Univerzita Karlova v Praze

Matematicko-fyzikální fakulta

BAKALÁŘSKÁ PRÁCE

Hana Tritová

Maximálně věrohodné odhady v časových řadách

Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce:	RNDr. Zbyněk Pawlas, Ph.D.
Studijní program:	Matematika
Studijní obor:	obecná matematika

Praha 2011

Tímto bych chtěla poděkovat vedoucímu práce, RNDr. Zbyňku Pawlasovi, Ph.D., za vstřícnost, trpělivost a cenné rady při psaní této práce, Kláře Holkové za psychickou podporu a připomínky a své rodině za perfektní zázemí při psaní této práce i po celou dobu studia.

Prohlašuji, že jsem tuto bakalářskou práci vypracovala 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 Hana Tritová

Název práce: Maximálně věrohodné odhady v časových řadách

Autor: Hana Tritová

Katedra: Katedra pravděpodobnosti a matematické statistiky (KPMS)

Vedoucí bakalářské práce: RNDr. Zbyněk Pawlas, Ph.D., KPMS

Abstrakt: Práce se zabývá maximálně věrohodnými odhady v časových řadách. Čtenář se seznámí se třemi základními modely časových řad: autoregresní posloupností (AR), posloupností klouzavých součtů (MA) a jejich kombinací (ARMA). Dále zjistí, jak vypadají jejich základní charakteristiky, např. střední hodnota nebo rozptyl. Pak zde nalezne odvození odhadů parametrů metodou maximální věrohodnosti – obecně a ve zmíněných modelech časových řad. Pro modely AR(1) a MA(1) jsou uvedeny ještě odhady metodou momentů a metodou nejmenších čtverců a závěr je věnován příkladům, které slouží ke srovnání všech tří metod.

Klíčová slova: maximálně věrohodný odhad, časová řada, autoregresní posloupnost, posloupnost klouzavých součtů

Title: Maximum Likelihood Estimators in Time Series

Author: Hana Tritová

Department: Department of Probability and Mathematical Statistics (KPMS)

Supervisor: RNDr. Zbyněk Pawlas, Ph.D., KPMS

Abstract: The thesis deals with maximum likelihood estimators in time series. The reader becomes familiar with three important models for time series: autoregressive model (AR), moving average model (MA) and autoregressive moving average (ARMA). Thereafter he can find out the form of their main characteristics, e.g. population mean and variance. Then there is the derivation of parameter estimates – generally and for mentioned models of times series. There are also stated two other methods for finding estimators of AR(1) and MA(1) parameters – method of moments and least squares method. The end is dedicated to examples which compares all three methods.

Keywords: maximum likelihood estimation, time series, autoregressive process, moving average process

Obsah

Úvod

Cílem této práce je odvodit maximálně věrohodné odhady parametrů pro konkrétní typy časových řad a získané výsledky srovnat s odhady pomocí jiných metod.

Maximálně věrohodný odhad (MLE) je jedním z odhadů parametrů, které se hojně používají ve statistice. Na úvod se tedy budeme věnovat tomu, jak takový odhad vytvořit. Pak se zaměříme na časové řady. Podrobněji se budeme věnovat třem lineárním modelům časových řad: posloupnostem klouzavých součtů, autoregresním posloupnostem a kombinaci těchto dvou modelů – ukážeme, jak vypadají a jaké mají základní vlastnosti. Na závěr uvedeme další dvě metody, pomocí kterých se dají odhadnout parametry (nejen) v časových řadách: metodu momentů a metodu nejmenších čtverců. Pro srovnání uvedeme pár příkladů.

1. Základní pojmy

1.1 Maximálně věrohodný odhad

Jak je uvedeno v [1], metoda maximální věrohodnosti je metoda konstrukce bodových odhadů. Nejprve tedy definujeme, co je bodový odhad.

Definice. Nechť $\mathcal{B}^k$ značí systém borelovských množin v $\mathbb{R}^k, \Theta \in \mathcal{B}^k$ a $\boldsymbol{\theta} \in \Theta$ je k-rozměrný vektor neznámých parametrů. Nechť $\mathbf{X} = (X_1, \ldots, X_n)^T$ je náhodný vektor a $\mathbf{x} = (x_1, \ldots, x_n)^T$ je n-tice uskutečněných pozorování. Borelovskou funkci g: $\Theta \to \mathbb{R}$ nazveme parametrickou funkcí. Pak *bodový odhad* parametrické funkce g($\boldsymbol{\theta}$) je jakákoli borelovská funkce $\Phi(\mathbf{x})$, jejíž funkční předpis nezávisí na $\boldsymbol{\theta}$.

Budeme předpokládat, že Θ je otevřená a distribuční funkci $F(\mathbf{x}; \boldsymbol{\theta})$, kde $\mathbf{x} = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$ a $\boldsymbol{\theta} \in \Theta$, lze vyjádřit ve tvaru

$$F(\mathbf{x}; \boldsymbol{\theta}) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f_{\mathbf{X}}(t_1, \ldots, t_n; \boldsymbol{\theta}) \, \mathrm{d}\nu(t_1, \ldots, t_n),$$

kde ν je σ-konečná míra na $(\mathbb{R}^n, \mathcal{B}^n)$ a $f_{\mathbf{X}}(\mathbf{x}; \boldsymbol{\theta})$ je nezáporná měřitelná funkce.

Definice. Odhad $\hat{\boldsymbol{\theta}}_n = \hat{\boldsymbol{\theta}}_n(x_1, \ldots, x_n)$ se nazývá *odhadem metodou maximální věrohodnosti*, pokud

$$\max_{\boldsymbol{\theta} \in \Theta} f_{\mathbf{X}}(x_1, \ldots, x_n; \boldsymbol{\theta}) = f_{\mathbf{X}}(x_1, \ldots, x_n; \hat{\boldsymbol{\theta}}_n). \tag{1.1}$$

Funkci $\boldsymbol{\theta} \mapsto f_{\mathbf{X}}(\mathbf{x}; \boldsymbol{\theta})$ nazýváme *věrohodnostní funkcí* a značíme $L(\mathbf{x}; \boldsymbol{\theta})$.

To jinými slovy znamená, že maximálně věrohodný odhad je taková hodnota $\boldsymbol{\theta}$, která maximalizuje sdruženou hustotu $f_{\mathbf{X}}(\mathbf{x}; \boldsymbol{\theta})$ vzhledem k míře ν. Jelikož řešení (1.1) je ekvivalentní řešení rovnice

$$\max_{\boldsymbol{\theta} \in \Theta} \log f_{\mathbf{X}}(x_1, \ldots, x_n; \boldsymbol{\theta}) = \log f_{\mathbf{X}}(x_1, \ldots, x_n; \hat{\boldsymbol{\theta}}_n),$$

tak nám k získání $\hat{\boldsymbol{\theta}}_n$ obvykle stačí vyřešit tzv. věrohodnostní rovnici, která má tvar

$$\frac{\partial}{\partial \boldsymbol{\theta}} \log f_{\mathbf{X}}(x_1, \ldots, x_n; \boldsymbol{\theta}) = 0. \tag{1.2}$$

Poznámka. Pokud jsou $X_1, \ldots, X_n$ nezávislé, pak platí

$$f_{\mathbf{X}}(x_1, \ldots, x_n; \boldsymbol{\theta}) = \prod_{i=1}^{n} f_{X_i}(x_i; \boldsymbol{\theta})$$

a věrohodnostní rovnice je tvaru

$$\frac{\partial}{\partial \boldsymbol{\theta}} \log \prod_{i=1}^{n} f_{X_i}(x_i; \boldsymbol{\theta}) = \frac{\partial}{\partial \boldsymbol{\theta}} \sum_{i=1}^{n} \log f_{X_i}(x_i; \boldsymbol{\theta}) = 0. \tag{1.3}$$

1.2 Časové řady

Časová řada je posloupnost pozorování, která jsou indexována podle času.

Definice. Řekneme, že náhodná posloupnost $\{X_t, t \in \mathbb{Z}\}$ s konečnými druhými momenty je *slabě stacionární*, pokud má konstantní střední hodnotu, tj. pro každé $t \in \mathbb{Z}$ je $\mathrm{E}X_t = \mu$, a autokovarianční funkce $R(s,t) = \mathrm{cov}(X_s, X_t)$ je funkcí pouze rozdílu $s - t$.

Poznámka. (i) Pokud řekneme, že posloupnost je stacionární, máme na mysli slabou stacionaritu.
(ii) Pokud zjišťujeme kovarianci mezi dvěma náhodnými procesy, pak pro vzniklou funkci používáme označení „kovarianční funkce“. Autokovarianční funkce je pojem, který se používá pro kovarianční funkci, která se vztahuje pouze k jednomu náhodnému procesu (počítáme kovariance pouze uvnitř našeho procesu).
(iii) Autokovarianční funkci stacionární posloupnosti definujeme jako funkci jedné proměnné vztahem $R(t) := R(t, 0)$. Její vlastnosti jsou pak mj. $R(-t) = \overline{R(t)}$ a $R(t) \geq 0$, viz [3].

Definice. Posloupnost $\{Y_t, t \in \mathbb{Z}\}$ stejně rozdělených náhodných veličin s nulovou střední hodnotou, rozptylem σ^2 a $\mathrm{cov}(Y_s, Y_t) = 0$ pro $s \neq t$ nazveme *bílým šumem*, často ho značíme $\mathrm{WN}(0, \sigma^2)$ (z anglického „white noise“).

Řekneme, že bílý šum je *gaussovský*, pokud náhodné veličiny $\{Y_t, t \in \mathbb{Z}\}$ jsou nezávislé a pro každé $t \in \mathbb{Z}$ platí, že Y_t má normální rozdělení se střední hodnotou 0 a rozptylem σ^2 (značíme $\mathrm{N}(0, \sigma^2)$).

Poznámka. Z definice bílého šumu vidíme, že se jedná o stacionární posloupnost.

Pomocí bílého šumu definujeme další tři důležité typy časových řad, kterým se dále budeme věnovat: posloupnosti klouzavých součtů, autoregresní posloupnosti a jejich kombinaci.

1.2.1 Posloupnosti klouzavých součtů

Definice. Nechť $\{Y_t, t \in \mathbb{Z}\}$ je $\mathrm{WN}(0, \sigma^2)$. Náhodnou posloupnost $\{X_t, t \in \mathbb{Z}\}$, která je definována předpisem

$$X_t = \mu + Y_t + b_1 Y_{t-1} + \cdots + b_p Y_{t-p}, \tag{1.4}$$

kde $\mu, b_1, \ldots, b_p$ jsou reálné konstanty, $b_p \neq 0$, nazveme *posloupností klouzavých součtů řádu p* a značíme $\mathrm{MA}(p)$.

Zřejmě $\mathrm{E}X_t = \mu$ pro každé $t \in \mathbb{Z}$. V [3], str. 51–52, můžeme nalézt odvození, že $\mathrm{cov}(X_s, X_{s+t})$ nezávisí na s, ale jen na t, tedy posloupnost $\{X_t, t \in \mathbb{Z}\}$ je slabě stacionární. Spočítejme rozptyl X_t a autokovarianční funkci posloupnosti $\{X_t, t \in \mathbb{Z}\}$:

$$\begin{aligned} \mathrm{var} X_t &= \mathrm{E}(X_t - \mathrm{E}X_t)^2 = \mathrm{E}(X_t - \mu)^2 = \mathrm{E}(Y_t + b_1 Y_{t-1} + \cdots + b_p Y_{t-p})^2 \\ &= \sigma^2 (1 + b_1^2 + \cdots + b_p^2), \qquad (1.5) \\ R(t) &= \mathrm{cov}(X_s, X_{s+t}) = \mathrm{E}(Y_s + b_1 Y_{s-1} + \cdots + b_p Y_{s-p})(Y_{s+t} + b_1 Y_{s+t-1} + \cdots + \\ &\qquad + b_p Y_{s+t-p}) = \sigma^2 \sum_{i=0}^{p-|t|} b_i b_{i+|t|}, \qquad (1.6) \end{aligned}$$

kde jsme položili $b_0 = 1$.

1.2.2 Autoregresní posloupnosti

Definice. Nechť $\{Y_t, t \in \mathbb{Z}\}$ je $\mathrm{WN}(0, \sigma^2)$. Náhodnou posloupnost $\{X_t, t \in \mathbb{Z}\}$, která je definována předpisem

$$X_t = c + \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \cdots + \varphi_m X_{t-m} + Y_t, \tag{1.7}$$

kde $c, \varphi_1, \ldots, \varphi_m$ jsou reálné konstanty, $\varphi_m \neq 0$, nazveme *autoregresní posloupností řádu m* a značíme $\mathrm{AR}(m)$.

Podle [2], str. 53, je posloupnost typu AR(1) slabě stacionární, pokud $|\varphi_1| < 1$. Zřejmě

$$\mu = \mathrm{E}X_t = \mathrm{E}(c + \varphi_1 X_{t-1} + Y_t) = c + \varphi_1 \mu,$$

tedy

$$\mu = \frac{c}{1 - \varphi_1}. \tag{1.8}$$

Nyní rozepíšeme X_t:

$$
\begin{aligned}
X_t &= c + Y_t + \varphi_1 X_{t-1} = c + Y_t + \varphi_1(c + Y_{t-1} + \varphi_1 X_{t-1}) = \cdots \\
&= c + Y_t + \varphi_1(c + Y_{t-1}) + \varphi_1^2(c + Y_{t-2}) + \varphi_1^3(c + Y_{t-3}) + \cdots \\
&= c(1 + \varphi_1 + \varphi_1^2 + \varphi_1^3 + \cdots) + Y_t + \varphi_1 Y_{t-1} + \varphi_1^2 Y_{t-2} + \varphi_1^3 Y_{t-3} + \cdots \\
&= \frac{c}{1-\varphi_1} + Y_t + \varphi_1 Y_{t-1} + \varphi_1^2 Y_{t-2} + \varphi_1^3 Y_{t-3} + \cdots, \qquad (1.9)
\end{aligned}
$$

kde v poslední rovnosti využijeme součtu geometrické řady.

Pomocí rovnic (1.8) a (1.9) můžeme spočítat rozptyl X_t:

$$
\begin{aligned}
\mathrm{var} X_t &= \mathrm{E}(X_t - \mathrm{E}X_t)^2 = \mathrm{E}(X_t - \mu)^2 = \\
&= \mathrm{E}(Y_t + \varphi_1 Y_{t-1} + \varphi_1^2 Y_{t-2} + \varphi_1^3 Y_{t-3} + \cdots)^2 = \\
&= \sigma^2(1 + \varphi_1^2 + \varphi_1^4 + \varphi_1^6 + \cdots) = \\
&= \frac{\sigma^2}{1-\varphi_1^2}, \qquad (1.10)
\end{aligned}
$$

kde jsme opět využili součtu geometrické řady.

Posloupnost typu AR(m) je podle [2], str. 58, slabě stacionární, pokud kořeny polynomu

$$1 - \varphi_1 z - \varphi_2 z^2 - \ldots - \varphi_m z^m = 0 \qquad (1.11)$$

leží mimo jednotkový kruh. Nadále budeme předpokládat, že je tato podmínka splněna. Pak pro střední hodnotu platí:

$$\mu = c + \varphi_1\mu + \varphi_2\mu + \cdots + \varphi_m\mu,$$

tedy

$$\mathrm{E}X_t = \mu = \frac{c}{1 - \varphi_1 - \varphi_2 - \cdots - \varphi_m}. \qquad (1.12)$$

Ještě spočítáme rozptyl X_t. Z rovnice (1.12) dosadíme za c do rovnice (1.7), převedeme μ, vynásobíme obě strany rovnosti výrazem $(X_t - \mu)$ a určíme střední hodnotu:

$$
\begin{aligned}
X_t &= \mu(1 - \varphi_1 - \varphi_2 - \cdots - \varphi_m) + \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \cdots + \varphi_m X_{t-m} + Y_t \\
X_t - \mu &= \varphi_1(X_{t-1} - \mu) + \varphi_2(X_{t-2} - \mu) + \cdots + \varphi_m(X_{t-m} - \mu) + Y_t \\
(X_t - \mu)^2 &= \varphi_1(X_{t-1} - \mu)(X_t - \mu) + \cdots + \varphi_m(X_{t-m} - \mu)(X_t - \mu) + Y_t(X_t - \mu) \\
\mathrm{var} X_t &= R(0) = \varphi_1 R(1) + \cdots + \varphi_m R(m) + \mathrm{E}Y_t(X_t - \mu) \\
&= \varphi_1 R(1) + \cdots + \varphi_m R(m) + \sigma^2, \qquad (1.13)
\end{aligned}
$$

kde $R(t)$ je autokovarianční funkce posloupnosti $\{X_t, t \in \mathbb{Z}\}$. K určení rozptylu tedy potřebujeme znát autokovarianční funkci, jejímž zjišťováním se v této práci nebudeme zabývat. Pro zájemce jsou metody k vypočítání autokovarianční funkce uvedeny v [3]. Jen ještě uvedeme, že pokud v předchozím postupu vynásobíme obě strany rovnosti výrazem $(X_{t-k} - \mu)$ pro $k \in \mathbb{N}$ místo výrazem $(X_t - \mu)$, dostaneme následující rovnici:

$$R(k) = \varphi_1 R(k-1) + \cdots + \varphi_m R(k-m). \qquad (1.14)$$

1.2.3 Smíšený model autoregrese a klouzavých součtů

Definice. Nechť $\{Y_t, t \in \mathbb{Z}\}$ je $\mathrm{WN}(0, \sigma^2)$. Náhodnou posloupnost $\{X_t, t \in \mathbb{Z}\}$, která je definována předpisem

$$X_t = c + \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \cdots + \varphi_m X_{t-m} + Y_t + b_1 Y_{t-1} + \cdots + b_p Y_{t-p}, \qquad (1.15)$$

kde $c, \varphi_1, \ldots, \varphi_m, b_1, \ldots, b_p$ jsou reálné konstanty, $\varphi_m \neq 0$ a $b_p \neq 0$, nazveme *posloupností typu* $\mathrm{ARMA}(m, p)$.

Modely $\mathrm{AR}(m)$ a $\mathrm{MA}(p)$ jsou zřejmě speciálními případy modelu $\mathrm{ARMA}(m, p)$.

Stejně jako u modelu $\mathrm{AR}(m)$ je posloupnost typu $\mathrm{ARMA}(m, p)$ slabě stacionární, pokud kořeny polynomu (1.11) leží vně jednotkového kruhu. V dalším budeme předpokládat, že je tato podmínka splněna.

Jelikož bílý šum má nulovou střední hodnotu, tak střední hodnota X_t u modelu $\mathrm{ARMA}(m, p)$ má stejný tvar jako střední hodnota X_t u modelu $\mathrm{AR}(m)$, tedy splňuje rovnici (1.12). Podobně jako u $\mathrm{AR}(m)$ budeme pro určení rozptylu X_t u modelu $\mathrm{ARMA}(m, p)$ potřebovat autokovarianční funkci, kterou nyní nebudeme dopočítávat – metody se dají opět dohledat v [3].

2. Odvození odhadů

Závěry v této kapitole vychází hlavně z poznatků uvedených v [2] a [3].

2.1 Odhady v modelech AR(1)

Uvažujme stacionární náhodnou posloupnost $\{X_t, t \in \mathbb{Z}\}$, která je definována předpisem

$$X_t = c + \varphi X_{t-1} + Y_t, \tag{2.1}$$

kde c je reálná konstanta, φ je neznámý parametr a $\{Y_t, t \in \mathbb{Z}\}$ je gaussovský bílý šum se střední hodnotou 0 a rozptylem σ^2. Pak vektor parametrů, které chceme odhadnout, je $\boldsymbol{\theta} = (\mu, \varphi, \sigma^2)^T$, kde μ je střední hodnota X_t, kterou můžeme dopočítat ze vztahu (1.8).

U metody maximální věrohodnosti chceme maximalizovat sdruženou hustotu $f_{\mathbf{X}}(\mathbf{x}; \boldsymbol{\theta})$, pokud existuje. Pak

$$\begin{aligned}
&f_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n; \boldsymbol{\theta}) = f_{X_n,X_{n-1},\ldots,X_2|X_1}(x_n, x_{n-1}, \ldots, x_2|x_1; \boldsymbol{\theta}) \cdot f_{X_1}(x_1; \boldsymbol{\theta})\\
&= f_{X_n,X_{n-1},\ldots,X_3|X_2,X_1}(x_n, x_{n-1}, \ldots, x_3|x_2, x_1; \boldsymbol{\theta}) \cdot f_{X_2|X_1}(x_2|x_1; \boldsymbol{\theta}) \cdot f_{X_1}(x_1; \boldsymbol{\theta})\\
&= f_{X_n,X_{n-1},\ldots,X_3|X_2}(x_n, x_{n-1}, \ldots, x_3|x_2; \boldsymbol{\theta}) \cdot f_{X_2|X_1}(x_2|x_1; \boldsymbol{\theta}) \cdot f_{X_1}(x_1; \boldsymbol{\theta})\\
&= \cdots = f_{X_1}(x_1; \boldsymbol{\theta}) \cdot \prod_{i=2}^{n} f_{X_i|X_{i-1}}(x_i|x_{i-1}; \boldsymbol{\theta}),
\end{aligned} \tag{2.2}$$

kde třetí rovnost platí, neboť $X_3, \ldots, X_n$ závisí na X_1 jen skrze X_2, tj. pokud známe hodnotu X_2, hodnotu X_1 již k určení $X_3, \ldots, X_n$ nepotřebujeme.

Jelikož Y_1 má rozdělení $\mathrm{N}(0, \sigma^2)$, tak X_1 má také normální rozdělení, a tedy existuje hustota. Víme, že střední hodnota X_1 je μ, a rozptyl určíme pomocí rovnice (1.10):

$$\mathrm{var} X_1 = \frac{\sigma^2}{1 - \varphi^2}.$$

Tedy X_1 má rozdělení $\mathrm{N}(\mu, \frac{\sigma^2}{1-\varphi^2})$ a hustota f_{X_1} má tvar

$$f_{X_1}(x_1; \boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi \frac{\sigma^2}{1-\varphi^2}}} \exp\left[\frac{-(x_1 - \mu)^2}{2\frac{\sigma^2}{1-\varphi^2}}\right]. \tag{2.3}$$

Nyní chceme určit podmíněnou hustotu $f_{X_2|X_1}(x_2|x_1;\boldsymbol{\theta})$. Protože

$$X_2 = c + \varphi X_1 + Y_2$$

a Y_2 má rozdělení $\mathrm{N}(0,\sigma^2)$, tak $X_2|X_1 = x_1$ má rozdělení $\mathrm{N}(c+\varphi x_1,\sigma^2)$. Za využití vztahu (1.8) je pak hledaná hustota tvaru

$$f_{X_2|X_1=x_1}(x_2|x_1;\boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi\sigma^2}}\,\mathrm{e}^{-\frac{(x_2-\mu(1-\varphi)-\varphi x_1)^2}{2\sigma^2}}.$$

Podobným způsobem pokračujeme dál a získáme podmíněnou hustotu

$$f_{X_3|X_2=x_2}(x_3|x_2;\boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi\sigma^2}}\,\mathrm{e}^{-\frac{(x_3-\mu(1-\varphi)-\varphi x_2)^2}{2\sigma^2}}.$$

Obecně pro $i = 2,\dots,n$ platí vztah

$$f_{X_i|X_{i-1}=x_{i-1}}(x_i|x_{i-1};\boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi\sigma^2}}\,\mathrm{e}^{-\frac{(x_i-\mu(1-\varphi)-\varphi x_{i-1})^2}{2\sigma^2}}. \tag{2.4}$$

Věrohodnostní rovnice (1.2) má s využitím vlastností logaritmu a rovnic (2.2), (2.3) a (2.4) tvar:

$$\frac{\partial}{\partial\boldsymbol{\theta}}\left(\frac{1}{2}\log(1-\varphi^2) - \frac{n}{2}\log\sigma^2 - \frac{n}{2}\log 2\pi - \frac{(x_1-\mu)^2}{2\frac{\sigma^2}{1-\varphi^2}} \right.$$
$$\left. - \sum_{i=2}^{n}\frac{(x_i-\mu(1-\varphi)-\varphi x_{i-1})^2}{2\sigma^2}\right) = 0.$$

Postupným derivováním podle jednotlivých parametrů a upravením výrazů dostáváme následující soustavu rovnic

$$\begin{aligned}
\mu(n-n\varphi+2\varphi) &= \sum_{i=1}^{n} x_i - \varphi\sum_{i=2}^{n-1} x_i,\\
\frac{\varphi}{1-\varphi^2} &= \frac{\varphi(x_1-\mu)^2}{\sigma^2} + \sum_{i=2}^{n}\frac{(x_{i-1}-\mu)(x_i-\mu(1-\varphi)-\varphi x_{i-1})}{\sigma^2},\\
n\sigma^2 &= (x_1-\mu)^2(1-\varphi^2) + \sum_{i=2}^{n}(x_i-\mu(1-\varphi)-\varphi x_{i-1})^2.
\end{aligned}$$

Pokud se nám podaří vypočítat odhad φ (budeme ho značit $\hat{\varphi}$), pak zbylé odhady parametrů již snadno dopočítáme z rovnic

$$\begin{aligned}
\hat{\mu} &= \frac{\sum_{i=1}^{n} x_i - \hat{\varphi}\sum_{i=2}^{n-1} x_i}{n-n\hat{\varphi}+2\hat{\varphi}},\\
\hat{\sigma}^2 &= \frac{(x_1-\hat{\mu})^2(1-\hat{\varphi}^2) + \sum_{i=2}^{n}(x_i-\hat{\mu}(1-\hat{\varphi})-\hat{\varphi}x_{i-1})^2}{n}.
\end{aligned}$$

Podle [2], str.122, neexistuje jednoduché vyjádření pro $\hat{\varphi}$ pomocí $x_1,\dots,x_n$ a pro určení hodnot odhadu tohoto parametru se používají iterační nebo numerické metody.

2.2 Odhady v modelech MA(1)

Uvažujme stacionární náhodnou posloupnost $\{X_t, t \in \mathbb{Z}\}$, která je definována předpisem

$$X_t = \mu + Y_t + b\,Y_{t-1}, \tag{2.5}$$

kde μ je reálná konstanta, b je neznámý parametr a $\{Y_t, t \in \mathbb{Z}\}$ je gaussovský bílý šum se střední hodnotou 0 a rozptylem σ^2. Pak vektor parametrů, které chceme odhadnout, je $\boldsymbol{\theta} = (\mu, b, \sigma^2)^T$.

Pro každé $t \in \mathbb{Z}$ má X_t rozdělení $\mathrm{N}(\mu, \sigma^2(1+b^2))$, tedy $\mathbf{X} = (X_1, X_2, \ldots, X_n)^T$ má n-rozměrné normální rozdělení se střední hodnotou $\boldsymbol{\mu} = (\mu, \ldots, \mu)^T$ a varianční maticí

$$\boldsymbol{\Gamma}_n = \sigma^2 \begin{pmatrix} 1+b^2 & b & 0 & \cdots & 0 \\ b & 1+b^2 & b & \cdots & 0 \\ 0 & b & 1+b^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1+b^2 \end{pmatrix}.$$

Věrohodnostní funkce je tvaru

$$L(\mathbf{x};\boldsymbol{\theta}) = f_{\mathbf{X}}(\mathbf{x};\boldsymbol{\theta}) = (2\pi)^{-\frac{n}{2}} (\det \boldsymbol{\Gamma}_n)^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Gamma}_n^{-1} (\mathbf{x}-\boldsymbol{\mu})\right]. \tag{2.6}$$

Budeme faktorizovat matici $\boldsymbol{\Gamma}_n$, tj. budeme hledat diagonální matici $\mathbf{D}_n$ a dolní trojúhelníkovou matici $\mathbf{A}$ takové, že $\boldsymbol{\Gamma}_n = \mathbf{A}\,\mathbf{D}_n\,\mathbf{A}^T$ a $\mathbf{A}$ má na diagonále samé jedničky. Pro zjednodušení výpočtů matici $\mathbf{D}_n$ rozepíšeme jako $\mathbf{D}_n = \sigma^2\,\mathbf{D}$. Prvky příslušné matice budeme značit malým písmenem s indexem pozice prvku. Vzhledem k tvaru $\boldsymbol{\Gamma}_n$ jsou prvky $a_{i+k,i}$ nulové pro $k \geq 2, i = 1, 2, \ldots, n$, tedy potřebujeme vypočítat jen prvky $a_{i+1,i}$ pro $i = 1, 2, \ldots, n-1$ a prvky $d_{i,i}$ pro $i = 1, 2, \ldots, n$. Snadno můžeme ukázat, že

$$\begin{aligned} \gamma_{1,1} &= d_{1,1}, \\ \gamma_{i+1,i} &= a_{i+1,i}\,d_{i,i}, \\ \gamma_{i+1,i+1} &= a_{i+1,i}^2\,d_{i,i} + d_{i+1,i+1} \end{aligned}$$

pro $i = 1, 2, \ldots, n$.

Tudíž

$$
\begin{aligned}
d_{1,1} &= \gamma_{1,1} = 1 + b^2,\\
a_{2,1} &= \frac{\gamma_{2,1}}{d_{1,1}} = \frac{b}{1+b^2}\,,\\
d_{2,2} &= \gamma_{2,2} - d_{1,1}\cdot a_{2,1}^2 = 1 + b^2 - \frac{b^2}{1+b^2} = \frac{1+b^2+b^4}{1+b^2}\,,\\
a_{3,2} &= \frac{\gamma_{3,2}}{d_{2,2}} = \frac{b(1+b^2)}{1+b^2+b^4}\,,
\end{aligned}
$$

obecně

$$
\begin{aligned}
d_{i,i} &= \frac{1+b^2+b^4+\cdots+b^{2i}}{1+b^2+b^4+\cdots+b^{2(i-1)}}\,,\\
a_{i+1,i} &= \frac{b(1+b^2+b^4+\cdots+b^{2(i-1)})}{1+b^2+b^4+\cdots+b^{2i}}\,,\\
d_{i+1,i+1} &= 1 + b^2 - \frac{b^2(1+b^2+b^4+\cdots+b^{2(i-1)})}{1+b^2+b^4+\cdots+b^{2i}} = \frac{1+b^2+b^4+\cdots+b^{2(i+1)}}{1+b^2+b^4+\cdots+b^{2i}}\,.
\end{aligned}
$$

Tedy $\boldsymbol{\Gamma}_n = \mathbf{A}\,\mathbf{D}_n\,\mathbf{A}^T$ a pro $(\boldsymbol{\Gamma}_n)^{-1}$ a determinant $\boldsymbol{\Gamma}_n$ platí

$$
\begin{aligned}
(\boldsymbol{\Gamma}_n)^{-1} &= (\mathbf{A}^T)^{-1}\mathbf{D}_n^{-1}\mathbf{A}^{-1},\\
\det \boldsymbol{\Gamma}_n &= \det \mathbf{A}\cdot \det \mathbf{D}_n \cdot \det \mathbf{A}^T = \det \mathbf{D}_n = \prod_{i=1}^{n} \sigma^2 d_{i,i}\\
&= (\sigma^2)^n (1+b^2+b^4+\cdots+b^{2n}),
\end{aligned}
$$

kde předposlední rovnost plyne z toho, že $\mathbf{D}_n$ je diagonální.

Označíme $\tilde{\mathbf{x}} = \mathbf{A}^{-1}(\mathbf{x}-\boldsymbol{\mu})$ a dosadíme do (2.6)

$$
L(\mathbf{x};\boldsymbol{\theta}) = (2\pi\sigma^2)^{-\frac{n}{2}}(1+b^2+b^4+\cdots+b^{2n})^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\,\tilde{\mathbf{x}}^T\,\mathbf{D}_n^{-1}\,\tilde{\mathbf{x}}\right].
$$

Inverzní matici k diagonální matici $\mathbf{D}_n$ získáme převrácením hodnot na diagonále, z čehož plyne vztah

$$
\tilde{\mathbf{x}}^T\,\mathbf{D}_n^{-1}\,\tilde{\mathbf{x}} = \sum_{i=1}^{n} \frac{\tilde{x}_i^2}{\sigma^2\, d_{i,i}},
$$

kde $\tilde{x}_i, i = 1, 2, \ldots, n$ dopočítáme ze vztahu $\mathbf{A}\tilde{\mathbf{x}} = \mathbf{x} - \boldsymbol{\mu}$:

$$
\begin{aligned}
\tilde{x}_1 &= x_1 - \mu,\\
a_{2,1}\tilde{x}_1 + \tilde{x}_2 &= x_2 - \mu,\\
&\vdots\\
a_{n,n-1}\tilde{x}_{n-1} + \tilde{x}_n &= x_n - \mu.
\end{aligned}
$$

Obecně pro $i = 2, 3, \ldots, n$ platí

$$\tilde{x}_i = x_i - \mu - \tilde{x}_{i-1} \cdot \frac{b(1 + b^2 + b^4 + \cdots + b^{2(i-2)})}{1 + b^2 + b^4 + \cdots + b^{2(i-1)}},$$

z čehož plyne

$$\tilde{x}_i = \sum_{j=1}^{i} (x_j - \mu)(-b)^{i-j} \cdot \frac{1 + b^2 + b^4 + \cdots + b^{2(j-1)}}{1 + b^2 + b^4 + \cdots + b^{2(i-1)}}.$$

Tedy věrohodnostní funkce je tvaru

$$L(\mathbf{x}; \boldsymbol{\theta}) = (2\pi\sigma^2)^{-\frac{n}{2}} (1 + b^2 + b^4 + \cdots + b^{2n})^{-\frac{1}{2}} \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^{n} \frac{\tilde{x}_i^2}{d_{i,i}} \right] \tag{2.7}$$

a s využitím vlastností logaritmu můžeme psát

$$\begin{aligned} &\log L(\mathbf{x}; \boldsymbol{\theta}) \\ &= -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log \sigma^2 - \frac{1}{2} \log(1 + b^2 + b^4 + \cdots + b^{2n}) - \frac{1}{2\sigma^2} \sum_{i=1}^{n} \frac{\tilde{x}_i^2}{d_{i,i}}. \end{aligned}$$

Nyní zderivujeme $\log L(\mathbf{x}; \boldsymbol{\theta})$ podle σ^2 a položíme rovno nule. Pak pro σ^2 získáme odhad

$$\hat{\sigma}^2 = \frac{1}{n} S(\hat{\mu}, \hat{b}),$$

kde

$$S(\mu, b) = \sum_{i=1}^{n} \frac{\tilde{x}_i^2}{d_{i,i}}$$

a hodnoty $\hat{\mu}$ a $\hat{b}$ jsou maximálně věrohodné odhady parametrů μ a b, které musíme určit numericky.

2.3 Odhady v modelech AR(m)

Uvažujme stacionární náhodnou posloupnost $\{X_t, t \in \mathbb{Z}\}$ definovanou jako v (1.7). V tomto případě odhadujeme vektor parametrů $\boldsymbol{\theta} = (\mu, \varphi_1, \varphi_2, \ldots, \varphi_m, \sigma^2)^T$, kde μ splňuje vztah (1.12). Označme

$$\sigma^2 \mathbf{V}_m = \begin{pmatrix} R(0) & R(1) & R(2) & \cdots & R(m-1) \\ R(1) & R(0) & R(1) & \cdots & R(m-2) \\ R(2) & R(1) & R(0) & \cdots & R(m-3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ R(m-1) & R(m-2) & R(m-3) & \cdots & R(0) \end{pmatrix}.$$

Pak $\mathbf{V}_m$ je regulární podle [3], věta 7.3, a $(X_1, X_2, \dots, X_m)^T$ má m-rozměrné normální rozdělení se střední hodnotou $\boldsymbol{\mu} = (\mu, \mu, \dots, \mu)^T$ a varianční maticí $\sigma^2\mathbf{V}_m$. Hustotu tedy můžeme zapsat jako

$$\begin{aligned}
&f_{X_m,X_{m-1},\dots,X_1}(\mathbf{x}_m;\boldsymbol{\theta})\\
&\qquad= (2\pi)^{-\frac{m}{2}}(\det(\sigma^2\mathbf{V}_m))^{-\frac{1}{2}}\exp\left[-\frac{1}{2\sigma^2}(\mathbf{x}_m-\boldsymbol{\mu})^T(\mathbf{V}_m)^{-1}(\mathbf{x}_m-\boldsymbol{\mu})\right]\\
&\qquad= (2\pi\sigma^2)^{-\frac{m}{2}}(\det\mathbf{V}_m)^{-\frac{1}{2}}\exp\left[-\frac{1}{2\sigma^2}(\mathbf{x}_m-\boldsymbol{\mu})^T(\mathbf{V}_m)^{-1}(\mathbf{x}_m-\boldsymbol{\mu})\right],
\end{aligned}$$

kde $\mathbf{x}_m = (x_1, x_2, \dots, x_m)^T$.

Pro každé pozorování X_t platí, že pokud známe m předchozích pozorování, pak X_t má podmíněné rozdělení $\mathrm{N}((c + \varphi_1 x_{t-1} + \varphi_2 x_{t-2} + \cdots + \varphi_m x_{t-m}), \sigma^2)$. Tedy pro $t = m+1, \dots, n$ platí

$$\begin{aligned}
&f_{X_t|X_{t-1},X_{t-2},\dots,X_{t-m}}(x_t|x_{t-1}, x_{t-2}, \dots, x_{t-m};\boldsymbol{\theta})\\
&\quad= \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{(x_t - c - \varphi_1 x_{t-1} - \varphi_2 x_{t-2} - \cdots - \varphi_m x_{t-m})^2}{2\sigma^2}\right],
\end{aligned}$$

kde c splňuje vztah (1.12).

Tudíž věrohodnostní funkce má tvar

$$\begin{aligned}
L(\mathbf{x};\boldsymbol{\theta}) &= f_{X_n,X_{n-1},\dots,X_1}(\mathbf{x};\boldsymbol{\theta})\\
&= f_{X_m,X_{m-1},\dots,X_1}(\mathbf{x}_m;\boldsymbol{\theta})\prod_{t=m+1}^{n} f_{X_t|X_{t-1},X_{t-2},\dots,X_{t-m}}(x_t|x_{t-1}, x_{t-2}, \dots, x_{t-m};\boldsymbol{\theta})\\
&= (2\pi\sigma^2)^{-\frac{n}{2}}(\det\mathbf{V}_m)^{-\frac{1}{2}}\exp\left[-\frac{1}{2\sigma^2}(\mathbf{x}_m-\boldsymbol{\mu})^T(\mathbf{V}_m)^{-1}(\mathbf{x}_m-\boldsymbol{\mu})\right]\\
&\quad\times\exp\left[-\sum_{t=m+1}^{n}\frac{(x_t - c - \varphi_1 x_{t-1} - \varphi_2 x_{t-2} - \cdots - \varphi_m x_{t-m})^2}{2\sigma^2}\right].
\end{aligned}$$

Pro zjednodušení výpočtů se často používá podmíněná metoda maximální věrohodnosti, v našem případě bude podmínkou prvních m pozorování. Pak

$$\begin{aligned}
&L(\mathbf{x};\boldsymbol{\theta})\\
&\quad= f_{X_n,X_{n-1},\dots,X_{m+1}|X_m,X_{m-1},\dots,X_1}(x_n, x_{n-1}, \dots, x_{m+1}|x_m, x_{m-1}, \dots, x_1;\boldsymbol{\theta})\\
&\quad= (2\pi\sigma^2)^{-\frac{n-m}{2}}\exp\left[-\sum_{t=m+1}^{n}\frac{(x_t - c - \varphi_1 x_{t-1} - \varphi_2 x_{t-2} - \cdots - \varphi_m x_{t-m})^2}{2\sigma^2}\right].
\end{aligned}$$

2.4 Odhady v modelech MA(p)

Mějme stacionární náhodnou posloupnost $\{X_t, t \in \mathbb{Z}\}$ definovanou jako v (1.4). Pak chceme odhadnout vektor parametrů $\boldsymbol{\theta} = (\mu, b_1, b_2, \ldots, b_p, \sigma^2)^T$. Budeme postupovat podobně jako u modelu MA(1) a dostaneme věrohodnostní funkci (2.6), kde $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T, \boldsymbol{\mu} = (\mu, \mu, \ldots, \mu)^T$ je n-rozměrný vektor a $\boldsymbol{\Gamma}_n$ je matice typu $n \times n$ a tvaru

$$\boldsymbol{\Gamma}_n = \begin{pmatrix} R(0) & R(1) & \cdots & R(p) & 0 & \cdots & 0 \\ R(1) & R(0) & \ddots & & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & & \ddots & 0 \\ R(p) & & \ddots & \ddots & \ddots & & R(p) \\ 0 & \ddots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \ddots & \ddots & R(1) \\ 0 & \cdots & 0 & R(p) & \cdots & R(1) & R(0) \end{pmatrix},$$

kde $R(t)$ značí autokovarianční funkci posloupnosti $\{X_t, t \in \mathbb{Z}\}$. Podle [3], věta 7.3, je $\boldsymbol{\Gamma}_n$ invertibilní, neboť $R(0) > 0$ a $R(k) \to 0$ pro $k \to \infty$.

Pro zjednodušení můžeme opět použít podmíněnou metodu maximální věrohodnosti, v tomto případě budeme podmiňovat jevem $Y_0 = Y_{-1} = \cdots = Y_{-p+1} = 0$. Dále budeme předpokládat, že kořeny polynomu $1 + b_1 z + b_2 z^2 + \cdots + b_p z^p$ leží vně jednotkového kruhu, aby podmíněná věrohodnostní funkce dobře aproximovala nepodmíněnou.

Podle [2],str. 130, je pak podmíněná věrohodnostní funkce tvaru

$$\begin{aligned} L(\mathbf{x}; \boldsymbol{\theta}) &= f_{X_n, X_{n-1}, \ldots, X_1 | Y_0 = Y_1 = \ldots, Y_{-p+1} = 0}(\mathbf{x} | y_0 = y_1 = \cdots = y_{-p+1} = 0; \boldsymbol{\theta}) \\ &= \left(2\pi\sigma^2\right)^{-\frac{n}{2}} \exp\left[-\sum_{t=1}^{n} \frac{y_t^2}{2\sigma^2}\right], \end{aligned}$$

kde

$$y_t = x_t - \mu - b_1 y_{t-1} - b_2 y_{t-2} - \cdots - b_p y_{t-p},$$

přičemž předpokládáme, že $y_0 = y_{-1} = \cdots = y_{-p+1} = 0$.

2.5 Odhady v modelech ARMA(m, p)

Uvažujme stacionární náhodnou posloupnost popsanou v (1.15). Chceme odhadnout vektor parametrů $\boldsymbol{\theta} = (\mu, \varphi_1, \varphi_2, \ldots, \varphi_m, b_1, b_2, \ldots, b_p, \sigma^2)^T$.

Nepodmíněnou metodou maximální věrohodnosti dostaneme stejný výsledek jako u modelu MA(p) – rozdíl se projeví jen v autokovarianční funkci.

Pokud budeme chtít použít podmíněnou metodu maximální věrohodnosti, pak předpokládáme, že kořeny polynomu $1 + b_1 z + b_2 z^2 + \cdots + b_p z^p$ leží mimo jednotkový kruh, a dále předpokládáme znalost hodnot $(x_0, x_{-1}, \ldots, x_{-m+1})^T$ a $(y_0, y_{-1}, \ldots, y_{-p+1})^T$. Pak pro $t = 1, 2, \ldots, n$ můžeme dopočítat

$$y_t = x_t - c - \varphi_1 x_{t-1} - \varphi_2 x_{t-2} - \cdots - \varphi_m x_{t-m} - b_1 y_{t-1} - b_2 y_{t-2} - \cdots - b_p y_{t-p},$$

kde c můžeme vyjádřit pomocí vztahu (1.12), a podmíněná věrohodnostní funkce má opět tvar

$$L(\mathbf{x}; \boldsymbol{\theta}) = \left(2\pi\sigma^2\right)^{-\frac{n}{2}} \exp\left[-\sum_{t=1}^{n} \frac{y_t^2}{2\sigma^2}\right].$$

3. Další metody odhadů

Nechť $\{X_t, t \in \mathbb{Z}\}$ je stacionární náhodná posloupnost se střední hodnotou μ. Pak podle [3], str. 127, považujeme za přirozený odhad μ tzv. výběrový průměr

$$\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i.$$

3.1 Metoda momentů

U metody momentů porovnáváme momenty s hodnotami jejich výběrových protějšků, v našem případě budeme používat výběrovou autokovarianční funkci, která je definovaná předpisem

$$\hat{R}(k) = \frac{1}{n} \sum_{i=1}^{n-k} (X_i - \overline{X}_n)(X_{i+k} - \overline{X}_n) \quad \text{pro } k = 0, 1, \ldots, n-1 \tag{3.1}$$

a pro $k < 0$ definujeme $\hat{R}(k) = \hat{R}(-k)$.

3.1.1 Metoda momentů pro AR(1)

Uvažujme model (2.1). Pak chceme odhadnout vektor parametrů $\boldsymbol{\theta} = (\varphi, \sigma^2)^T$. Podle (1.13) a (1.14) platí

$$R(0) = \varphi R(1) + \sigma^2,$$
$$R(1) = \varphi R(0),$$

kde $R(t)$ je autokovarianční funkce posloupnosti $\{X_t, t \in \mathbb{Z}\}$. Nahradíme autokovarianční funkci jejím výběrovým protějškem a za podmínky $\hat{R}(0) > 0$ získáme odhady

$$\hat{\varphi} = \frac{\hat{R}(1)}{\hat{R}(0)},$$
$$\hat{\sigma}^2 = \hat{R}(0) - \hat{\varphi}\hat{R}(1).$$

3.1.2 Metoda momentů pro MA(1)

Nyní uvažujme model (2.5). Budeme odhadovat vektor parametrů $\boldsymbol{\theta} = (b, \sigma^2)^T$. Podle (1.5) a (1.6) platí

$$\begin{aligned} R(0) &= \sigma^2(1+b^2), \\ R(1) &= \sigma^2 b. \end{aligned} \tag{3.2}$$

K vyřešení této soustavy použijeme autokorelační funkci definovanou předpisem

$$r(k) = \frac{R(k)}{R(0)}.$$

Místo autokovarianční funkce budeme uvažovat výběrovou autokovarianční funkci, čímž dostaneme výběrovou autokorelační funkci

$$\hat{r}(k) = \frac{\hat{R}(k)}{\hat{R}(0)}.$$

Ze soustavy (3.2) pak dostáváme rovnici

$$\hat{r}(1) = \frac{\hat{R}(1)}{\hat{R}(0)} = \frac{b}{1+b^2}.$$

Odtud již můžeme vypočítat odhad pro b

$$\hat{b}_{1,2} = \frac{1 \pm \sqrt{1 - 4(\hat{r}(1))^2}}{2\hat{r}(1)}, \tag{3.3}$$

přičemž tato řešení jsou reálná, jen pokud $|\hat{r}(1)| \leq \frac{1}{2}$.

Podle [3], str. 144, za momentové odhady považujeme následující hodnoty:

1) Pokud $|\hat{r}(1)| < \frac{1}{2}$ a pro skutečnou hodnotu parametru b platí $|b| < 1$, pak

$$\begin{aligned} \hat{b} &= \frac{1 - \sqrt{1 - 4(\hat{r}(1))^2}}{2\hat{r}(1)}, \\ \hat{\sigma}^2 &= \frac{\hat{R}(0)}{1 + \hat{b}^2}. \end{aligned}$$

2) Pokud $|\hat{r}(1)| = \frac{1}{2}$, pak

$$\begin{aligned} \hat{b} &= \frac{1}{2\hat{r}(1)} = \frac{\hat{r}(1)}{|\hat{r}(1)|}, \\ \hat{\sigma}^2 &= \frac{1}{2}\hat{R}(0). \end{aligned}$$

3) Pokud $|\hat{r}(1)| > \frac{1}{2}$, pak rovnice (3.3) nemají reálné řešení a za odhady považujeme hodnoty získáné v případu 2).

3.2 Metoda nejmenších čtverců

3.2.1 Metoda nejmenších čtverců pro AR(1)

Opět uvažujeme model (2.1). Jak název metody napovídá, postupujeme minimalizací součtu čtverců – v našem případě hledáme

$$\min_{\varphi} \sum_{i=2}^{n} (X_i - \varphi X_{i-1} - c)^2.$$

Za využití vztahu (1.8) vede hledání tohoto minima k řešení rovnice

$$\sum_{i=2}^{n} (X_i - \varphi X_{i-1} - \mu(1-\varphi))(X_{i-1} - \mu) = 0.$$

Odtud již snadno vyjádříme odhad φ

$$\hat{\varphi} = \frac{\sum_{i=2}^{n} (X_i - \mu)(X_{i-1} - \mu)}{\sum_{i=2}^{n} (X_{i-1} - \mu)^2}.$$

Jak bylo uvedeno na začátku kapitoly, za odhad parametru μ budeme považovat výběrový průměr. Pak můžeme odhady parametrů φ a σ^2 zapsat následovně

$$\hat{\varphi} = \frac{\sum_{i=2}^{n} (X_i - \hat{\mu})(X_{i-1} - \hat{\mu})}{\sum_{i=2}^{n} (X_{i-1} - \hat{\mu})^2},$$

$$\hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=2}^{n} (X_i - \hat{\varphi} X_{i-1} - \hat{\mu}(1-\hat{\varphi}))^2 .$$

Vzhledem k podobnosti vyjádření odhadu parametru φ metodou momentů a metodou nejmenších čtverců lze předpokládat, že tyto odhady budou mít v praxi podobné hodnoty.

3.2.2 Metoda nejmenších čtverců pro MA(1)

Uvažujeme model (2.5) a parametry b a σ^2 budeme odhadovat metodou podmíněných nejmenších čtverců. Rozepíšeme Y_t do sumy:

$$\begin{aligned} Y_t &= X_t - \mu - bY_{t-1} = X_t - \mu - b(X_{t-1} - \mu) + b^2 Y_{t-2} = \cdots \\ &= \sum_{i=0}^{t-1} (-b)^i (X_{t-i} - \mu) + (-b)^t Y_0. \end{aligned}$$

Chceme podmiňovat jevem $Y_0 = 0$. Z předchozího vyjádření je vidět, že pokud touto podmínkou nechceme přijít o důležitá data, musí skutečná hodnota parametru b splňovat $|b| < 1$.

Podle [3], str. 146, hledáme odhad parametru b za předpokladu $Y_0 = 0$ jako řešení, které minimalizuje součet čtverců

$$\sum_{t=1}^{n} Y_t^2 = \sum_{t=1}^{n} \Big(\sum_{i=0}^{t-1} (-b)^i (X_{t-i} - \mu) \Big)^2.$$

Za odhad parametru μ opět považujeme výběrový průměr, tedy hledáme hodnotu b, která minimalizuje součet

$$\sum_{t=1}^{n} \Big(\sum_{i=0}^{t-1} (-b)^i (X_{t-i} - \hat{\mu}) \Big)^2.$$

Tato hodnota musí být určena numericky, a pokud ji označíme $\hat{b}$, pak odhad parametru σ^2 je

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{t=1}^{n} \Big(\sum_{i=0}^{t-1} (-\hat{b})^i (X_{t-i} - \hat{\mu}) \Big)^2.$$

4. Porovnání odhadů

V této kapitole uvedeme pro porovnání odhadů parametrů zmíněnými metodami několik příkladů. Pro zadané parametry příslušného modelu a počet pozorování $n = 50$ určíme střední kvadratickou chybu odhadů parametrů pomocí uvedených metod na základě 100 nezávislých opakování s nezměněnými parametry.

Zavedeme označení MM pro metodu momentů, LS pro metodu nejmenších čtverců (z anglického „least squares") a střední kvadratickou chybu odhadu parametru budeme značit jako parametr s vlnkou, např. $\tilde{\mu}$.

Výpočty jsou prováděny pomocí programu Mathematica (viz [4]).

Začneme odhady v modelech AR(1).

Příklad 1. Nechť $\mu = 0$, $\sigma^2 = 1$ a $\varphi = 0$, tedy členy náhodné posloupnosti $\{X_t, t \in \mathbb{Z}\}$ jsou nezávislé.

	$\tilde{\mu}$	$\tilde{\varphi}$	$\tilde{\sigma}^2$
MLE	0,0192467	0,0167068	0,0400955
MM	0,0192328	0,0161351	0,0401618
LS	0,0192328	0,016641	0,0400343

Tabulka 1. Střední kvadratické chyby odhadů parametrů $\mu = 0$, $\sigma^2 = 1$ a $\varphi = 0$ v modelu AR(1)

Příklad 2. Nechť $\mu = 0$, $\sigma^2 = 1$ a $\varphi = 0{,}2$.

	$\tilde{\mu}$	$\tilde{\varphi}$	$\tilde{\sigma}^2$
MLE	0,0308068	0,0216238	0,0400978
MM	0,0307676	0,0212977	0,0402084
LS	0,0307676	0,0217	0,040093

Tabulka 2. Střední kvadratické chyby odhadů parametrů $\mu = 0$, $\sigma^2 = 1$ a $\varphi = 0{,}2$ v modelu AR(1)

Příklad 3. Nechť $\mu = 0$, $\sigma^2 = 1$ a $\varphi = 0{,}8$.

	$\tilde{\mu}$	$\tilde{\varphi}$	$\tilde{\sigma}^2$
MLE	0,429203	0,0168929	0,0417255
MM	0,434898	0,0199234	0,0508741
LS	0,434898	0,0181188	0,0413954

Tabulka 3. Střední kvadratické chyby odhadů parametrů $\mu = 0$, $\sigma^2 = 1$ a $\varphi = 0{,}8$ v modelu AR(1)

Z příkladů 1.–3. vidíme, že s rostoucí hodnotou φ roste chyba odhadů parametru μ, tj. pro silněji korelovaná data se odhad μ zhoršuje.

V tomto množství dat je rozdíl mezi středními kvadratickými chybami odhadů metodou maximální věrohodnosti a odhadů zbylými metodami jen zhruba v řádu tisícin. Pokud uvážíme i početní náročnost, pak je jistě lepší použít metodu momentů nebo metodu nejmenších čtverců.

Nyní budeme odhadovat parametry modelu MA(1).

Příklad 4. Nechť $\mu = 0$, $\sigma^2 = 1$ a $b = 0{,}2$.

	$\tilde{\mu}$	$\tilde{b}$	$\tilde{\sigma}^2$
MLE	0,0265716	0,0242992	0,0379697
MM	0,026246	0,0219939	0,0385169
LS	0,026246	0,0237	0,0377884

Tabulka 4. Střední kvadratické chyby odhadů parametrů $\mu = 0$, $\sigma^2 = 1$ a $b = 0{,}2$ v modelu MA(1)

Příklad 5. Nechť $\mu = 0$, $\sigma^2 = 1$ a $b = 0{,}8$.

	$\tilde{\mu}$	$\tilde{b}$	$\tilde{\sigma}^2$
MLE	0,0551705	0,0318243	0,0539633
MM	0,0568682	0,0793134	0,0608907
LS	0,0568682	0,015481	0,0376756

Tabulka 5. Střední kvadratické chyby odhadů parametrů $\mu = 0$, $\sigma^2 = 1$ a $b = 0{,}8$ v modelu MA(1)

Příklad 6. Nechť $\mu = 0$, $\sigma^2 = 1$ a $b = -0{,}5$.

	$\tilde{\mu}$	$\tilde{b}$	$\tilde{\sigma}^2$
MLE	0,00343956	0,0393226	0,0475764
MM	0,0035013	0,0555001	0,0702636
LS	0,0035013	0,0345117	0,0473093

Tabulka 6. Střední kvadratické chyby odhadů parametrů $\mu = 0$, $\sigma^2 = 1$ a $b = -0{,}5$ v modelu MA(1)

Z příkladů 4.–6. můžeme usuzovat, že metoda momentů je pro model MA(1) vhodná pouze, pokud je hodnota parametru b blízko nuly.

Metoda maximální věrohodnosti nedává výrazně lepší (při těchto simulacích často ani lepší) výsledky než metoda nejmenších čtverců a uvážíme-li početní náročnost, pak při tomto množství dat je pro nás nejvýhodnější použít metodu nejmenších čtverců.

Závěr

Cílem této práce bylo odvodit maximálně věrohodné odhady parametrů pro konkrétní typy časových řad a získané výsledky srovnat s odhady pomocí jiných metod.

Nejprve jsme ukázali, jak se obecně postupuje při určování odhadu parametrů metodou maximální věrohodnosti. Pak jsme definovali tři základní lineární modely časových řad: posloupnost klouzavých součtů, autoregresní posloupnost a jejich kombinaci. U zmíněných modelů jsme uvedli jejich základní vlastnosti jako jsou střední hodnota, rozptyl či autokovarianční funkce.

V druhé kapitole jsme přistoupili k určování odhadů parametrů časových řad metodou maximální věrohodnosti. Podrobně jsme ukázali, jak odvodit odhady parametrů v modelech AR(1) a MA(1). Zjistili jsme, že již u těchto jednoduchých modelů je k dopočítání kýžených odhadů nutné použít numerické metody. Dále jsme nastínili, jak získat odhady parametrů v modelech AR(m), MA(p) a ARMA(m, p) a zmínili se o podmíněné metodě maximální věrohodnosti. Ta je na rozdíl od nepodmíněné metody maximální věrohodnosti jednodušší k počítání, ovšem často na úkor přesnosti výsledků.

Třetí kapitola se věnuje odhadům parametrů v modelech AR(1) a MA(1) metodou momentů a metodou nejmenších čtverců. Tentokrát musíme dopočítat řešení numericky pouze u podmíněné metody nejmenších čtverců, kterou využíváme k určení odhadů parametrů v modelu MA(1).

Na závěr je uvedeno po třech příkladech pro modely AR(1) a MA(1), které srovnávají všechny tři metody. Z tohoto srovnání obecně vyšla nejlépe metoda nejmenších čtverců, přičemž jsme hleděli na přesnost odhadů a početní náročnost jejich určení.

Seznam použitých zkratek

$\mathbf{A}$	matice
$\mathbf{A}^T$	transponovaná matice k $\mathbf{A}$
$\mathbf{A}^{-1}$	inverzní matice k $\mathbf{A}$
$a_{i,j}$	prvek $\mathbf{A}$ na pozici (i,j)
AR(m)	autoregresní posloupnost řádu m
ARMA(m,p)	smíšený model autoregrese a klouzavých součtů
$\mathcal{B}^k$	systém borelovských množin v $\mathbb{R}^k$
$\det \mathbf{A}$	determinant matice $\mathbf{A}$
$L(\mathbf{x};\boldsymbol{\theta})$	věrohodnostní funkce
LS	metoda nejmenších čtverců
MA(p)	posloupnost klouzavých součtů řádu p
MLE	maximálně věrohodný odhad
MM	metoda momentů
$\tilde{\mu}$	střední kvadratická chyba odhadu parametru μ
$\mathbb{N}$	množina přirozených čísel
N$(0,\ \sigma^2)$	normální rozdělení se střední hodnotou 0 a rozptylem σ^2
$\mathbb{R}^k$	k-rozměrný vektorový prostor reálných čísel
$\mathbf{x}$	sloupcový vektor
$\hat{\theta}$	odhad parametru θ
WN $(0,\sigma^2)$	bílý šum (s nulovou střední hodnotou a rozptylem σ^2)
$\mathbb{Z}$	množina celých čísel

Seznam použité literatury

[1] DUPAČ, Václav – HUŠKOVÁ, Marie. *Pravděpodobnost a matematická statistika.* 1. vydání. Praha: Karolinum. Univerzita Karlova. Matematicko-fyzikální fakulta. Katedra pravděpodobnosti a matematické statistiky. 2005. v, 164 s. ISBN 80-246-0009-9.

[2] HAMILTON, James Douglas. *Time Series Analysis.* 1. vydání. Princeton (New Jersey): Princeton University Press, c 1994. xiv, 800 s. ISBN 0-691-04289-6.

[3] PRÁŠKOVÁ, Zuzana. *Základy náhodných procesů II.* 1. vydání. Praha: Karolinum. Univerzita Karlova. 2007. 152 s. ISBN 978-80-246-0971-3.

[4] Wolfram Research, Inc. *Wolfram Mathematica* [počítačový program]. Ver. 7.0.0, Champaign, 2008.