA modelu vyžaduje pro adekvátní popis dané časové řady použití modelu vysokého řádu. Proto je užitečné do modelu zahrnout autoregresní členy i členy modelu klouzavých součtů zároveň, čímž vznikne tzv. smíšený model ARMA. Potřebný počet parametrů se v takovém případě často sníží.

Definice 18. Řekneme, že časová řada $\{Z_t\}$ se řídí modelem ARMA řádů p a q , pokud platí

$$Z_t = \varphi_1 Z_{t-1} + \dots + \varphi_p Z_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}, \quad (1.10)$$

$$\text{tj. } \Phi(B) Z_t = \Theta(B) \varepsilon_t, \quad (1.11)$$

kde ε_t je bílý šum, $\varphi_1, \dots, \varphi_p, \theta_1, \dots, \theta_q \in \mathbb{R}$ jsou parametry ($\varphi_p \neq 0, \theta_q \neq 0$),

$$\Phi(B) = 1 - \varphi_1 B - \dots - \varphi_p B^p$$

je autoregresní operátor a

$$\Theta(B) = 1 + \theta_1 B + \dots + \theta_q B^q$$

je operátor klouzavých součtů.

Postačující podmínka stacionarity modelu ARMA(p, q) odpovídá podmínce stacionarity modelu AR(p), tedy že všechny kořeny polynomu

$$\Phi(z) = 1 - \varphi_1 z - \dots - \varphi_p z^p$$

leží vně jednotkového kruhu v komplexní rovině. Stejně tak podmínka invertibility modelu ARMA(p, q) odpovídá podmínce invertibility modelu MA(q), tedy že všechny kořeny polynomu

$$\Theta(z) = 1 + \theta_1 z + \dots + \theta_q z^q$$

leží vně jednotkového kruhu v komplexní rovině.

Přenásobením rovnice (1.10) členem Z_{t-k} a přechodem ke středním hodnotám dostaneme

$$\begin{aligned} E(Z_t \cdot Z_{t-k}) &= \varphi_1 E(Z_{t-1} \cdot Z_{t-k}) + \dots + \varphi_p E(Z_{t-p} \cdot Z_{t-k}) + \\ &+ E(\varepsilon_t \cdot Z_{t-k}) + \theta_1 E(\varepsilon_{t-1} \cdot Z_{t-k}) + \dots + \theta_q E(\varepsilon_{t-q} \cdot Z_{t-k}). \end{aligned}$$

Dále využijeme, že platí

$$E(\varepsilon_{t-i} \cdot Z_{t-k}) = 0 \quad \text{pro } k > i,$$

neboť z (1.11) je

$$Z_{t-k} = \Phi^{-1}(B) \Theta(B) \varepsilon_{t-k}$$

lineární proces vytvořený z hodnot $\varepsilon_{t-k}, \varepsilon_{t-k-1}, \dots$ a náhodné veličiny ε_{t-i} a Z_{t-k} jsou pro $k > i$ nekorelované. Lze tedy psát

$$\gamma_k = \varphi_1 \gamma_{k-1} + \dots + \varphi_p \gamma_{k-p} \quad \text{pro } k > p$$

a po vydělení rozptylem γ_0 dostaneme diferenční rovnici pro autokovarianční funkci modelu ARMA(p, q)

$$\rho_k = \varphi_1 \rho_{k-1} + \dots + \varphi_p \rho_{k-p} \quad \text{pro } k > q.$$

Ta nemá bod useknutí a dle [4], str. 337 tvoří lineární kombinaci klesajících geometrických posloupností a sinusoid různých frekvencí s geometricky klesajícími amplitudami kromě hodnot $\rho_0, \rho_1, \dots, \rho_{q-p}$ (v případě, že $p \leq q$).

Parciální autokovarianční funkce dle téhož zdroje taktéž nemá bod useknutí a je omezena lineární kombinací klesajících geometrických posloupností a sinusoid různých frekvencí s geometricky klesajícími amplitudami kromě hodnot $\rho_{00}, \rho_{11}, \dots, \rho_{p-q, p-q}$ (v případě, že $p \geq q$).

Poznámka. Uvedené stacionární modely lze zobecnit i pro nenulovou (v čase neměnnou) střední hodnotu $E Z_t = \mu$. Například model ARMA(p, q) se střední hodnotou μ pak má tvar

$$Z_t - \mu = \varphi_1 (Z_{t-1} - \mu) + \dots + \varphi_p (Z_{t-p} - \mu) + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q},$$

neboli

$$Z_t = \alpha + \varphi_1 Z_{t-1} + \dots + \varphi_p Z_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}, \quad (1.12)$$

kde $\alpha = \mu \cdot (1 - \varphi_1 - \dots - \varphi_p)$.

1.6 Výstavba modelu

V praxi se obvykle pro pozorovanou časovou řadu provádí konstrukce modelu v následujících třech krocích:

1. identifikace modelu – pro pozorovanou časovou řadu hledáme typ modelu (AR, MA, ARMA) a jeho řád, například MA(2);
2. odhad parametrů modelu – například v modelu MA(2) odhadujeme parametry θ_1, θ_2 ;
3. diagnostika modelu – ověřujeme shodu identifikovaného a odhadnutého modelu s daty a případně jeho předpovědní schopnosti.

Metody odhadu parametrů v ARMA modelu jsou popsány v knize [20], kap. 7. Jedná se například o metodu nejmenších čtverců, metodu maximální věrohodnosti či momentovou metodu. Pro jejich praktickou aplikaci se používají vhodné softwarové produkty. Stejný zdroj uvádí i různé diagnostické nástroje, pomocí kterých se ověřuje vhodnost modelu pomocí posouzení splnění jeho předpokladů. Kontrolujeme zejména nulovost střední hodnoty vypočteného bílého šumu, jeho konstantní rozptyl, nekorelovanost a normalitu. V další části práce se budeme podrobněji věnovat identifikačnímu kroku výstavby modelu.

Kapitola 2

Identifikace modelů jednorozměrných časových řad

Při konstrukci modelu časové řady je klíčová volba správného typu modelu a jeho řádu. Prvním krokem je vždy grafická prohlídka analyzovaných dat a zjištění, zda řada v čase nemění svůj charakter a má konstantní úroveň. V případě, že střední hodnota není nulová, ale v čase se nemění, lze řadu nejprve centrovat nebo zvolit model s konstantní střední hodnotou, např. (1.12). Pokud je řada nestacionární, obvykle se provádí vhodná transformace, a to zejména diferencování a transformace homogenizující rozptyl, například *Boxova-Coxova transformace*.

2.1 Boxova-Coxova transformace

Tvar Boxovy-Coxovy transformace je

$$Z_t^{(\lambda)} = \begin{cases} \frac{(Z_t + c)^\lambda - 1}{\lambda} & \text{pro } \lambda \neq 0, \\ \ln(Z_t + c) & \text{pro } \lambda = 0, \end{cases}$$

kde $c > 0$ je úroňový parametr volený tak, aby $Z_t + c > 0$, a $\lambda \in \mathbb{R}$ je typový parametr klíčový pro tvar transformace. Volbu logaritmické transformace pro $\lambda = 0$ dostaneme z toho, že

$$\lim_{\lambda \rightarrow 0} \frac{(Z_t + c)^\lambda - 1}{\lambda} = \ln(Z_t + c).$$

Způsob odhadu parametru λ je popsán v [20], str. 84, ale v praxi se obvykle stanovuje následujícím postupem (viz [4], str. 317):

1. řada z_1, \dots, z_n se rozdělí do krátkých segmentů stejné délky,
2. v j -tém segmentu se vypočte výběrový průměr $\bar{z}(j)$ a výběrová směrodatná odchylka $s_z(j)$,
3. graficky se znázorní body $(\bar{z}(j), s_z(j))$,
4. body se přibližně proloží hladkou křivkou a podle jejího tvaru se zvolí λ dle tabulky 2.1.

Kromě čtyř možností volby typového parametru uvedených v tabulce 2.1 připouští zdroj [20] také volbu $\lambda = -1$ s transformací typu $1/Z_t$.

Tvar křivky	Volba λ	Typ transformace
konstanta	1	Z_t
konkávní	$\frac{1}{2}$	$\sqrt{Z_t}$
přímka	0	$\ln Z_t$
konvexní	$-\frac{1}{2}$	$1/\sqrt{Z_t}$

Tabulka 2.1: Volba typového parametru λ v Boxově-Coxově transformaci

Boxova-Coxova transformace nejen homogenizuje rozptyl, aby byl přibližně konstantní v čase, ale také sešikmené rozdělení řady přibližuje k symetrickému, často dokonce normálnímu rozdělení, které je užitečné pro konstrukci předpovědních intervalů a linearizuje model časové řady. Boxova-Coxova transformace se doporučuje provést ještě před případným diferencováním.

2.2 Transformace diferencováním

Před vlastní volbou modelu je vhodné ověřit nutný řád diferencování analyzované řady. Vykreslíme odhadnutý korelogram a parciální korelogram původní řady (po případné Boxově-Coxově transformaci) a rozhodneme o potřebě a řádu diferencování podle následujících pravidel (dle [20], str. 105):

1. V případě, že odhadnutá ACF velmi pomalu klesá a odhadnutá PACF má bod useknutí $k_0 = 1$, provedeme diferencování. Diference prvního řádu dostaneme jako $\Delta Z_t = (1 - B) Z_t = Z_t - Z_{t-1}$. Jde o velmi subjektivní metodu, kterou lze nahradit provedením statistického testu na jednotkový kořen, např. *Dickeyova-Fullerova testu*, viz dále.
2. V některých případech je potřeba vyšší řád diferencování $(1 - B)^d Z_t$ pro $d > 1$. Řád d obvykle bývá 0, 1 nebo 2. Diferencování se doporučuje volit co nejjednodušší, aby nedošlo k tzv. přediferencování, které vede k rostoucímu rozptylu transformované řady, viz [4], str. 320.

Možnost stacionarizace časové řady pomocí diferencování souvisí s tím, že příslušný autoregresní operátor modelu má (téměř) jednotkový kořen. Podle prohlídky odhadnutého korelogramu však nelze vždy jednoznačně rozlišit nestacionární model náhodné procházky s jednotkovým kořenem $Z_t = Z_{t-1} + \varepsilon_t$ např. od stacionárního AR(1) modelu $Z_t = 0.95Z_{t-1} + \varepsilon_t$ s téměř jednotkovým kořenem. Proto je vhodné kromě prohlídky ACF a PACF použít také statistický test na jednotkový kořen.

2.2.1 Dickeyův-Fullerův test

Uvažujeme model AR(1)

$$Z_t = \varphi_1 Z_{t-1} + \varepsilon_t, \quad t = 1, 2, \dots$$

kde $Z_0 = 0$. Pokud $|\varphi_1| = 1$, existuje jednotkový kořen autoregresního polynomu a proces je nestacionární. Dickey a Fuller navrhli v člancích [5] a [6] tři varianty testu na jednotkový kořen, souhrnně označované jako τ -testy:

$$\begin{aligned}
\tau\text{-test: } H_0: Z_t &= Z_{t-1} + \varepsilon_t \quad \text{proti } H_1: Z_t = \varphi_1 Z_{t-1} + \varepsilon_t, & \varphi_1 < 1, \\
\tau_\mu\text{-test: } H_0: Z_t &= Z_{t-1} + \varepsilon_t \quad \text{proti } H_1: Z_t = \mu + \varphi_1 Z_{t-1} + \varepsilon_t, & \varphi_1 < 1, \\
\tau_\tau\text{-test: } H_0: Z_t &= Z_{t-1} + \varepsilon_t \quad \text{proti } H_1: Z_t = \mu + \beta t + \varphi_1 Z_{t-1} + \varepsilon_t, & \varphi_1 < 1.
\end{aligned}$$

Jde o jednostranné testy, ve kterých v alternativě uvažujeme $\varphi_1 < 1$. V nulové hypotéze vystupuje model náhodné procházky a v alternativě postupně stacionární model AR(1), stacionární model AR(1) s nenulovou konstantní střední hodnotou $\mu_Z = \frac{\mu}{1-\varphi_1}$ a stacionární model AR(1) s lineárním trendem $\mu + \beta t$. Nulová hypotéza lze pro všechny tři varianty testu napsat jako

$$H_0: \Delta Z_t = \delta Z_{t-1} + \varepsilon_t, \quad \text{kde } \delta = 0 \quad (2.1)$$

a alternativa jako

$$H_1: \Delta Z_t = \mu + \beta t + \delta Z_{t-1} + \varepsilon_t, \quad \text{kde } \delta < 0,$$

přičemž $\delta = \varphi_1 - 1$, pro τ -test $\mu = \beta = 0$ a pro τ_μ -test $\beta = 0$. U τ_μ -testu a τ_τ -testu nás zajímá pouze to, zda $\delta < 0$. Na významnosti parametrů μ a β nezáleží.

Jelikož tedy testujeme významnost parametru δ v regresním modelu (2.1), testová statistika pro všechny tři verze testu odpovídá testové statistice t -testu

$$DF = \frac{\hat{\delta}}{\hat{\sigma}(\hat{\delta})}, \quad (2.2)$$

kde $\hat{\delta}$ značí odhad parametru δ metodou nejmenších čtverců a $\hat{\sigma}(\hat{\delta})$ je odhad směrodatné odchylky odhadu parametru δ .

Způsob výpočtu těchto odhadů pro pozorování Z_1, \dots, Z_n se pro jednotlivé varianty testu liší. Pro τ -test uvažujeme regresní model

$$\begin{pmatrix} \Delta Z_2 \\ \vdots \\ \Delta Z_n \end{pmatrix} = \begin{pmatrix} Z_1 \\ \vdots \\ Z_{n-1} \end{pmatrix} \cdot \delta + \begin{pmatrix} \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix},$$

ve kterém odhad parametru δ dostaneme jako

$$\hat{\delta} = \frac{\sum_{t=2}^n (\Delta Z_t) Z_{t-1}}{\sum_{t=2}^n Z_{t-1}^2}$$

a odhad jeho směrodatné odchylky jako

$$\hat{\sigma}(\hat{\delta}) = \sqrt{\frac{\sum_{t=2}^n (\Delta Z_t - \hat{\delta} Z_{t-1})^2}{(n-2) \sum_{t=2}^n Z_{t-1}^2}}.$$

V případě τ_μ -testu pracujeme s regresním modelem

$$\begin{pmatrix} \Delta Z_2 \\ \vdots \\ \Delta Z_n \end{pmatrix} = \begin{pmatrix} 1 & Z_1 \\ \vdots & \vdots \\ 1 & Z_{n-1} \end{pmatrix} \cdot \begin{pmatrix} \mu \\ \delta \end{pmatrix} + \begin{pmatrix} \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix},$$

pro který vycházejí odhady parametrů

$$\hat{\mu} = \frac{\bar{z}_{(0)} \sum_{t=2}^n Z_{t-1}^2 - \bar{z}_{(-1)} \sum_{t=2}^n (\Delta Z_t) Z_{t-1}}{\sum_{t=2}^n Z_{t-1}^2 - \bar{z}_{(-1)} \sum_{t=2}^n Z_{t-1}},$$

$$\hat{\delta} = \frac{\sum_{t=2}^n (\Delta Z_t - \bar{z}_{(0)}) (Z_{t-1} - \bar{z}_{(-1)})}{\sum_{t=2}^n (Z_{t-1} - \bar{z}_{(-1)})^2},$$

kde

$$\bar{z}_{(0)} = \frac{\sum_{t=2}^n \Delta Z_t}{n-1}, \quad \bar{z}_{(-1)} = \frac{\sum_{t=2}^n Z_{t-1}}{n-1}$$

a odhad směrodatné odchylky parametru δ potřebný pro testovou statistiku (2.2)

$$\hat{\sigma}(\hat{\delta}) = \sqrt{\frac{\sum_{t=2}^n (\Delta Z_t - \hat{\mu} - \hat{\delta} Z_{t-1})^2}{(n-3) \sum_{t=2}^n (Z_{t-1} - \bar{z}_{(-1)})^2}}.$$

Odvození testové statistiky τ_τ -testu probíhá obdobně s využitím regresního modelu

$$\begin{pmatrix} \Delta Z_2 \\ \vdots \\ \Delta Z_n \end{pmatrix} = \begin{pmatrix} 1 & 2 & Z_1 \\ \vdots & \vdots & \vdots \\ 1 & n & Z_{n-1} \end{pmatrix} \cdot \begin{pmatrix} \mu \\ \beta \\ \delta \end{pmatrix} + \begin{pmatrix} \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}.$$

Testová statistika DF za platnosti nulové hypotézy nemá t -rozdělení jako v případě klasického t -testu, ale nestandardní rozdělení, pro které bylo nutné kvantily $\hat{\tau}_\alpha(n)$ napočítat simulačně pro různé délky řad n . Kvantily pro jednotlivé typy testu jsou uvedeny v tabulkách 2.2, 2.3 a 2.4 převzatých z [7], str. 642. Kritický obor má tvar

$$DF \leq \hat{\tau}_\alpha(n).$$

2.2.2 Rozšířený Dickeyův-Fullerův test

Dickeyův-Fullerův test je použitelný jen pro model AR(1), jehož reziduální složka ε_t je nezávislý bílý šum. Pokud řád autokorelace řady Z_t je větší než 1, proměnná ΔZ_t v modelu (2.1) stále obsahuje autokorelovanost. Potom test vykazuje pravděpodobnost chyby prvního druhu větší, než je zvolená hladina α . Tento problém řeší tzv. *rozšířený Dickeyův-Fullerův test (ADF)* zmíněný v [4], str. 355. Nulová hypotéza má tvar

$$H_0: \Delta Z_t = \delta_0 Z_{t-1} + \delta_1 \Delta Z_{t-1} + \dots + \delta_p \Delta Z_{t-p} + \varepsilon_t, \quad \text{kde } \delta_0 = 0$$

a testová statistika i kritické obory pro test nulovosti parametru δ_0 jsou stejné jako u předchozího Dickeyova-Fullerova testu. Řád p autoregrese je třeba stanovit před použitím testu. K jeho volbě se doporučuje využít informační kritéria popsána v kapitole 2.5.

Délka řady n	α								
	0,01	0,025	0,05	0,10	0,50	0,90	0,95	0,975	0,99
25	-2,65	-2,26	-1,95	-1,60	-0,47	0,92	1,33	1,70	2,15
50	-2,62	-2,25	-1,95	-1,61	-0,49	0,91	1,31	1,66	2,08
100	-2,60	-2,24	-1,95	-1,61	-0,50	0,90	1,29	1,64	2,04
250	-2,58	-2,24	-1,95	-1,62	-0,50	0,89	1,28	1,63	2,02
500	-2,58	-2,23	-1,95	-1,62	-0,50	0,89	1,28	1,62	2,01
∞	-2,58	-2,23	-1,95	-1,62	-0,51	0,89	1,28	1,62	2,01

Tabulka 2.2: Kvantily testové statistiky τ -testu

Délka řady n	α								
	0,01	0,025	0,05	0,10	0,50	0,90	0,95	0,975	0,99
25	-3,75	-3,33	-2,99	-2,64	-1,53	-0,37	0,00	0,34	0,71
50	-3,59	-3,23	-2,93	-2,60	-1,55	-0,41	-0,04	0,28	0,66
100	-3,50	-3,17	-2,90	-2,59	-1,56	-0,42	-0,06	0,26	0,63
250	-3,45	-3,14	-2,88	-2,58	-1,56	-0,42	-0,07	0,24	0,62
500	-3,44	-3,13	-2,87	-2,57	-1,57	-0,44	-0,07	0,24	0,61
∞	-3,42	-3,12	-2,86	-2,57	-1,57	-0,44	-0,08	0,23	0,60

Tabulka 2.3: Kvantily testové statistiky τ_μ -testu

Délka řady n	α								
	0,01	0,025	0,05	0,10	0,50	0,90	0,95	0,975	0,99
25	-4,38	-3,95	-3,60	-3,24	-2,14	-1,14	-0,81	-0,50	-0,15
50	-4,16	-3,80	-3,50	-3,18	-2,16	-1,19	-0,87	-0,58	-0,24
100	-4,05	-3,73	-3,45	-3,15	-2,17	-1,22	-0,90	-0,62	-0,28
250	-3,98	-3,69	-3,42	-3,13	-2,18	-1,23	-0,92	-0,64	-0,31
500	-3,97	-3,67	-3,42	-3,13	-2,18	-1,24	-0,93	-0,65	-0,32
∞	-3,96	-3,67	-3,41	-3,13	-2,18	-1,25	-0,94	-0,66	-0,32

Tabulka 2.4: Kvantily testové statistiky τ_τ -testu

Poznámka. Nevýhodou Dickeyova-Fullerova testu je, že má někdy slabou rozlišovací schopnost. Proto se často provádí současně s *KPSS-testem* (viz [11]), jehož nulová hypotéza a alternativa jsou opačné, než u rozšířeného Dickeyova-Fullerova testu. Nulovou hypotézou je tedy stacionarita testované řady. Podle výsledku obou testů se pak rozhodne takto:

- pokud H_0^{ADF} se zamítá a H_0^{KPSS} nelze zamítnout, je potvrzena stacionarita,
- pokud H_0^{ADF} nelze zamítnout a H_0^{KPSS} se zamítá, je potvrzena nestacionarita a přítomnost jednotkového kořene autoregresního polynomu,
- ve zbývajících dvou případech se výsledek považuje za neprůkazný.

2.3 Identifikace pomocí tvaru ACF a PACF

Po provedení potřebných transformací lze přistoupit k vlastní identifikaci vhodného modelu. Jednou z možností je identifikace podle tvaru odhadnuté autokorelační a parciální autokorelační funkce. Jde o subjektivní metodu, která spočívá v prohlídce odhadu korelogramu a parciálního korelogramu transformované časové řady a porovnání s jejich teoretickými protějšky. Rozhodujeme se zejména podle bodu useknutí k_0 , za kterým je funkce již nulová, (či jeho neexistence) a podle tvaru korelogramu. Základní vlastnosti autokorelační a parciální autokorelační funkce stacionárních a invertibilních procesů $\text{AR}(p)$, $\text{MA}(q)$ a $\text{ARMA}(p, q)$ shrnuje tabulka 2.5 převzatá z [4], str. 339. Písmenem U se zde označuje křivka ve tvaru lineární kombinace klesajících geometrických posloupností a sinusoid s geometricky klesajícími amplitudami.

Při ověřování bodu useknutí je třeba stanovit, jak blízko nule musí odhadnutá autokorelační funkce být, abychom mohli s předem danou spolehlivostí tvrdit, že $\rho_k = 0$. Lze testovat hypotézu, že přirozené číslo k_0 je bod useknutí, pomocí následujících aproximací (viz [4], str. 340):

- *Bartlettova aproximace* pro autokorelační funkci s asymptotickým kritickým oborem na hladině významnosti 5 %

$$|r_k| \geq 2\sqrt{\frac{1}{n} \left(1 + 2 \sum_{j=1}^{k_0} r_j^2 \right)} \quad \text{pro nějaké } k > k_0, \quad (2.3)$$

- *Quenouilleova aproximace* pro parciální autokorelační funkci s kritickým oborem na hladině významnosti 5 %

$$|r_{kk}| \geq 2\sqrt{\frac{1}{n}} \quad \text{pro nějaké } k > k_0. \quad (2.4)$$

V [20] na str. 105 se pro úspěšnou identifikaci doporučuje délka pozorované řady $n \geq 50$ a sestavení korelogramu pro $k = 1, \dots, \frac{n}{4}$.

	AR(p)	MA(q)	ARMA(p, q)
ρ_k	ve tvaru křivky U ; neexistuje k_0	$k_0 = q$	ve tvaru křivky U po prvních $q - p$ hodnotách; neexistuje k_0
ρ_{kk}	$k_0 = p$	omezená křivkou U ; neexistuje k_0	omezená křivkou U po prvních $p - q$ hodnotách; neexistuje k_0

Tabulka 2.5: Vlastnosti ACF a PACF stacionárních a invertibilních procesů AR, MA a ARMA

2.4 Identifikace pomocí speciálních funkcí

K identifikaci vhodného modelu lze také využít některé speciální funkce jako inverzní autokorelační funkci nebo rozšířenou výběrovou autokorelační funkci, které jsou popsány např. v kapitolách 6.3 a 6.4 knihy [20].

2.4.1 Inverzní autokorelační funkce

Uvažujme stacionární a invertibilní model ARMA(p, q)

$$\Phi(B) Z_t = \Theta(B) \varepsilon_t, \quad (2.5)$$

kde $\Phi(B)$ je autoregresní operátor a $\Theta(B)$ je operátor klouzavých součtů. Rovnici (2.5) lze psát ve tvaru lineárního procesu (1.3) jako

$$Z_t = \frac{\Theta(B)}{\Phi(B)} \varepsilon_t = \psi(B) \varepsilon_t.$$

Vytvořující funkci autokovarianční funkce tohoto modelu dostaneme z (1.4) jako

$$\begin{aligned} \gamma(B) &= \sum_{k=-\infty}^{\infty} \gamma_k B^k = \\ &= \sigma_\varepsilon^2 \psi(B) \psi(B^{-1}) = \\ &= \sigma_\varepsilon^2 \cdot \frac{\Theta(B) \Theta(B^{-1})}{\Phi(B) \Phi(B^{-1})}. \end{aligned} \quad (2.6)$$

Předpokládejme, že $|\gamma(B)| > 0$ pro všechna $|B| \leq 1$ a označme

$$\begin{aligned} \gamma^{(I)}(B) &= \frac{1}{\gamma(B)} = \sum_{k=-\infty}^{\infty} \gamma_k^{(I)} B^k = \\ &= \frac{1}{\sigma_\varepsilon^2} \cdot \frac{\Phi(B) \Phi(B^{-1})}{\Theta(B) \Theta(B^{-1})}. \end{aligned} \quad (2.7)$$

Proces, jehož vytvořující funkce autokovarianční funkce je $\gamma^{(I)}(B)$, nazýváme *inverzním procesem* k procesu (2.5). $\gamma^{(I)}(B)$ nazýváme *vytvořující funkcí inverzní autokovarianční funkce* procesu $\{Z_t\}$.

Definice 19. Vytvořující funkci inverzní autokorelační funkce definujeme jako

$$\rho^{(I)}(B) = \frac{\gamma^{(I)}(B)}{\gamma_0^{(I)}}.$$

Definice 20. Inverzní autokorelační funkci (IACF) jako funkci zpoždění k definujeme jako

$$\rho_k^{(I)} = \frac{\gamma_k^{(I)}}{\gamma_0^{(I)}}.$$

Hodnota $\gamma_k^{(I)}$ představuje koeficient u B^k a B^{-k} v řadě $\gamma^{(I)}(B)$. Hodnota $\rho_k^{(I)}$ představuje koeficient u B^k a B^{-k} v řadě

$$\rho^{(I)}(B) = \sum_{k=-\infty}^{\infty} \rho_k^{(I)} B^k.$$

Z rovnic (2.6) a (2.7) dostaneme, že pokud se časová řada $\{Z_t\}$ řídí např. modelem ARMA(p, q), pak inverzní proces je ARMA(q, p) apod. Tento a další vztahy mezi procesy AR, MA a ARMA a jejich inverzními procesy shrnuje tabulka 2.6.

Proces	Inverzní proces
AR(p)	MA(p)
MA(q)	AR(q)
ARMA(p, q)	ARMA(q, p)

Tabulka 2.6: Procesy AR, MA a ARMA a jejich inverzní procesy

Pokud je $\{Z_t\}$ proces MA(q) s bodem useknutí autokorelační funkce v bodě q , pak jeho inverzní autokorelační funkce je omezená křivkou U a nemá bod useknutí. Je-li $\{Z_t\}$ proces AR(p) s autokorelační funkcí ve tvaru křivky U bez bodu useknutí, pak jeho inverzní autokorelační funkce má bod useknutí $k_0 = p$. Inverzní autokorelační funkce tedy má vlastnosti, které odpovídají parciální autokorelační funkci, a lze ji taktéž použít k identifikaci vhodného modelu.

Jako příklad uveďme inverzní autokorelační funkci modelu AR(p). Ta odpovídá autokorelační funkci modelu klouzavých součtů MA(p) dle (1.6) a má tvar

$$\rho_k^{(I)} = \begin{cases} \frac{\varphi_k + \varphi_1 \varphi_{k+1} + \dots + \varphi_{p-k} \varphi_p}{1 + \varphi_1^2 + \dots + \varphi_p^2} & \text{pro } k = 1, \dots, p, \\ 0 & \text{pro } k > p. \end{cases}$$

Její odhad lze získat nahrazením parametrů φ_i odhady $\hat{\varphi}_i$ získanými metodou nejmenších čtverců, tedy jako

$$\hat{\rho}_k^{(I)} = r_k^{(I)} = \begin{cases} \frac{\hat{\varphi}_k + \hat{\varphi}_1 \hat{\varphi}_{k+1} + \dots + \hat{\varphi}_{p-k} \hat{\varphi}_p}{1 + \hat{\varphi}_1^2 + \dots + \hat{\varphi}_p^2} & \text{pro } k = 1, \dots, p, \\ 0 & \text{pro } k > p. \end{cases}$$

Podobně jako u ACF a PACF, při určování bodu useknutí je důležité určit, jak malou hodnotu musí odhadnutá inverzní autokorelační funkce mít, abychom ji mohli

považovat za nulovou. Pro testování bodu useknutí k_0 lze použít Quenouilleovu aproximaci (2.4) s kritickým oborem na hladině významnosti 5 %

$$|r_k^{(I)}| \geq 2\sqrt{\frac{1}{n}} \quad \text{pro nějaké } k > k_0. \quad (2.8)$$

2.4.2 Rozšířená výběrová autokorelační funkce

V případě modelu $MA(q)$ je identifikace podle bodu useknutí $k_0 = q$ odhadnuté ACF poměrně snadná. Stejně tak u $AR(p)$ nebývá složité identifikovat model podle bodu useknutí $k_0 = p$ odhadnuté PACF a IACF. Obtížnější situace však nastává u modelů $ARMA(p, q)$, jejichž ACF, PACF i IACF nemají bod useknutí a tvarem odpovídají křivce U . Alternativou, pomocí které lze najít řády modelu $ARMA$, je *rozšířená výběrová autokorelační funkce (ESACF)*.

Uvažujme model $ARMA(p, q)$

$$(1 - \varphi_1 B - \dots - \varphi_p B^p) Z_t = (1 + \theta_1 B + \dots + \theta_q B^q) \varepsilon_t. \quad (2.9)$$

Platí, že

$$Y_t = (1 - \varphi_1 B - \dots - \varphi_p B^p) Z_t$$

se řídí modelem $MA(q)$

$$Y_t = (1 + \theta_1 B + \dots + \theta_q B^q) \varepsilon_t.$$

Jednou z možností je nejprve odhadnout model $AR(p)$ metodou nejmenších čtverců a pomocí ACF odhadnutých reziduí identifikovat řád modelu $MA(q)$. Tím dostaneme řády p, q modelu $ARMA(p, q)$. Takový postup však může vést k chybné identifikaci, jelikož odhady $\hat{\varphi}_i$ autoregresních parametrů φ_i metodou nejmenších čtverců nejsou konzistentní, pokud se časová řada $\{Z_t\}$ řídí modelem $ARMA(p, q)$ pro $q > 0$. Podrobné zdůvodnění uvádí [20], str. 147-149.

Pro nalezení konzistentních odhadů parametrů φ_i předpokládejme, že máme k dispozici n pozorování časové řady, která se řídí modelem (2.9). Pro model $AR(p)$

$$Z_t = \sum_{i=1}^p \varphi_i Z_{t-i} + e_t, \quad t = p+1, \dots, n,$$

kde e_t je chybová složka, najdeme odhady metodou nejmenších čtverců $\hat{\varphi}_i^{(0)}$ parametrů $\varphi_i, i = 1, \dots, p$, a spočítáme rezidua

$$\hat{e}_t^{(0)} = Z_t - \sum_{i=1}^p \hat{\varphi}_i^{(0)} Z_{t-i}.$$

Odhady $\hat{\varphi}_i^{(0)}$ nejsou konzistentní a odhadnutá rezidua $\hat{e}_t^{(0)}$ netvoří bílý šum. Proto autor [20] navrhuje následující iterační postup.

- 1. iterační krok

Uvažujeme $AR(p)$ model s doplněným regresorem $\hat{e}_{t-1}^{(0)}$ ve tvaru

$$Z_t = \sum_{i=1}^p \varphi_i^{(1)} Z_{t-i} + \beta_1^{(1)} \hat{e}_{t-1}^{(0)} + e_t^{(1)}, \quad t = p+2, \dots, n.$$

Spočítáme odhady metodou nejmenších čtverců parametrů $\hat{\varphi}_i^{(1)}, i = 1, \dots, p$ a $\hat{\beta}_1^{(1)}$. Odhady $\hat{\varphi}_i^{(1)}$ jsou konzistentní pouze pro $q = 1$. Pro $q > 1$ jsou opět nekonzistentní a odhadnutá rezidua $\hat{e}_t^{(1)}$ netvoří bílý šum.

- *2. iterační krok*

Uvažujeme AR(p) model

$$Z_t = \sum_{i=1}^p \varphi_i^{(2)} Z_{t-i} + \beta_1^{(2)} \hat{e}_{t-1}^{(1)} + \beta_2^{(2)} \hat{e}_{t-2}^{(0)} + e_t^{(2)}, \quad t = p+3, \dots, n.$$

Spočítáme odhady parametrů $\hat{\varphi}_i^{(2)}, i = 1, \dots, p, \hat{\beta}_1^{(2)}$ a $\hat{\beta}_2^{(2)}$ a odhadnutá rezidua $\hat{e}_t^{(2)}$. Odhady $\hat{\varphi}_i^{(2)}$ metodou nejmenších čtverců jsou konzistentní pouze pro $q = 2$.

⋮

- *q-tý iterační krok*

Uvažujeme AR(p) model

$$Z_t = \sum_{i=1}^p \varphi_i^{(q)} Z_{t-i} + \sum_{i=1}^q \beta_i^{(q)} \hat{e}_{t-i}^{(q-i)} + e_t^{(q)}, \quad t = p+q+1, \dots, n,$$

kde $\hat{e}_t^{(j)} = Z_t - \sum_{i=1}^p \hat{\varphi}_i^{(j)} Z_{t-i} - \sum_{i=1}^j \hat{\beta}_i^{(j)} \hat{e}_{t-i}^{(j-i)}$ jsou odhadnutá rezidua z j -tého iteračního kroku a $\hat{\varphi}_i^{(j)}$ a $\hat{\beta}_i^{(j)}$ jsou odhady metodou nejmenších čtverců z j -tého iteračního kroku. Odhady $\hat{\varphi}_i^{(q)}$ již jsou konzistentní, tedy platí

$$\hat{\varphi}_i^{(q)} \xrightarrow{P} \varphi_i, \quad i = 1, \dots, p.$$

V průběhu identifikace vhodného modelu jsou však řády p, q modelu ARMA neznámé a je třeba je odhadnout. Předpokládejme, že $\hat{\varphi}_i^{(j)}, i = 1, \dots, m$, jsou odhady metodou nejmenších čtverců z j -tého iteračního kroku modelu AR(m) pro časovou řadu $\{Z_t\}$, která se řídí modelem ARMA.

Definice 21. Pro časovou řadu $\{Z_t\}$ definujeme m -tou rozšířenou výběrovou autokorelační funkci (ESACF) $r_j^{(m)}$ jako výběrovou autokorelační funkci transformované řady

$$Y_t^{(m,j)} = \left(1 - \hat{\varphi}_1^{(j)} B - \dots - \hat{\varphi}_m^{(j)} B^m\right) Z_t, \quad t = m+1, \dots, n. \quad (2.10)$$

Hodnoty ESACF lze uspořádat do tabulky 2.7. Prvek $r_j^{(m)}$ značí výběrovou autokorelační funkci řady $\{Y_t^{(m,j)}\}$ pro zpoždění j . Čísla řádků tabulky odpovídají řádu m modelu AR a čísla sloupců řádu j modelu MA. Hodnoty $r_j^{(0)}$ v prvním řádku tabulky tvoří klasickou odhadnutou autokorelační funkci časové řady $\{Z_t\}$ dle definice 12.

Rozšířená výběrová autokorelační funkce $r_j^{(m)}$ je funkce počtu pozorování n . Jak je ukázáno v [18], pro proces ARMA(p, q) platí následující konvergence v pravděpodobnosti pro $n \rightarrow \infty$

$$r_j^{(m)} \xrightarrow{P} \begin{cases} 0 & \text{pro } 0 \leq m - p < j - q, \\ X \neq 0 & \text{jinak,} \end{cases}$$

kde $m = 1, 2, \dots$ a $j = 1, 2, \dots$. Konkrétně například pro model ARMA(1,1) platí

$$r_j^{(m)} \xrightarrow{P} \begin{cases} 0 & \text{pro } 1 \leq m < j, \\ X \neq 0 & \text{jinak,} \end{cases}$$

což můžeme přehledně zapsat do tabulky 2.8. Vidíme, že nulové hodnoty asymptotické ESACF tvoří trojúhelník s vrcholem na pozici (1,1). Pro obecný model ARMA(p, q) bude vrchol tohoto trojúhelníka ležet na pozici (p, q).

		MA(j)					
		0	1	2	3	4	...
AR(m)	0	$r_1^{(0)}$	$r_2^{(0)}$	$r_3^{(0)}$	$r_4^{(0)}$	$r_5^{(0)}$...
	1	$r_1^{(1)}$	$r_2^{(1)}$	$r_3^{(1)}$	$r_4^{(1)}$	$r_5^{(1)}$...
	2	$r_1^{(2)}$	$r_2^{(2)}$	$r_3^{(2)}$	$r_4^{(2)}$	$r_5^{(2)}$...
	3	$r_1^{(3)}$	$r_2^{(3)}$	$r_3^{(3)}$	$r_4^{(3)}$	$r_5^{(3)}$...
	⋮			⋮			

Tabulka 2.7: Rozšířená výběrová autokorelační funkce

		MA					
		0	1	2	3	4	...
AR	0	X	X	X	X	X	...
	1	X	0	0	0	0	...
	2	X	X	0	0	0	...
	3	X	X	X	0	0	...
	4	X	X	X	X	0	...
⋮				⋮			

Tabulka 2.8: Asymptotická ESACF modelu ARMA(1,1)

Rozšířená výběrová autokorelační funkce tedy může být vhodným nástrojem pro identifikaci smíšeného modelu ARMA. V praxi máme k dispozici jen konečný počet pozorování, a proto hodnota $r_j^{(m)}$ pro $0 \leq m - p < j - q$ nemusí být přesně nula. Lze však testovat nulovost $r_j^{(m)}$ pomocí Bartlettovy aproximace (2.3). Pro účely identifikace modelu pak můžeme pro danou časovou řadu sestavit tabulku ESACF, ve které symbol X značí hodnoty $r_j^{(m)}$ splňující

$$\left| r_j^{(m)} \right| \geq 2 \sqrt{\frac{1}{n - m - j} \left(1 + 2 \sum_{t=1}^{j-1} \left(r_t^{(m)} \right)^2 \right)}$$

a 0 označuje ostatní hodnoty. V tabulce pak hodnoty 0 naznačují tvar trojúhelníku, pomocí jehož vrcholu určíme řady p, q modelu ARMA. Lze použít i méně přesný postup s dolní mezí $2 \sqrt{\frac{1}{n - m - j}}$.

2.5 Identifikace pomocí informačních kritérií

K identifikaci modelu lze použít také informační kritéria. Jde o běžně používaný přístup, který umožňuje rutinní identifikaci modelu bez subjektivního pohledu badatele. K identifikaci modelu ARMA se přistupuje jako ke hledání odhadů \hat{p}, \hat{q} řádů modelu p, q pomocí optimalizace

$$(\hat{p}, \hat{q}) = \underset{k,l}{\operatorname{argmin}} A(k, l),$$

kde $A(k, l)$ je vhodné kritérium, přes předem danou síť hodnot $k = 0, 1, \dots, K$ a $l = 0, 1, \dots, L$. Je tedy nutné odhadnout $(K + 1)(L + 1)$ modelů ARMA a pro každý z nich vypočítat příslušné kritérium. Představíme si několik informačních kritérií založených na reziduiích modelu. Poznamenejme, že definice jednotlivých kritérií se v literatuře různí.

2.5.1 Akaikeho informační kritérium

Pro statistický model s M parametry definuje autor [20] (str. 153) *Akaikeho informační kritérium (AIC)* jako

$$\text{AIC}(M) = -2 \ln [\text{maximální věrohodnost}] + 2M.$$

Předpokládejme, že máme k dispozici n pozorování časové řady $\{Z_t\}$, která se řídí modelem ARMA

$$(1 - \varphi_1 B - \dots - \varphi_p B^p) Z_t = (1 + \theta_1 B + \dots + \theta_q B^q) \varepsilon_t, \quad (2.11)$$

kde $\{\varepsilon_t\}$ je posloupnost nezávislých stejně rozdělených náhodných veličin s normálním rozdělením s nulovou střední hodnotou a konstantním rozptylem σ_ε^2 . Označme $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_n)^\top$, $\boldsymbol{\Phi} = (\varphi_1, \dots, \varphi_p)^\top$, $\boldsymbol{\Theta} = (\theta_1, \dots, \theta_q)^\top$.

Používá se nepodmíněná metoda maximální věrohodnosti (viz [20], str. 138). Logaritmickou věrohodnostní funkci máme ve tvaru

$$\ln L(\boldsymbol{\Phi}, \boldsymbol{\Theta}, \sigma_\varepsilon^2) = -\frac{n}{2} \ln(2\pi\sigma_\varepsilon^2) - \frac{S(\boldsymbol{\Phi}, \boldsymbol{\Theta})}{2\sigma_\varepsilon^2}, \quad (2.12)$$

kde

$$S(\boldsymbol{\Phi}, \boldsymbol{\Theta}) = \sum_{t=-\infty}^n [\mathbb{E}(\varepsilon_t | \boldsymbol{\Phi}, \boldsymbol{\Theta}, \mathbf{Z})]^2 \quad (2.13)$$

je součet čtverců podmíněných středních hodnot ε_t při daných $\boldsymbol{\Phi}, \boldsymbol{\Theta}$ a vektoru $\mathbf{Z} = (Z_1, \dots, Z_n)^\top$. Po získání odhadů parametrů $\hat{\boldsymbol{\Phi}}$ a $\hat{\boldsymbol{\Theta}}$, které maximalizují (2.12), odhad $\hat{\sigma}_\varepsilon^2$ rozptylu σ_ε^2 dostaneme jako

$$\hat{\sigma}_\varepsilon^2 = \frac{S(\hat{\boldsymbol{\Phi}}, \hat{\boldsymbol{\Theta}})}{n} \quad (2.14)$$

a odhad logaritmické věrohodnostní funkce podle 2.12 přejde na tvar

$$\ln \hat{L} = -\frac{n}{2} \ln \hat{\sigma}_\varepsilon^2 - \frac{n}{2} (1 + \ln 2\pi). \quad (2.15)$$

Jelikož druhý sčítanec ve (2.15) je konstantní, tvar AIC kritéria lze psát jako

$$\text{AIC}(M) = n \ln \hat{\sigma}_\varepsilon^2 + 2M.$$

Poznámka. Ukažme nyní možný postup výpočtu součtu čtverců (2.13). V praxi se obvykle nahrazuje výrazem

$$S(\Phi, \Theta) = \sum_{t=-N}^n [\text{E}(\varepsilon_t | \Phi, \Theta, \mathbf{Z})]^2, \quad (2.16)$$

přičemž dle [20], str. 139 se volí takové N , pro které je přírůstek $|\text{E}(Z_t | \Phi, \Theta, \mathbf{Z}) - \text{E}(Z_{t-1} | \Phi, \Theta, \mathbf{Z})|$ menší než předem zvolená hodnota pro $t \leq -N - 1$. Uvažujme model ARMA(1,1) ve tvaru

$$Z_t = \varphi_1 Z_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1}. \quad (2.17)$$

V [13] je na str. 188-189 uveden postup výpočtu (2.16), který využívá tzv. zpětné vyjádření modelu časové řady, tedy v tomto případě

$$Z_t = \varphi_1 Z_{t+1} + e_t + \theta_1 e_{t+1}, \quad (2.18)$$

kde $\{e_t\}$ je bílý šum s nulovou střední hodnotou a konstantním rozptylem σ_e^2 . Označíme-li $[X] = \text{E}(X | \Phi, \Theta, \mathbf{Z})$, potom z (2.17) a (2.18) postupně dostaneme

$$[\varepsilon_t] = [Z_t] - \varphi_1 [Z_{t-1}] - \theta_1 [\varepsilon_{t-1}], \quad (2.19)$$

$$[Z_t] = \varphi_1 [Z_{t+1}] + [e_t] + \theta_1 [e_{t+1}]. \quad (2.20)$$

V prvním kroku vypočítáme

$$[e_t] = [Z_t] - \varphi_1 [Z_{t+1}] - \theta_1 [e_{t+1}] \quad \text{pro } t = 1, 2, \dots, n-1,$$

přičemž volíme $[e_n] = 0$ a $[e_t] = 0$ pro $t \leq 0$. Dále lze získat střední hodnoty pozdějších hodnot časové řady jako

$$[Z_t] = \varphi_1 [Z_{t+1}] + \theta_1 [e_{t+1}] \quad \text{pro } t = -N, -N+1, \dots, 0$$

a naposledy podmíněné střední hodnoty bílého šumu

$$[\varepsilon_t] = [Z_t] - \varphi_1 [Z_{t-1}] - \theta_1 [\varepsilon_{t-1}] \quad \text{pro } t = -N+1, -N+2, \dots, n,$$

kde uvažujeme startovací hodnotu $[\varepsilon_{-N}] = [Z_{-N}]$. Hodnoty $[\varepsilon_t]$ poté dosazujeme do (2.16).

2.5.2 Akaikeho Bayesovské informační kritérium

Autor [16] uvedl, že AIC má tendenci nadhodnocovat řády autoregrese. Akaike proto v [1] a [2] provedl rozšíření na *Akaikeho Bayesovské informační kritérium* (ABIC) definované jako

$$\text{ABIC}(M) = n \ln \hat{\sigma}_\varepsilon^2 - (n - M) \ln \left(1 - \frac{M}{n}\right) + M \ln n + M \ln \left[\left(\frac{\hat{\sigma}_Z^2}{\hat{\sigma}_\varepsilon^2} - 1 \right) / M \right],$$

kde $\hat{\sigma}_\varepsilon^2$ je maximálně věrohodný odhad rozptylu bílého šumu σ_ε^2 , M je počet parametrů a $\hat{\sigma}_Z^2 = \hat{\gamma}_0$ je odhad rozptylu pozorované řady $\{Z_t\}$. V [1] bylo poté simulacemi zjištěno, že ABIC méně nadhodnocuje řády autoregrese než AIC.

2.5.3 Další informační kritéria

Mezi další používaná kritéria vhodná k identifikaci modelu s M parametry patří:

- *Schwartzovo Bayesovské kritérium (BIC)* definované v [20], str. 153 ve tvaru

$$\text{BIC}(M) = n \ln \hat{\sigma}_\varepsilon^2 + M \ln n,$$

- *Parzenovo informační kritérium (CAT – criterion for autoregressive transfer functions)* pro řád p autoregresního modelu definované dle [14] jako

$$\text{CAT}(p) = \begin{cases} -\left(1 + \frac{1}{n}\right) & \text{pro } p = 0, \\ \frac{1}{n} \sum_{j=1}^p \frac{1}{\hat{\sigma}_j^2} - \frac{1}{\hat{\sigma}_p^2} & \text{pro } p = 1, 2, 3, \dots, \end{cases}$$

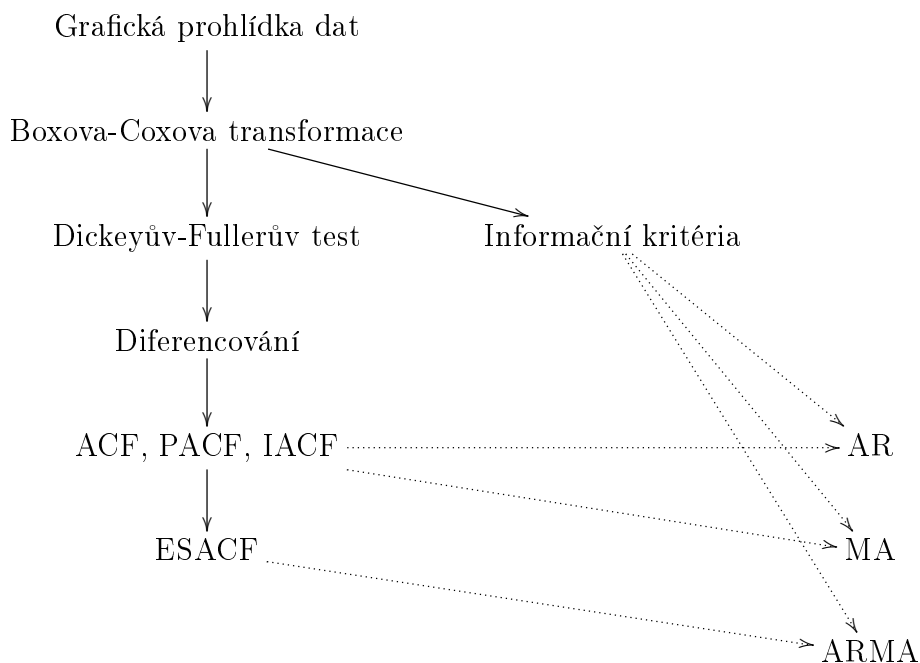
kde $\hat{\sigma}_j^2$ je nestranný odhad σ_ε^2 v modelu $\text{AR}(j)$ pro danou časovou řadu a n je počet pozorování,

- *Hannanovo-Quinnovo informační kritérium (HQ)* dle [8] ve tvaru

$$\text{HQ}(M) = n \ln \hat{\sigma}_\varepsilon^2 + 2M \ln \ln n.$$

2.6 Postup identifikace modelu

Na závěr shrňme varianty postupu při identifikaci modelu jednorozměrné časové řady. Po grafické prohlídce dat a provedení nutných transformací odhadneme autokorelační, parciální autokorelační a inverzní autokorelační funkci. Podle jejich tvaru lze identifikovat autoregresní model či model klouzavých součtů. Řád smíšeného modelu ARMA však z těchto funkcí rozpoznat nedokážeme, proto se nabízí sestavit tabulku rozšířené výběrové autokorelační funkce a podle souřadnic vrcholu trojúhelníka nulových hodnot určit řád modelu ARMA. Další variantou jsou informační kritéria, pro jejichž použití není třeba řadu diferencovat. Informačními kritérii totiž lze vybírat model z třídy procesů ARIMA, které jsou obecně nestacionární. Tyto modely jsou podrobně popsány např. v [20], str. 67-82. Schematicky lze identifikační proceduru znázornit následovně:



Kapitola 3

Simulace jednorozměrných časových řad

Nyní přistoupíme k ověření fungování výše popsaných identifikačních kritérií na simulovaných jednorozměrných časových řadách. Nejprve podrobně ukážeme postup identifikace modelu pro tři simulované řady dle modelů AR, MA a ARMA pomocí jednotlivých kritérií. Z informačních kritérií budeme používat AIC a BIC.

3.1 Simulace procesu AR(1)

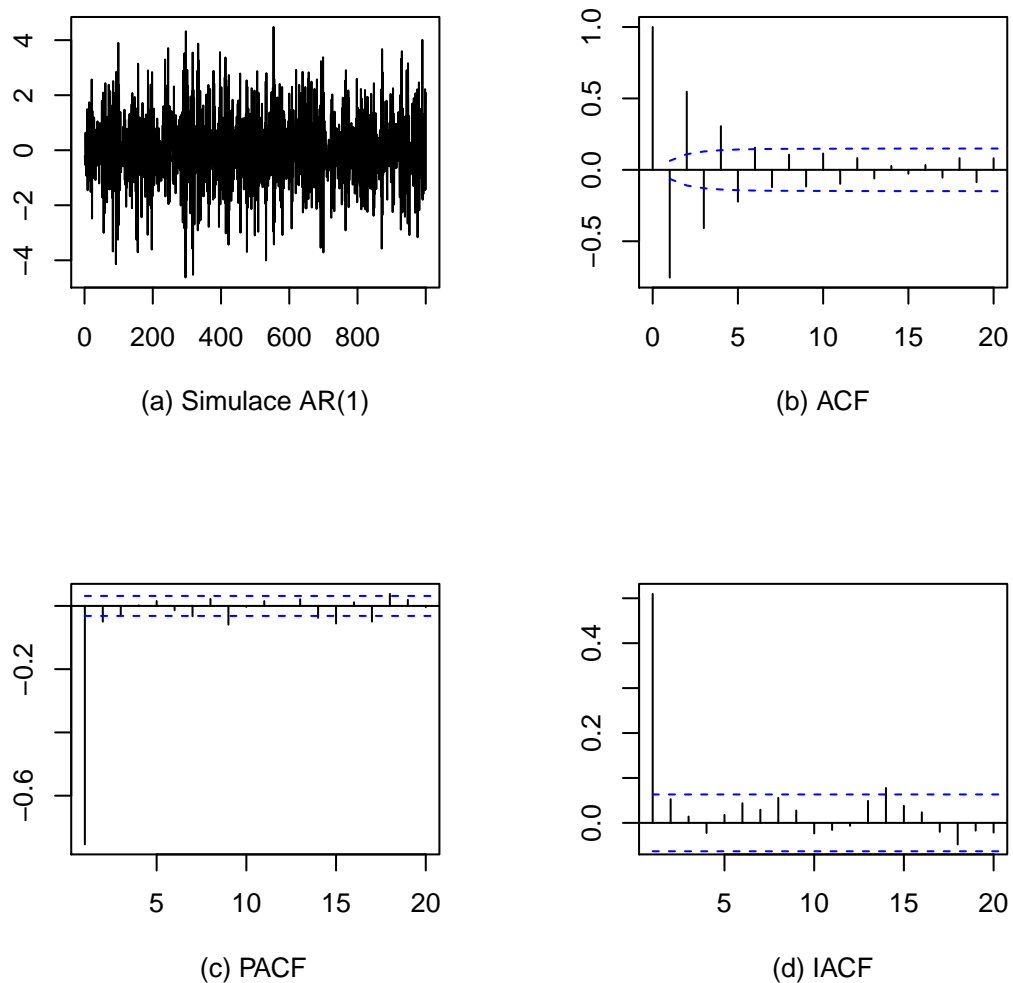
Budeme generovat časovou řadu délky $n = 1000$, která se řídí modelem AR(1) s parametrem $\varphi_1 = -0.75$ ve tvaru

$$Z_t = -0.75Z_{t-1} + \varepsilon_t,$$

přičemž rozdělení bílého šumu volíme normální s jednotkovým rozptylem, tedy $\varepsilon_t \sim N(0,1)$. Tento proces je invertibilní a vzhledem k tomu, že kořen polynomu $\Phi(z) = 1 + 0,75z$ leží vně jednotkového kruhu v komplexní rovině, také stacionární. V obrázku 3.1 je zobrazena trajektorie procesu a odhadnuté autokorelační a parciální autokorelační funkce spolu s mezemi pro test jejich nulovosti sestavenými dle aproximací (2.3) a (2.4). Tvar ACF odpovídá křivce U a PACF je useknutá v bodě $k_0 = 1$. Podle odhadnutých autokorelačních funkcí bychom tedy identifikovali model AR(1). Odhadnuté hodnoty inverzní autokorelační funkce jsou zobrazené v grafu 3.1 (d) společně s mezemi dle aproximace (2.8). Bod useknutí $k_0 = 1$ odhadnuté IACF opět vede na model AR(1).

Rozšířená výběrová autokorelační funkce je užitečná zejména pro smíšené modely ARMA, ale její fungování vyzkoušíme i na procesu AR(1). Nulové body v tabulce 3.1 sice netvoří přesný trojúhelník, ale naznačují jeho tvar s pravděpodobným vrcholem na pozici (1,0). Odhad $r_1^{(1)} \doteq -0,0666$ na souřadnici (1,0) je navíc jen nepatrně menší než dolní mez $-0,0633$. Z hlediska ESACF bychom zvolili model AR(1).

Vypočtené hodnoty informačních kritérií jsou uvedeny v tabulce 3.2. Na základě AIC identifikujeme model AR(2) a podle BIC model AR(1). AIC v tomto případě nadhodnotilo řád autoregrese, jak bylo zmíněno v kapitole 2.5.2, nicméně odlišnost hodnot AIC pro AR(1) a AR(2) je zanedbatelná.



Obrázek 3.1: Simulovaná časová řada AR(1) a odhadnuté autokorelační funkce

		MA								
		0	1	2	3	4	5	6	7	8
AR	0	X	X	X	X	X	X	X	X	X
	1	X	0	0	0	0	0	0	0	0
	2	X	0	0	0	0	0	0	0	0
	3	0	X	0	0	0	0	0	X	0
	4	X	X	X	0	0	0	0	X	0
	5	X	X	X	0	0	0	0	0	0
	6	X	X	X	X	X	0	0	0	0

Tabulka 3.1: Rozšířená výběrová autokorelační funkce v modelu AR(1)

	AIC	BIC
AR(1)	2817,791	2827,607
AR(2)	2817,292	2832,015
MA(1)	3101,641	3111,456
ARMA(2,1)	2818,747	2838,387
ARMA(2,2)	2820,230	2844,769

Tabulka 3.2: Hodnoty informačních kritérií

3.2 Simulace procesu MA(1)

Přistoupíme k simulaci časové řady délky $n = 1000$ dle modelu MA(1) s parametrem $\theta_1 = 0,6$ a bílým šumem $\varepsilon_t \sim N(0,1)$. Příslušný zápis modelu je

$$Z_t = \varepsilon_t + 0,6\varepsilon_{t-1}.$$

Tento proces je stacionární a jelikož kořen polynomu $\Theta(z) = 1 + 0,6z$ leží vně jednotkového kruhu v komplexní rovině, je také invertibilní. V obrázku 3.2 vidíme trajektorii procesu a odhadnuté autokorelační funkce. ACF je useknutá v bodě $k_0 = 1$ a PACF má tvar křivky U . Na základě těchto dvou odhadnutých funkcí pro uvažovanou časovou řadu identifikujeme model MA(1).

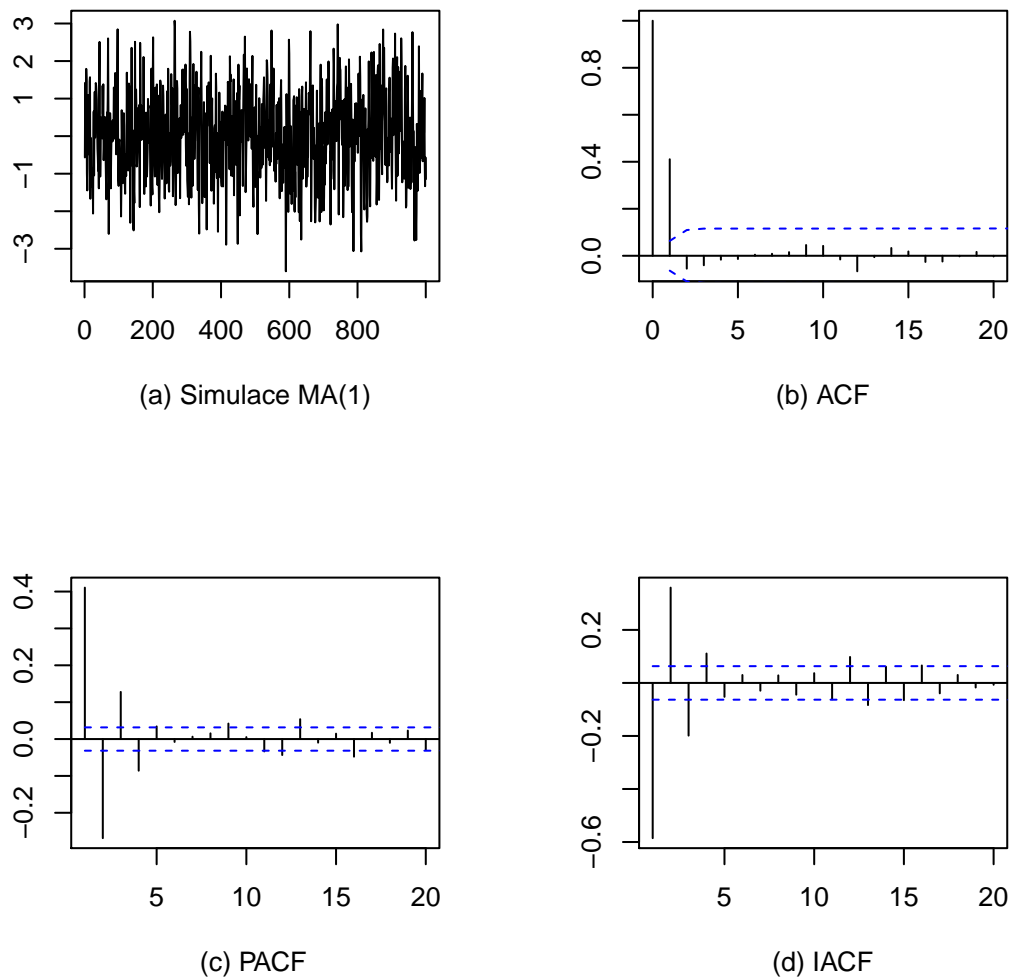
Odhadnutá inverzní autokorelační funkce zobrazená v grafu 3.1 (d) má tvar připomínající křivku U , což odpovídá chování teoretické IACF modelu klouzavých součtů. V tomto případě však odhadnutá IACF neumožní určení přesného řádu modelu, jako tomu bylo u autoregresního procesu.

Správnou identifikaci modelu poskytla i rozšířená výběrová autokorelační funkce, jejíž hodnoty jsou uvedené v tabulce 3.3. Vrchol trojúhelníka nulových hodnot je v bodě na pozici (0,1), což vede k identifikaci modelu MA(1).

Přístup pomocí Akaikeho informačního kritéria a Schwartzova Bayesovského informačního kritéria dal v obou případech správný výsledek. Nejmenší hodnota obou kritérií byla softwarem vypočtena pro model MA(1), viz tabulka 3.4.

	MA								
	0	1	2	3	4	5	6	7	8
	0	X	0	0	0	0	0	0	0
	1	X	X	0	0	0	0	0	0
	2	X	X	X	0	0	0	0	0
AR	3	X	X	X	X	0	0	0	0
	4	X	X	X	X	0	0	0	0
	5	X	X	0	0	0	0	0	0
	6	X	X	X	0	X	0	X	0

Tabulka 3.3: Rozšířená výběrová autokorelační funkce v modelu MA(1)



Obrázek 3.2: Simulovaná časová řada MA(1) a odhadnuté autokorelační funkce

	AIC	BIC
AR(1)	2923,370	2933,185
MA(1)	2825,504	2835,319
MA(2)	2827,098	2841,821
ARMA(1,1)	2827,155	2841,878
ARMA(1,2)	2827,860	2847,491
ARMA(2,2)	2829,066	2853,605

Tabulka 3.4: Hodnoty informačních kritérií

3.3 Simulace procesu ARMA(1,1)

Dále provedeme simulaci časové řady délky $n = 1000$, která se řídí modelem ARMA(1,1) s parametry $\varphi_1 = 0,9$ a $\theta_1 = 0,6$. Volíme normální rozdělení bílého šumu $\varepsilon_t \sim N(0,1)$. Model zapíšeme rovnicí

$$Z_t = 0,9Z_{t-1} + \varepsilon_t + 0,6\varepsilon_{t-1}.$$

Kořen polynomu $\Phi(z) = 1 - 0,9z$ leží vně jednotkového kruhu, tedy proces je stacionární. Navíc i kořen polynomu $\Theta(z) = 1 + 0,6z$ leží vně jednotkového kruhu v komplexní rovině, proto je proces invertibilní. Simulovaná časová řada je znázorněna v obrázku 3.3 spolu s odhadnutými autokorelačními funkcemi. V tomto případě odhadnuté ACF, PACF i IACF odpovídají křivce U a nelze podle nich jednoznačně identifikovat příslušný model.

Jasný závěr naopak získáme z rozšířené výběrové autokorelační funkce. Trojúhelník nulových hodnot v tabulce 3.5 má vrchol v bodě na pozici (1,1). Podle ESACF tedy identifikujeme model ARMA(1,1).

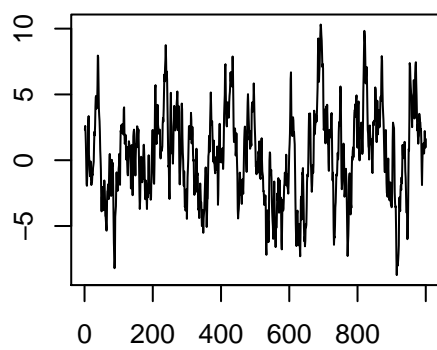
Stejný výsledek dávají i použitá informační kritéria. Jako nejvhodnější model byl podle AIC i BIC zvolen smíšený model ARMA(1,1), viz tabulka 3.6.

		MA								
		0	1	2	3	4	5	6	7	8
AR	0	X	X	X	X	X	X	X	X	X
	1	X	0	0	0	0	0	0	0	0
	2	X	X	0	0	0	0	0	0	0
	3	X	X	X	0	0	0	0	0	0
	4	X	X	X	X	0	0	0	0	0
	5	X	X	X	X	0	0	0	0	0
	6	X	X	0	X	0	X	0	0	0

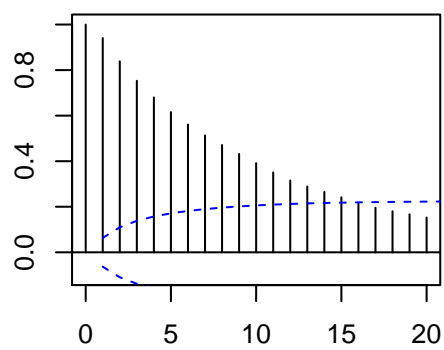
Tabulka 3.5: Rozšířená výběrová autokorelační funkce v modelu ARMA(1,1)

	AIC	BIC
AR(1)	3119,225	3129,040
MA(1)	4123,296	4133,112
ARMA(1,1)	2849,951	2864,674
ARMA(1,2)	2851,855	2871,486
ARMA(2,1)	2851,870	2871,501
ARMA(2,2)	2853,871	2878,410

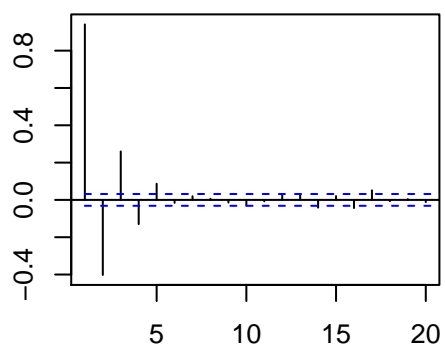
Tabulka 3.6: Hodnoty informačních kritérií



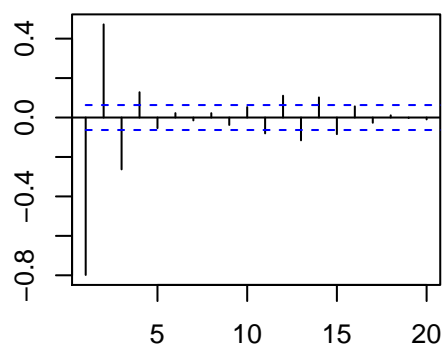
(a) Simulace ARMA(1,1)



(b) ACF



(c) PACF



(d) IACF

Obrázek 3.3: Simulovaná časová řada ARMA(1,1) a odhadnuté autokorelační funkce

3.4 Porovnání identifikačních kritérií

V předchozí části jsme podrobně ukázali použití jednotlivých identifikačních kritérií na simulované jednorozměrné časové řady. Ověřili jsme, že identifikace modelu pomocí grafů autokorelačních funkcí i podle tabulky s hodnotami rozšířené výběrové autokorelační funkce velmi závisí na subjektivním pohledu badatele. Například bod useknutí odhadnutých autokorelačních funkcí není vždy jednoznačný či nulové hodnoty ESACF nemusí tvořit přesně trojúhelník, jak jsme viděli v tabulce 3.1. Potom může být nejasné, který model bychom pro danou časovou řadu identifikovali. Naproti tomu přístup pomocí informačního kritéria vždy dá jednoznačný výsledek, jelikož vybíráme minimum z předem daného počtu hodnot. Proto je vhodné použít odhady zmíněných autokorelačních funkcí pro orientační posouzení chování časové řady a finální výběr modelu provést pomocí informačních kritérií.

Dále se budeme věnovat oběma použitým informačním kritériím – AIC a BIC. Porovnáme jejich fungování na 100 simulovaných časových řadách dle modelů AR(1), MA(1) a ARMA(1,1) délky $n = 1000$. U každého z nich budeme sledovat, pro kolik procent simulovaných řad byl identifikován správný model dle Akaikeho kritéria a dle Schwartzova Bayesovského kritéria.

Pro model AR(1) volíme shodně s kapitolou 3.1 tvar

$$Z_t = -0.75Z_{t-1} + \varepsilon_t.$$

Tabulka 3.7 shrnuje, v kolika případech byly pomocí AIC a BIC identifikovány uvedené modely. Ukázalo se, že Akaikeho kritérium bylo v identifikaci autoregresního modelu prvního řádu mnohem méně přesné než Bayesovské kritérium. Potvrdil se fakt, že Akaikeho informační kritérium má tendenci nadhodnocovat řád autoregrese. Ke správné identifikaci vedlo jen v 74 % případů. Ve zbylých případech by AIC identifikovalo autoregresní řád druhého řádu či smíšené modely ARMA, a to dokonce až do řádů (3,2). Úspěšnější bylo Schwartzovo Bayesovské identifikační kritérium, které správně rozpoznalo model AR(1) u 96 % simulací.

	AIC	BIC
AR(1)	74	96
AR(2)	8	3
ARMA(1,1)	5	1
ARMA(1,2)	5	0
ARMA(1,3)	1	0
ARMA(2,1)	5	0
ARMA(2,2)	1	0
ARMA(3,2)	1	0

Tabulka 3.7: Počty identifikovaných modelů dle AIC a BIC

Stejný postup jsme opakovali také pro simulované řady dle modelu klouzavých součtů řádu 1 ve tvaru

$$Z_t = \varepsilon_t + 0,6\varepsilon_{t-1}.$$

Počty jednotlivých identifikovaných modelů jsou uvedeny v tabulce 3.8. V tomto případě bylo Akaikeho kritérium ještě méně úspěšné než u modelu AR(1). Správně identifikovalo model klouzavých součtů řádu 1 jen v 69 % případů. Naproti tomu Schwartzovo Bayesovské informační kritérium bylo bezchybné u 98 % simulovaných řad.

	AIC	BIC
MA(1)	69	98
MA(2)	2	0
ARMA(1,1)	15	2
ARMA(1,2)	8	0
ARMA(2,1)	4	0
ARMA(2,2)	2	0

Tabulka 3.8: Počty identifikovaných modelů dle AIC a BIC

Na závěr jsme generovali časové řady podle smíšeného modelu ARMA(1,1) ve tvaru

$$Z_t = 0,9Z_{t-1} + \varepsilon_t + 0,6\varepsilon_{t-1}.$$

V tabulce 3.9 vidíme, že Schwartzovo Bayesovské informační kritérium bylo opět úspěšnější – správně identifikovalo model ARMA(1,1) u 98 % řad. Přístup pomocí Akaikeho kritéria vedl k tomuto modelu v 83 % případů.

	AIC	BIC
ARMA(1,1)	83	98
ARMA(1,2)	10	1
ARMA(2,2)	6	1
ARMA(3,1)	1	0

Tabulka 3.9: Počty identifikovaných modelů dle AIC a BIC

U autoregresního modelu, modelu klouzavých součtů i smíšeného modelu spolehlivěji identifikovalo model Schwartzovo Bayesovské informační kritérium. Akaikeho kritérium častěji nadhodnocovalo řady modelů. Nejméně úspěšné bylo u modelu MA(1), naopak největšího procenta správně identifikovaných modelů dosáhlo u modelu ARMA(1,1).

3.5 Dickeyův-Fullerův test

Důležitým krokem při identifikaci modelu časové řady je ověření její stacionarity. To lze provést pomocí testu na jednotkový kořen, například Dickeyova-Fullerova testu. Nyní přistoupíme k ověření jeho fungování na simulovaných časových řadách. Budeme generovat vždy 100 realizací časových řad délek $n = 100, 250, 500, 1000$, které odpovídají nulové hypotéze náhodné procházky a alternativám jednotlivých τ -testů z kapitoly 2.2.1. Pomocí funkce `ur.df` z balíčku `urca` v softwaru R budeme následně testovat stacionaritu simulovaných časových řad. Rozhodnutí o zamítnutí či nezamítnutí nulové hypotézy učiníme porovnáním testové statistiky s příslušným kvantilem z tabulek 2.2, 2.3, 2.4.

Nejprve budeme generovat časové řady odpovídající nulové hypotéze

$$H_0: Z_t = Z_{t-1} + \varepsilon_t,$$

tedy modelu náhodné procházky s bílým šumem $\varepsilon_t \sim N(0,1)$ Tabulka 3.10 uvádí, kolik procent ze 100 testů skončilo zamítnutím nulové hypotézy pro různé hladiny významnosti. Vidíme, že procento zamítnutých hypotéz se blíží zvolené hladině α , v některých případech se jí přímo rovná.

Délka řady n	α		
	0,01	0,05	0,10
100	0,00	0,05	0,13
250	0,00	0,05	0,09
500	0,03	0,07	0,14
1000	0,01	0,04	0,07

Tabulka 3.10: Procento zamítnutých nulových hypotéz Dickeyova-Fullerova testu

Dále budeme generovat časové řady dle alternativních hypotéz jednotlivých variant Dickeyova-Fullerova testu:

$$\begin{aligned} \tau\text{-test: } H_1: Z_t &= \varphi_1 Z_{t-1} + \varepsilon_t, & \varphi_1 < 1, \\ \tau_\mu\text{-test: } H_1: Z_t &= \mu + \varphi_1 Z_{t-1} + \varepsilon_t, & \varphi_1 < 1, \\ \tau_\tau\text{-test: } H_1: Z_t &= \mu + \beta t + \varphi_1 Z_{t-1} + \varepsilon_t, & \varphi_1 < 1. \end{aligned}$$

Ve všech třech případech volíme dvě různé hodnoty parametru $\varphi_1 = 0,95; 0,98$. Pro τ_μ -test a τ_τ -test volíme $\mu = 5$ a $\beta = 0,2$. Budeme pozorovat, kolik procent ze sta testů skončilo zamítnutím nulové hypotézy, a podle toho určíme dosaženou sílu testu. Její hodnoty jsou uvedeny v tabulkách 3.11, 3.12, 3.13. Největší sílu z použitých testů má τ -test a nejmenší τ_τ -test. Pro časové řady krátkých délek však bylo dosaženo velmi malé síly u všech tří variant testu. U volby autoregresního parametru $\varphi_1 = 0,98 \approx 1$ se projevilo, že generovaná řada je velmi blízká náhodné procházce, což vede k menší síle testu.

Délka řady n	$\varphi_1 = 0,95$			$\varphi_1 = 0,98$		
	α			α		
	0,01	0,05	0,10	0,01	0,05	0,10
100	0,07	0,35	0,61	0,02	0,11	0,28
250	0,54	0,93	0,98	0,06	0,32	0,58
500	0,99	1,00	1,00	0,40	0,77	0,93
1000	1,00	1,00	1,00	0,91	1,00	1,00

Tabulka 3.11: Dosažená síla τ -testu

Délka řady n	$\varphi_1 = 0,95$			$\varphi_1 = 0,98$		
	α			α		
	0,01	0,05	0,10	0,01	0,05	0,10
100	0,02	0,20	0,32	0,01	0,05	0,12
250	0,16	0,52	0,72	0,02	0,13	0,27
500	0,83	0,98	0,99	0,07	0,37	0,61
1000	0,99	0,99	0,99	0,53	0,85	0,91

Tabulka 3.12: Dosažená síla τ_μ -testu

Délka řady n	$\varphi_1 = 0,95$			$\varphi_1 = 0,98$		
	α			α		
	0,01	0,05	0,10	0,01	0,05	0,10
100	0,03	0,11	0,18	0,00	0,04	0,12
250	0,07	0,29	0,47	0,00	0,06	0,13
500	0,40	0,85	0,91	0,04	0,14	0,35
1000	0,98	0,98	0,98	0,27	0,65	0,80

Tabulka 3.13: Dosažená síla τ_τ -testu

Kapitola 4

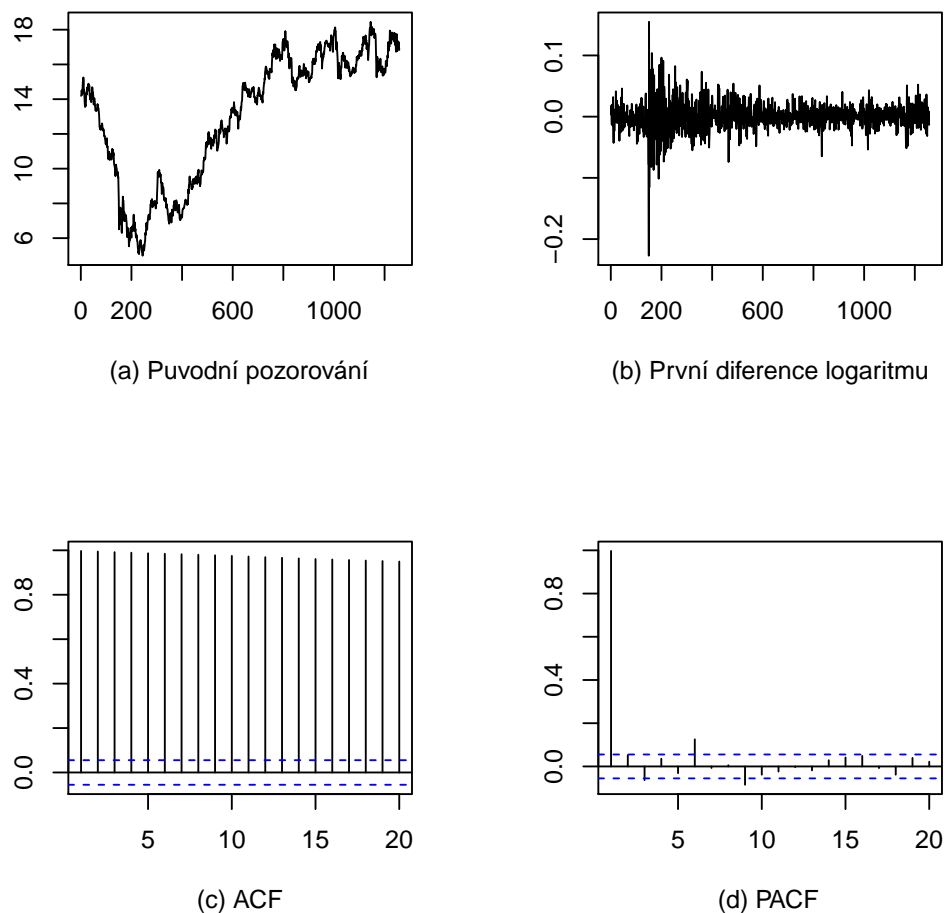
Softwarové zpracování jednorozměrné časové řady

Na závěr první části práce, která se zabývá jednorozměrnými časovými řadami, ukážeme postup identifikace modelu na reálných finančních datech. Použijeme denní časovou řadu kurzů akcií Bank of America Corporation o délce $n = 1258$ pozorování (od 3.1.2011 do 31.12.2015). Jde o zavírací ceny uvedené v amerických dolarech. Data byla získána z webové stránky finance.yahoo.com.

Při identifikaci modelu pro uvedenou časovou řadu se budeme řídit postupem popsaným v kapitole 2.6. Z obrázku 4.1 (a) usuzujeme, že řada je nestacionární ve střední hodnotě i rozptylu, a proto přistoupíme k potřebným transformacím. Nabízí se Boxova-Coxova transformace, o jejímž tvaru lze rozhodnout například podle metody popsané v kapitole 2.1. V našem případě tento postup dává nejednoznačný výsledek, který velmi závisí na volbě délky segmentů. Z grafu 4.1 (a) se však lze domnívat, že rozptyl řady roste s její vzrůstající hodnotou, a proto přistoupíme k logaritmické transformaci, která patří v oblasti financí k nejčastějším.

V dalším kroku je nutné časovou řadu stacionarizovat a centrovat, například pomocí diferencování. O potřebě diferencování se můžeme rozhodnout podle odhadnuté autokorelační funkce logaritmované řady na obrázku 4.1 (c), která velmi pomalu klesá, a odhadnuté parciální autokorelační funkce v grafu (d), která je useknutá v bodě 1, jak bylo popsáno v kapitole 2.2. Alternativně je možné provést Dickeyův-Fullerův test, jehož testová statistika pro naše data vychází 0,0672. Nulovou hypotézu nestacionarity tedy na hladině 5 % nezamítáme, protože testová statistika je větší než příslušný kvantil $-1,95$ (viz tabulka 2.2). Dále budeme pracovat s řadou prvních diferencí, která je na obrázku 4.1 (b). Tu již považujeme za stacionární, neboť Dickeyův-Fullerův test pro tuto řadu zamítá na hladině 5% nulovou hypotézu nestacionarity s hodnotou testové statistiky $-38,0031$.

Po provedení potřebných transformací přistoupíme k vlastní volbě vhodného modelu časové řady. První možností je prohlídka odhadnutých autokorelačních funkcí na obrázku 4.2. Hodnoty ACF, PACF a IACF jsou statisticky významné i pro větší zpoždění a vzhledem k absenci bodu useknutí tedy připouštíme možnost použití smíšeného modelu ARMA.



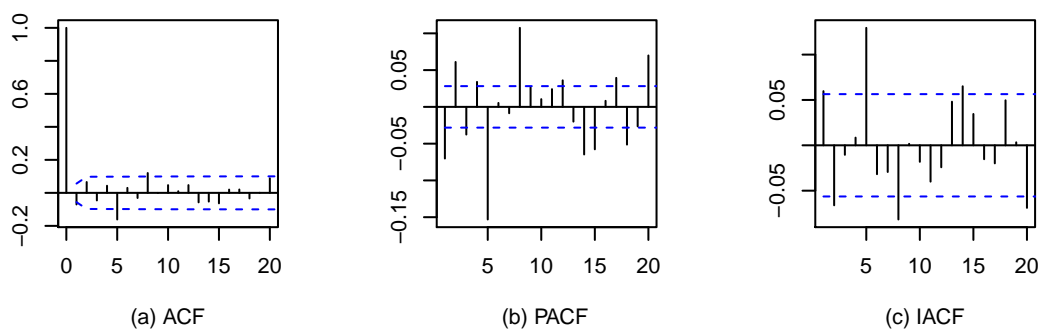
Obrázek 4.1: Řada kurzů akcií, její transformace a autokorelační funkce

Řády smíšeného modelu ARMA se pokusíme určit podle rozšířené výběrové autokorelační funkce v tabulce 4.1. Vidíme, že její nulové hodnoty mnohem méně tvoří trojúhelník než v případě simulovaných časových řad v kapitole 3. Přesto bychom jako nejpravděpodobnější vrchol případného trojúhelníka označili bod (1,1), který vede na identifikaci modelu ARMA(1,1).

Jako poslední možnost využijeme informační kritéria. V tabulce 4.2 jsou uvedeny jejich vypočtené hodnoty. V tomto případě bychom na základě AIC i BIC zvolili model ARMA(1,1) ve shodě s rozšířenou výběrovou autokorelační funkcí.

Dalšími kroky výstavby modelu jsou odhad parametrů a diagnostika, kterými se tato práce podrobněji nezabývá. Jsou podrobně zpracovány např. v kapitole 7 knihy [20]. Uvedme pouze výsledný model spolu s odhadnutými parametry, který má tvar

$$Z_t = -0,9096Z_{t-1} + \varepsilon_t + 0,8472\varepsilon_{t-1}.$$



Obrázek 4.2: Odhadnuté autokorelační funkce

		MA								
		0	1	2	3	4	5	6	7	8
AR	0	X	X	0	0	X	0	0	X	0
	1	X	0	0	0	X	X	0	X	0
	2	X	X	0	0	X	0	0	X	0
	3	X	0	X	0	X	0	0	X	0
	4	X	X	X	X	X	0	0	X	0
	5	0	0	X	X	0	0	0	0	0
	6	X	0	X	X	0	X	0	0	0

Tabulka 4.1: Rozšířená výběrová autokorelační funkce

	AIC	BIC
AR(1)	-5932,090	-5921,817
ARMA(1,1)	-5952,107	-5936,698
ARMA(1,2)	-5950,403	-5929,857
ARMA(2,1)	-5950,390	-5929,844
ARMA(2,2)	-5948,109	-5922,426

Tabulka 4.2: Hodnoty informačních kritérií

Kapitola 5

Lineární modely pro mnohorozměrné časové řady

V ekonomii a financích často sledujeme vývoj a vzájemné ovlivňování v čase několika různých veličin. Mohou to být například časové řady cen akcií několika společností nebo indexu spotřebitelských cen v různých zemích. Jejich vzájemné vztahy lze popisovat pomocí modelů mnohorozměrných časových řad. Místo náhodných veličin Z_t tedy budeme uvažovat m -rozměrné náhodné vektory $\mathbf{Z}_t = (Z_{1,t}, Z_{2,t}, \dots, Z_{m,t})^\top$, jejichž složky tvoří jednorozměrné časové řady. V následující části zobecníme pojmy a modely uvedené v kapitole 1 pro mnohorozměrný případ.

5.1 Základní pojmy

Definice 22. *Nechť náhodné veličiny Z_1, \dots, Z_t jsou definovány na stejném pravděpodobnostním prostoru (Ω, \mathcal{F}, P) . Potom vektor $\mathbf{Z}_t = (Z_1, \dots, Z_t)^\top$ nazveme náhodným vektorem.*

Definice 23. *Nechť (Ω, \mathcal{F}, P) je pravděpodobnostní prostor, nechť $T \subset \mathbb{R}$. Nechť pro každé $t \in T$ je \mathbf{Z}_t m -rozměrný náhodný vektor. Potom množinu reálných náhodných vektorů $\{\mathbf{Z}_t, t \in T\}$ definovaných na (Ω, \mathcal{F}, P) nazveme m -rozměrný náhodný proces.*

Definice 24. *m -rozměrný náhodný proces $\{\mathbf{Z}_t, t \in T\}$, kde $T \subset \mathbb{Z}$, nazveme m -rozměrnou časovou řadou.*

Definice 25. *Časovou řadu $\{\mathbf{Z}_t, t \in T\}$ nazveme silně stacionární, pokud je pravděpodobnostní rozdělení daného náhodného procesu invariantní vůči posunům v čase, tedy pro libovolné $n \in \mathbb{N}$, pro všechna $t_1, \dots, t_n \in T$ a pro libovolné $k \in \mathbb{Z}$ takové, že $t_i + k \in T$ pro $1 \leq i \leq n$, platí*

$$\mathcal{L}(\mathbf{Z}_{t_1}, \dots, \mathbf{Z}_{t_n}) = \mathcal{L}(\mathbf{Z}_{t_1+k}, \dots, \mathbf{Z}_{t_n+k}).$$

Definice 26. *Časovou řadu $\{\mathbf{Z}_t, t \in T\}$ s konečnými druhými momenty nazveme slabě stacionární, pokud pro všechna $s, t \in T$ a libovolné $k \in \mathbb{Z}$ takové, že $s+k, t+k \in T$, platí*

$$\begin{aligned} E(\mathbf{Z}_t) &= \boldsymbol{\mu} = \text{konst}, \\ \text{cov}(\mathbf{Z}_s, \mathbf{Z}_t) &= \text{cov}(\mathbf{Z}_{s+k}, \mathbf{Z}_{t+k}). \end{aligned}$$

Stejně jako v jednorozměrném případě budeme pod pojmem stacionarita rozumět slabou stacionaritu. Z její definice také vyplývá, že $\text{var}(\mathbf{Z}_t) = \Sigma_Z = \text{konst.}$ Pro vektor středních hodnot $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_m)^\top$ platí $\mu_i = \mathbb{E}(Z_{i,t})$. Nadále budeme uvažovat pouze stacionární časové řady.

Korelovanost vícerozměrných časových řad se popisuje pomocí maticových funkcí, které nyní definujeme dle [20], str. 333.

Definice 27. Maticovou autokovarianční funkci pro zpoždění k definujeme jako

$$\begin{aligned} \Gamma_k &= \text{cov}(\mathbf{Z}_t, \mathbf{Z}_{t+k}) = \mathbb{E}(\mathbf{Z}_t - \boldsymbol{\mu})(\mathbf{Z}_{t+k} - \boldsymbol{\mu})^\top = \\ &= \begin{pmatrix} \gamma_{11}(k) & \gamma_{12}(k) & \cdots & \gamma_{1m}(k) \\ \gamma_{21}(k) & \gamma_{22}(k) & \cdots & \gamma_{2m}(k) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{m1}(k) & \gamma_{m2}(k) & \cdots & \gamma_{mm}(k) \end{pmatrix}, \quad k = \dots, -1, 0, 1, \dots \end{aligned}$$

Pro $i \neq j$ označuje $\gamma_{ij}(k) = \mathbb{E}(Z_{i,t} - \mu_i)(Z_{j,t+k} - \mu_j)$ vzájemnou kovarianční funkci časových řad $Z_{i,t}$ a $Z_{j,t}$.

Prvky na diagonále uvedené matice, tedy v případě $i = j$, odpovídají autokovariančním funkcím jednorozměrných časových řad (viz definice 5). Tedy $\gamma_{ii}(k)$ je autokovarianční funkce časové řady $Z_{i,t}$. Jako speciální případ maticové autokovarianční funkce pro $k = 0$ dostaneme varianční matici $\Gamma_0 = \Sigma_Z$.

Definice 28. Maticovou autokorelační funkci pro zpoždění k definujeme jako

$$\boldsymbol{\rho}_k = D^{-1}\Gamma_k D^{-1} = \begin{pmatrix} \rho_{11}(k) & \rho_{12}(k) & \cdots & \rho_{1m}(k) \\ \rho_{21}(k) & \rho_{22}(k) & \cdots & \rho_{2m}(k) \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{m1}(k) & \rho_{m2}(k) & \cdots & \rho_{mm}(k) \end{pmatrix}, \quad k = \dots, -1, 0, 1, \dots,$$

kde D je diagonální matice $D = \text{diag}\{\sqrt{\gamma_{11}(0)}, \sqrt{\gamma_{22}(0)}, \dots, \sqrt{\gamma_{mm}(0)}\}$. Pro $i \neq j$ označuje

$$\rho_{ij}(k) = \frac{\gamma_{ij}(k)}{\sqrt{\gamma_{ii}(0) \cdot \gamma_{jj}(0)}}$$

vzájemnou korelační funkci časových řad $Z_{i,t}$ a $Z_{j,t}$.

Matice D v předchozí definici má na diagonále směrodatné odchylky časových řad $Z_{1,t}, \dots, Z_{m,t}$. Prvky na diagonále matice $\boldsymbol{\rho}_k$ jsou autokorelační funkce jednorozměrných časových řad (viz definice 6), tedy $\rho_{ii}(k)$ je autokorelační funkce časové řady $Z_{i,t}$.

Poznamenejme, že autokovarianční a autokorelační maticová funkce nejsou sudé funkce, jako tomu bylo v jednorozměrném případě. Platí totiž

$$\gamma_{ij}(k) = \mathbb{E}(Z_{i,t} - \mu_i)(Z_{j,t+k} - \mu_j) = \mathbb{E}(Z_{j,t+k} - \mu_j)(Z_{i,t} - \mu_i) = \gamma_{ji}(-k),$$

a tedy $\Gamma_k = \Gamma_{-k}^\top$ a $\boldsymbol{\rho}_k = \boldsymbol{\rho}_{-k}^\top$. Ze stejného vztahu také dostaneme, že Γ_k a $\boldsymbol{\rho}_k$ jsou symetrické matice pouze pro $k = 0$. Nadále budeme uvažovat jen $k \in \mathbb{N}_0$.

Předpokládejme, že máme k dispozici pozorování $\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_n$ mnohorozměrné časové řady. Z nich můžeme spočítat odhady střední hodnoty, autokovarianční a autokorelační maticové funkce následujícím způsobem.

Definice 29. Pro stacionární časovou řadu definujeme odhad střední hodnoty jako

$$\hat{\boldsymbol{\mu}} = \bar{\mathbf{Z}} = \frac{1}{n} \sum_{t=1}^n \mathbf{Z}_t.$$

Vektor $\bar{\mathbf{Z}}$ má složky $\bar{Z}_1, \dots, \bar{Z}_m$, kde

$$\bar{Z}_i = \frac{1}{n} \sum_{t=1}^n Z_{i,t}.$$

Definice 30. Odhad maticové autokovarianční funkce definujeme jako

$$\hat{\boldsymbol{\Gamma}}_k = \mathbf{C}_k = \frac{1}{n} \sum_{t=1}^{n-k} (\mathbf{Z}_t - \bar{\mathbf{Z}}) (\mathbf{Z}_{t+k} - \bar{\mathbf{Z}})^\top, \quad k = 0, 1, \dots, n-1.$$

Definice 31. Odhad maticové autokorelační funkce definujeme jako matici $\hat{\boldsymbol{\rho}}_k = \mathbf{R}_k$ s prvky

$$\hat{\rho}_{ij}(k) = r_{ij}(k) = \frac{\sum_{t=1}^{n-k} (Z_{i,t} - \bar{Z}_i) (Z_{j,t+k} - \bar{Z}_j)}{\sqrt{\sum_{t=1}^n (Z_{i,t} - \bar{Z}_i)^2 \sum_{t=1}^n (Z_{j,t} - \bar{Z}_j)^2}} = \frac{c_{ij}(k)}{\sqrt{c_{ii}(0) \cdot c_{jj}(0)}}$$

pro $k = 0, 1, \dots, n-1$, kde $c_{ij}(k)$ jsou prvky matice \mathbf{C}_k .

Podle hodnot vzájemné korelační funkce $\rho_{ij}(k)$ rozlišujeme několik typů lineární závislosti mezi řadami $\{Z_{i,t}\}$ a $\{Z_{j,t}\}$ v čase (dle [4], str. 420):

- Mezi řadami $\{Z_{i,t}\}$ a $\{Z_{j,t}\}$ neexistuje žádná lineární závislost (tj. jsou vzájemně nekorelované), pokud $\rho_{ij}(k) = \rho_{ji}(k) = 0$ pro všechna $k \geq 0$.
- Řady $\{Z_{i,t}\}$ a $\{Z_{j,t}\}$ jsou současně nekorelované, pokud $\rho_{ij}(0) = 0$.
- Řady $\{Z_{i,t}\}$ a $\{Z_{j,t}\}$ jsou nesoučasně nekorelované, když $\rho_{ij}(k) = \rho_{ji}(k) = 0$ pro všechna $k > 0$.
- Existuje jednosměrná závislost řady $\{Z_{i,t}\}$ na $\{Z_{j,t}\}$, pokud $\rho_{ij}(k) = 0$ pro všechna $k > 0$, ale $\rho_{ji}(l) \neq 0$ pro nějaké $l > 0$. V takovém případě $\{Z_{j,t}\}$ nezávisí na žádné minulé hodnotě $\{Z_{i,t}\}$, ale $\{Z_{i,t}\}$ závisí na nějaké minulé hodnotě $\{Z_{j,t}\}$.
- Existuje zpětná vazba mezi řadami $\{Z_{i,t}\}$ a $\{Z_{j,t}\}$, pokud $\rho_{ij}(k) \neq 0$ pro nějaké $k > 0$ a $\rho_{ji}(l) \neq 0$ pro nějaké $l > 0$.

Uveďme také zobecnění definic bílého šumu a lineárního procesu pro mnohorozměrný případ.

Definice 32. Časovou řadu $\{\boldsymbol{\varepsilon}_t\}$ nazveme mnohorozměrný bílý šum, pokud $\{\boldsymbol{\varepsilon}_t\}$ je posloupnost náhodných vektorů splňující

$$\begin{aligned} \mathbf{E}(\boldsymbol{\varepsilon}_t) &= \mathbf{0}, \\ \mathbf{E}(\boldsymbol{\varepsilon}_s \boldsymbol{\varepsilon}_t^\top) &= \begin{cases} \boldsymbol{\Sigma}_\varepsilon & \text{pro } s = t, \\ \mathbf{0} & \text{pro } s \neq t, \end{cases} \end{aligned} \quad (5.1)$$

kde $\boldsymbol{\Sigma}_\varepsilon$ je symetrická pozitivně definitní matice a $\mathbf{0}$ je matice s nulovými prvky.

Vztah (5.1) znamená, že složky vektorů bílého šumu jsou nesoučasně nekorelované, ale ve stejném čase mohou být korelované s rozptylovou maticí Σ_ε . V případě, že je matice Σ_ε diagonální, pak jsou složky bílého šumu současně nekorelované.

Definice 33. *Mnohorozměrný lineární proces definujeme jako nekonečnou časovou řadu*

$$\mathbf{Z}_t = \varepsilon_t + \Psi_1 \varepsilon_{t-1} + \Psi_2 \varepsilon_{t-2} + \dots = \sum_{j=0}^{\infty} \Psi_j \varepsilon_{t-j}, \quad (5.2)$$

kde ε_t je mnohorozměrný bílý šum a Ψ_j jsou maticové parametry.

Analogicky s jednorozměrným případem (viz definice 15), mnohorozměrný lineární proces představuje lineární kombinaci členů posloupnosti mnohorozměrného bílého šumu, ve které $\Psi_0 = \mathbf{I}$ je jednotková matice, a lze ho zapsat pomocí operátoru zpětného posunutí jako

$$\mathbf{Z}_t = (1 + \Psi_1 \mathbf{B} + \Psi_2 \mathbf{B}^2 + \dots) \varepsilon_t = \Psi(\mathbf{B}) \varepsilon_t,$$

kde $\Psi(\mathbf{B}) = \sum_{j=0}^{\infty} \Psi_j \mathbf{B}^j$. Předpokládejme, že matice Ψ_j mají prvky $\{\psi_{k,l,j}\}$, kde $k, l \in \{1, \dots, m\}$. Proces (5.2) je stacionární, jestliže platí

$$\sum_{j=0}^{\infty} \psi_{k,l,j}^2 < \infty \quad \forall k, l \in \{1, \dots, m\}.$$

Pokud je vícerozměrný lineární proces invertibilní, lze ho zapsat ve tvaru

$$\mathbf{Z}_t = \Pi_1 \mathbf{Z}_{t-1} + \Pi_2 \mathbf{Z}_{t-2} + \dots + \varepsilon_t,$$

$$\text{tj. } \varepsilon_t = \mathbf{Z}_t - \Pi_1 \mathbf{Z}_{t-1} - \dots = \Pi(\mathbf{B}) \mathbf{Z}_t,$$

kde $\Pi(\mathbf{B}) = \mathbf{I} - \sum_{j=1}^{\infty} \Pi_j \mathbf{B}^j$ a Π_j jsou maticové koeficienty s prvky $\{\pi_{k,l,j}\}$, kde $k, l \in \{1, \dots, m\}$. Podmínkou invertibility procesu (5.2) je

$$\sum_{j=0}^{\infty} |\pi_{k,l,j}| < \infty \quad \forall k, l \in \{1, \dots, m\}.$$

Poznámka. Ukažme podrobněji, jak funguje operátor zpětného posunutí ve vícerozměrném případě. Jeho použitím na náhodný vektor \mathbf{Z}_t dostaneme

$$\mathbf{B}\mathbf{Z}_t = \begin{pmatrix} B & 0 & \cdots & 0 \\ 0 & B & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & B \end{pmatrix} \begin{pmatrix} Z_{1,t} \\ Z_{2,t} \\ \vdots \\ Z_{m,t} \end{pmatrix} = \begin{pmatrix} Z_{1,t-1} \\ Z_{2,t-1} \\ \vdots \\ Z_{m,t-1} \end{pmatrix} = \mathbf{Z}_{t-1}.$$

5.2 Vektorový model klouzavých součtů

Definice 34. Řekneme, že časová řada $\{\mathbf{Z}_t\}$ se řídí vektorovým modelem klouzavých součtů řádu q , pokud platí

$$\mathbf{Z}_t = \boldsymbol{\varepsilon}_t + \boldsymbol{\Theta}_1 \boldsymbol{\varepsilon}_{t-1} + \dots + \boldsymbol{\Theta}_q \boldsymbol{\varepsilon}_{t-q} = \boldsymbol{\Theta}(\mathbf{B}) \boldsymbol{\varepsilon}_t,$$

kde $\boldsymbol{\varepsilon}_t$ je m -rozměrný bílý šum, $\boldsymbol{\Theta}_1, \dots, \boldsymbol{\Theta}_q$ jsou matice o rozměrech $m \times m$ ($\boldsymbol{\Theta}_q \neq \mathbf{0}$) a

$$\boldsymbol{\Theta}(\mathbf{B}) = \mathbf{I} + \boldsymbol{\Theta}_1 \mathbf{B} + \dots + \boldsymbol{\Theta}_q \mathbf{B}^q$$

je operátor klouzavých součtů.

Vektorový model klouzavých součtů VMA(q) je vždy stacionární. Označme $|\boldsymbol{\Theta}(\mathbf{B})|$ determinant maticového polynomu $\boldsymbol{\Theta}(\mathbf{B})$. Za podmínky, že všechny kořeny polynomu $|\boldsymbol{\Theta}(\mathbf{B})|$ leží vně jednotkového kruhu v komplexní rovině, je proces VMA(q) invertibilní. Jeho vektor středních hodnot je $\mathbf{E}(\mathbf{Z}_t) = \boldsymbol{\mu} = \mathbf{0}$. Maticová autokovarianční funkce má tvar

$$\begin{aligned} \boldsymbol{\Gamma}_k &= \mathbf{E}(\boldsymbol{\varepsilon}_t + \boldsymbol{\Theta}_1 \boldsymbol{\varepsilon}_{t-1} + \dots + \boldsymbol{\Theta}_q \boldsymbol{\varepsilon}_{t-q}) (\boldsymbol{\varepsilon}_{t+k}^\top + \boldsymbol{\varepsilon}_{t+k-1}^\top \boldsymbol{\Theta}_1^\top + \dots + \boldsymbol{\varepsilon}_{t+k-q}^\top \boldsymbol{\Theta}_q^\top) = \\ &= \begin{cases} \sum_{j=0}^{q-k} \boldsymbol{\Theta}_j \boldsymbol{\Sigma}_\varepsilon \boldsymbol{\Theta}_{j+k}^\top & \text{pro } k = 0, 1, \dots, q, \\ \mathbf{0} & \text{pro } k > q, \end{cases} \end{aligned} \quad (5.3)$$

kde $\boldsymbol{\Theta}_0 = \mathbf{I}$, a je useknutá v bodě q . Odtud dostaneme také vztah pro maticovou autokorelační funkci

$$\boldsymbol{\rho}_k = \mathbf{0} \quad \text{pro } k > q.$$

Uveďme jako příklad model VMA(1)

$$\mathbf{Z}_t = \boldsymbol{\varepsilon}_t + \boldsymbol{\Theta}_1 \boldsymbol{\varepsilon}_{t-1},$$

který lze například pro $m = 2$ zapsat maticově jako

$$\begin{pmatrix} Z_{1,t} \\ Z_{2,t} \end{pmatrix} = \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix} + \begin{pmatrix} \theta_{11} & \theta_{12} \\ \theta_{21} & \theta_{22} \end{pmatrix} \begin{pmatrix} \varepsilon_{1,t-1} \\ \varepsilon_{2,t-1} \end{pmatrix}$$

nebo pomocí rovnic

$$\begin{aligned} Z_{1,t} &= \varepsilon_{1,t} + \theta_{11} \varepsilon_{1,t-1} + \theta_{12} \varepsilon_{2,t-1}, \\ Z_{2,t} &= \varepsilon_{2,t} + \theta_{21} \varepsilon_{1,t-1} + \theta_{22} \varepsilon_{2,t-1}. \end{aligned}$$

Z tohoto zápisu vidíme, že mimodiagonální prvky matice $\boldsymbol{\Theta}_1$ vyjadřují vzájemnou časovou závislost mezi řadami $\{\varepsilon_{1,t}\}$ a $\{\varepsilon_{2,t}\}$. Jak uvádí [20] na str. 344, podmínka invertibility procesu VMA(1) je ekvivalentní podmínce, že vlastní čísla matice $\boldsymbol{\Theta}_1$ leží uvnitř jednotkového kruhu.

Maticová autokovarianční funkce (5.3) modelu VMA(1) se zjednoduší na

$$\boldsymbol{\Gamma}_k = \begin{cases} \boldsymbol{\Sigma}_\varepsilon + \boldsymbol{\Theta}_1 \boldsymbol{\Sigma}_\varepsilon \boldsymbol{\Theta}_1^\top & \text{pro } k = 0, \\ \boldsymbol{\Sigma}_\varepsilon \boldsymbol{\Theta}_1^\top & \text{pro } k = 1, \\ \mathbf{0} & \text{pro } k > 1. \end{cases} \quad (5.4)$$

5.3 Vektorový autoregresní model

Definice 35. Řekneme, že časová řada $\{\mathbf{Z}_t\}$ se řídí vektorovým autoregresním modelem řádu p , pokud platí

$$\mathbf{Z}_t = \Phi_1 \mathbf{Z}_{t-1} + \dots + \Phi_p \mathbf{Z}_{t-p} + \varepsilon_t, \quad (5.5)$$

$$\text{tj. } \mathbf{Z}_t - \Phi_1 \mathbf{Z}_{t-1} - \dots - \Phi_p \mathbf{Z}_{t-p} = \Phi(\mathbf{B}) \mathbf{Z}_t = \varepsilon_t,$$

kde ε_t je m -rozměrný bílý šum, Φ_1, \dots, Φ_p jsou matice o rozměru $m \times m$ ($\Phi_p \neq \mathbf{0}$) a

$$\Phi(\mathbf{B}) = \mathbf{I} - \Phi_1 \mathbf{B} - \dots - \Phi_p \mathbf{B}^p$$

je autoregresní operátor.

Vektorový autoregresní model VAR(p) je vždy invertibilní a za podmínky, že všechny kořeny polynomu $|\Phi(\mathbf{B})|$ leží vně jednotkového kruhu v komplexní rovině, také stacionární s nulovou střední hodnotou.

Ukažme nyní, jak vypadá model VAR(1). Jeho tvar je

$$\mathbf{Z}_t = \Phi_1 \mathbf{Z}_{t-1} + \varepsilon_t. \quad (5.6)$$

Nechť například $m = 2$. Potom model (5.6) zapíšeme maticově jako

$$\begin{pmatrix} Z_{1,t} \\ Z_{2,t} \end{pmatrix} = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix} \begin{pmatrix} Z_{1,t-1} \\ Z_{2,t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix}$$

nebo pomocí rovnic

$$Z_{1,t} = \phi_{11} Z_{1,t-1} + \phi_{12} Z_{2,t-1} + \varepsilon_{1,t}, \quad (5.7)$$

$$Z_{2,t} = \phi_{21} Z_{1,t-1} + \phi_{22} Z_{2,t-1} + \varepsilon_{2,t}. \quad (5.8)$$

Z rovnic 5.7, 5.8 a z tvaru kovarianční matice 5.9 je patrná souvislost mimo-diagonálních prvků matice Φ_1 se závislostí mezi řadami $\{Z_{1,t}\}$ a $\{Z_{2,t}\}$ v čase. Pokud je navíc matice Σ_ε diagonální, pak platí následující vztahy:

- Pokud $\phi_{12} \neq 0$ a $\phi_{21} \neq 0$, pak existuje zpětná vazba mezi řadami $\{Z_{1,t}\}$ a $\{Z_{2,t}\}$.
- Pokud $\phi_{12} \neq 0$ a zároveň $\phi_{21} = 0$, existuje jednosměrná závislost řady $\{Z_{1,t}\}$ na $\{Z_{2,t}\}$.
- Pokud $\phi_{12} = \phi_{21} = 0$, pak jsou řady $\{Z_{1,t}\}$ a $\{Z_{2,t}\}$ nesoučasně nekorelované.

Jak je ukázáno v [20] na str. 339-340, podmínka stacionarity procesu VAR(1) je ekvivalentní podmínce, že vlastní čísla matice Φ_1 leží uvnitř jednotkového kruhu. Maticovou autokovarianční funkci modelu (5.6) dostaneme jako

$$\begin{aligned} \Gamma_k &= \mathbb{E}(\mathbf{Z}_t \mathbf{Z}_{t+k}^\top) = \\ &= \mathbb{E}(\mathbf{Z}_t (\Phi_1 \mathbf{Z}_{t+k-1} + \varepsilon_{t+k})^\top) = \\ &= \mathbb{E}(\mathbf{Z}_t \mathbf{Z}_{t+k-1}^\top \Phi_1^\top + \mathbf{Z}_t \varepsilon_{t+k}^\top) = \\ &= \begin{cases} \Gamma_{-1} \Phi_1^\top + \Sigma_\varepsilon & \text{pro } k = 0, \\ \Gamma_{k-1} \Phi_1^\top = \Gamma_0 (\Phi_1^\top)^k & \text{pro } k \geq 1. \end{cases} \end{aligned} \quad (5.9)$$

V poslední rovnosti pro $k \geq 1$ jsme využili, že $E(\mathbf{Z}_t \boldsymbol{\varepsilon}_{t+k}^\top) = \mathbf{0}$, neboť $\mathbf{Z}_t = \boldsymbol{\Phi}^{-1}(\mathbf{B}) \boldsymbol{\varepsilon}_t$ je lineární proces vytvořený z hodnot $\boldsymbol{\varepsilon}_t, \boldsymbol{\varepsilon}_{t-1}, \dots$ a náhodné vektory $\boldsymbol{\varepsilon}_{t+k}$ a \mathbf{Z}_t jsou pro $k \geq 1$ nekorelované.

Pro obecný model VAR(p) lze obdobně jako v jednorozměrném případě sestavit soustavu *Yule-Walkerových rovnic*. Pokud vynásobíme vztah (5.5) hodnotou \mathbf{Z}_{t-k}^\top pro $k > 0$ a přejdeme ke středním hodnotám, dostaneme rovnici

$$\boldsymbol{\Gamma}_k = \boldsymbol{\Gamma}_{k-1} \boldsymbol{\Phi}_1^\top + \dots + \boldsymbol{\Gamma}_{k-p} \boldsymbol{\Phi}_p^\top,$$

která pro $k = 1, \dots, p$ tvoří příslušnou soustavu.

5.4 Vektorový smíšený model VARMA

Definice 36. Řekneme, že časová řada $\{\mathbf{Z}_t\}$ se řídí vektorovým smíšeným modelem VARMA řádů p a q , pokud platí

$$\mathbf{Z}_t = \boldsymbol{\Phi}_1 \mathbf{Z}_{t-1} + \dots + \boldsymbol{\Phi}_p \mathbf{Z}_{t-p} + \boldsymbol{\varepsilon}_t + \boldsymbol{\Theta}_1 \boldsymbol{\varepsilon}_{t-1} + \dots + \boldsymbol{\Theta}_q \boldsymbol{\varepsilon}_{t-q},$$

$$\text{tj. } \boldsymbol{\Phi}(\mathbf{B}) \mathbf{Z}_t = \boldsymbol{\Theta}(\mathbf{B}) \boldsymbol{\varepsilon}_t,$$

kde $\boldsymbol{\varepsilon}_t$ je m -rozměrný bílý šum, $\boldsymbol{\Phi}_1, \dots, \boldsymbol{\Phi}_p, \boldsymbol{\Theta}_1, \dots, \boldsymbol{\Theta}_q$ jsou matice o rozměru $m \times m$ ($\boldsymbol{\Phi}_p \neq \mathbf{0}$ a $\boldsymbol{\Theta}_q \neq \mathbf{0}$),

$$\boldsymbol{\Phi}(\mathbf{B}) = \mathbf{I} - \boldsymbol{\Phi}_1 \mathbf{B} - \dots - \boldsymbol{\Phi}_p \mathbf{B}^p$$

je autoregresní operátor a

$$\boldsymbol{\Theta}(\mathbf{B}) = \mathbf{I} + \boldsymbol{\Theta}_1 \mathbf{B} + \dots + \boldsymbol{\Theta}_q \mathbf{B}^q$$

je operátor klouzavých součtů.

Postačující podmínka stacionarity modelu VARMA(p, q) obdobně jako v jednorozměrném případě odpovídá podmínce stacionarity modelu VAR(p) a podmínka invertibility modelu VARMA(p, q) odpovídá podmínce invertibility modelu VMA(q).

Jako příklad uveďme model VARMA(1,1), který má tvar

$$\mathbf{Z}_t = \boldsymbol{\Phi}_1 \mathbf{Z}_{t-1} + \boldsymbol{\varepsilon}_t + \boldsymbol{\Theta}_1 \boldsymbol{\varepsilon}_{t-1}. \quad (5.10)$$

Mějme $m = 2$. Potom lze ekvivalentně psát

$$\begin{pmatrix} Z_{1,t} \\ Z_{2,t} \end{pmatrix} = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix} \begin{pmatrix} Z_{1,t-1} \\ Z_{2,t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix} + \begin{pmatrix} \theta_{11} & \theta_{12} \\ \theta_{21} & \theta_{22} \end{pmatrix} \begin{pmatrix} \varepsilon_{1,t-1} \\ \varepsilon_{2,t-1} \end{pmatrix},$$

neboli

$$\begin{aligned} Z_{1,t} &= \phi_{11} Z_{1,t-1} + \phi_{12} Z_{2,t-1} + \varepsilon_{1,t} + \theta_{11} \varepsilon_{1,t-1} + \theta_{12} \varepsilon_{2,t-1}, \\ Z_{2,t} &= \phi_{21} Z_{1,t-1} + \phi_{22} Z_{2,t-1} + \varepsilon_{2,t} + \theta_{21} \varepsilon_{1,t-1} + \theta_{22} \varepsilon_{2,t-1}. \end{aligned}$$

Přenásobením rovnice (5.10) zpožděnou hodnotou \mathbf{Z}_{t-k}^\top dostaneme

$$\mathbf{Z}_t \mathbf{Z}_{t-k}^\top = \Phi_1 \mathbf{Z}_{t-1} \mathbf{Z}_{t-k}^\top + \varepsilon_t \mathbf{Z}_{t-k}^\top + \Theta_1 \varepsilon_{t-1} \mathbf{Z}_{t-k}^\top$$

a po přechodu ke středním hodnotám

$$\Gamma_{-k} = \Phi_1 \Gamma_{-k+1} + E(\varepsilon_t \mathbf{Z}_{t-k}^\top) + \Theta_1 E(\varepsilon_{t-1} \mathbf{Z}_{t-k}^\top).$$

S využitím vztahu $\Gamma_k = \Gamma_{-k}^\top$ lze psát

$$\Gamma_k^\top = \Phi_1 \Gamma_{k-1}^\top + E(\varepsilon_t \mathbf{Z}_{t-k}^\top) + \Theta_1 E(\varepsilon_{t-1} \mathbf{Z}_{t-k}^\top),$$

odkud dostaneme vyjádření maticové autokovarianční funkce modelu (5.10)

$$\Gamma_k = \begin{cases} \Gamma_1^\top \Phi_1^\top + \Sigma_\varepsilon + (\Phi_1 + \Theta_1) \Sigma_\varepsilon \Theta_1^\top & \text{pro } k = 0, \\ \Gamma_0^\top \Phi_1^\top + \Sigma_\varepsilon \Theta_1^\top & \text{pro } k = 1, \\ \Gamma_{k-1} \Phi_1^\top & \text{pro } k \geq 2. \end{cases}$$

Kapitola 6

Identifikace modelů mnohorozměrných časových řad

Postup identifikace modelů mnohorozměrných časových řad je obdobný jednorozměrnému případu. Po grafické prohlídce se provedou případné transformace vedoucí ke stacionarizaci řady a homogenizaci jejího rozptylu. Dále lze použít některou z identifikačních metod, jako jsou odhady maticových korelačních funkcí nebo informační kritéria.

6.1 Identifikace pomocí maticové autokorelační funkce

Nejprve se podíváme na identifikaci modelu pro mnohorozměrnou časovou řadu pomocí maticové autokorelační funkce ρ_k . Lze ji odhadnout z dat dle definice 31 a porovnat její chování s teoretickými protějšky. Body useknutí k_0 jednotlivých složek $r_{ij}(k)$ odhadnuté matice lze dle [20], str. 350 testovat pomocí Bartlettovy aproximace s asymptotickým kritickým oborem na hladině významnosti 5 %

$$|r_{ij}(k)| \geq 2\sqrt{\frac{1}{n-k} \left(1 + 2 \sum_{l=1}^{k_0} r_{ii}(l)r_{jj}(l)\right)} \quad \text{pro nějaké } |k| > k_0.$$

V případě, že řada $\{\mathbf{Z}_t\}$ je vícerozměrný bílý šum, uvedený kritický obor se zjednoduší na

$$|r_{ij}(k)| \geq 2\sqrt{\frac{1}{n-k}} \quad \text{pro nějaké } |k| > k_0, \quad (6.1)$$

přičemž jmenovatel $(n-k)$ uvedených zlomků se pro delší časové řady nahrazuje výrazem n . Autoři [17] navrhli využít zjednodušený kritický obor (6.1) a nahradit jednotlivé prvky odhadnutých maticových autokorelačních funkcí $r_{ij}(k)$ symboly

$$\begin{aligned} + & \quad \text{pro } r_{ij}(k) > \frac{2}{\sqrt{n}}, \\ - & \quad \text{pro } r_{ij}(k) < -\frac{2}{\sqrt{n}}, \\ \cdot & \quad \text{jinak.} \end{aligned}$$

V dalším kroku lze na jejich základě stanovit bod useknutí maticové autokorelační funkce. Uvedený postup je vhodný pro identifikaci vektorového modelu klouzavých součtů VMA(q), neboť jeho teoretická maticová autokorelační funkce má bod useknutí $k_0 = q$.

6.2 Identifikace pomocí parciální korelační maticové funkce

Podobně jako u jednorozměrných časových řad lze i vektorové procesy studovat z hlediska parciální korelace jejich složek. Za tímto účelem definujeme dle [20] (str. 356-358) parciální korelační maticovou funkci pro zpoždění k , pomocí které se popisuje korelace náhodných vektorů \mathbf{Z}_t a \mathbf{Z}_{t+k} po vyloučení lineární závislosti na vektorech $\mathbf{Z}_{t+1}, \dots, \mathbf{Z}_{t+k-1}$ ležících v čase mezi nimi.

Definice 37. *Zavedme vektory*

$$\begin{aligned} \mathbf{u}_{k-1,t+k} &= \mathbf{Z}_{t+k} - \boldsymbol{\alpha}_{k-1,1}\mathbf{Z}_{t+k-1} - \boldsymbol{\alpha}_{k-1,2}\mathbf{Z}_{t+k-2} - \dots - \boldsymbol{\alpha}_{k-1,k-1}\mathbf{Z}_{t+1} \\ \mathbf{v}_{k-1,t} &= \mathbf{Z}_t - \boldsymbol{\beta}_{k-1,1}\mathbf{Z}_{t+1} - \boldsymbol{\beta}_{k-1,2}\mathbf{Z}_{t+2} - \dots - \boldsymbol{\beta}_{k-1,k-1}\mathbf{Z}_{t+k-1}, \end{aligned}$$

kde $\boldsymbol{\alpha}_{k-1,s}, \boldsymbol{\beta}_{k-1,s}$ pro $s = 1, \dots, k-1$ jsou matice minimalizující $\mathbb{E}|\mathbf{u}_{k-1,t+k}|^2$, resp. $\mathbb{E}|\mathbf{v}_{k-1,t}|^2$. Dále označme matice

$$\begin{aligned} V_{\mathbf{u}}(k) &= \text{var}(\mathbf{u}_{k-1,t+k}), \\ V_{\mathbf{v}}(k) &= \text{var}(\mathbf{v}_{k-1,t}), \\ V_{\mathbf{vu}}(k) &= \text{cov}(\mathbf{v}_{k-1,t}, \mathbf{u}_{k-1,t+k}), \end{aligned}$$

diagonální matici $D_{\mathbf{u}}(k)$, jejíž i -tý diagonální prvek je odmocnina i -tého diagonálního prvku matice $V_{\mathbf{u}}(k)$ a diagonální matici $D_{\mathbf{v}}(k)$, která je odvozena stejným způsobem z matice $V_{\mathbf{v}}(k)$. Parciální korelační maticovou funkci pro zpoždění $k = 1, 2, \dots$ definujeme jako

$$\boldsymbol{\phi}_k = [D_{\mathbf{v}}(k)]^{-1}V_{\mathbf{vu}}(k)[D_{\mathbf{u}}(k)]^{-1}. \quad (6.2)$$

Autor [20] na str. 359-361 uvádí rekurentní postup výpočtu matice $\boldsymbol{\phi}_k$. Pro $k = 1$ máme

$$\begin{aligned} V_{\mathbf{u}}(1) &= V_{\mathbf{v}}(1) = \boldsymbol{\Gamma}_0, \\ V_{\mathbf{vu}}(1) &= \boldsymbol{\Gamma}_1, \\ \boldsymbol{\alpha}_{1,1} &= \boldsymbol{\Gamma}_1^\top (\boldsymbol{\Gamma}_0)^{-1}, \\ \boldsymbol{\beta}_{1,1} &= \boldsymbol{\Gamma}_1 (\boldsymbol{\Gamma}_0)^{-1} \end{aligned}$$

a pro $k \geq 2$ a $s = 1, \dots, k-1$

$$\begin{aligned} V_{\mathbf{u}}(k) &= \boldsymbol{\Gamma}_0 - \sum_{s=1}^{k-1} \boldsymbol{\alpha}_{k-1,s} \boldsymbol{\Gamma}_s, \\ V_{\mathbf{v}}(k) &= \boldsymbol{\Gamma}_0 - \sum_{s=1}^{k-1} \boldsymbol{\beta}_{k-1,s} \boldsymbol{\Gamma}_s^\top, \\ V_{\mathbf{vu}}(k) &= \boldsymbol{\Gamma}_k - \sum_{s=1}^{k-1} \boldsymbol{\Gamma}_{k-s} \boldsymbol{\alpha}_{k-1,s}^\top, \\ \boldsymbol{\alpha}_{k,k} &= V_{\mathbf{vu}}^\top(k) (V_{\mathbf{v}}(k))^{-1}, \\ \boldsymbol{\alpha}_{k,s} &= \boldsymbol{\alpha}_{k-1,s} - \boldsymbol{\alpha}_{k,k} \boldsymbol{\beta}_{k-1,k-s}, \\ \boldsymbol{\beta}_{k,k} &= V_{\mathbf{vu}}(k) (V_{\mathbf{u}}(k))^{-1}, \\ \boldsymbol{\beta}_{k,s} &= \boldsymbol{\beta}_{k-1,s} - \boldsymbol{\beta}_{k,k} \boldsymbol{\alpha}_{k-1,k-s}. \end{aligned}$$

Matice $D_u(k)$ a $D_v(k)$ vypočítáme dle definice 37 a parciální korelační maticovou funkci získáme dosazením do (6.2).

Uvedený rekurentní postup je zároveň metodou pro výpočet odhadu $\hat{\phi}_k$ funkce ϕ_k , pokud bychom místo Γ_s brali $\hat{\Gamma}_s$ dle definice 30 pro $s = 0, \dots, k-1$. Tyto odhady jsou užitečné pro identifikaci autoregresního modelu $\text{VAR}(p)$, neboť jeho teoretická parciální korelační maticová funkce je useknutá v bodě $k_0 = p$. Pokud označíme prvky odhadnuté matice $\hat{\phi}_k$ jako $r_{ij}(kk)$, pak můžeme jejich body useknutí testovat pomocí Quenouilleovy aproximace s kritickým oborem na hladině významnosti 5 %

$$|r_{ij}(kk)| \geq 2\sqrt{\frac{1}{n}} \quad \text{pro nějaké } |k_0| > k.$$

Poznámka. Alternativně lze k určení řádu modelu $\text{VAR}(p)$ využít také *parciální autoregresní maticovou funkci*. V knize [20] na str. 351-355 je definována jako maticový koeficient $\Phi_{k,k}$ v lineárním regresním modelu

$$\mathbf{Z}_{t+k} = \Phi_{k,1}\mathbf{Z}_{t+k-1} + \Phi_{k,2}\mathbf{Z}_{t+k-2} + \dots + \Phi_{k,k}\mathbf{Z}_t + \mathbf{e}_{k,t+k}.$$

Stejný zdroj uvádí, že pro model $\text{VAR}(p)$ mají parciální autoregresní maticová funkce a parciální korelační maticová funkce stejný bod useknutí $k_0 = p$.

6.3 Identifikace pomocí informačních kritérií

Předchozí postupy vyžadují subjektivní zásah badatele, proto často vedou k různým výsledkům a je obtížné je automatizovat. Navíc neumožňují identifikaci smíšeného modelu $\text{VARMA}(p,q)$. Stejně jako v jednorozměrném případě se však i zde nabízí přístup pomocí informačních kritérií. Ten spočívá ve výběru nejmenší z předem daného počtu hodnot, a vede tak k jednoznačnému závěru. Uveďme nyní definice nejpoužívanějších informačních kritérií pro vektorový proces $\text{VARMA}(p,q)$ dle [12], str. 147-150:

- *Akaikeho informační kritérium (AIC)* ve tvaru

$$\text{AIC} = \ln |\hat{\Sigma}_\varepsilon| + \frac{2(p+q)m^2}{n},$$

kde n je počet pozorování, $|\hat{\Sigma}_\varepsilon|$ je determinant odhadnuté varianční matice bílého šumu a m je rozměr vektorového procesu \mathbf{Z}_t ,

- *Schwartzovo Bayesovské informační kritérium (BIC)* ve tvaru

$$\text{BIC} = \ln |\hat{\Sigma}_\varepsilon| + \frac{(p+q)m^2 \ln n}{n},$$

- *Hannanovo-Quinnovo informační kritérium (HQ)* ve tvaru

$$\text{HQ} = \ln |\hat{\Sigma}_\varepsilon| + \frac{2(p+q)m^2 \ln \ln n}{n}.$$

Poznamenejme, že kritéria AIC, BIC a HQ jsou přímou analogií svých jednorozměrných verzí uvedených v kapitolách 2.5.1 a 2.5.3.

Kapitola 7

Simulace mnohorozměrných časových řad

V této kapitole ukážeme fungování identifikačních kritérií na simulovaných mnohorozměrných časových řadách. Využijeme jak autokorelační maticovou funkci a parciální korelační maticovou funkci, tak Akaikeho a Schwartzovo Bayesovské informační kritérium. K výpočtům použijeme program R.

7.1 Simulace procesu VAR(1)

Nejprve budeme generovat dvourozměrnou časovou řadu délky $n = 1000$, která se řídí vektorovým autoregresním modelem řádu 1 s maticovým parametrem

$$\Phi_1 = \begin{pmatrix} 0,8 & 0,3 \\ 0,1 & 0,6 \end{pmatrix}$$

a normálním bílým šumem s jednotkovou varianční maticí, tj. $\varepsilon_t \sim \mathbf{N}(\mathbf{0}, \mathbf{I})$. Model lze zapsat jako

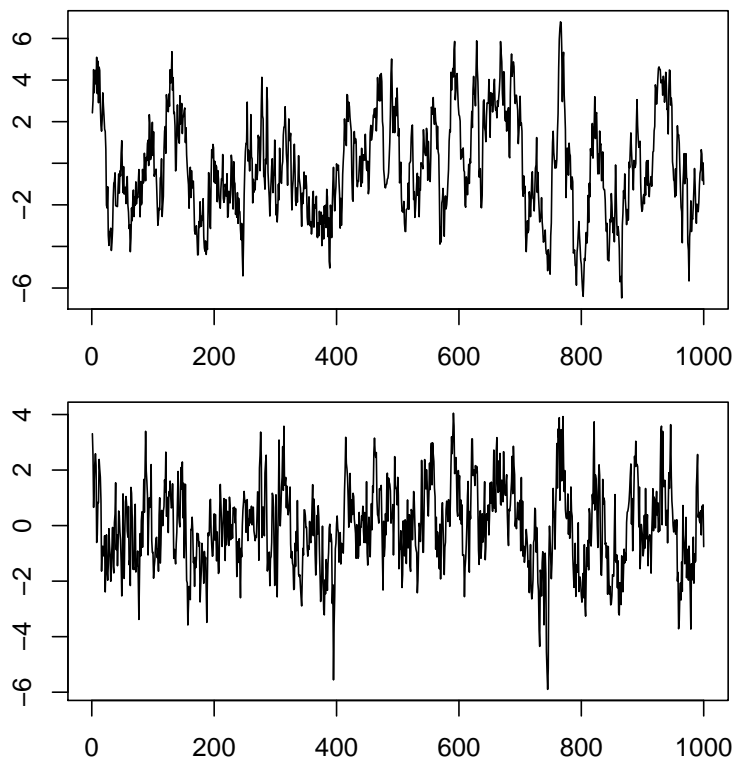
$$\begin{pmatrix} Z_{1,t} \\ Z_{2,t} \end{pmatrix} = \begin{pmatrix} 0,8 & 0,3 \\ 0,1 & 0,6 \end{pmatrix} \begin{pmatrix} Z_{1,t-1} \\ Z_{2,t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix}.$$

Autoregresní proces je invertibilní a vzhledem k tomu, vlastní čísla matice Φ_1 (0,9 a 0,5) leží uvnitř jednotkového kruhu v komplexní rovině, také stacionární. Trajektorie obou složek simulované časové řady jsou zobrazeny v grafu 7.1.

Odhadnutá autokorelační maticová funkce nemá bod useknutí, naproti tomu výběrová parciální korelační maticová funkce je useknutá v bodě $k_0 = 1$, jak je vidět v odhadnutých korelogramech na obrázcích 7.2 a 7.3. Tyto odhady autokorelačních funkcí tedy vedou na identifikaci modelu VAR(1).

Pomocí Akaikeho informačního kritéria i Schwartzova Bayesovského informačního kritéria bychom taktéž identifikovali vektorový autoregresní model řádu 1. Hodnoty obou těchto kritérií totiž vycházejí pro model VAR(1) nejmenší, jak je vidět v tabulce 7.1.

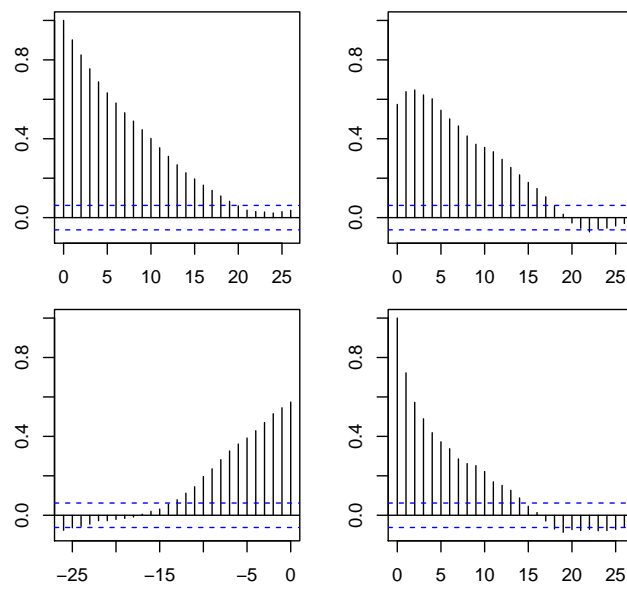
Poznámka. Pro úplnost uveďme význam jednotlivých grafů na obrázku 7.2. Horní levý graf zobrazuje odhadnutou autokorelační funkci $\rho_{11}(k)$ jednorozměrné časové řady $\{Z_{1,t}\}$ a dolní pravý graf odhadnutou autokorelační funkci $\rho_{22}(k)$ časové řady $\{Z_{2,t}\}$. Ve zbylých dvou grafech je vykreslen odhad vzájemné autokorelační funkce $\rho_{12}(k)$ obou těchto řad, přičemž platí $\rho_{12}(k) = \rho_{21}(-k)$.



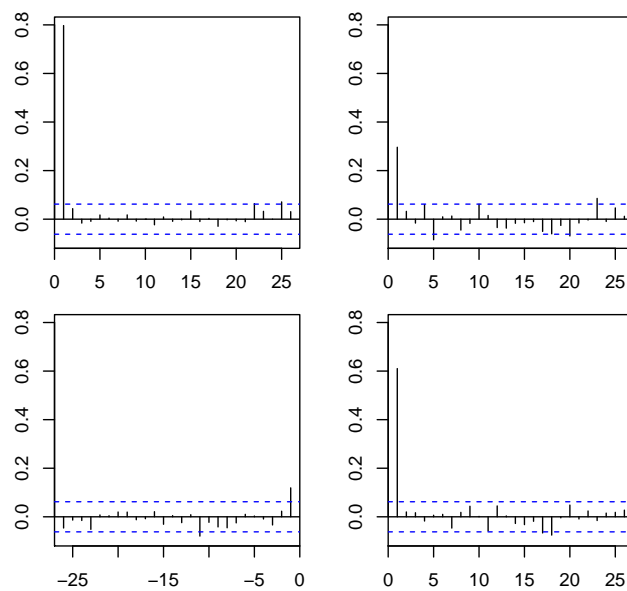
Obrázek 7.1: Simulovaná dvourozměrná časová řada VAR(1)

	AIC	BIC
VAR(1)	0,009	0,029
VAR(2)	0,016	0,055
VMA(1)	1,142	1,161
VARMA(1,1)	0,014	0,053
VARMA(1,2)	0,022	0,081

Tabulka 7.1: Hodnoty informačních kritérií



Obrázek 7.2: Odhadnutá autokorelační maticová funkce



Obrázek 7.3: Odhadnutá parciální korelační maticová funkce

7.2 Simulace procesu VMA(1)

Nyní přistoupíme k simulaci dvourozměrné časové řady délky $n = 1000$ dle vektorového modelu klouzavých součtů řádu 1 s maticovým parametrem

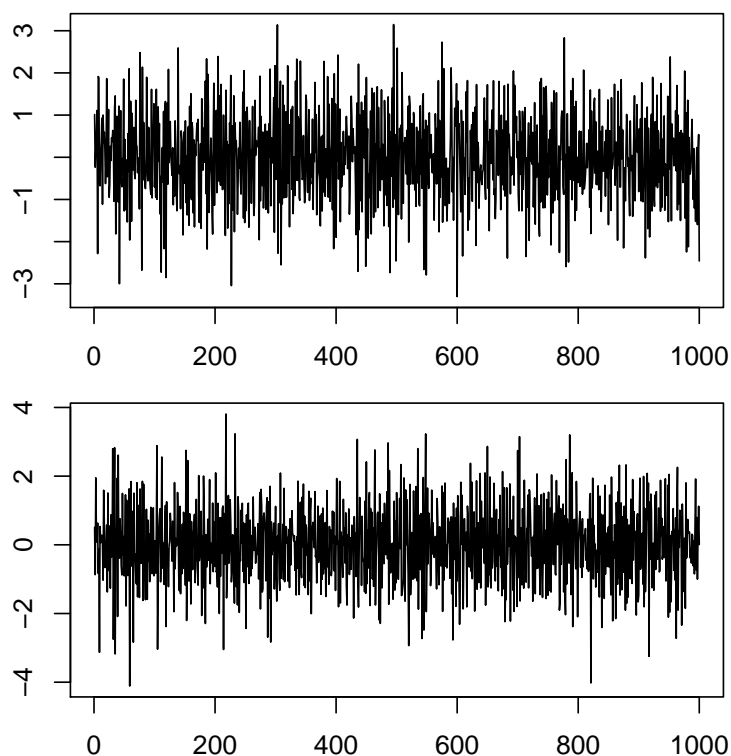
$$\Theta_1 = \begin{pmatrix} 0,4 & 0,2 \\ 0,3 & 0,6 \end{pmatrix}.$$

Pro model lze zvolit zápis

$$\begin{pmatrix} Z_{1,t} \\ Z_{2,t} \end{pmatrix} = \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix} + \begin{pmatrix} 0,4 & 0,2 \\ 0,3 & 0,6 \end{pmatrix} \begin{pmatrix} \varepsilon_{1,t-1} \\ \varepsilon_{2,t-1} \end{pmatrix}.$$

Bílý šum byl generován z rozdělení $N(\mathbf{0}, \mathbf{I})$.

Proces klouzavých součtů je stacionární a navíc také invertibilní, jelikož vlastní čísla matice Θ_1 (0,76 a 0,24) leží uvnitř jednotkového kruhu v komplexní rovině. Graf 7.4 zobrazuje trajektorie obou složek simulované časové řady.

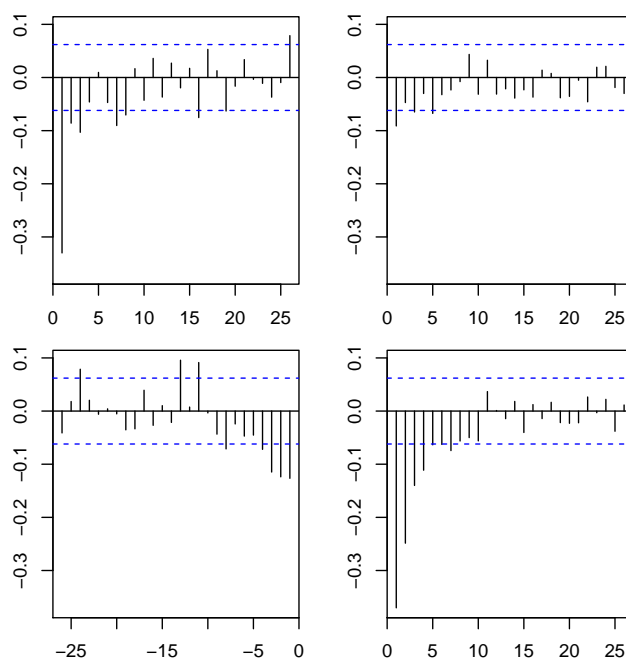


Obrázek 7.4: Simulovaná dvourozměrná časová řada VMA(1)

Uveďme v tomto případě namísto odhadnutého korelogramu druhý obvyklý způsob vyobrazení hodnot odhadnuté maticové autokorelační funkce zmíněný v kapitole 6.1, a to pomocí matic se symboly $+$, $-$ a \cdot . Tabulka 7.2 zobrazuje odhady maticové ACF pro zpoždění $k = 1, \dots, 8$ společně s příslušnými symboly, na základě kterých určíme bod useknutí. Pokud bychom ignorovali hodnotu $\hat{\rho}_{11}(3) \doteq -0,07$, která je jen nepatrně menší než příslušná mez $-\frac{2}{\sqrt{n}} = -\frac{2}{\sqrt{1000}} \doteq -0,063$, za bod useknutí odhadnuté maticové autokorelační funkce lze považovat $k_0 = 1$. Naproti tomu jednotlivé grafy odhadnuté parciální korelační maticové funkce na obrázku 7.5 připomínají spíše křivku U . Proto bychom se na základě těchto odhadnutých funkcí přiklonili k modelu VMA(1).

k	1	2	3	4
$\hat{\rho}_k$	$\begin{pmatrix} -0,35 & -0,17 \\ -0,19 & -0,40 \\ - & - \\ - & - \end{pmatrix}$	$\begin{pmatrix} 0,06 & 0,04 \\ -0,01 & -0,05 \\ \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} -0,07 & -0,04 \\ -0,01 & 0,02 \\ - & \cdot \\ \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} 0,03 & 0,03 \\ 0,02 & -0,01 \\ \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}$
k	5	6	7	8
$\hat{\rho}_k$	$\begin{pmatrix} 0,02 & -0,04 \\ 0,00 & 0,01 \\ \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} -0,04 & 0,01 \\ -0,02 & -0,01 \\ \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} -0,04 & 0,00 \\ 0,02 & -0,02 \\ \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} -0,01 & 0,02 \\ -0,03 & 0,01 \\ \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}$

Tabulka 7.2: Odhadnutá autokorelační maticová funkce



Obrázek 7.5: Odhadnutá parciální korelační maticová funkce

Ke správné identifikaci modelu vedl také přístup pomocí Akaikeho a Schwarzova Bayesovského informačního kritéria. V tabulce 7.3 vidíme, že obě kritéria nabyla své minimální hodnoty právě pro model VMA(1).

	AIC	BIC
VMA(1)	-0,055	-0,036
VMA(2)	-0,050	-0,010
VARMA(1,1)	-0,051	-0,012
VARMA(1,2)	-0,046	0,012
VARMA(2,1)	-0,046	0,013

Tabulka 7.3: Hodnoty informačních kritérií

7.3 Simulace procesu VARMA(1,1)

Na závěr budeme generovat dvourozměrnou časovou řadu délky $n = 1000$ dle modelu VARMA(1,1) s maticovými parametry

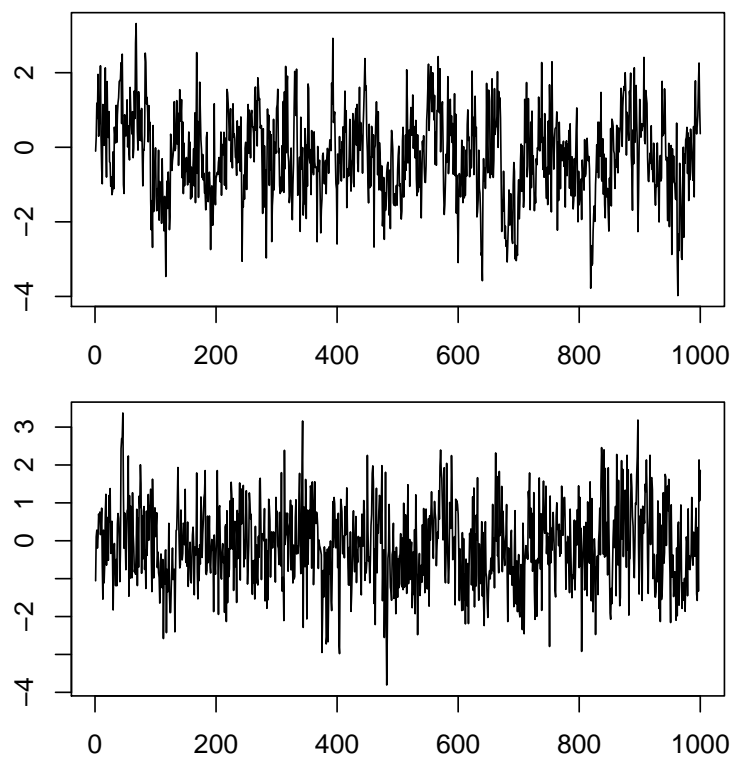
$$\Phi_1 = \begin{pmatrix} 0,8 & 0,1 \\ 0,2 & 0,7 \end{pmatrix} \quad \text{a} \quad \Theta_1 = \begin{pmatrix} 0,5 & 0,3 \\ 0,1 & 0,6 \end{pmatrix}$$

a bílým šumem $\varepsilon_t \sim N(\mathbf{0}, \mathbf{I})$. Tvar uvažovaného modelu je

$$\begin{pmatrix} Z_{1,t} \\ Z_{2,t} \end{pmatrix} = \begin{pmatrix} 0,8 & 0,1 \\ 0,2 & 0,7 \end{pmatrix} \begin{pmatrix} Z_{1,t-1} \\ Z_{2,t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix} + \begin{pmatrix} 0,5 & 0,3 \\ 0,1 & 0,6 \end{pmatrix} \begin{pmatrix} \varepsilon_{1,t-1} \\ \varepsilon_{2,t-1} \end{pmatrix}.$$

Vlastní čísla matice Φ_1 jsou 0,9 a 0,6. Jelikož obě leží uvnitř jednotkového kruhu v komplexní rovině, proces je stacionární. To samé platí i pro vlastní čísla matice Θ_1 (0,73 a 0,37), a proto je proces také invertibilní. Graf 7.6 ukazuje trajektorie obou složek simulované časové řady.

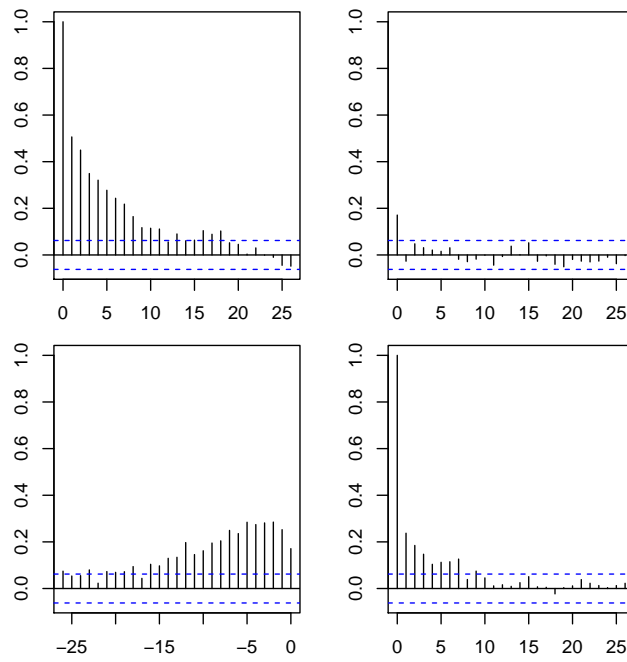
Grafy odhadnuté autokorelační maticové funkce i parciální korelační maticové funkce na obrázcích 7.7 a 7.8 připomínají křivku U , a proto nevedou k jednoznačné identifikaci vhodného modelu. Nabízejí se tedy informační kritéria AIC a BIC, která se v tomto případě neshodují. Akaikeho kritérium nadhodnotilo řády modelu a jako nejvhodnější zvolilo model VARMA(2,2). Naproti tomu BIC správně identifikovalo model VARMA(1,1), viz tabulka 7.4.



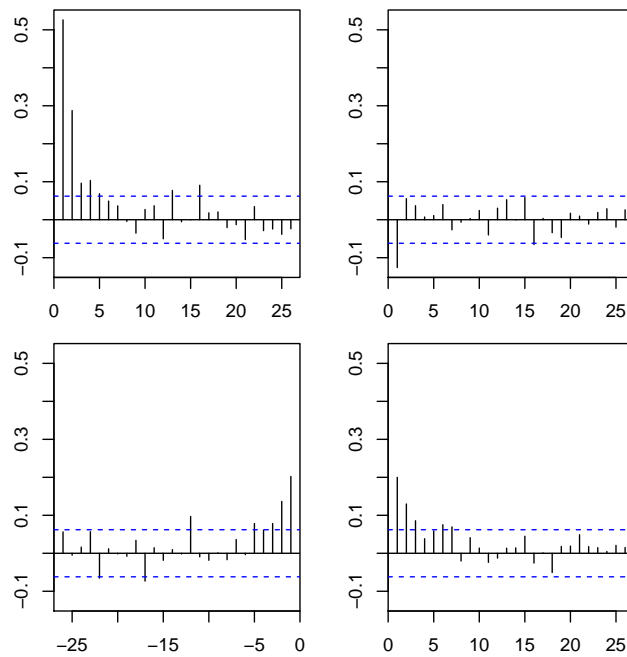
Obrázek 7.6: Simulovaná dvourozměrná časová řada VARMA(1,1)

	AIC	BIC
VAR(1)	0,0142	0,0338
VAR(2)	-0,1146	-0,0753
VMA(1)	0,1658	0,1855
VMA(2)	0,0282	0,0675
VARMA(1,1)	-0,1780	-0,1387
VARMA(1,2)	-0,1737	-0,1148
VARMA(2,1)	-0,1739	-0,1150
VARMA(2,2)	-0,1828	-0,1043

Tabulka 7.4: Hodnoty informačních kritérií



Obrázek 7.7: Odhadnutá autokorelační maticová funkce



Obrázek 7.8: Odhadnutá parciální korelační maticová funkce

7.4 Porovnání identifikačních kritérií

V předchozím textu jsme na simulovaných dvourozměrných časových řadách ukázali možné přístupy k identifikaci modelu. Ukázalo se, že identifikace pomocí odhadnutých maticových autokorelačních funkcí je mírně komplikovanější než u jednorozměrných časových řad. Za bod useknutí maticové autokorelační funkce se totiž považuje takový bod, za kterým jsou nulové všechny prvky této matice. U časových řad větších rozměrů tak může být volba bodu useknutí již velmi subjektivní záležitost. Jeho nalezení lze provést dvěma způsoby – podle grafu příslušné maticové funkce nebo podle tabulky se symboly $+$, $-$ a \cdot . Druhá možnost je pro časové řady velkých rozměrů přehlednější, ale její nevýhodou je, že v tabulce na první pohled nepoznáme, o kolik jsou hodnoty zastoupené symboly $+$ a $-$ větší či menší než příslušná mez. Tuto informaci nám dávají grafy maticových autokorelačních funkcí, a proto je výhodné oba přístupy kombinovat.

Identifikace modelu pomocí informačních kritérií funguje na stejném principu jako v jednorozměrném případě. Spočívá ve výběru minima z předem daného počtu hodnot, a proto vždy vede k jednoznačnému výsledku.

Podívejme nyní se podrobněji na praktické použití Akaikeho a Schwartzova Bayesovského informačního kritéria. Podobně jako v kapitole 3.4 porovnáme jejich fungování na 100 simulovaných dvourozměrných časových řadách dle modelů VAR(1), VMA(1) a VARMA(1,1) délky $n = 1000$. Budeme pozorovat, které kritérium vedlo ke správné identifikaci příslušného modelu ve větším počtu případů.

Tabulka 7.5 ukazuje počty identifikovaných modelů pro simulovanou časovou řadu dle modelu VAR(1) ve tvaru

$$\begin{pmatrix} Z_{1,t} \\ Z_{2,t} \end{pmatrix} = \begin{pmatrix} 0,8 & 0,3 \\ 0,1 & 0,6 \end{pmatrix} \begin{pmatrix} Z_{1,t-1} \\ Z_{2,t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix}.$$

Vidíme, že Akaikeho kritérium správně určilo model v 83 % případů, naproti tomu Schwartzovo Bayesovské kritérium bylo tentokrát bezchybné.

	AIC	BIC
VAR(1)	83	100
VAR(2)	5	0
VARMA(1,1)	4	0
VARMA(1,2)	3	0
VARMA(2,1)	4	0
VARMA(2,2)	1	0

Tabulka 7.5: Počty identifikovaných modelů dle AIC a BIC

Dále jsme generovali 100 dvourozměrných časových řad dle modelu klouzavých součtů ve tvaru

$$\begin{pmatrix} Z_{1,t} \\ Z_{2,t} \end{pmatrix} = \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix} + \begin{pmatrix} 0,4 & 0,2 \\ 0,3 & 0,6 \end{pmatrix} \begin{pmatrix} \varepsilon_{1,t-1} \\ \varepsilon_{2,t-1} \end{pmatrix}.$$

Schwartzovo Bayesovské kritérium opět vedlo u každé řady k identifikaci správného modelu. Akaikeho kritérium tentokrát bylo méně úspěšné než v případě autoregresního modelu. Rozpoznalo model VMA(1) jen v 66 % případů. Podrobnější výsledky jsou uvedeny v tabulce 7.6.

	AIC	BIC
VMA(1)	66	100
VMA(2)	7	0
VARMA(1,1)	6	0
VARMA(1,2)	13	0
VARMA(2,1)	6	0
VARMA(2,2)	2	0

Tabulka 7.6: Počty identifikovaných modelů dle AIC a BIC

Nakonec jsme generovali dvourozměrné časové řady podle smíšeného modelu VARMA(1,1) ve tvaru

$$\begin{pmatrix} Z_{1,t} \\ Z_{2,t} \end{pmatrix} = \begin{pmatrix} 0,8 & 0,1 \\ 0,2 & 0,7 \end{pmatrix} \begin{pmatrix} Z_{1,t-1} \\ Z_{2,t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix} + \begin{pmatrix} 0,5 & 0,3 \\ 0,1 & 0,6 \end{pmatrix} \begin{pmatrix} \varepsilon_{1,t-1} \\ \varepsilon_{2,t-1} \end{pmatrix}.$$

Podle tabulky 7.7 Akaikeho kritérium identifikovalo správný model VARMA(1,1) u 71 % simulací, ve zbylých případech byly nadhodnoceny řady modelu. Schwartzovo Bayesovské informační kritérium opět u všech časových řad identifikovalo model správně.

	AIC	BIC
VARMA(1,1)	71	100
VARMA(1,2)	4	0
VARMA(2,1)	6	0
VARMA(2,2)	19	0

Tabulka 7.7: Počty identifikovaných modelů dle AIC a BIC

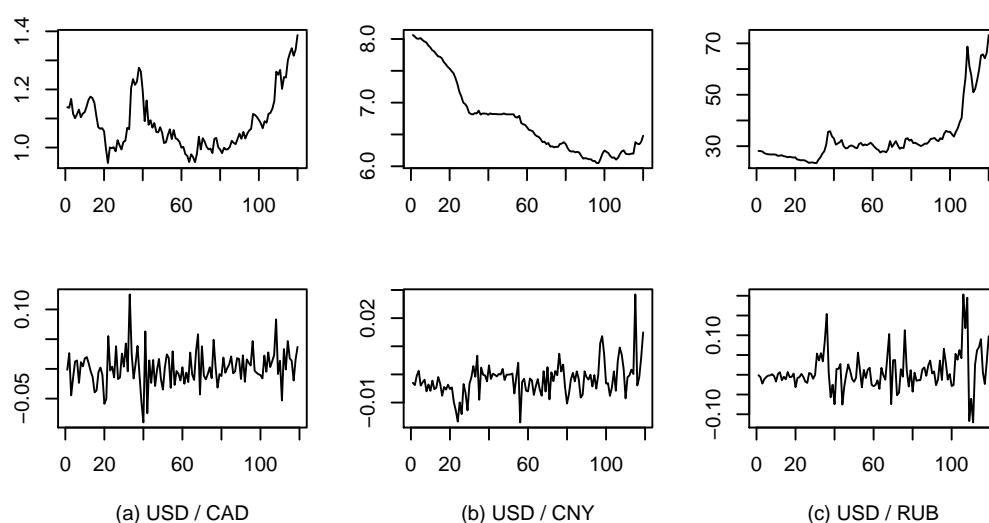
Ověřili jsme, že i v případě mnohorozměrných časových řad je přesnější Bayesovské než Akaikeho informační kritérium, které má tendenci nadhodnocovat řady modelu. Identifikace pomocí BIC byla tentokrát u 100 % simulací úspěšná. Co se týče AIC, největšího procenta správných identifikací dosáhlo u modelu VAR(1), naopak nejméně přesné bylo v identifikaci modelu VMA(1).

Kapitola 8

Softwarové zpracování mnohorozměrné časové řady

Na závěr přistoupíme k ukázce identifikace modelu reálné mnohorozměrné finanční časové řady. Bude se jednat o trojrozměrnou řadu měsíčních měnových kurzů USD¹/CAD², USD/CNY³ a USD/RUB⁴ za období od ledna 2006 do prosince 2015 čítající $n = 120$ pozorování, vždy z posledního dne každého měsíce. Data byla získána z webové stránky www.quandl.com. Jednotlivé složky analyzované řady jsou vykresleny v obrázku 8.1. Pro dosažení stacionarity jsme vypočetli první diference logaritmů každé ze složek.

U většího počtu časových řad se samozřejmě nabízí možnost modelovat je odděleně každou zvlášť. Výhodnější však je nahlížet na ně jako na mnohorozměrnou časovou řadu a modelovat je společně. Díky tomuto přístupu lze sledovat jejich vzájemné vztahy a vazby v čase. Jedna z řad může například ovlivňovat ostatní s určitým časovým zpožděním. Jak uvádějí autoři [17], další výhodou je větší přesnost předpovědí budoucích hodnot.



Obrázek 8.1: Řady měnových kurzů před a po stacionarizaci

¹ americký dolar ² kanadský dolar ³ čínský jüan ⁴ ruský rubl

Odhadnuté autokorelační funkce trojrozměrné časové řady v tabulce 8.1 a na obrázku 8.2 jsou již poměrně komplikovaným nástrojem. Pokud se podíváme pouze na autokorelační funkce jednorozměrných časových řad (v tabulce 8.1 se jedná o diagonální prvky matic a na obrázku 8.2 jsou to tři grafy na hlavní diagonále), mohli bychom předpokládat možné použití autoregresního modelu nízkého řádu vzhledem k bodům useknutí jednotlivých parciálních korelačních funkcí.

Dvě použitá informační kritéria – Akaikeho a Schwartzovo Bayesovské – se v tomto případě neshodují. Jak vidíme v tabulce 8.2, nejnižší hodnota AIC vychází pro model VARMA(1,1), zatímco BIC vyšlo nejmenší pro model VAR(1). Vzhledem k větší úspěšnosti BIC v kapitole 7 zvolíme model VAR(1), který má také jednodušší interpretaci a méně parametrů oproti VARMA(1,1). Uvádíme odhad maticového parametru Φ_1 a p -hodnoty testů nulovosti jednotlivých odhadnutých prvků matice.

$$\hat{\Phi}_1 = \begin{pmatrix} -0,23 & 0,10 & 0,15 \\ 0,01 & 0,36 & 0,01 \\ -0,28 & -0,66 & 0,43 \end{pmatrix} \quad p\text{-hodnoty: } \begin{pmatrix} 0,03 & 0,82 & 0,03 \\ 0,49 & <0,01 & 0,43 \\ 0,10 & 0,36 & <0,01 \end{pmatrix}$$

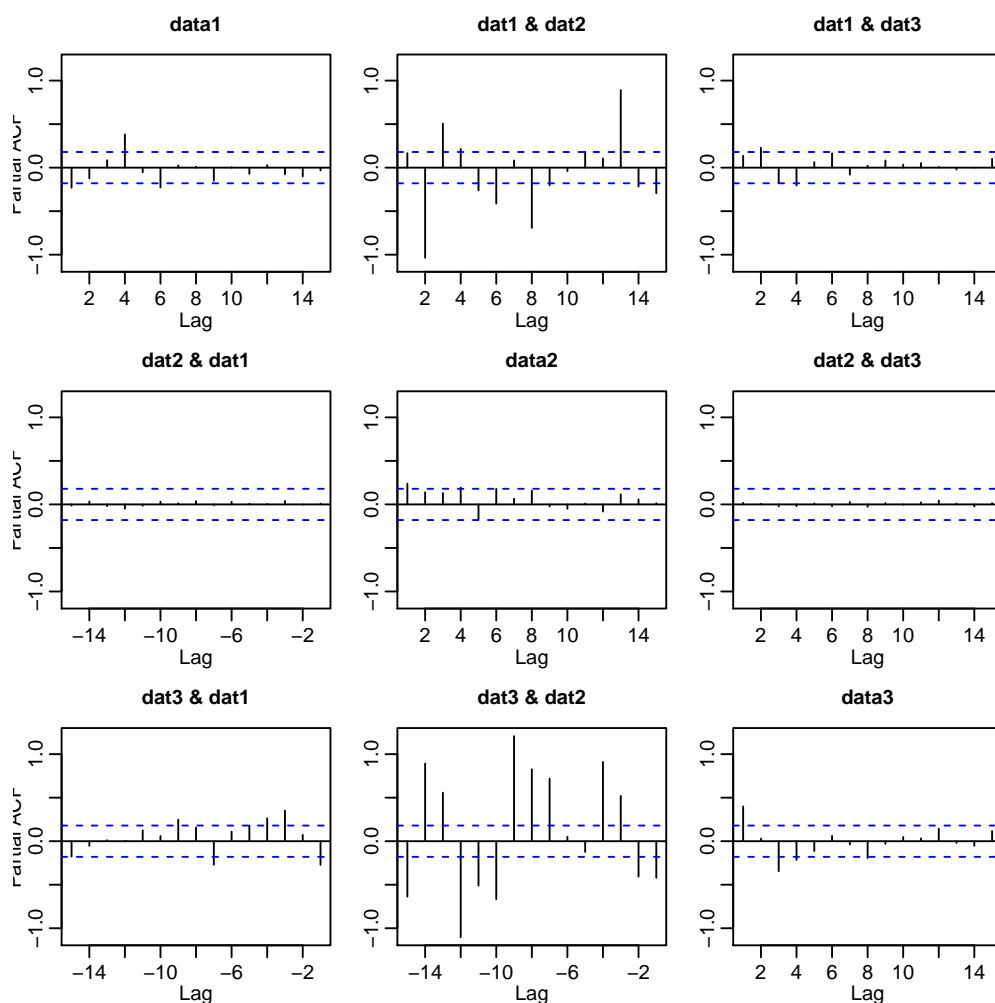
V matici p -hodnot vidíme, že na hladině 5 % nezamítáme nulovost odhadů na diagonále. Jedná se o odhady parametrů, které charakterizují závislost jednotlivých jednorozměrných řad na svých vlastních zpožděných hodnotách. Z mimodiagonálních prvků byla zamítnuta nulovost pouze odhadu $\hat{\phi}_{13} \doteq 0,15$, který souvisí se závislostí měnového kurzu USD/CAD na zpožděných hodnotách kurzu USD/RUB. Z důvodu nezamítnutí nulovosti odhadů ostatních parametrů nastavíme jejich hodnoty pevně na nulu a vypočítáme odhady zbylých parametrů znovu. Ty se proto mohou lišit od předešlých. Výsledný tvar modelu je

$$\begin{pmatrix} Z_{1,t} \\ Z_{2,t} \\ Z_{3,t} \end{pmatrix} = \begin{pmatrix} -0,13 & 0 & 0,11 \\ 0 & 0,41 & 0 \\ 0 & 0 & 0,29 \end{pmatrix} \begin{pmatrix} Z_{1,t-1} \\ Z_{2,t-1} \\ Z_{3,t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \varepsilon_{3,t} \end{pmatrix}. \quad (8.1)$$

Mohli bychom také pracovat s oběma alternativními modely VAR(1) a VARMA(1,1) a finální volbu provést až v průběhu verifikace jejich porovnáním z hlediska splnění předpokladů. Odhadu parametrů a verifikační části výstavby modelu se v této práci blíže nevěnujeme. Touto problematikou včetně použitého testu nulovosti odhadů parametrů se podrobně zabývá kapitola 14 knihy [20].

k	1	2	3	4
$\hat{\rho}_k$	$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & + & + \\ \cdot & \cdot & + \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot & + \\ \cdot & + & + \\ \cdot & \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} + & \cdot & \cdot \\ \cdot & + & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$
k	5	6	7	8
$\hat{\rho}_k$	$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & + & + \\ \cdot & \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$

Tabulka 8.1: Odhadnutá autokorelační maticová funkce



Obrázek 8.2: Odhadnutá parciální korelační maticová funkce

	AIC	BIC
VAR(1)	-23,823	-23,612
VAR(2)	-23,832	-23,412
VMA(1)	-23,747	-23,537
VMA(2)	-23,831	-23,410
VARMA(1,1)	-23,970	-23,549
VARMA(1,2)	-23,855	-23,225
VARMA(2,1)	-23,842	-23,211
VARMA(2,2)	-23,892	-22,051

Tabulka 8.2: Hodnoty informačních kritérií

Závěr

Cílem práce bylo popsat vybrané matematické modely vhodné k analýze finančních časových řad a jejich identifikační kritéria a dále ověřit fungování kritérií na simulovaných časových řadách a reálných datech z finanční praxe.

V první polovině práce byly představeny jednorozměrné lineární modely AR-MA a různé přístupy k jejich identifikaci. Při analýze reálných dat je prvním krokem obvykle grafická prohlídka časové řady a provedení potřebných transformací jako např. Boxova-Coxova transformace nebo diferencování. O potřebě diferencování se lze rozhodnout pomocí Dickeyova-Fullerova testu, jehož fungování bylo ověřeno na simulovaných časových řadách. Ukázalo se, že pro krátké časové řady má test malou sílu, ale pro delší řady může být poměrně spolehlivým ukazatelem nutnosti diferencování i v případě autoregresní posloupnosti s téměř jednotkovým kořenem.

Z vlastních identifikačních kritérií jsme se zabývali autokorelační funkcí, parciální autokorelační funkcí a inverzní autokorelační funkcí, které jsou vhodné k identifikaci modelů AR a MA, nicméně model ARMA z nich rozpoznat nedokážeme. Alternativou je rozšířená výběrová autokorelační funkce, podle které je možné určit i řád modelu ARMA. Tato kritéria založená na korelační struktuře vyžadují subjektivní posouzení situace a často poskytují nejednoznačný výsledek. Proto bychom doporučili použít je spíše pro orientační pohled na chování časové řady a konečnou volbu modelu a jeho řádu provést pomocí informačních kritérií. Jejich fungování bylo taktéž ověřeno pomocí simulovaných časových řad. Ukázalo se, že Schwartzovo Bayesovské kritérium bylo u všech typů modelů přesnější než Akaikeho kritérium, které častěji nadhodnocovalo řády modelu. Identifikační kritéria jednorozměrných modelů ARMA byla následně aplikována na reálnou finanční časovou řadu. Šlo o kurzy akcií Bank of America Corporation.

Ve druhé polovině práce jsme se zabývali mnohorozměrnými lineárními modely ARMA a jejich identifikačními kritérii, a to maticovými autokorelačními funkcemi a informačními kritérii. Maticové autokorelační funkce jsou mnohem méně přehledné než autokorelační funkce v jednorozměrném případě. Proto bychom je doporučili pouze jako orientační nástroj a pro finální volbu modelu bychom dali přednost informačním kritériím. Ta jsme také použili k identifikaci modelů simulovaných časových řad a zjistili jsme, že Akaikeho kritérium bylo opět méně spolehlivé než Schwartzovo Bayesovské. K ukázce identifikace mnohorozměrného modelu pro reálná data jsme použili trojrozměrnou řadu měsíčních měnových kurzů USD/CAD, USD/CNY a USD/RUB. Zvolený model naznačuje, že se jednotlivé časové řady navzájem výrazně neovlivňují.

Seznam použité literatury

- [1] AKAIKE, H. *A Bayesian extension of the minimum AIC procedure*. Annals of the Institute of Statistical Mathematics, Vol. 30, Part A (Dec., 1978), pp. 9-14.
- [2] AKAIKE, H. *A Bayesian extension of the minimum AIC procedure of autoregressive model fitting*. Biometrika, Vol. 66, No. 2. (Aug., 1979), pp. 237-242.
- [3] BOX, G. E. P., JENKINS, G. M., REINSEL, G. C. *Time series analysis: forecasting and control*. 4th ed. New Jersey, John Wiley 2008. ISBN 978-0-470-27284-8.
- [4] CIPRA, T. *Finanční ekonometrie*. Ekopress, Praha 2008. ISBN 978-80-86929-43-9.
- [5] DICKEY, D. A., FULLER, W. A. *Distribution of the estimators for autoregressive time series with a unit root*. Journal of the American Statistical Association, Vol. 74, No. 366 (Jun., 1979), pp. 427-431.
- [6] DICKEY, D. A., FULLER, W. A. *Likelihood ratio statistics for autoregressive time series with a unit root*. Econometrica, Vol. 49, No. 4 (Jul., 1981), pp. 1057-1072.
- [7] FULLER, W. A. *Introduction to statistical time series*. 2nd ed. Wiley, New York 1996. ISBN 0-471-55239-9.
- [8] HANNAN, E. J., QUINN, B. G. *The determination of the order of an autoregression*. Journal of the Royal Statistical Society. Series B (Methodological), Vol. 41, No. 2 (1979), pp. 190-195.
- [9] HYNDMAN, R. *forecast: Forecasting functions for time series and linear models*. R package version 7.1, 2016.
- [10] CHAN, K., RIPLEY, B. *TSA: Time series analysis*. R package version 1.01, 2012.
- [11] KWIATKOWSKI, D., PHILLIPS, P. C. B., SCHMIDT, P., SHIN, Y. *Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?*. Journal of Econometrics, Vol. 54, No. 1-3 (1992), pp. 159-178.
- [12] LÜTKEPOHL, H. *New introduction to multiple time series analysis*. Berlin, Springer-Verlag 2005. ISBN 3-540-40172-5.

- [13] MILLS, T. *A very British affair. Six Britons and the development of time series analysis during the 20th century.* Palgrave Macmillan UK, 2013. ISBN 978-1-137-29126-4.
- [14] PARZEN, E. *Multiple time series modeling: Determining the order of approximating autoregressive schemes.* Multivariate Analysis IV, (1977), pp. 283-295.
- [15] PFAFF, B., ZIVOT, E., STIGLER, M. *urca: Unit root and cointegration tests for time series data.* R package version 1.2-9, 2016.
- [16] SHIBATA, R. *Selection of the order of an autoregressive model by Akaike's information criterion.* Biometrika, Vol. 63, No. 1 (Apr., 1976), pp. 117-126.
- [17] TIAO, G. C., BOX, G. E. P. *Modeling multiple times series with applications.* Journal of the American Statistical Association, Vol. 76, No. 376 (Dec., 1981), pp. 802-816.
- [18] TSAY, R. S., TIAO, G. C. *Consistent estimates od autoregressive parameters and extended sample autocorrelation function for stationary and non-stationary ARIMA models.* Journal of the American Statistical Association, Vol. 79, No. 385 (Mar., 1984), pp. 84-96.
- [19] TSAY, R. S. *MTS: All-purpose toolkit for analyzing multivariate time series (MTS) and estimating multivariate volatility models.* R package version 0.33, 2015.
- [20] WEI, W. W. S. *Time series analysis: univariate and multivariate methods.* Addison-Wesley, Redwood City 1990. ISBN 0-201-15911-2.

Seznam obrázků

3.1	Simulovaná časová řada AR(1) a odhadnuté autokorelační funkce	28
3.2	Simulovaná časová řada MA(1) a odhadnuté autokorelační funkce	30
3.3	Simulovaná časová řada ARMA(1,1) a odhadnuté autokorelační funkce	32
4.1	Řada kurzů akcií, její transformace a autokorelační funkce	38
4.2	Odhadnuté autokorelační funkce	39
7.1	Simulovaná dvourozměrná časová řada VAR(1)	52
7.2	Odhadnutá autokorelační maticová funkce	53
7.3	Odhadnutá parciální korelační maticová funkce	53
7.4	Simulovaná dvourozměrná časová řada VMA(1)	54
7.5	Odhadnutá parciální korelační maticová funkce	55
7.6	Simulovaná dvourozměrná časová řada VARMA(1,1)	57
7.7	Odhadnutá autokorelační maticová funkce	58
7.8	Odhadnutá parciální korelační maticová funkce	58
8.1	Řady měnových kurzů před a po stacionarizaci	61
8.2	Odhadnutá parciální korelační maticová funkce	63

Seznam tabulek

2.1	Volba typového parametru λ v Boxově-Coxově transformaci	13
2.2	Kvantily testové statistiky τ -testu	16
2.3	Kvantily testové statistiky τ_μ -testu	16
2.4	Kvantily testové statistiky τ_τ -testu	16
2.5	Vlastnosti ACF a PACF stacionárních a invertibilních procesů AR, MA a ARMA	18
2.6	Procesy AR, MA a ARMA a jejich inverzní procesy	19
2.7	Rozšířená výběrová autokorelační funkce	22
2.8	Asymptotická ESACF modelu ARMA(1,1)	22
3.1	Rozšířená výběrová autokorelační funkce v modelu AR(1)	28
3.2	Hodnoty informačních kritérií	29
3.3	Rozšířená výběrová autokorelační funkce v modelu MA(1)	29
3.4	Hodnoty informačních kritérií	30
3.5	Rozšířená výběrová autokorelační funkce v modelu ARMA(1,1)	31
3.6	Hodnoty informačních kritérií	31
3.7	Počty identifikovaných modelů dle AIC a BIC	33
3.8	Počty identifikovaných modelů dle AIC a BIC	34
3.9	Počty identifikovaných modelů dle AIC a BIC	34
3.10	Procento zamítnutých nulových hypotéz Dickeyova-Fullerova testu	35
3.11	Dosažená síla τ -testu	36
3.12	Dosažená síla τ_μ -testu	36
3.13	Dosažená síla τ_τ -testu	36
4.1	Rozšířená výběrová autokorelační funkce	39
4.2	Hodnoty informačních kritérií	39
7.1	Hodnoty informačních kritérií	52
7.2	Odhadnutá autokorelační maticová funkce	55
7.3	Hodnoty informačních kritérií	56
7.4	Hodnoty informačních kritérií	57
7.5	Počty identifikovaných modelů dle AIC a BIC	59
7.6	Počty identifikovaných modelů dle AIC a BIC	60
7.7	Počty identifikovaných modelů dle AIC a BIC	60
8.1	Odhadnutá autokorelační maticová funkce	62
8.2	Hodnoty informačních kritérií	63