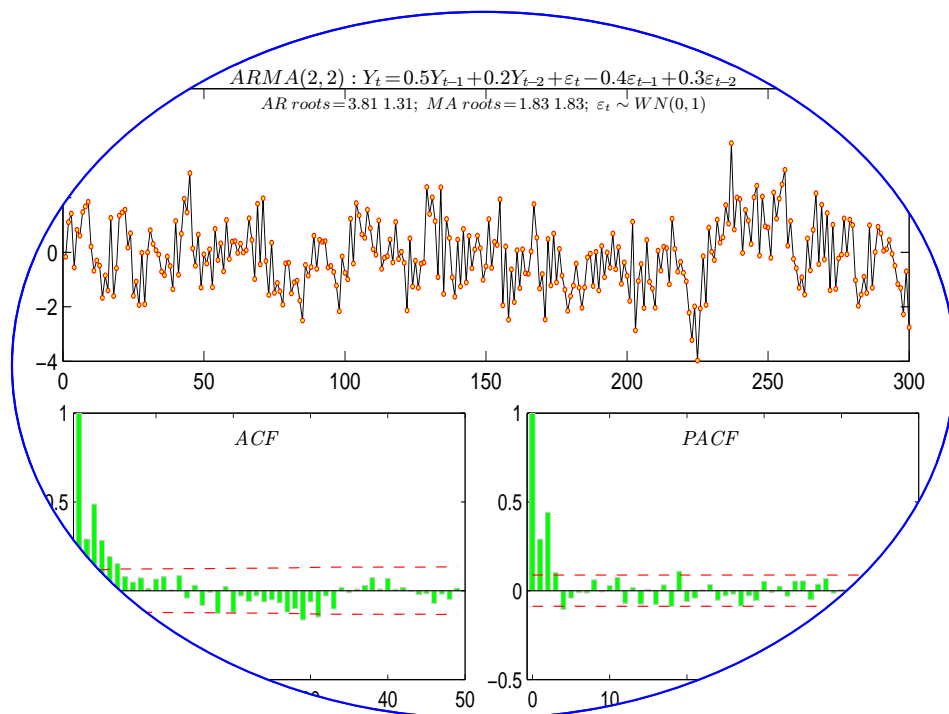


MARIE FORBELSKÁ

STOCHASTICKÉ MODELOVÁNÍ JEDNOROZMĚRNÝCH ČASOVÝCH ŘAD II



BRNO 2007



Obsah

Seznam použitých symbolů	1
1 Box-Jenkinsova metodologie	3
1.1 Úvod	3
1.2 Základní pojmy	4
1.2.1 Operátor zpětného posunutí	4
1.2.2 Lineární proces	4
1.2.3 Lineární filtry	8
1.2.4 Autokovarianční a autokorelační generující funkce stacionárních procesů	10
1.2.5 Definice ARMA procesu	12
1.2.6 Kauzalita	13
1.2.7 Invertibilita	22
1.2.8 Vícenásobná reprezentace $MA(q)$ procesů	28
1.2.9 Společné kořeny polynomů $\Phi(z)$ a $\Theta(z)$	29
1.2.10 Nutná a postačující podmínka kauzality a invertibility ARMA procesu.	29
1.2.11 Střední hodnota, rozptyl, autokovarianční a autokorelační funkce procesů $ARMA(p, q)$	30
1.2.12 Spektrální hustota $ARMA(p, q)$ procesů	32
1.3 Stacionární procesy a nejlepší lineární predikce	35
1.4 Rekurentní metody pro výpočet nejlepší lineární predikce	38
1.4.1 Důsledek Durbin-Levinsonova algoritmu	43
1.5 Parciální autokorelační funkce (PACF)	44
1.6 Inovační algoritmus	45
1.6.1 Jednokroková nejlepší lineární predikce v $AR(p)$	50
1.6.2 Vícekroková nejlepší lineární predikce v $AR(p)$	51
1.6.3 PACF pro $AR(p)$, $MA(q)$ a $ARMA(p, q)$	52
1.6.4 Jednokroková nejlepší lineární predikce v $MA(q)$	53
1.6.5 Nejlepší lineární predikce v $ARMA(p, q)$	54
1.7 Výstavba modelů v B-J metodologii	57
1.7.1 Odhady v ARMA procesech	57
1.7.2 Yuleovy-Walkerovy rovnice a odhad parametrů v $AR(p)$	61

1.7.3	Předběžné odhady v $AR(p)$ a Durbin-Levinsův algoritmus.	62
1.7.4	Předběžné odhady v $MA(q)$ a inovační algoritmus.	63
1.7.5	Předběžné odhady v $ARMA(p, q)$ procesu.	64
1.7.6	Maximálně věrohodné odhady.	66
1.8	Výstavba modelů a predikce v ARIMA procesech	71
1.8.1	Procesy nestacionární ve střední hodnotě	71
1.8.2	Procesy nestacionární v rozptylu	73
1.8.3	Volba řádu modelu	81
1.8.4	Verifikace modelu – analýza reziduí	83
1.9	Modelování sezónnosti pomocí SARIMA modelů	84
2	Dynamické lineární modely	89
2.1	Motivační příklad	89
2.2	Stavově-prostorové modely	91
2.3	Stacionární stavově-prostorové modely	94
2.4	Nejlepší lineární predikce pomocí projekce náhodných vektorů druhého řádu	96
2.5	Kalmanův iterační proces	105
	Literatura	107

Seznam použitých symbolů

\mathbb{N}	množina přirozených čísel
\mathbb{Z}	množina celých čísel
\mathbb{R}	množina reálných čísel
\mathbb{R}^m	reálný m -rozměrný euklidovský prostor
a, b, \dots	reálná čísla
$\mathbf{a}, \mathbf{b}, \dots$	vektory reálných čísel (sloupcové)
$\ \mathbf{a}\ $	euklidovská norma vektoru \mathbf{a} , tj. $\ \mathbf{a}\ = \sqrt{\mathbf{a}'\mathbf{a}}$
$\mathbf{A}, \mathbf{B}, \dots$	matice
\mathbf{I}_n	jednotková matice typu $(n \times n)$
$\mathbf{0}$	nulový vektor nebo nulová matice
$\mathbf{1}_n$	n -rozměrný sloupcový vektor samých jedniček
$h(\mathbf{A})$	hodnota matice
\mathbf{A}'	transponovaná matice
\mathbf{A}^{-1}	inverzní matice
\mathbf{A}^-	pseudoinverzní matice, tj. $\mathbf{A}^- = \mathbf{A}\mathbf{A}^- \mathbf{A}$
$ \mathbf{A} , \det(\mathbf{A})$	determinant matice
$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}$	bloková matice
$\ln x$	přirozený logaritmus se základem e
Ω	prostor elementárních jevů
\mathcal{A}	σ -algebra
\mathcal{B}^m	borelovská σ -algebra v m -rozměrném euklidovském prostoru
I_A	indikátor množiny A
$A \cup B$	sjednocení množin
$A \cap B$	průnik množin
$A - B$	rozdíl množin
\emptyset	prázdná množina
$\langle x, y \rangle$	skalární součin
$P(A)$	pravděpodobnost jevu A
(Ω, \mathcal{A}, P)	pravděpodobnostní prostor
X, Y, \dots	náhodné veličiny
x, y, \dots	realizace náhodných veličin
$\mathbf{X}, \mathbf{Y}, \dots$	náhodný vektor
$\mathbf{x}, \mathbf{y}, \dots$	realizace náhodných vektorů

EX	střední hodnota náhodné veličiny X
$C(X, Y)$	kovariance náhodných veličin X a Y
DX	rozptyl náhodné veličiny X
$C(\mathbf{X}, \mathbf{Y})$	kovarianční matice náhodných vektorů \mathbf{X} a \mathbf{Y}
$D\mathbf{X}$	varianční matice náhodného vektoru \mathbf{X}
$X \sim \mathcal{L}(\mu, \sigma^2)$	náhodná veličina X s rozdělením se střední hodnotou μ a rozptylem σ^2
$X \sim N(\mu, \sigma^2)$	náhodná veličina X se řídí normálním rozdělením se střední hodnotou μ a rozptylem σ^2
$Rs(a, b)$	rozdělení rovnoměrně spojitě na intervalu (a, b) , $a, b \in \mathbb{R}$, $a < b$
$\chi^2(n)$	χ^2 -rozdělení o n stupních volnosti
$N_m(\boldsymbol{\mu}, \boldsymbol{\Sigma})$	m -rozměrné normální rozdělení se střední hodnotou $\boldsymbol{\mu}$ a kovarianční maticí $\boldsymbol{\Sigma}$
$\{X_t, t \in T\}$	náhodný proces, T je neprázdná množina
$\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$	bílý šum, kde $E\varepsilon_t = 0$, $D\varepsilon_t = \sigma_\varepsilon^2$, $C(\varepsilon_t, \varepsilon_s) = 0$, ($s \neq t$)
$\varepsilon_t \sim IID(0, \sigma_\varepsilon^2)$	<i>IID</i> (independent identical defined) proces, bílý šum s nezávislými náhodnými veličinami
$X_{t_0} = \text{l.i.m.}_{t \rightarrow t_0} X_t$	limita podle kvadratického středu v bodě t_0 , anglicky <i>limit in the mean</i> , t.j. $E X_t - X_{t_0} ^2 \rightarrow 0$ pro $t \rightarrow t_0$
$L^2(\Omega, \mathcal{A}, P)$,	Hilbertův prostor reálných náhodných veličin druhého řádu, tj. veličin s konečnými druhými momenty, se skalárním součinem $\langle X, Y \rangle = EXY$
$L^2_{\mathbb{C}}(\Omega, \mathcal{A}, P)$	Hilbertův prostor komplexních náhodných veličin druhého řádu se skalárním součinem $\langle X, Y \rangle = EX\bar{Y}$
B	operátor zpětného chodu, tj. $BX_t = X_{t-1}$
Δ	diferenční operátor, tj. $\Delta X_t = X_t - X_{t-1}$
Δ_L	diferenční operátor o délce $L > 0$, $L \in \mathbb{N}$, tj. $\Delta_L X_t = X_t - X_{t-L}$
$f_X(\omega)$	spektrální hustota náhodného procesu $\{X_t, t \in T\}$
$\gamma(t)$	autokovarianční funkce stacionárního procesu
$\rho(t)$	autokorelační funkce stacionárního procesu
$\alpha(t)$	parciální autokorelační funkce stacionárního procesu
$ARMA(p, q)$	ARMA proces řádu p a q ($p, q \in \mathbb{N}$)
$AR(p)$	AR (autokorelační) proces řádu p ($p \in \mathbb{N}$)
$MA(q)$	MA proces řádu q ($q \in \mathbb{N}$)
$ARIMA(p, d, q)$	smíšený proces řádu p , d a q ($p, d, q \in \mathbb{N}$)

Kapitola 1

Box-Jenkinsova metodologie

1.1 Úvod

V předchozím jsme podali výklad analýzy jednorozměrných časových řad založený na dekompozičním principu. Pokoušeli jsme se vyčlenit (a posléze modelovat) deterministickou část časové řady (deterministický trend, deterministickou sezónní složku) a stochastickou část jsme hlouběji neanalyzovali - chápali jsme ji jako část zbytkovou.

Ve skutečnosti však také v reziduální (zbytkové) části existují jisté systematickosti, jež jsou velmi významné, a je proto nutné se jimi zabývat. Box-Jenkinsova metodologie (zkráceně B-J metodologie) ztotožňuje systematickou část časové řady s částí deterministickou a je založena na myšlence, že časová řada může být chápána jako řada stochastického charakteru. Zásluha Boxe a Jenkinse nespočívá v objevení principů, které budeme popisovat, ale ve vytvoření konkrétního postupu, jak tyto principy prakticky využívat.

Téměř každý kvalitní statistický programový paket tuto metodologii obsahuje (SAS, BMDP, SPSS, STATGRAPHICS, RATS,)

Výhody B-J metodologie:

- je flexibilní a rychle se adaptuje na změnu v charakteru modelovaného procesu
- v mnoha případech dává nejlepší výsledky (vzhledem k MSE-střední čtvercové chybě)

Nevýhody B-J metodologie:

- musí být dostatečně dlouhé realizace
- ztrácí se možnost jednoduché interpretace výsledných modelů (lze těžko přesvědčit zadavatele, že řadu lze modelovat pomocí náhodných šoků. Jediným argumentem jsou zde často jen kvalitní předpovědi získané pomocí těchto modelů).

1.2 Základní pojmy

V dalším budeme uvažovat centrované stacionární náhodné posloupnosti $\{Y_t, t \in \mathbb{Z}\}$, kde $Y_t \in L^2(\Omega, \mathcal{A}, P)$, což je Hilbertův prostor reálných náhodných veličin s konečnými druhými momenty, ve kterém dvě náhodné veličiny X a Y považujeme za ekvivalentní, pokud

$$P(X = Y) = 1.$$

1.2.1 Operátor zpětného posunutí

Definice 1.2.1.

Nechť $\{Y_t, t \in \mathbb{Z}\}$ je posloupnost náhodných veličin. **Operátor zpětného posunutí** je definován pomocí výrazu

$$BY_t = Y_{t-1},$$

přičemž jej lze aplikovat několikanásobně jako

$$B^j Y_t = Y_{t-j}.$$

1.2.2 Lineární proces

Než zavedeme pojem lineárního procesu, vyslovme větu, která zabezpečuje jeho korektnost.

Věta 1.2.2.

Nechť $\{\varepsilon_t, t \in \mathbb{Z}\} \sim WN(0, \sigma_\varepsilon^2)$ je bílý šum, dále mějme posloupnost reálných čísel $\{\psi_j\}_{j=0}^\infty$ takovou, že $\sum_{j=0}^\infty \psi_j^2 < \infty$. Pak řada $\sum_{j=0}^\infty \psi_j \varepsilon_j$ konverguje podle kvadratického středu, tj. existuje náhodná veličina $Y \in L^2(\Omega, \mathcal{A}, P)$ a platí

$$Y = \text{l.i.m.}_{N \rightarrow \infty} \sum_{j=0}^N \psi_j \varepsilon_j.$$

DŮKAZ. Víme, že bílý šum $\varepsilon_t \in L^2(\Omega, \mathcal{A}, P)$. Pro libovolná přirozená čísla

$k, N \in \mathbb{N}$ platí

$$\begin{aligned} \left\| \sum_{j=0}^{N+k} \psi_j \varepsilon_j - \sum_{t=0}^N \psi_t \varepsilon_t \right\|^2 &= E \left| \sum_{j=0}^{N+k} \psi_j \varepsilon_j - \sum_{t=0}^N \psi_t \varepsilon_t \right|^2 = E \left| \sum_{j=N+1}^{N+k} \psi_j \varepsilon_j \right|^2 \\ &= E \left(\sum_{j=N+1}^{N+k} \psi_j \varepsilon_j \right) \left(\sum_{h=N+1}^{N+k} \psi_h \varepsilon_h \right) \\ &= \sum_{j=N+1}^{N+k} \sum_{h=N+1}^{N+k} \psi_j \psi_h E \underbrace{\varepsilon_j \varepsilon_h}_{\text{nekorel.}} = \sigma_\varepsilon^2 \sum_{j=N+1}^{N+k} \psi_j^2 \xrightarrow{N \rightarrow \infty} 0 \end{aligned}$$

Posloupnost částečných součtů je tedy cauchyovská, tj. existuje k ní limita

$$Y = \text{l.i.m.}_{N \rightarrow \infty} \sum_{j=0}^N \psi_j \varepsilon_j. \quad \blacksquare$$

Definice 1.2.3.

Mějme $\{\varepsilon_t, t \in \mathbb{Z}\} \sim WN(0, \sigma_\varepsilon^2)$ a posloupnost reálných čísel $\{\psi_j\}_{j=0}^\infty$ takovou, že $\sum_{j=0}^\infty \psi_j^2 < \infty$, pak **lineární proces** je definován vztahem

$$Y_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}.$$

Počítejme postupně střední hodnotu, rozptyl a autokovarianční funkci lineárního procesu a přesvědčeme se, že lineární proces je stacionární.

$$\begin{aligned} \boxed{EY_t} &= E \left(\sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} \right) = \sum_{j=0}^{\infty} \psi_j \underbrace{E\varepsilon_{t-j}}_{=0} = \boxed{0} \\ \boxed{DY_t} &= D \left(\sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} \right) \stackrel{\text{nekorel.}}{=} \sum_{j=0}^{\infty} \psi_j^2 \underbrace{D\varepsilon_{t-j}}_{=\sigma_\varepsilon^2} = \boxed{\sigma_\varepsilon^2 \sum_{j=0}^{\infty} \psi_j^2} = \sigma_Y^2 \\ \boxed{\gamma(t)} &= C(Y_s, Y_{s+t}) = EY_s Y_{s+t} = E \left(\sum_{j=0}^{\infty} \psi_j \varepsilon_{s-j} \right) \left(\sum_{h=0}^{\infty} \psi_h \varepsilon_{s+t-h} \right) \\ &= \sum_{j=0}^{\infty} \sum_{h=0}^{\infty} \psi_j \psi_h E \underbrace{\varepsilon_{s-j} \varepsilon_{s+t-h}}_{\text{nekorel.}} = \left| \begin{array}{cc} s-j & = s+t-h \\ h & = j+t \end{array} \right| \\ &= \boxed{\sigma_\varepsilon^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+t}}. \end{aligned}$$

Ze Schwarzovy nerovnosti dostaneme

$$|\gamma(t)| = |C(Y_s, Y_{s+t})| = \left| \sigma_\varepsilon^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+t} \right| \leq \sqrt{DY_s DY_{s+t}} = \gamma(0) = \sigma_\varepsilon^2 \sum_{j=0}^{\infty} \psi_j^2 < \infty.$$

Podmínka stacionarity je tedy podmínka

$$\sum_{j=0}^{\infty} \psi_j^2 < \infty.$$

Pokud zavedeme funkci

$$\Psi(z) = \sum_{j=0}^{\infty} \psi_j z^j,$$

pak podmínka

$$\sum_{j=0}^{\infty} \psi_j^2 < \infty$$

implikuje, že funkce

$\Psi(z)$ je holomorfní uvnitř kružnice $|z| < 1$.

Takže podmínku stacionarity lze vyslovit i pomocí podmínky

$$\Psi(z) \text{ je holomorfní pro } |z| < 1, \text{ přičemž } \sum_{j=0}^{\infty} \psi_j^2 < \infty.$$

Oba požadavky budou splněny, pokud bude platit

$$\Psi(z) \text{ je holomorfní uvnitř a na jednotkové kružnici.}$$

Lineární proces lze ještě zobecnit takto:

Definice 1.2.4.

Mějme $\{\varepsilon_t, t \in \mathbb{Z}\} \sim WN(0, \sigma_\varepsilon^2)$ a posloupnost reálných čísel $\{\psi_j\}_{j=-\infty}^{\infty}$ takovou, že $\sum_{j=-\infty}^{\infty} \psi_j^2 < \infty$, pak **zobecněný lineární proces** je definován vztahem

$$Y_t = \sum_{j=-\infty}^{\infty} \psi_j \varepsilon_{t-j}.$$

Pro takto definovaný zobecněný lineární proces dokážeme obdobným způsobem jak pro obyčejný lineární proces spočítat

$$EY_t = 0, \quad DY_t = \sigma_\varepsilon^2 \sum_{j=-\infty}^{\infty} \psi_j^2 \quad \text{a} \quad \gamma(t) = \sigma_\varepsilon^2 \sum_{j=-\infty}^{\infty} \psi_j \psi_{j+t}.$$

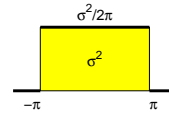
Na závěr tohoto odstavce počítejme ještě **spektrální hustotu zobecněného lineárního procesu.**

Nejprve odvodíme spektrální hustotu bílého šumu, a to pomocí jeho autokovarianční funkce $\gamma_\varepsilon(t)$

$$f_\varepsilon(\omega) = \frac{1}{2\pi} \sum_{t=-\infty}^{\infty} e^{-it\omega} \gamma_\varepsilon(t) = \frac{1}{2\pi} \sum_{t=-\infty}^{\infty} e^{-it\omega} \sigma_\varepsilon^2 \delta(t) = \begin{cases} \frac{\sigma_\varepsilon^2}{2\pi} & \omega \in \langle -\pi, \pi \rangle, \\ 0 & \text{jinak} \end{cases}$$

kde

$$\delta(t) = \begin{cases} 1 & t = 0, \\ 0 & \text{jinak.} \end{cases}$$



Pak pomocí autokovarianční funkce zobecněného lineárního procesu počítáme spektrální hustotu pro $\omega \in \langle -\pi, \pi \rangle$

$$\begin{aligned} f_Y(\omega) &= \frac{1}{2\pi} \sum_{t=-\infty}^{\infty} e^{-it\omega} \gamma(t) = \frac{1}{2\pi} \sum_{t=-\infty}^{\infty} e^{-it\omega} \left(\sigma_\varepsilon^2 \sum_{j=-\infty}^{\infty} \psi_j \psi_{j+t} \right) \\ &= \underbrace{\frac{\sigma_\varepsilon^2}{2\pi}}_{f_\varepsilon(\omega)} \sum_{j=-\infty}^{\infty} \psi_j \sum_{t=-\infty}^{\infty} \psi_{j+t} e^{-it\omega} = f_\varepsilon(\omega) \sum_{j=-\infty}^{\infty} \psi_j e^{ij\omega} \sum_{t=-\infty}^{\infty} \psi_{j+t} e^{-i(j+t)\omega} \\ &= f_\varepsilon(\omega) \left| \sum_{t=-\infty}^{\infty} \psi_j e^{-it\omega} \right|^2 = f_\varepsilon(\omega) \left| \sum_{t=-\infty}^{\infty} \psi_j e^{it\omega} \right|^2 \quad \text{neboť } |z|^2 = z \cdot \bar{z} \end{aligned}$$

Pokud položíme $\Psi(z) = \sum_{j=-\infty}^{\infty} \psi_j z^j$, pak můžeme psát

$$f_Y(\omega) = f_\varepsilon(\omega) |\Psi(e^{-i\omega})|^2 = \frac{\sigma_\varepsilon^2}{2\pi} |\Psi(e^{-i\omega})|^2 \quad \left(= \frac{\sigma_\varepsilon^2}{2\pi} |\Psi(e^{i\omega})|^2 \right).$$

1.2.3 Lineární filtry

Věta 1.2.5.

Nechť $\{X_t, t \in \mathbb{Z}\}$ je (centrovaná) stacionární náhodná posloupnost a $\{\psi_j\}_{j=-\infty}^{\infty}$ je absolutně konvergentní posloupnost reálných čísel (tj. $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$). Pak platí $Y_t = \sum_{j=-\infty}^{\infty} \psi_j X_{t-j} \in L^2(\Omega, \mathcal{A}, P)$ tj. $\{Y_t, t \in \mathbb{Z}\}$ je stacionární náhodná posloupnost.

DŮKAZ. Je zřejmé, že stačí dokázat existenci náhodných veličin

$$Y_t^{(1)} = \sum_{j=-\infty}^{-1} \psi_j X_{t-j} \quad \text{a} \quad Y_t^{(2)} = \sum_{j=0}^{\infty} \psi_j X_{t-j},$$

protože pak bude platit $Y_t = Y_t^{(1)} + Y_t^{(2)}$. Označme

$$\gamma_X(h) = EX_t E_{t+|h|} \quad \text{a} \quad \gamma_X(0) = \sigma_X^2 > 0.$$

Pak pro libovolná přirozená čísla $k, N \in \mathbb{N}$ platí

$$\begin{aligned} \left\| \sum_{j=0}^{N+k} \psi_j X_{t-j} - \sum_{h=0}^N \psi_h X_{t-h} \right\|^2 &= E \left| \sum_{j=0}^{N+k} \psi_j X_{t-j} - \sum_{h=0}^N \psi_h X_{t-h} \right|^2 \\ &= E \left| \sum_{j=N+1}^{N+k} \psi_j X_{t-j} \right|^2 = E \left(\sum_{j=N+1}^{N+k} \psi_j X_{t-j} \right) \left(\sum_{h=N+1}^{N+k} \psi_h X_{t-h} \right) \\ &= \sum_{j=N+1}^{N+k} \sum_{h=N+1}^{N+k} \psi_j \psi_h \underbrace{EX_{t-j} X_{t-h}}_{|\gamma(j-h)| \leq \gamma_X(0) = \sigma_X^2} \\ &\leq \sigma_X^2 \sum_{j=N+1}^{N+k} \sum_{h=N+1}^{N+k} |\psi_j| |\psi_h| = \underbrace{\sigma_X^2}_{< \infty} \underbrace{\left(\sum_{j=N+1}^{N+k} |\psi_j| \right)}_{\rightarrow 0} \xrightarrow{N \rightarrow \infty} 0. \end{aligned}$$

Posloupnost částečných součtů je tedy cauchyovská (podle kvadratického středu), tj. existuje k ní limita

$$Y_t^{(1)} = \text{l.i.m.}_{N \rightarrow \infty} \sum_{j=0}^N \psi_j X_{t-j}, \quad \equiv \quad Y_t^{(1)} \in L^2(\Omega, \mathcal{A}, P),$$

tj. $Y_t^{(1)}$ má nulovou střední hodnotu a konečný rozptyl a je tedy stacionární. Podobně se dokáže i existence stacionární náhodné posloupnosti $Y_t^{(2)}$. ■

Definice 1.2.6.

Nechť $\{X_t, t \in \mathbb{Z}\}$ je stacionární náhodná posloupnost a $\{\psi_j\}_{j=-\infty}^{\infty}$ je absolutně konvergentní posloupnost reálných čísel (tj. $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$). Pak

$$Y_t = \sum_{j=-\infty}^{\infty} \psi_j X_{t-j} \quad \text{nazveme } \mathbf{lineárním filtrem} \text{ procesu } \{X_t, t \in \mathbb{Z}\}.$$

Věta 1.2.7.

Mějme centrovanou stacionární náhodnou posloupnost $\{X_t, t \in \mathbb{Z}\}$ se spektrální hustotou $f_X(\omega)$. Nechť $\{\psi_j\}_{j=-\infty}^{\infty}$ je absolutně konvergentní posloupnost reálných čísel (tj. $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$). Pak náhodná posloupnost $Y_t = \sum_{j=-\infty}^{\infty} \psi_j X_{t-j}$ je stacionární se spektrální hustotou

$$f_Y(\omega) = f_X(\omega) |\Psi(e^{-i\omega})|^2, \quad \text{kde} \quad \Psi(z) = \sum_{j=-\infty}^{\infty} \psi_j z^j \quad |z| \leq 1$$

se nazývá **generující funkce filtru** a

$$\psi(\omega) = \Psi(e^{-i\omega}) \quad \mathbf{přenosová funkce filtru}.$$

DŮKAZ. Stacionaritu jsme dokázali v předchozí větě. Nyní počítejme autokovarianční funkci.

$$\begin{aligned} \gamma_Y(t) &= C(Y_s, Y_{s+t}) = C\left(\sum_{j=-\infty}^{\infty} \psi_j X_{s-j}, \sum_{h=-\infty}^{\infty} \psi_h X_{s+t-h}\right) \\ &= \sum_{j=-\infty}^{\infty} \sum_{h=-\infty}^{\infty} \psi_j \psi_h C(X_{s-j}, X_{s+t-h}) = \sum_{j=-\infty}^{\infty} \sum_{h=-\infty}^{\infty} \psi_j \psi_h \gamma_X(t+j-h) \\ &= \sum_{j=-\infty}^{\infty} \sum_{h=-\infty}^{\infty} \psi_j \psi_h \int_{-\pi}^{\pi} e^{i(t+j-h)\omega} f_X(\omega) d\omega \\ &= \int_{-\pi}^{\pi} e^{it\omega} \left(\sum_{j=-\infty}^{\infty} \psi_j e^{ij\omega}\right) \left(\sum_{h=-\infty}^{\infty} \psi_h e^{-ih\omega}\right) f_X(\omega) d\omega \\ &= \int_{-\pi}^{\pi} e^{it\omega} \left|\sum_{j=-\infty}^{\infty} \psi_j e^{ij\omega}\right|^2 f_X(\omega) d\omega = \int_{-\pi}^{\pi} e^{it\omega} \left|\sum_{h=-\infty}^{\infty} \psi_h e^{-ih\omega}\right|^2 f_X(\omega) d\omega. \end{aligned}$$

Označme

$$\Psi(z) = \sum_{j=-\infty}^{\infty} \psi_j z^j \quad \text{pro} \quad |z| \leq 1.$$

Pak, protože platí

$$\gamma_Y(t) = \int_{-\pi}^{\pi} e^{it\omega} f_Y(\omega) d\omega,$$

dostaneme

$$f_Y(\omega) = f_X(\omega) |\Psi(e^{-i\omega})|^2 = f_X(\omega) |\Psi(e^{i\omega})|^2.$$

■

1.2.4 Autokovarianční a autokorelační generující funkce stationárních procesů

Při vyšetřování vlastností autokovarianční a autokorelační funkce stationárních procesů je užitečné zavést následující transformaci

$$G(z) = \sum_{k=-\infty}^{\infty} \gamma(k) z^k \quad \text{resp.} \quad R(z) = \sum_{k=-\infty}^{\infty} \rho(k) z^k,$$

které se nazývají **autokovarianční**, resp. **autokorelační**, **generující funkce**.

Všimněme si ještě vztahu mezi spektrální hustotou a autokovarianční generující funkcí

$$\boxed{f(\omega)} = \frac{1}{2\pi} \sum_{t=-\infty}^{\infty} \gamma(t) e^{-it\omega} = \boxed{\frac{1}{2\pi} G(e^{-i\omega})}.$$

BÍLÝ ŠUM

PŘÍKLAD 1.2.8.

Mějme bílý šum $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$, s autokovarianční funkcí $\gamma_\varepsilon(k) = \begin{cases} \sigma_\varepsilon^2 & k = 0, \\ 0 & \text{jinak} \end{cases}$

takže

$$\boxed{G_\varepsilon(z)} = \sum_{k=-\infty}^{\infty} \gamma(k) z^k = \boxed{\sigma_\varepsilon^2} \quad \text{a} \quad R_\varepsilon(z) = \sum_{k=-\infty}^{\infty} \rho(k) z^k = 1.$$

ZOBECNĚNÝ LINEÁRNÍ PROCES

PŘÍKLAD 1.2.9.

Mějme posloupnost reálných čísel $\{\psi_j\}_{j=-\infty}^{\infty}$ takovou, že $\sum_{j=-\infty}^{\infty} \psi_j^2 < \infty$, pak **zobecněný lineární proces** je definován vztahem

$$X_t = \sum_{j=-\infty}^{\infty} \psi_j \varepsilon_{t-j}$$

a má autokovarianční funkci tvaru

$$\gamma_X(k) = \sigma_\varepsilon^2 \sum_{j=-\infty}^{\infty} \psi_j \psi_{j+k}.$$

Odtud autokovarianční generující funkce a spektrální hustota jsou rovny

$$\begin{aligned} \boxed{G_X(z)} &= \sum_{k=-\infty}^{\infty} \gamma(k) z^k = \sigma_\varepsilon^2 \sum_{k=-\infty}^{\infty} z^k \sum_{j=-\infty}^{\infty} \psi_j \psi_{j+k} = \sigma_\varepsilon^2 \sum_{k=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \psi_j z^{-j} \psi_{j+k} z^{j+k} \\ &= \left| \begin{array}{l} \text{subst.} \\ h = j + k \end{array} \right| = \underbrace{\sigma_\varepsilon^2}_{=G_\varepsilon(z)} \sum_{h=-\infty}^{\infty} \psi_h z^h \sum_{j=-\infty}^{\infty} \psi_h z^{-j} = \boxed{G_\varepsilon(z) \Psi_X(z) \Psi_X(z^{-1})}. \\ \boxed{f_X(\omega)} &= \frac{1}{2\pi} G_X(e^{-i\omega}) = \underbrace{\frac{1}{2\pi} G_\varepsilon(e^{-i\omega})}_{f_\varepsilon(\omega)} \Psi_X(e^{-i\omega}) \Psi_X(e^{i\omega}) = \boxed{f_\varepsilon(\omega) |\Psi_X(e^{-i\omega})|^2}, \end{aligned}$$

kde $\boxed{\Psi_X(z)} = \sum_{j=-\infty}^{\infty} \psi_j z^j$ pro $|z| \leq 1$ se nazývá **generující funkce filtru** a $\boxed{\psi_X(\omega)} = \Psi_X(e^{-i\omega})$ **přenosová funkce filtru** (zobecněný lineární proces je také filtrem).

LINEÁRNÍ FILTR

PŘÍKLAD 1.2.10.

Mějme posloupnost reálných čísel $\{\psi_j\}_{j=-\infty}^{\infty}$ takovou, že $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$. Dále necht $\{X_t, t \in \mathbb{Z}\}$ je centrovaná stacionární náhodná posloupnost. Pak **lineární filtr** je definován vztahem

$$Y_t = \sum_{j=-\infty}^{\infty} \psi_j X_{t-j}$$

a má autokovarianční funkci

$$\begin{aligned} \boxed{\gamma_Y(k)} &= EY_s Y_{s+k} = E \left(\sum_{j=-\infty}^{\infty} \psi_j X_{s-j} \sum_{h=-\infty}^{\infty} \psi_h X_{s+k-h} \right) \\ &= \sum_{j=-\infty}^{\infty} \sum_{h=-\infty}^{\infty} \psi_j \psi_h E X_{s-j} X_{s+k-h} \\ &= \boxed{\sum_{j=-\infty}^{\infty} \sum_{h=-\infty}^{\infty} \psi_j \psi_h \gamma_X(k+j-h)}, \end{aligned}$$

takže

$$\begin{aligned}
 \boxed{G_Y(z)} &= \sum_{k=-\infty}^{\infty} \gamma(k) z^k = \sum_{k=-\infty}^{\infty} z^k \sum_{j=-\infty}^{\infty} \sum_{h=-\infty}^{\infty} \psi_j \psi_h \gamma_X(k+j-h) \\
 &= \sum_{k=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \sum_{h=-\infty}^{\infty} \psi_j z^{-j} \psi_h z^h \gamma_X(k+j-h) z^{k+j-h} = \left| \begin{array}{l} \text{subst.} \\ s = k+j-h \end{array} \right| \\
 &= \sum_{h=-\infty}^{\infty} \psi_h z^h \sum_{j=-\infty}^{\infty} \psi_j z^{-j} \underbrace{\sum_{s=-\infty}^{\infty} \gamma_X(s) z^s}_{=G_X(z)} = \boxed{G_X(z) \Psi_Y(z) \Psi_Y(z^{-1})}
 \end{aligned}$$

$$\boxed{f_Y(\omega)} = \frac{1}{2\pi} G_Y(e^{-i\omega}) = \frac{1}{2\pi} \underbrace{G_X(e^{-i\omega})}_{f_X(\omega)} \Psi_Y(e^{-i\omega}) \Psi_Y(e^{i\omega}) = \boxed{f_X(\omega) |\Psi_Y(e^{-i\omega})|^2},$$

kde $\Psi_Y(z) = \sum_{j=-\infty}^{\infty} \psi_j z^j$ pro $|z| \leq 1$ se nazývá **generující funkce filtru** a $\psi_Y(\omega) = \Psi_Y(e^{-i\omega})$ **přenosová funkce filtru**.

1.2.5 Definice ARMA procesu

Definice 1.2.11.

ARMA proces řádu p, q je definován vztahem

$$\boxed{Y_t - \varphi_1 Y_{t-1} - \dots - \varphi_p Y_{t-p} = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}}, \text{ kde}$$

$$\varepsilon_t \sim WN(0, \sigma_\varepsilon^2),$$

přičemž pomocí operátoru zpětného chodu lze psát

$$Y_t \sim ARMA(p, q) : \Phi(B)Y_t = \Theta(B)\varepsilon_t,$$

kde

$$\Phi(B) = 1 - \varphi_1 B - \dots - \varphi_p B^p \quad (\varphi_0 \equiv 1)$$

a

$$\Theta(B) = 1 + \theta_1 B + \dots + \theta_q B^q \quad (\theta_0 \equiv 1).$$

Řekneme, že $\{Y_t, t \in \mathbb{Z}\}$ je ARMA(p, q) se **střední hodnotou** μ , jestliže $\{Y_t - \mu\}$ je ARMA(p, q) proces.

Speciální případy ARMA procesů nazýváme:

Autoregresní proces (AR proces): $Y_t \sim AR(p) \sim ARMA(p, 0)$, tj. $\boxed{q = 0}$

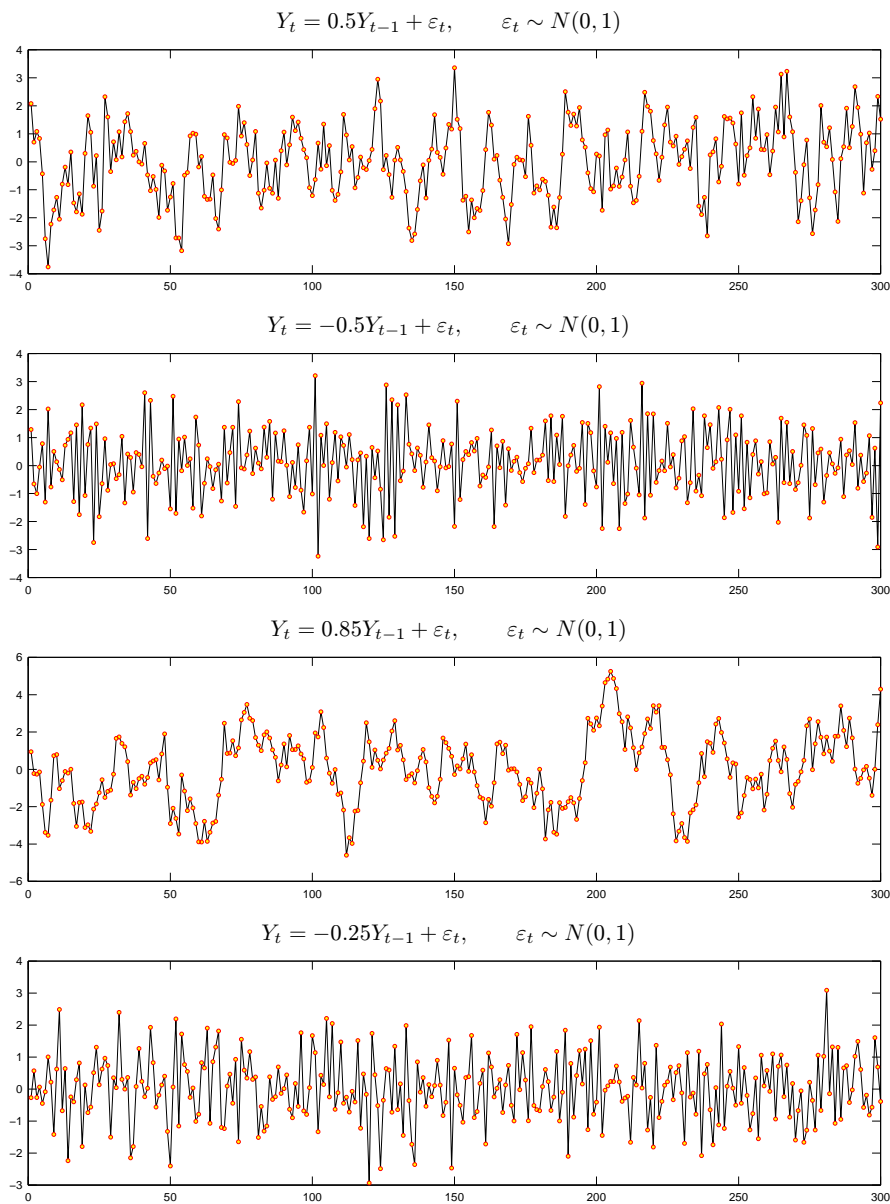
Proces klouzavých součtů (MA proces): $Y_t \sim MA(q) \sim ARMA(0, q)$, tj.

$$\boxed{p = 0}$$

1.2.6 Kauzalita

Dříve než zavedeme pojem kauzality, všimněme si blíže $AR(1)$ procesu.

AUTOREGRESNÍ PROCESY PRVNÍHO ŘÁDU.



Obrázek 1.1: Ukázky autoregresních procesů 1. řádu

Pro autoregresní proces prvního řádu $\boxed{Y_t - \phi_1 Y_{t-1} = \varepsilon_t}$ postupně v k krocích upravujeme

$$\begin{aligned} \boxed{Y_t} &= \varphi_1 Y_{t-1} + \varepsilon_t = \varphi_1 (\varphi_1 Y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t = \varphi_1^2 Y_{t-2} + \varphi_1 \varepsilon_{t-1} + \varepsilon_t \\ &= \varphi_1^2 (\varphi_1 Y_{t-3} + \varepsilon_{t-2}) + \varphi_1 \varepsilon_{t-1} + \varepsilon_t = \varphi_1^3 Y_{t-3} + \varphi_1^2 \varepsilon_{t-2} + \varphi_1 \varepsilon_{t-1} + \varepsilon_t \\ &\vdots \\ &= \varphi_1^k (\varphi_1 Y_{t-k-1} + \varepsilon_{t-k}) + \sum_{j=0}^{k-1} \varphi_1^j \varepsilon_{t-j} = \boxed{\varphi_1^{k+1} Y_{t-k-1} + \sum_{j=0}^k \varphi_1^j \varepsilon_{t-j}} \end{aligned}$$

(1) Uvažujme nejprve případ, kdy $\boxed{|\varphi_1| < 1}$ a $\{Y_t, t \in \mathbb{Z}\}$ je stacionární, tj. $Y_t \in L^2(\Omega, \mathcal{A}, P)$ a $EY_t^2 < \infty$, pak

$$\begin{aligned} \left\| Y_t - \sum_{j=0}^k \varphi_1^j \varepsilon_{t-j} \right\|^2 &= \left\| \varphi_1^{k+1} Y_{t-k-1} \right\|^2 = E \left| \varphi_1^{k+1} Y_{t-k-1} \right|^2 \\ &= \underbrace{\varphi_1^{2k+2}}_{\rightarrow 0} \underbrace{E |Y_{t-k-1}|^2}_{=\sigma_Y^2 < \infty} \xrightarrow{k \rightarrow \infty} 0 \end{aligned}$$

tj. $\sum_{j=0}^{\infty} \varphi_1^j \varepsilon_{t-j}$ konverguje podle kvadratického středu k Y_t a můžeme psát

$$\boxed{Y_t = \sum_{j=0}^{\infty} \varphi_1^j \varepsilon_{t-j}}.$$

Pak dokážeme spočítat

$$\boxed{EY_t} = E \sum_{j=0}^{\infty} \varphi_1^j \varepsilon_{t-j} = \sum_{j=0}^{\infty} \varphi_1^j E \varepsilon_{t-j} = \boxed{0}$$

$$\boxed{DY_t} = D \sum_{j=0}^{\infty} \varphi_1^j \varepsilon_{t-j} \stackrel{\text{nekorel.}}{=} \sum_{j=0}^{\infty} \varphi_1^{2j} D \varepsilon_{t-j} = \sigma_\varepsilon^2 \sum_{j=0}^{\infty} \varphi_1^{2j} = \boxed{\frac{\sigma_\varepsilon^2}{1 - \varphi_1^2}}$$

$$\begin{aligned} \boxed{\gamma(t)} &= C(Y_s, Y_{s+|t|}) = E(Y_s \cdot Y_{s+|t|}) = E \left(\sum_{j=0}^{\infty} \varphi_1^j \varepsilon_{s-j} \right) \left(\sum_{h=0}^{\infty} \varphi_1^h \varepsilon_{s+|t|-h} \right) \\ &= \sum_{j=0}^{\infty} \sum_{h=0}^{\infty} \varphi_j \varphi_h E \underbrace{\varepsilon_{s-j} \varepsilon_{s+|t|-h}}_{\text{nekorel.}} = \begin{vmatrix} s-j & = & s+|t|-h \\ h & = & j+|t| \end{vmatrix} \\ &= \sigma_\varepsilon^2 \sum_{j=0}^{\infty} \varphi_1^j \varphi_1^{j+|t|} = \boxed{\frac{\varphi_1^{|t|}}{1 - \varphi_1^2} \sigma_\varepsilon^2}. \end{aligned}$$

Autokorelační funkce (ACF) je pak tvaru

$$\boxed{\rho(t)} = \frac{\gamma(t)}{\gamma(0)} = \boxed{\varphi_1^{|t|}}$$

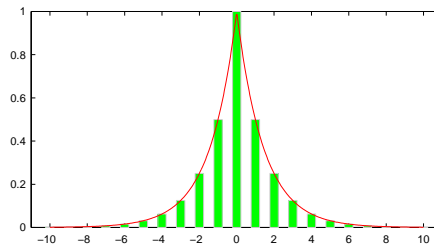
Pomocí generující funkce filtru

$$\boxed{\Psi_{AR(1)}(z)} = \sum_{j=0}^{\infty} \varphi_1^j y^j = \boxed{\frac{1}{1 - \varphi_1 z}}$$

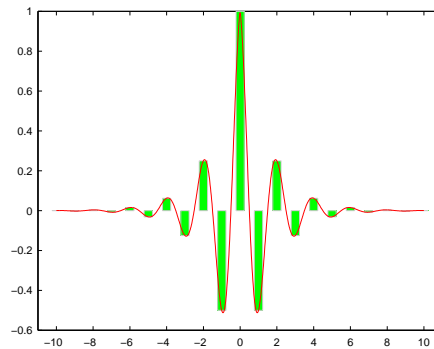
pro $|z| < 1$ a $|\varphi_1| < 1$ dokážeme snadno spočítat i spektrální hustotu

$$\begin{aligned} \boxed{f_{AR(1)}(\omega)} &= \frac{\sigma_\varepsilon^2}{2\pi} |\Psi_{AR(1)}(e^{-i\omega})|^2 \\ &= \frac{\sigma_\varepsilon^2}{2\pi} \frac{1}{|1 - \varphi_1 e^{-i\omega}|^2} \\ &= \frac{\sigma_\varepsilon^2}{2\pi} \frac{1}{|e^{-i\omega}(e^{i\omega} - \varphi_1)|^2} \\ &= \boxed{\frac{\sigma_\varepsilon^2}{2\pi} \frac{1}{|e^{i\omega} - \varphi_1|^2}} \end{aligned}$$

ACF pro $Y_t = 0.5Y_{t-1} + \varepsilon_t$

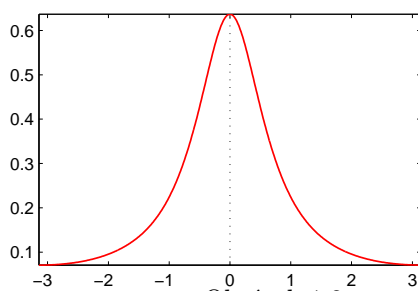


ACF pro $Y_t = -0.5Y_{t-1} + \varepsilon_t$

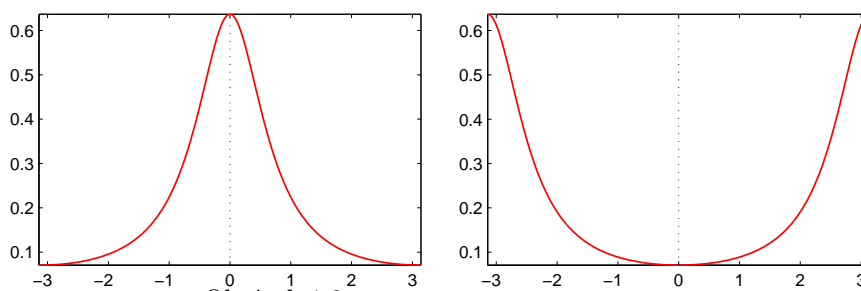


Obrázek 1.2: Ukázky autokorelačních funkcí

$f_{AR(1)}(\omega)$ pro $Y_t = 0.5Y_{t-1} + \varepsilon_t$



$f_{AR(1)}(\omega)$ pro $Y_t = -0.5Y_{t-1} + \varepsilon_t$



Obrázek 1.3: Ukázky spektrálních hustot

(2) Dále řešme případ, kdy $|\varphi_1| > 1$. Použijeme-li vztah

$$Y_{t-1} = \frac{1}{\varphi_1} Y_t - \frac{1}{\varphi_1} \varepsilon_t$$

a postupně v k krocích budeme upravovat

$$\begin{aligned}
 \boxed{Y_t} &= \frac{1}{\varphi_1} Y_{t+1} - \frac{1}{\varphi_1} \varepsilon_{t+1} = \frac{1}{\varphi_1} \left(\frac{1}{\varphi_1} Y_{t+2} - \frac{1}{\varphi_1} \varepsilon_{t+2} \right) - \frac{1}{\varphi_1} \varepsilon_{t+1} \\
 &= \frac{1}{\varphi_1^2} Y_{t+2} - \frac{1}{\varphi_1^2} \varepsilon_{t+2} - \frac{1}{\varphi_1} \varepsilon_{t+1} = \frac{1}{\varphi_1^2} \left(\frac{1}{\varphi_1} Y_{t+3} - \frac{1}{\varphi_1} \varepsilon_{t+3} \right) - \frac{1}{\varphi_1^2} \varepsilon_{t+2} - \frac{1}{\varphi_1} \varepsilon_{t+1} \\
 &= \frac{1}{\varphi_1^3} Y_{t+3} - \frac{1}{\varphi_1^3} \varepsilon_{t+3} - \frac{1}{\varphi_1^2} \varepsilon_{t+2} - \frac{1}{\varphi_1} \varepsilon_{t+1} \\
 &\vdots \\
 &= \boxed{\frac{1}{\varphi_1^{k+1}} Y_{t+k+1} - \sum_{j=0}^k \frac{1}{\varphi_1^{k+1}} \varepsilon_{t-j}},
 \end{aligned}$$

stejně jako v předchozím případě $-\sum_{j=0}^{\infty} \frac{1}{\varphi_1^j} \varepsilon_{t-j}$ konverguje podle kvadratického středu k Y_t . Avšak vidíme, že Y_t zde vyjadřujeme pomocí budoucích hodnot $\{\varepsilon_s, s > t\}$. Tím porušujeme přirozenou podmínku, že Y_t je na budoucnosti nezávislá a říkáme, že není kauzální.

- (3) V případě, že platí $|\varphi_1| = 1$, pak $AR(1)$ není stacionární, jde o tzv. **náhodnou procházku**.

Nyní již můžeme zavést pojem kauzality.

Definice 1.2.12.

$ARMA$ proces $Y_t \sim ARMA(p, q)$ se nazývá **kauzální**, jestliže existuje absolutně konvergentní posloupnost reálných čísel $\Psi = \{\psi_j\}_{j=0}^{\infty}$, (tj. $\sum_{j=0}^{\infty} |\psi_j| < \infty$) tak, že

$$Y_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}, \quad \text{tj. zkráceně} \quad Y_t \sim MA(\infty) : Y_t = \Psi(B) \varepsilon_t.$$

POZNÁMKA 1.2.13.

Protože platí $\sum_{j=0}^{\infty} \psi_j^2 \leq \left(\sum_{j=0}^{\infty} |\psi_j| \right)^2 < \infty$, pak kauzální proces $Y_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$

je **lineárním procesem**. Protože lineární proces je stacionárním procesem, je **kauzální ARMA proces** $Y_t \sim ARMA(p, q)$, kde $\sum_{j=0}^{\infty} |\psi_j| < \infty$, také **stacionárním procesem**.

AUTOREGRESNÍ PROCES p -TÉHO ŘÁDU: $AR(p) : \Phi(B)Y_t = \varepsilon_t$

Mějme polynom $\Phi(z) = \varphi_0 - \varphi_1 z - \dots - \varphi_p z^p$ a necht' $\frac{1}{\lambda_j}$ jsou jeho kořeny, tj. $\Phi\left(\frac{1}{\lambda_j}\right) = 0$. Pak platí

$$\varphi_0 - \varphi_1 z - \dots - \varphi_p z^p = \varphi_p \prod_j \left(z - \frac{1}{\lambda_j}\right) = \varphi_0 \prod_j (1 - \lambda_j z),$$

v našem případě $\varphi_0 = 1$ a $\varphi_p \neq 0$.

Provedme tedy rozklad polynomu $\Phi(z)$ na součin kořenových činitelů

$$\Phi(z) = (1 - \lambda_1 z)^{p_1} \dots (1 - \lambda_k z)^{p_k},$$

kde

$$z_{01} = \frac{1}{\lambda_1}, \dots, z_{0k} = \frac{1}{\lambda_k}$$

jsou **rozdílné** (reálné či komplexní) **kořeny polynomu** $\Phi(z)$, p_1, \dots, p_k je jejich násobnost (přičemž platí $p_1 + \dots + p_k = p$).

Budeme hledat takovou absolutně konvergentní posloupnost čísel

$$\Psi = \{\psi_j\}_{j=0}^{\infty},$$

tak aby $Y_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} = \Psi(B)\varepsilon_t$ byl **kauzální proces**. Takže postupně odvozuje

$$\varepsilon_t = \Phi(B) \underbrace{Y_t}_{=\Psi(B)\varepsilon_t} = \Phi(B)\Psi(B)\varepsilon_t,$$

tj.

$$\Phi(B)\Psi(B) = 1 \quad \text{nebo} \quad \Phi(z)\Psi(z) = 1 \quad \text{čili} \quad \Psi(z) = \frac{1}{\Phi(z)}.$$

Z věty o rozkladu na částečné zlomky dostáváme (pokud pro názornost předpokládáme, že všechny kořeny jsou jednoduché)

$$\frac{1}{\Phi(z)} = \frac{1}{(1 - \lambda_1 z) \dots (1 - \lambda_p z)} = \frac{c_1}{1 - \lambda_1 z} + \dots + \frac{c_p}{1 - \lambda_p z}$$

pro vhodná c_1, \dots, c_p .

Pokud pro $k = 1, \dots, p$ platí

$$|\lambda_k z| < 1,$$

můžeme psát

$$\frac{c_k}{1 - \lambda_k z} = c_k \sum_{j=0}^{\infty} (\lambda_k z)^j$$

a dokázali jsme najít konvergentní řadu

$$\Psi(z) = \sum_{j=0}^{\infty} \psi_j z^j = \sum_{k=1}^p c_k \sum_{j=0}^{\infty} \lambda_k^j z^j = \sum_{j=0}^{\infty} (c_1 \lambda_1^j + \dots + c_p \lambda_p^j) z^j,$$

přičemž

$$\psi_j = c_1 \lambda_1^j + \dots + c_p \lambda_p^j,$$

neboť $\Psi(z) = \frac{1}{\Phi(z)}$ je holomorfní pro $|z| \leq 1$ pouze když

$$|\lambda_1| < 1, \dots, |\lambda_p| < 1 \quad \Leftrightarrow \quad \underbrace{\left| \frac{1}{\lambda_1} \right| > 1, \dots, \left| \frac{1}{\lambda_p} \right| > 1}_{|z_{01}| > 1}$$

tedy **všechny kořeny polynomu $\Phi(z)$ musí ležet vně jednotkové kružnice.**

Tím jsme ukázali, že existuje řešení

$$Y_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$$

tzv. *stochastické diferencní rovnice*

$$\boxed{Y_t - \varphi_1 Y_{t-1} - \dots - \varphi_p Y_{t-p} = \varepsilon_t \quad \varepsilon_t \sim WN(0, \sigma_\varepsilon^2)} \quad (1.1)$$

a tímto řešením je **kauzální autoregresní posloupnost řádu p** . Protože Y_t je lineární proces, je toto řešení **stacionární**.

Podmínka týkající se kořenů polynomu $\Phi(z)$ je podstatná. Lze ukázat, že v případě, kdy alespoň jeden kořen polynomu $\Phi(z)$ leží uvnitř nebo na hranici jednotkové kružnice, neexistuje kauzální posloupnost $\{Y_t, t \in \mathbb{Z}\}$ splňující stochastickou diferencní rovnici (1.1). Snadno se dá ukázat, že toto řešení je jediné.

STŘEDNÍ HODNOTA, ROZPTYL, AUTOKOVARIANCE A AUTOKORELACE $AR(p)$

Pro kauzální $AR(p)$ procesy počítejme nejprve

$$\boxed{EY_t} = E \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} = \sum_{j=0}^{\infty} \psi_j E\varepsilon_{t-j} = \boxed{0}$$

Abychom mohli spočítat rozptyl kauzálního $AR(p)$ procesu, nejprve rovnici

$$Y_t = \varphi_1 Y_{t-1} + \dots + \varphi_p Y_{t-p} + \varepsilon_t$$

vynásobíme výrazem Y_t a spočítáme střední hodnoty obou stran, tj.

$$EY_t^2 = \varphi_1 EY_{t-1}Y_t + \dots + \varphi_p EY_{t-p}Y_t + E\varepsilon_t Y_t. \quad (\text{A}_1)$$

Protože

$$EY_t = 0,$$

pak autokovarianční funkce je rovna

$$\gamma(j) = C(Y_t, Y_{t-j}) = E(Y_t - EY_t)(Y_{t-j} - EY_{t-j}) = EY_t Y_{t-j}$$

a rozptyl

$$\gamma(0) = EY_t^2 = DY_t.$$

Dále spočtěme

$$\boxed{EY_t \varepsilon_t} = E \left(\sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} \right) \varepsilon_t = \sum_{j=0}^{\infty} \psi_j E \varepsilon_{t-j} \varepsilon_t = \sum_{j=0}^{\infty} \psi_j \sigma_\varepsilon^2 \delta(j) = \boxed{\sigma_\varepsilon^2},$$

kde

$$\delta(j) = \begin{cases} 1 & j = 0, \\ 0 & \text{jinak.} \end{cases}$$

Vraťme se k rovnici (A₁), pak po dosazení $EY_t \varepsilon_t = \sigma_\varepsilon^2$ a $\gamma(0) = EY_t^2$ dostaneme

$$\gamma(0) = \varphi_1 \gamma(1) + \dots + \varphi_p \gamma(p) + \sigma^2 \quad (\text{A}_2)$$

Podělme obě strany rovnice (A₂) výrazem $\gamma(0) > 0$ a protože pro autokorelaci platí $\rho(k) = \frac{\gamma(k)}{\gamma(0)}$, dostaneme

$$\underbrace{\rho(0)}_{=1} = \varphi_1 \rho(1) + \dots + \varphi_p \rho(p) + \frac{\sigma_\varepsilon^2}{\gamma(0)}$$

a odtud již plyne, že $\boxed{DY_t} = \gamma(0) = \frac{\sigma_\varepsilon^2}{1 - \varphi_1 \rho(1) - \dots - \varphi_p \rho(p)}$.

Při výpočtu autokovariance (nebo autokorelace ACF) budeme předpokládat, že $k > 0$, neboť $\gamma(0) = DY_t$ již jsme spočítali. Rovnici

$$Y_t = \varphi_1 Y_{t-1} + \dots + \varphi_p Y_{t-p} + \varepsilon_t$$

vynásobíme výrazem Y_{t-k} a spočítáme střední hodnoty obou stran, tj.

$$EY_t Y_{t-k} = \varphi_1 EY_{t-1} Y_{t-k} + \dots + \varphi_p EY_{t-p} Y_{t-k} + E\varepsilon_t Y_{t-k}. \quad (\text{A}_3)$$

Připomeňme, že s využitím vztahu

$$EY_t = 0,$$

je

$$\gamma(k) = C(Y_t, Y_{t-k}) = E(Y_t - EY_t)(Y_{t-k} - EY_{t-k}) = EY_t Y_{t-k}.$$

Spočtěme

$$EY_{t-k}\varepsilon_t = E \left(\sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j-k} \right) \varepsilon_t = \sum_{j=0}^{\infty} \psi_j E\varepsilon_{t-j-k} \varepsilon_t = \sum_{j=0}^{\infty} \psi_j \sigma^2 \underbrace{\delta_{j+k}}_{=0} = 0.$$

Vraťme se k rovnici (A₃), pak po dosazení $EY_t\varepsilon_t = 0$ a $\gamma(k) = EY_tY_{t-k}$ dostaneme

$$\boxed{\gamma(k) = \varphi_1\gamma(k-1) + \dots + \varphi_p\gamma(k-p)} \quad (\text{A}_4)$$

Poděleme obě strany rovnice (A₄) věyrazem $\gamma(0)$ a protoĹe $\rho(k) = \frac{\gamma(k)}{\gamma(0)}$, dostaneme tzv. **Yuleovy-Walkerovy rovnice**.

$$\boxed{\rho(k) = \varphi_1\rho(k-1) + \dots + \varphi_p\rho(k-p)} \quad k \geq 1 \quad (\text{A}_5)$$

EXPLICITNÍ VYJĀDŘENÍ AUTOKORELAČNÍ FUNKCE PROCESU $AR(p)$

Při explicitním vyjĀdřeni autokorelační funkce procesu vyjdeme z Yuleo-WalkerovĹch rovnic

$$\rho(k) = \varphi_1\rho(k-1) + \dots + \varphi_p\rho(k-p) \quad k \geq 1.$$

OznaĹme

$$B\rho(k) = \rho(k-1), \quad \text{přičemĹ} \quad \rho(0) = 1 \quad \text{a} \quad \rho(-j) = \rho(j)$$

a hledejme řešení tzv. **homogenní diferenní rovnice**

$$\rho(k) - \varphi_1\rho(k-1) - \dots - \varphi_p\rho(k-p) = 0 \quad k \geq 1 \quad \text{tj.} \quad \boxed{\Phi(B)\rho(k) = 0}.$$

POZNĀMKA: ŘEŠENÍ HOMOGENNÍ DIFERENČNÍ ROVNICE

Mějme polynom $\Phi(z) = \varphi_0 - \varphi_1z - \dots - \varphi_pz^p$ a nechtě $\frac{1}{\lambda_j}$ jsou jeho kořeny, tj. $\Phi\left(\frac{1}{\lambda_j}\right) = 0$. Pak platí

$$\varphi_0 - \varphi_1z - \dots - \varphi_pz^p = \varphi_p \prod_j \left(z - \frac{1}{\lambda_j} \right) = \varphi_0 \prod_j (1 - \lambda_j z),$$

v našem pŕípadě $\varphi_0 = 1$ a $\varphi_p \neq 0$.

1. Nechtě $\frac{1}{\lambda_j}$ je kořen polynomu $\Phi(z)$, pak $\boxed{\lambda_j^k}$ je řešením $\Phi(B)\rho(k) = 0$.

Důkaz:

$$\begin{aligned} \Phi(B)\lambda_j^k &= (1 - \varphi_1 B - \dots - \varphi_p B^p)\lambda_j^k = \lambda_j^k - \varphi_1\lambda_j^{k-1} - \dots - \varphi_p\lambda_j^{k-p} \\ &= \lambda_j^k \left(1 - \varphi_1 \frac{1}{\lambda_j} - \dots - \varphi_p \frac{1}{\lambda_j^p} \right) = \Phi\left(\frac{1}{\lambda_j}\right) \lambda_j^k = 0. \end{aligned}$$

nebo ekvivalentně: jestliĹe uvaĹujeme faktorizaci $\Phi(B) = \varphi_0 \prod_i (1 - \lambda_i B)$, tak mezi faktory je i člen $(1 - \lambda_j B)$ a platí

$$(1 - \lambda_j B)\lambda_j^k = \lambda_j^k - \lambda_j B(\lambda_j^k) = \lambda_j^k - \lambda_j \cdot \lambda_j^{k-1} = 0.$$

2. Necht' $\frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_p}$ jsou různé jednoduché kořeny, pak

$$c_1 \lambda_1^k + \dots + c_p \lambda_p^k$$

jsou řešením homogenní diferenční rovnice a c_1, \dots, c_p jsou konstanty, které jsou určeny **počátečními podmínkami**.

3. Je-li kořen $\frac{1}{\lambda_j}$ **dvojnásobný kořen**, pak

$$\boxed{\lambda_j^k} \quad \text{a} \quad \boxed{k \lambda_j^k}$$

jsou řešeními $\Phi(B)\rho(k) = 0$.

Důkaz: Díky faktorizaci můžeme psát $\Phi(B) = (1 - \lambda_j B)^2 \prod_{k \neq j} (1 - \lambda_k B)$. Pak

$$(1 - \lambda_j B)^2 \lambda_j^k = (1 - 2\lambda_j B + \lambda_j^2 B^2) \lambda_j^k = \lambda_j^k - 2\lambda_j \lambda_j^{k-1} + \lambda_j^2 \lambda_j^{k-2} = 0$$

$$\begin{aligned} (1 - \lambda_j B)^2 k \lambda_j^k &= (1 - 2\lambda_j B + \lambda_j^2 B^2) k \lambda_j^k \\ &= k \lambda_j^k - 2\lambda_j (k-1) \lambda_j^{k-1} + \lambda_j^2 (k-2) \lambda_j^{k-2} \\ &= k \lambda_j^k - 2k \lambda_j^k + k \lambda_j^k + 2\lambda_j^k - 2\lambda_j^k = 0 \end{aligned}$$

4. Analogicky dostaneme: je-li kořen $\frac{1}{\lambda_j}$ r -tého řádu, pak

$$\boxed{\lambda_j^k, k \lambda_j^k, \dots, k^{r-1} \lambda_j^k}$$

jsou řešeními $\Phi(B)\rho(k) = 0$.

Shrneme-li tedy předchozí, za předpokladu, že $\frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_m}$ jsou různé kořeny s násobnostmi p_1, \dots, p_m , přičemž $p = p_1 + \dots + p_m$, pak řešení homogenní diferenční rovnice $\Phi(B)\rho(k) = 0$ je tvaru

$$\rho(k) = \sum_{j=1}^m \left(\sum_{s=0}^{p_j-1} c_{js} k^s \right) \lambda_j^k,$$

kde c_{js} jsou konstanty, které jsou určeny počátečními podmínkami. Dále položíme $\lambda_j = r_j e^{i\theta_j}$. Pak máme

$$\rho(k) = \sum_{j=1}^m \left(\sum_{s=0}^{p_j-1} c_{js} k^s \right) r_j^k e^{ik\theta_j},$$

Vzhledem k tomu, že platí

$$|\lambda_j| = r_j < 1,$$

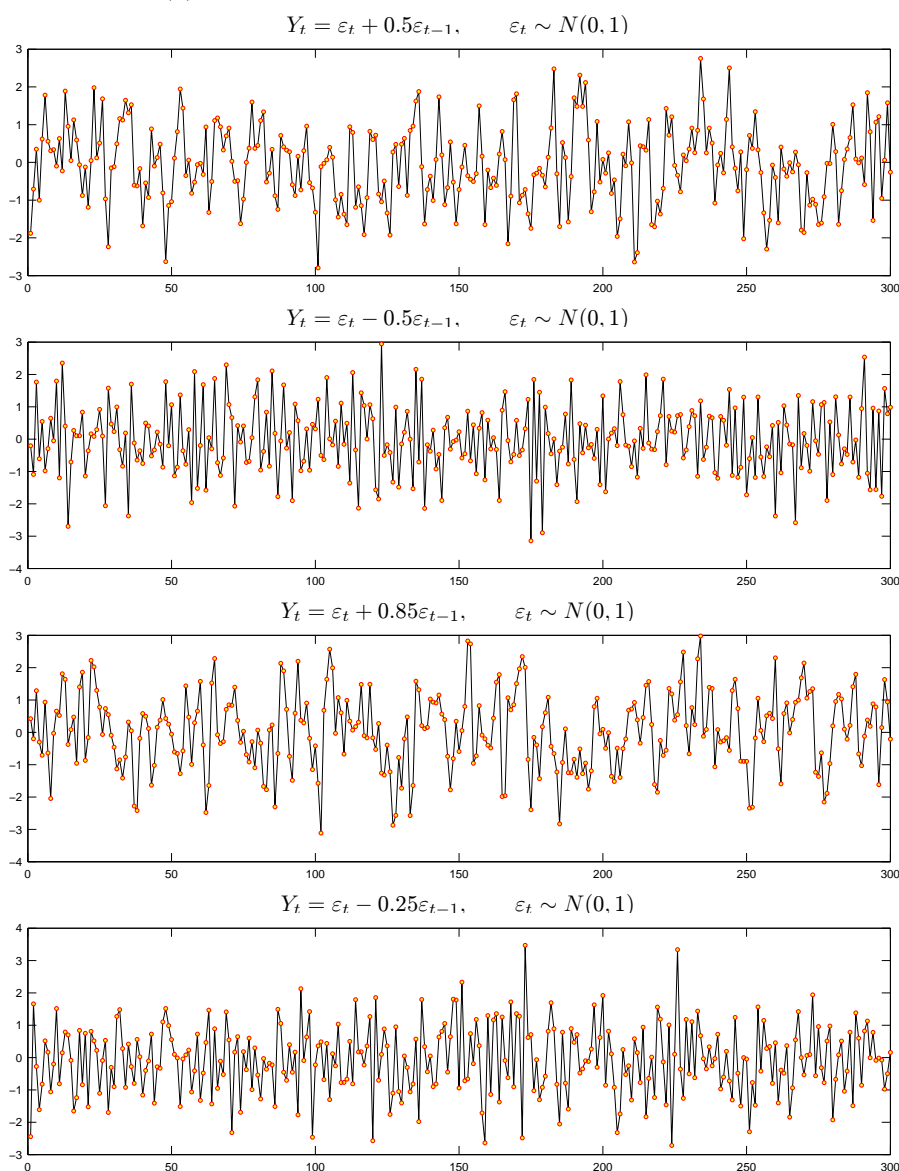
dostáváme odtud, že a $\rho(k)$ klesá pro $k \rightarrow \infty$ exponenciálně k nule, tj.

$$\boxed{\rho(k) \xrightarrow[k \rightarrow \infty]{} 0},$$

což je velmi důležitá identifikační vlastnost autoregresních $AR(p)$ procesů.

1.2.7 Invertibilita

Víme, že *kauzální* autoregresní proces konečného řádu p lze vyjádřit pomocí MA procesu nekonečného řádu, tj. $AR(p) \equiv MA(\infty)$. Zajímá nás, za jakých podmínek můžeme MA proces konečného řádu vyjádřit pomocí autoregresního procesu nekonečného řádu, tj. $MA(q) \equiv AR(\infty)$. Nejprve si všimneme jednoduchého případu, a to $MA(1)$ procesu.



Obrázek 1.4: Ukázky MA procesů prvního řádu

MA PROCES PRVNÍHO ŘÁDU:

$$Y_t \sim MA(1) : Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1}, \quad \varepsilon_t \sim WN(0, \sigma_\varepsilon^2).$$

Nejprve označme $\theta_1 = -\theta$ a předpokládejme, že

$$|\theta_1| = |\theta| < 1.$$

Využijeme-li vztahu

$$Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} \Rightarrow \varepsilon_t = Y_t + \theta \varepsilon_{t-1},$$

můžeme postupně upravovat

$$\begin{aligned} \varepsilon_t Y_t + \theta \varepsilon_{t-1} &= Y_t + \theta (Y_{t-1} + \theta \varepsilon_{t-2}) \\ &= Y_t + \theta Y_{t-1} + \theta^2 \varepsilon_{t-2} \\ &\vdots \\ &= \sum_{j=0}^k \theta^j Y_{t-j} + \theta^{k+1} \varepsilon_{t-k-1} \end{aligned}$$

a

$$\begin{aligned} \left\| \varepsilon_t - \sum_{j=0}^k \theta^j Y_{t-j} \right\|^2 &= E \left| \varepsilon_t - \sum_{j=0}^k \theta^j Y_{t-j} \right|^2 \\ &= E |\theta^{k+1} \varepsilon_{t-k-1}|^2 \\ &= \theta^{2(k+1)} \sigma_\varepsilon^2 \xrightarrow[k \rightarrow \infty]{} 0, \end{aligned}$$

tedy

$$\varepsilon_t = \sum_{j=0}^{\infty} \theta^j Y_{t-j} \quad \text{pro} \quad |\theta| < 1.$$

Dále využijme vztah

$$\varepsilon_{t-1} = \frac{1}{\theta_1} Y_t - \frac{1}{\theta_1} \varepsilon_t$$

a označme

$$\theta_1 = -\theta.$$

Za předpokládejme, že platí

$$|\theta_1| = |\theta| > 1,$$

můžeme opět postupně upravovat

$$\begin{aligned}\varepsilon_t &= \frac{1}{\theta} Y_{t+1} + \frac{1}{\theta} \varepsilon_{t+1} \\ &= \frac{1}{\theta} Y_{t+1} + \frac{1}{\theta} \left(\frac{1}{\theta} Y_{t+2} + \frac{1}{\theta} \varepsilon_{t+2} \right) \\ &\vdots \\ &= \sum_{j=1}^{k+1} \frac{1}{\theta^j} Y_{t+j} + \frac{1}{\theta^{k+1}} \varepsilon_{t+k+1}.\end{aligned}$$

I když posloupnost

$$\sum_{j=1}^N \frac{1}{\theta^j} Y_{t+j}$$

konverguje pro $N \rightarrow \infty$ také k ε_t , tento rozvoj nemá praktický smysl, neboť ε_t je vyjádřena pomocí budoucích hodnot $\{Y_s, s > t\}$.

Dále si všimněme dalšího důležitého faktu, že pokud platí

$$\boxed{|\theta_1| > 1},$$

a uvažujeme-li dva procesy

$$\begin{aligned}(1) \quad & Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} \quad \varepsilon_t \sim WN(0, \sigma^2), \\ (2) \quad & X_t = \eta_t + \frac{1}{\theta_1} \eta_{t-1} \quad \eta_t \sim WN(0, \theta_1^2 \sigma^2),\end{aligned}$$

pak oba dva procesy mají stejné první a druhé momenty, neboť

$$EY_t = E(\varepsilon_t + \theta_1 \varepsilon_{t-1}) = E\varepsilon_t + \theta_1 E\varepsilon_{t-1} = \boxed{0},$$

$$EX_t = E\left(\eta_t + \frac{1}{\theta_1} \eta_{t-1}\right) = E\eta_t + \frac{1}{\theta_1} E\eta_{t-1} = \boxed{0},$$

$$\begin{aligned}\gamma_Y(k) &= EY_t Y_{t+k} = E(\varepsilon_t + \theta_1 \varepsilon_{t-1})(\varepsilon_{t+k} + \theta_1 \varepsilon_{t+k-1}) \\ &= E\varepsilon_t \varepsilon_{t+k} + \theta_1 E\varepsilon_t \varepsilon_{t+k-1} + \theta_1 E\varepsilon_{t-1} \varepsilon_{t+k} + \theta_1^2 E\varepsilon_{t-1} \varepsilon_{t+k-1} \\ &= \left| \begin{array}{cccc} \text{pokud:} & t = t+k & t = t+k-1 & t-1 = t+k & t-1 = t+k-1 \\ \text{pak:} & k=0 & k=1 & k=-1 & k=0 \end{array} \right| \\ &= \begin{cases} \sigma^2 + \theta_1^2 \sigma^2 = \boxed{\sigma^2(1 + \theta_1^2)} & k=0 \\ \boxed{\theta_1 \sigma^2} & k = \pm 1 \\ \boxed{0} & \text{jinak} \end{cases}\end{aligned}$$

$$\begin{aligned}
\gamma_X(k) &= EX_t X_{t+k} = E\left(\eta_t + \frac{1}{\theta_1} \eta_{t-1}\right) \left(\eta_{t+k} + \frac{1}{\theta_1} \eta_{t+k-1}\right) \\
&= E\eta_t \varepsilon_{t+k} + \frac{1}{\theta_1} E\eta_t \varepsilon_{t+k-1} + \frac{1}{\theta_1} E\eta_{t-1} \eta_{t+k} + \frac{1}{\theta_1^2} E\eta_{t-1} \eta_{t+k-1} \\
&= \begin{cases} \theta_1^2 \sigma^2 + \frac{1}{\theta_1^2} \theta_1^2 \sigma^2 = \boxed{\sigma^2(1 + \theta_1^2)} & k = 0 \\ \frac{1}{\theta_1} \theta_1^2 \sigma^2 = \boxed{\theta_1 \sigma^2} & k = \pm 1 \\ \boxed{0} & \text{jinak.} \end{cases}
\end{aligned}$$

I když obě invertibilní i neinvertibilní MA reprezentace generují procesy se stejnými momenty prvního a druhého řádu, z praktických důvodů dáváme přednost procesu invertibilnímu, neboť nepozorovatelné veličiny ε_t můžeme odhadnout pomocí přítomných a minulých hodnot pozorovatelných veličin $\{X_s, s < t\}$, kdežto u neinvertibilních MA reprezentací nepozorovatelné veličiny ε_t neodhadneme, neboť nemáme ještě k dispozici budoucí hodnoty $\{Y_s, s > t\}$. Nyní již můžeme podat definici invertibility.

Definice 1.2.14.

ARMA proces $Y_t \sim ARMA(p, q)$ se nazývá **invertibilní**, jestliže existuje absolutně konvergentní posloupnost reálných čísel $\pi = \{\pi_j\}_{j=0}^{\infty}$ (tj. $\sum_{j=0}^{\infty} |\pi_j| < \infty$) tak, že

$$\varepsilon_t = \sum_{j=0}^{\infty} \pi_j Y_{t-j}, \quad \text{tj. zkráceně} \quad Y_t \sim AR(\infty) : \varepsilon_t = \pi(B)Y_t.$$

Nyní vyšetřeme, za jakých podmínek je invertibilní MA proces řádu q .

MA PROCES ŘÁDU q :

$$Y_t \sim MA(q) : Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} \quad \varepsilon_t \sim WN(0, \sigma_\varepsilon^2).$$

Naprostο analogickým postupem jako v případě kauzálního $AR(p)$ procesu, lze ukázat, že všechny kořeny $\Theta(z)$ musí ležet vně jednotkového kruhu. Provedme tedy nejprve rozklad polynomu $\Theta(z) = 1 + \theta_1 z + \dots + \theta_q z^q$ na součin kořenových činitelů

$$\Theta(z) = (1 - \lambda_1 z)^{r_1} \dots (1 - \lambda_k z)^{r_k},$$

kde $z_{01} = \frac{1}{\lambda_1}, \dots, z_{0k} = \frac{1}{\lambda_k}$ jsou rozdílné (reálné či komplexní) kořeny polynomu $\Theta(z)$, r_1, \dots, r_k je jejich násobnost (přičemž platí $r_1 + \dots + r_k = q$). Nyní budeme hledat taková absolutně konvergentní $\pi = \{\pi_j\}_{j=0}^{\infty}$ (tj. $\sum_{j=0}^{\infty} |\pi_j| < \infty$), aby

$$\varepsilon_t = \sum_{j=0}^{\infty} \psi_j Y_{t-j}$$

byl **invertibilní proces**. Pokud použijeme operátor zpětného chodu, můžeme psát:

$$\Theta(B)Y_t = \varepsilon_t,$$

přitom hledáme $\pi(B)$ takové, aby platilo

$$\pi(B)\Theta(B) = 1 \quad \text{nebo} \quad \pi(z)\Theta(z) = 1 \quad \text{čili} \quad \pi(z) = \frac{1}{\Theta(z)}.$$

Z věty o rozkladu na částečné zlomky dostáváme (pokud pro názornost předpokládáme, že všechny kořeny jsou jednoduché)

$$\frac{1}{\Theta(z)} = \frac{1}{(1 - \lambda_1 z) \dots (1 - \lambda_p z)} = \frac{c_1}{1 - \lambda_1 z} + \dots + \frac{c_p}{1 - \lambda_p z}$$

pro vhodná c_1, \dots, c_p . Pokud pro $k = 1, \dots, p$ platí $|\lambda_k z| < 1$, můžeme psát

$$\frac{c_k}{1 - \lambda_k z} = c_k \sum_{j=0}^{\infty} (\lambda_k z)^j$$

a dokázali jsme najít konvergentní řadu

$$\pi(z) = \sum_{j=0}^{\infty} \pi_j z^j = \sum_{k=1}^p c_k \sum_{j=0}^{\infty} \lambda_k^j z^j = \sum_{j=0}^{\infty} (c_1 \lambda_1^j + \dots + c_p \lambda_p^j) z^j,$$

přičemž

$$\pi_j = c_1 \lambda_1^j + \dots + c_p \lambda_p^j,$$

neboť $\pi(z) = \frac{1}{\Theta(z)}$ je holomorfní pro $|z| \leq 1$ právě když

$$|\lambda_1| < 1, \dots, |\lambda_p| < 1 \quad \Leftrightarrow \quad \underbrace{\left| \frac{1}{\lambda_1} \right| > 1, \dots, \left| \frac{1}{\lambda_p} \right| > 1}_{|z_{01}| > 1, \dots, |z_{0p}| > 1}$$

tedy **všechny kořeny polynomu $\Theta(z)$ musí ležet vně jednotkového kruhu**.

Na závěr tohoto odstavce ještě spočítejme střední hodnotu, rozptyl, autokovarianční funkci a také spektrální hustotu $MA(q)$ procesu. Protože $MA(q)$ proces je lineárním procesem, je vždy slabě stacionární, proto můžeme počítat

$$\begin{aligned} \overline{EY_t} &= E(\varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}) = 0 \\ \overline{DY_t} &= D(\varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}) = \sigma_\varepsilon^2 (1 + \theta_1^2 + \dots + \theta_q^2) \\ \gamma(t) &= C(Y_s, Y_{s+t}) = EY_s Y_{s+t} = \sum_{j=0}^q \sum_{h=0}^q \theta_j \theta_s E\varepsilon_{s-j} \varepsilon_{s+t-h} \\ &= \left| \begin{array}{cc} s-j & = & s+t-h \\ h & = & j+t \end{array} \right| \\ &= \sigma_\varepsilon^2 \sum_{j=0}^q \theta_j \theta_{j+t} \end{aligned}$$

Protože

$$\theta_0 = 1 \quad \text{a} \quad \theta_j = 0 \quad \text{pro} \quad j > q,$$

dostáváme

$$\boxed{\gamma(t)} = \begin{cases} \sigma_\varepsilon^2(1 + \theta_1^2 + \dots + \theta_{q-2}^2 + \theta_{q-1}^2 + \theta_q^2) & \text{pro } t=0 \\ \sigma_\varepsilon^2(\theta_1 + \theta_1\theta_2 + \dots + \theta_{q-2}\theta_{q-1} + \theta_{q-1}\theta_q) & t=1 \\ \sigma_\varepsilon^2(\theta_2 + \theta_1\theta_3 + \dots + \theta_{q-2}\theta_q) & t=2 \\ \vdots & \vdots \\ \sigma_\varepsilon^2(\theta_{q-1} + \theta_1\theta_q) & t=q-1 \\ \sigma_\varepsilon^2\theta_q & t=q \\ 0 & \text{jinak} \end{cases}$$

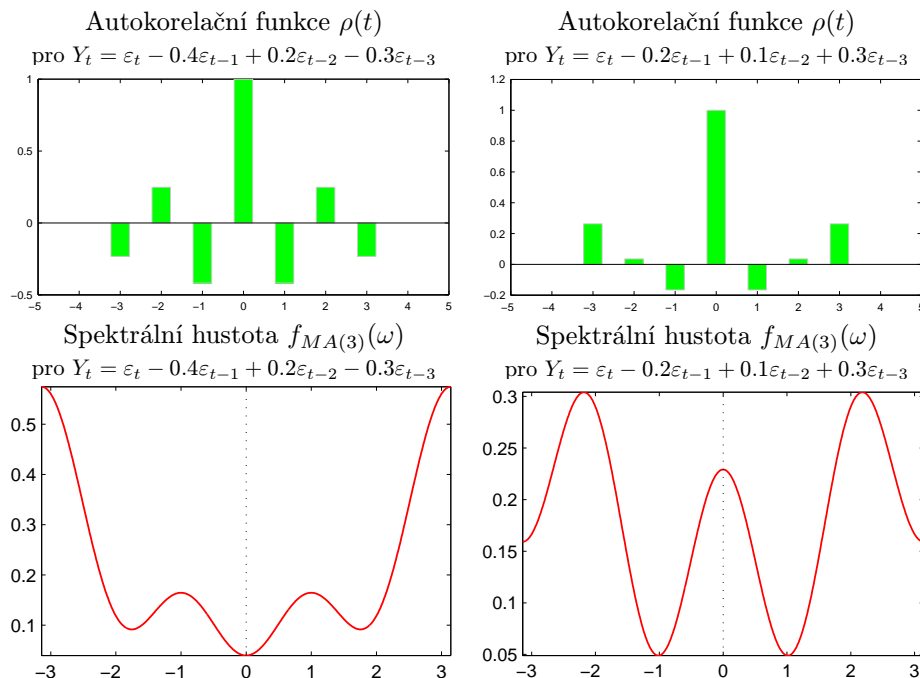
Autokorelační funkce je pak rovna

$$\rho(t) = \begin{cases} 1 & t=0 \\ \frac{1}{1+\theta_1^2+\dots+\theta_q^2} \sum_{j=0}^{q-t} \theta_j\theta_{j+t} & 1 \leq t \leq q, \theta_0 \equiv 1 \\ 0 & \text{jinak} \end{cases}$$

tedy, pro $t > q$ je autokorelační funkce nulová, což je velmi důležitá identifikační vlastnost $MA(q)$ procesů.

Díky tomu, že $MA(q)$ proces je lineárním procesem, spektrální hustota je rovna

$$\boxed{f_Y(\omega) = \frac{\sigma_\varepsilon^2}{2\pi} |\Theta(e^{-i\omega})|^2}.$$



Obrázek 1.5: Ukázky autokorelačních funkcí a spektrálních hustot pro $MA(3)$ procesy.

1.2.8 Vícenásobná reprezentace $MA(q)$ procesů

Mějme MA proces řádu q :

$$Y_t \sim MA(q) : Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} \quad \varepsilon_t \sim WN(0, \sigma_\varepsilon^2).$$

Provedme tedy rozklad polynomu $\Theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q$ na součin kořenových činitelů

$$\Phi(z) = \prod_j (1 - \lambda_j z),$$

Pak (protože $MA(q)$ proces je lineárním procesem) **autokovarianční generující funkce** je rovna

$$G_Y(z) = \Theta(z)\Theta(z^{-1})\sigma_\varepsilon^2.$$

Dále platí

$$\begin{aligned} (1 - \lambda_j z)(1 - \lambda_j z^{-1}) &= 1 - \lambda_j z - \lambda_j z^{-1} + \lambda_j^2 = \lambda_j^2 (\lambda_j^{-2} - \lambda_j^{-1} z - \lambda_j^{-1} z^{-1} + 1) \\ &= \lambda_j^2 \left(1 - \frac{1}{\lambda_j} z\right) \left(1 - \frac{1}{\lambda_j} z^{-1}\right) \end{aligned}$$

Tudíž

$$\begin{aligned} G_Y(z) &= \sigma_\varepsilon^2 \Theta(z)\Theta(z^{-1}) = \sigma_\varepsilon^2 \prod_j (1 - \lambda_j z) \prod_j (1 - \lambda_j z^{-1}) \\ &= \underbrace{\sigma_\varepsilon^2 \prod_j \lambda_j^2}_{\sigma_*^2} \underbrace{\prod_j \left(1 - \frac{1}{\lambda_j} z\right)}_{\Theta_*(z)} \underbrace{\prod_j \left(1 - \frac{1}{\lambda_j} z^{-1}\right)}_{\Theta_*(z^{-1})} = \sigma_*^2 \Theta_*(z)\Theta_*(z^{-1}) \end{aligned}$$

Takže proces

$$Y_t^* \sim MA(q) : Y_t^* = \varepsilon_t^* + \theta_1^* \varepsilon_{t-1}^* + \cdots + \theta_q^* \varepsilon_{t-q}^* \quad \varepsilon_t^* \sim WN(0, \sigma_*^2)$$

má stejnou autokovarianční generující funkci

$$G_Y(z) = \sigma_*^2 \Theta_*(z)\Theta_*(z^{-1})$$

a jsou proto z hlediska prvních dvou momentů nerozlišitelné.

Obecně můžeme dostat 2^q různých procesů s funkcí

$$\Phi_{*s}(z) = \prod_{j=1}^q (1 - \lambda_j^{\pm 1} z) \quad s = 1, \dots, 2^q$$

Mezi všemi těmito procesy **pouze jediný je invertibilní**, a to ten, pro kterého platí

$$\lambda_j^{invert} = \begin{cases} \lambda_j & |\lambda_j| < 1, \\ \lambda_j^{-1} & |\lambda_j| \geq 1. \end{cases}$$

Takže podmínka invertibility zajišťuje identifikovatelnost $MA(q)$ procesu z hlediska prvních dvou momentů. Dříve než uvedeme nutnou a postačující podmínku pro kauzalitu a invertibilitu $ARMA(p, q)$ procesů, vyšetřeme problematiku společných kořenů $\Phi(z)$ a $\Theta(z)$.

1.2.9 Společné kořeny polynomů $\Phi(z)$ a $\Theta(z)$

Mějme

$$Y_t \sim ARMA(p, q) : Y_t - \varphi_1 Y_{t-1} - \dots - \varphi_p Y_{t-p} = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \dots + \theta_q \varepsilon_{t-q},$$

kde $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$ a předpokládáme že $\Phi(z)$ a $\Theta(z)$ **mají společný kořen** $\boxed{\frac{1}{\lambda}}$.
Pak můžeme psát

$$\begin{aligned}\Phi(z) &= (1 - \lambda z)(1 - \varphi_1^* z - \dots - \varphi_{p-1}^* z^{p-1}) = (1 - \lambda z)\Phi^*(z) \\ \Theta(z) &= (1 - \lambda z)(1 + \theta_1^* z + \dots + \theta_{q-1}^* z^{q-1}) = (1 - \lambda z)\Theta^*(z)\end{aligned}$$

tj.

$$(1 - \lambda B)\Phi^*(B)Y_t = (1 - \lambda B)\Theta^*(B)\varepsilon_t.$$

Pokud obě strany rovnice vydělíme výrazem $(1 - \lambda B)$, dostaneme

$$Y_t \sim ARMA(p-1, q-1) : \Phi^*(B)Y_t = \Theta^*(B)\varepsilon_t.$$

Takže podmínka, že

$$\boxed{\Phi(z) \text{ a } \Theta(z) \text{ nemají společné kořeny}}$$

zajišťuje, že řady ARMA procesů nelze již snižovat.

1.2.10 Nutná a postačující podmínka kauzality a invertibility ARMA procesu.

V předchozích odstavcích jsme ukázali, že platí

- (1) $Y_t \sim AR(p) : \Phi(B)Y_t = \varepsilon_t \wedge \Phi(z) \neq 0$ pro $\forall z \in \mathbb{C} \wedge |z| \leq 1 \Leftrightarrow AR(p)$ je **kauzální**;
- (2) $Y_t \sim MA(q) : Y_t = \Theta(B)\varepsilon_t \wedge \Theta(z) \neq 0$ pro $\forall z \in \mathbb{C} \wedge |z| \leq 1 \Leftrightarrow MA(q)$ je **invertibilní**.

Naprosto analogickým způsobem lze dokázat obecnější tvrzení:

Věta 1.2.15.

Nechť $\Phi(B)$ a $\Theta(B)$ nemají společné kořeny. Pak

- (i) $Y_t \sim ARMA(p, q) : \Phi(B)Y_t = \Theta(B)\varepsilon_t$ je **kauzální** $\Leftrightarrow \Phi(z) \neq 0$ pro $\forall z \in \mathbb{C} \wedge |z| \leq 1$.
- (ii) $Y_t \sim ARMA(p, q) : \Phi(B)Y_t = \Theta(B)\varepsilon_t$ je **invertibilní** $\Leftrightarrow \Theta(z) \neq 0$ pro $\forall z \in \mathbb{C} \wedge |z| \leq 1$.

Znamená to tedy, že $Y_t \sim ARMA(p, q)$ je kauzálním a invertibilním ARMA procesem, jestliže všechny kořeny polynomů $\Phi(z)$ a $\Theta(z)$ leží vně jednotkového

kruhu a koeficienty ψ_j a π_j jsou určeny ze vztahů

$$\Psi(z) = \sum_{j=0}^{\infty} \psi_j z^j = \frac{\Phi(z)}{\Theta(z)} \quad \text{pro } |z| \leq 1$$

$$\pi(z) = \sum_{j=0}^{\infty} \pi_j B^j = \frac{\Theta(z)}{\Phi(z)} \quad \text{pro } |z| \leq 1.$$

V dalším budeme uvažovat pouze takové

$$Y_t \sim ARMA(p, q) : \Phi(B)Y_t = \Theta(B)\varepsilon_t$$

procesy, které splňují následující podmínky

(P1) $\Phi(B)$ a $\Theta(B)$ nemají společné kořeny.

(P2) $Y_t \sim ARMA(p, q)$ je kauzální.

(P3) $Y_t \sim ARMA(p, q)$ je invertibilní.

1.2.11 Střední hodnota, rozptyl, autokovarianční a autokorelační funkce procesů $ARMA(p, q)$

STŘEDNÍ HODNOTA

Vzhledem ke kauzalitě $ARMA(p, q)$ procesu můžeme počítat

$$\boxed{EY_t} = E \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} = \sum_{j=0}^{\infty} \psi_j E\varepsilon_{t-j} = \boxed{0}$$

ROZPTYL

Při odvození rozptylu nejprve rovnici

$$Y_t = \varphi_1 Y_{t-1} + \dots + \varphi_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$$

vynásobme výrazem Y_t a spočtěme střední hodnoty obou stran, tj.

$$EY_t^2 = \varphi_1 EY_{t-1}Y_t + \dots + \varphi_p EY_{t-p}Y_t + E\varepsilon_t Y_t + \theta_1 E\varepsilon_{t-1}Y_t + \dots + \theta_q E\varepsilon_{t-q}Y_t. \quad (\text{A}_6)$$

Spočtěme pro $i = 0, 1, \dots, q$

$$E\varepsilon_{t-i}Y_t = E\varepsilon_{t-i} \left(\sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} \right) = \sum_{j=0}^{\infty} \psi_j E\varepsilon_{t-i}\varepsilon_{t-j} = \psi_i \sigma_\varepsilon^2 \quad (\text{přičemž } \psi_0 = 1).$$

Po dosazení do rovnice (A₆) dostaneme

$$\gamma(0) - \varphi_1 \gamma(1) - \dots - \varphi_p \gamma(p) = \sigma_\varepsilon^2 (1 + \theta_1 \psi_1 + \dots + \theta_q \psi_q) \quad (\text{A}_7).$$

Podělme obě strany rovnice (A₇) výrazem $\gamma(0)$. Vzhledem k tomu, že $\rho(k) = \frac{\gamma(k)}{\gamma(0)}$, dostaneme

$$1 - \varphi_1\rho(1) - \dots - \varphi_p\rho(p) = \frac{\sigma_\varepsilon^2(1 + \theta_1\psi_1 + \dots + \theta_q\psi_q)}{\gamma(0)}$$

takže

$$\boxed{DY_t = \gamma(0) = \frac{\sigma^2(1 + \theta_1\psi_1 + \dots + \theta_q\psi_q)}{1 - \varphi_1\rho(1) - \dots - \varphi_p\rho(p)}}.$$

AUTOKOVARIANČNÍ A AUTOKORELAČNÍ FUNKCE (ACF)

Při výpočtu autokovariance rovnici

$$Y_t - \varphi_1Y_{t-1} - \dots - \varphi_pY_{t-p} = \varepsilon_t + \theta_1\varepsilon_{t-1} + \dots + \theta_q\varepsilon_{t-q}$$

vynásobíme výrazem Y_{t-k} a spočítáme střední hodnoty obou stran, takže dostaneme

$$\gamma(k) - \varphi_1\gamma(k-1) - \dots - \varphi_p\gamma(k-p) = EY_{t-k}\varepsilon_t + \theta_1EY_{t-k}\varepsilon_{t-1} + \dots + \theta_qEY_{t-k}\varepsilon_{t-q} \quad (\text{A}_8).$$

Nejprve je třeba si uvědomit, že pro $s \geq 0$ platí

$$E\varepsilon_t Y_{t-s} = E\left(\varepsilon_t \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-s-j}\right) = \sum_{j=0}^{\infty} \psi_j E\varepsilon_t \varepsilon_{t-s-j} = 0.$$

Spočtíme pro $i = 0, 1, \dots, q$

$$\begin{aligned} E\varepsilon_{t-i} Y_{t-k} &= E\varepsilon_{t-i} \left(\sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j-k} \right) = \sum_{j=0}^{\infty} \psi_j E\varepsilon_{t-i} \varepsilon_{t-j-k} \\ &= \begin{vmatrix} t-i & = & t-j-k \\ j & = & i-k \end{vmatrix} \\ &= \begin{cases} \sigma_\varepsilon^2 \psi_{i-k} & k \leq i \text{ přičemž } \psi_0 = 1 & \Rightarrow & k \leq q \\ 0 & k > i \text{ neboť } \psi_j = 0 \text{ pro } j < 0 \end{cases} \end{aligned}$$

Uvážíme-li, že $\psi_j = 0$ pro $j < 0$, potom pro $\boxed{0 \leq k \leq \max(p, q+1)}$ platí

$$\boxed{\gamma(k) - \varphi_1\gamma(k-1) - \dots - \varphi_p\gamma(k-p) = \sigma_\varepsilon^2(\theta_k + \theta_{k+1}\psi_1 + \dots + \theta_q\psi_{q-k})} \quad (\text{A}_9)$$

a pro $\boxed{k > \max(p, q+1)}$ platí

$$\boxed{\gamma(k) - \varphi_1\gamma(k-1) - \dots - \varphi_p\gamma(k-p) = 0}.$$

Podělme obě strany rovnice (A₉) výrazem $\gamma(0)$. Dostaneme

$$\rho(k) - \varphi_1\rho(k-1) - \dots - \varphi_p\rho(k-p) = \frac{\sigma_\varepsilon^2(\theta_k + \theta_{k+1}\psi_1 + \dots + \theta_q\psi_{q-k})}{\gamma(0)}$$

resp.

$$\rho(k) - \varphi_1\rho(k-1) - \dots - \varphi_p\rho(k-p) = 0.$$

Nechť např.

$$q+1 > p.$$

Pak máme více rovnic pro určení počátečních p podmínek. V tomto případě prvních $q-p+1$ autokovariančních koeficientů jsou určeny z prvních $q-p+1$ podmínek.

Obecné řešení homogenní diferenční rovnice

$$\rho(k) - \varphi_1\rho(k-1) - \dots - \varphi_p\rho(k-p) = 0 \quad \text{tj.} \quad \Phi(B)\rho(k) = 0$$

je tvaru

$$\rho(k) = \sum_{j=1}^m \left(\sum_{s=0}^{p_j-1} c_{js} k^s \right) \lambda_j^k,$$

kde

$$\frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_m}$$

jsou různé kořeny s násobnostmi

$$p_1, \dots, p_m,$$

přičemž

$$p = p_1 + \dots + p_m$$

a c_{js} je právě p konstant, které jsou určeny počátečními podmínkami.

1.2.12 Spektrální hustota $ARMA(p, q)$ procesů

Věta 1.2.16 (Spektrální hustota $ARMA(p, q)$ procesů).

Nechť $\Phi(B)Y_t = \Theta(B)\varepsilon_t$ je kauzální a invertibilní $ARMA(p, q)$ proces, přičemž $\Phi(z)$ a $\Theta(z)$ nemají společné kořeny. Pak spektrální hustota $ARMA(p, q)$ procesu je rovna

$$f_Y(\omega) = \frac{\sigma_\varepsilon^2}{2\pi} \frac{|\Theta(e^{-i\omega})|^2}{|\Phi(e^{-i\omega})|^2} \quad \text{pro} \quad \omega \in \langle -\pi, \pi \rangle.$$

DŮKAZ. Protože všechny kořeny $\Phi(z)$ a $\Theta(z)$ leží vně jednotkového kruhu, $ARMA(p, q)$ proces je kauzální a invertibilní.

Kauzalita značí, že existuje absolutně konvergentní posloupnost reálných čísel $\psi = \{\psi_j\}_{j=0}^{\infty}$ (tj. $\sum_{j=0}^{\infty} |\psi_j| < \infty$) taková, že platí

$$Y_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} \quad \text{kde} \quad \varepsilon_t \sim WN(0, \sigma_\varepsilon^2).$$

Víme, že spektrální hustota bílého šumu je rovna

$$f_\varepsilon(\omega) = \frac{\sigma_\varepsilon^2}{2\pi} \quad \text{kde} \quad \omega \in \langle -\pi, \pi \rangle.$$

Protože Y_t je lineárním procesem, víme, že má spektrální hustotu

$$f_Y(\omega) = |\Psi(e^{-i\omega})|^2 f_\varepsilon(\omega) = |\Psi(e^{-i\omega})|^2 \frac{\sigma_\varepsilon^2}{2\pi} \quad \text{kde} \quad \omega \in \langle -\pi, \pi \rangle.$$

Také $\Theta(B)\varepsilon_t$ jakožto lineární proces má spektrální hustotu tvaru

$$|\Theta(e^{-i\omega})|^2 \frac{\sigma_\varepsilon^2}{2\pi} \quad \text{pro} \quad \omega \in \langle -\pi, \pi \rangle.$$

Rovněž $\Phi(B)Y_t$ jakožto lineární filtr má také spektrální hustotu, a ta je rovna

$$|\Theta(e^{-i\omega})|^2 f_Y(\omega) \quad \text{pro} \quad \omega \in \langle -\pi, \pi \rangle.$$

Protože platí

$$\Phi(B)Y_t = \Theta(B)\varepsilon_t,$$

musí také platit

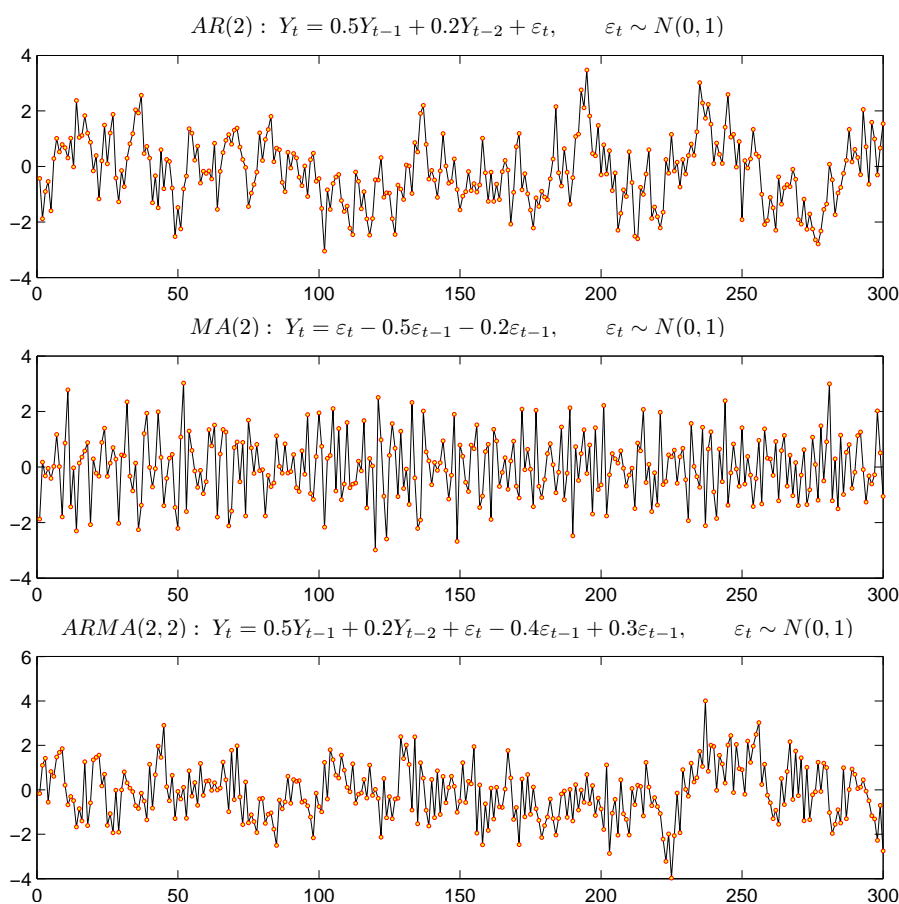
$$|\Theta(e^{-i\omega})|^2 f_Y(\omega) = |\Theta(e^{-i\omega})|^2 \frac{\sigma_\varepsilon^2}{2\pi} \quad \text{pro} \quad \omega \in \langle -\pi, \pi \rangle.$$

Odtud již dostáváme tvrzení věty

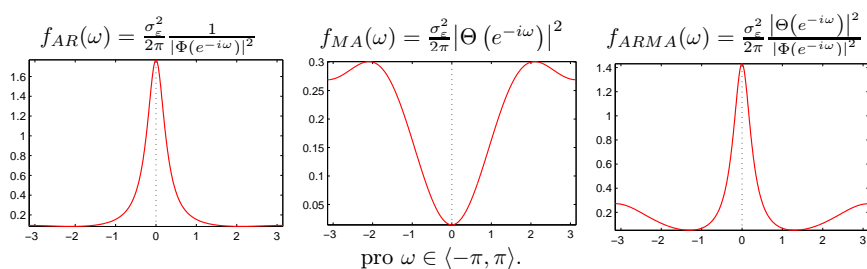
$$f_Y(\omega) = \frac{\sigma_\varepsilon^2}{2\pi} \frac{|\Theta(e^{-i\omega})|^2}{|\Phi(e^{-i\omega})|^2} \quad \text{pro} \quad \omega \in \langle -\pi, \pi \rangle.$$

■

Na závěr tohoto odstavce jsou vykresleny příklady tří realizací AR, MA a ARMA procesů spolu s jejich teoretickými spektrálními hustotami.



Obrázek 1.6: Ukázky realizací AR, MA a ARMA procesů



Obrázek 1.7: Ukázky spektrálních hustot AR, MA a ARMA procesů.

1.3 Stacionární procesy a nejlepší lineární predikce

Nechť $\{Y_t, t \in \mathbb{Z}\} \in L^2(\Omega, \mathcal{A}, P)$ je stacionární proces se střední hodnotou μ_Y a autokovarianční funkcí $\gamma_Y(t)$. Pak náhodný proces $\{Y_t - \mu_Y, t \in \mathbb{Z}\}$ má nulovou střední hodnotu (tj. je centrován) a má stejnou autokovarianční funkci $\gamma_Y(t)$.

Uvažujme **nejlepší lineární predikci** \hat{Y}_t pomocí Y_{t-1}, \dots, Y_{t-n} ($n \geq 1$), která je ortogonální projekcí

$$\hat{Y}_t = P_{\overline{\text{sp}}\{1, Y_{t-1}, \dots, Y_{t-n}\}}(Y_t).$$

Lze snadno ukázat, že platí

$$\hat{Y}_t = P_{\overline{\text{sp}}\{1, Y_{t-1}, \dots, Y_{t-n}\}}(Y_t) = \mu_Y + P_{\overline{\text{sp}}\{Y_{t-1}, \dots, Y_{t-n}\}}(Y_t).$$

Takže bez újmy na obecnosti můžeme dále uvažovat pouze centrované stacionární procesy $\{Y_t, t \in \mathbb{Z}\}$, pro které platí

$$\hat{Y}_t = P_{\overline{\text{sp}}\{1, Y_{t-1}, \dots, Y_{t-n}\}}(Y_t) = P_{\overline{\text{sp}}\{Y_{t-1}, \dots, Y_{t-n}\}}(Y_t).$$

Nejprve definujeme jednokrokovou predikci.

Definice 1.3.1.

Nechť $\{Y_t, t \in \mathbb{Z}\} \in L^2(\Omega, \mathcal{A}, P)$ je centrovaný stacionární proces. Označme pro $n \geq 1$

$$\mathcal{M}_n = \overline{\text{sp}}\{Y_1, \dots, Y_n\}.$$

Pak **jednokroková (lineární) predikce** je definována vztahem

$$\hat{Y}_{n+1} = \hat{Y}_{n+1|n} = \begin{cases} 0 (= \mu_Y) & n = 0, \\ P_{\overline{\text{sp}}\{Y_1, \dots, Y_n\}}(Y_{n+1}) = P_{\mathcal{M}_n}(Y_{n+1}) & n \geq 1. \end{cases}$$

Protože pro $n \geq 1$ $\hat{Y}_{n+1} \in \mathcal{M}_n$, pak platí

$$\hat{Y}_{n+1} = \phi_{n,1}Y_n + \dots + \phi_{n,n}Y_1$$

a $\phi_{n,1}, \dots, \phi_{n,n}$ minimalizují

$$\|Y_{n+1} - \hat{Y}_{n+1}\|^2 = E|Y_{n+1} - \hat{Y}_{n+1}|^2.$$

Podle projekční věty pro každé $X \in L^2(\Omega, \mathcal{A}, P)$ a pro každé $Y \in \mathcal{M}_n$ platí

$$\langle X - \hat{X}, Y \rangle = \langle X, Y \rangle - \langle \hat{X}, Y \rangle = 0 \Rightarrow \langle X, Y \rangle = \langle \hat{X}, Y \rangle \text{ což je } \boxed{EXY = E\hat{X}Y},$$

takže jestliže pro $j = 1, \dots, n$ položíme $X = Y_{n+1}$ a $Y = Y_{n+1-j}$, pak musí platit

$$\begin{aligned} EY_{n+1}Y_{n+1-j} &= E\widehat{Y}_{n+1}Y_{n+1-j} \\ \gamma(j) &= E\left(Y_{n+1-j} \sum_{i=1}^n \phi_{n,i} Y_{n+1-i}\right) = \sum_{i=1}^n \phi_{n,i} EY_{n+1-i}Y_{n+1-j} \\ &= \sum_{i=1}^n \phi_{n,i} \gamma(i-j) \end{aligned}$$

což lze maticově zapsat takto

$$\begin{pmatrix} \gamma(1) \\ \gamma(2) \\ \vdots \\ \gamma(n) \end{pmatrix} = \begin{pmatrix} \gamma(0) & \gamma(1) & \cdots & \gamma(n-1) \\ \gamma(1) & \gamma(0) & \cdots & \gamma(n-2) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma(n-1) & \gamma(n-2) & \cdots & \gamma(0) \end{pmatrix} \begin{pmatrix} \phi_{n,1} \\ \phi_{n,2} \\ \vdots \\ \phi_{n,n} \end{pmatrix}$$

tj.

$$\boldsymbol{\gamma}_n = \boldsymbol{\Gamma}_n \boldsymbol{\phi}_n.$$

Projekční věta zaručuje existenci právě jednoho řešení $\widehat{Y}_{n+1} \in \mathcal{M}_n$ pro nějaké $\boldsymbol{\phi}_n \in \mathbb{R}^n$ (kterých obecně může být více, jejich výsledkem je však pouze jediné \widehat{Y}_{n+1}). Jestliže $\boldsymbol{\Gamma}_n$ je regulární, máme právě jediné $\boldsymbol{\phi}_n \in \mathbb{R}^n$ a platí

$$\boldsymbol{\phi}_n = \boldsymbol{\Gamma}_n^{-1} \boldsymbol{\gamma}_n.$$

Následující věta dává postačující podmínku k tomu, aby pro každé $n \in \mathbb{N}$ byla $\boldsymbol{\Gamma}_n$ regulární maticí.

Věta 1.3.2.

Jestliže platí

$$\gamma(0) > 0 \quad a \quad \gamma(h) \xrightarrow{h \rightarrow \infty} 0,$$

pak kovarianční matice

$$\boldsymbol{\Gamma}_n = (\gamma(i-j))_{i,j=1}^n$$

je regulární pro každé $n \in \mathbb{N}$.

DŮKAZ. Tento důkaz se provádí sporem, viz Brockwel, Davis (1987), str. 160-161. ■

Důsledek 1.3.3.

Označme

$$\mathbf{Y}_n = (Y_n, \dots, Y_1)'$$

Jestliže platí $\gamma(0) > 0$ a $\gamma(h) \xrightarrow{h \rightarrow \infty} 0$, pak nejlepší lineární predikce \widehat{Y}_{n+1} náhodné veličiny Y_{n+1} je tvaru

$$\widehat{Y}_{n+1} = \phi_{n,1}Y_n + \dots + \phi_{n,n}Y_1 \quad \text{tj.} \quad \widehat{Y}_{n+1} = \phi_n' \mathbf{Y}_n \quad \text{přičemž} \quad \phi_n = \mathbf{\Gamma}_n^{-1} \boldsymbol{\gamma}_n.$$

Střední kvadratická chyba je rovna

$$v_n = MSE(\widehat{Y}_{n+1}) = E(Y_{n+1} - \widehat{Y}_{n+1})^2 = \gamma(0) - \boldsymbol{\gamma}_n' \mathbf{\Gamma}_n^{-1} \boldsymbol{\gamma}_n. \quad (1.2)$$

DŮKAZ. Tvrzení týkající se tvaru nejlepší lineární predikce a vektoru ϕ_n plynou z předchozích poznámek a předešlé věty. Zbývá vypočítat střední kvadratickou chybu.

$$E(Y_{n+1} - \widehat{Y}_{n+1})^2 = E(Y_{n+1} - \phi_n' \mathbf{Y}_n)^2 = EY_{n+1}^2 - 2E(\phi_n' \mathbf{Y}_n Y_{n+1}) + E(\phi_n' \mathbf{Y}_n)^2.$$

Nejprve počítejme

$$E\mathbf{Y}_n Y_{n+1} = (EY_n Y_{n+1}, EY_{n-1} Y_{n+1}, \dots, EY_1 Y_{n+1})' = (\gamma(1), \gamma(2), \dots, \gamma(n))' = \boldsymbol{\gamma}_n.$$

Dále si všimněme, že lze psát

$$(\phi_n' \mathbf{Y}_n)^2 = \phi_n' \mathbf{Y}_n \mathbf{Y}_n' \phi_n$$

a počítejme

$$\begin{aligned} E\mathbf{Y}_n \mathbf{Y}_n' &= E \left[\begin{pmatrix} Y_n \\ Y_{n-1} \\ \vdots \\ Y_1 \end{pmatrix} (Y_n, \dots, Y_1) \right] = \begin{pmatrix} EY_n^2 & EY_n Y_{n-1} & \dots & EY_n Y_1 \\ EY_{n-1} Y_n & EY_{n-1}^2 & \dots & EY_{n-1} Y_1 \\ \vdots & \vdots & \ddots & \vdots \\ EY_1 Y_n & EY_1 Y_{n-1} & \dots & EY_1^2 \end{pmatrix} \\ &= \begin{pmatrix} \gamma(0) & \gamma(1) & \dots & \gamma(n-1) \\ \gamma(1) & \gamma(0) & \dots & \gamma(n-2) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma(n-1) & \gamma(n-2) & \dots & \gamma(0) \end{pmatrix} = \mathbf{\Gamma}_n \end{aligned}$$

Takže můžeme pokračovat ve výpočtu střední kvadratické chyby

$$\begin{aligned} E(Y_{n+1} - \widehat{Y}_{n+1})^2 &= EY_{n+1}^2 - 2\phi_n' E\mathbf{Y}_n Y_{n+1} + \phi_n' E\mathbf{Y}_n \mathbf{Y}_n' \phi_n \\ &= \gamma(0) - 2\phi_n' \boldsymbol{\gamma}_n + \phi_n' \mathbf{\Gamma}_n \phi_n \\ &= \gamma(0) - 2\boldsymbol{\gamma}_n' \mathbf{\Gamma}_n^{-1} \boldsymbol{\gamma}_n + \boldsymbol{\gamma}_n' \mathbf{\Gamma}_n^{-1} \mathbf{\Gamma}_n \mathbf{\Gamma}_n^{-1} \boldsymbol{\gamma}_n = \gamma(0) - \boldsymbol{\gamma}_n' \mathbf{\Gamma}_n^{-1} \boldsymbol{\gamma}_n. \quad \blacksquare \end{aligned}$$

Nyní definujme h -krokovou (lineární) predikci.

Definice 1.3.4.

Nechť $\{Y_t, t \in \mathbb{Z}\} \in L^2(\Omega, \mathcal{A}, P)$ je centrováný stacionární proces. Označme pro $n \geq 1$

$$\mathcal{M}_n = \overline{\text{sp}}\{Y_1, \dots, Y_n\}.$$

Pak h -kroková predikce je definována vztahem

$$\hat{Y}_{n+h} = \hat{Y}_{n+h|n} = \begin{cases} 0 (= \mu_Y) & n, h = 0, \\ P_{\overline{\text{sp}}\{Y_1, \dots, Y_n\}}(Y_{n+h}) = P_{\mathcal{M}_n}(Y_{n+h}) & n, h \geq 1. \end{cases}$$

Obdobným způsobem jako u jednokrokové predikce můžeme odvodit, že jestliže platí $\gamma(0) > 0$ a $\gamma(h) \xrightarrow{h \rightarrow \infty} 0$, pak nejlepší lineární h -kroková predikce \hat{Y}_{n+h} náhodné veličiny Y_{n+h} je tvaru

$$\hat{Y}_{n+h} = \phi_{n,1}^{(h)} Y_n + \dots + \phi_{n,n}^{(h)} Y_1 \quad \text{tj.} \quad \hat{Y}_{n+h} = \left(\phi_n^{(h)}\right)' \mathbf{Y}_n$$

přičemž

$$\phi_n^{(h)} = \mathbf{\Gamma}_n^{-1} \gamma_n^{(h)} \quad \text{a} \quad \gamma_n^{(h)} = (\gamma(h), \gamma(h+1), \dots, \gamma(h+n-1))'.$$

Střední kvadratická chyba je rovna

$$v_n^{(h)} = \text{MSE}(\hat{Y}_{n+h}) = E(Y_{n+1} - \hat{Y}_{n+1})^2 = \gamma(0) - \left(\phi_n^{(h)}\right)' \mathbf{\Gamma}_n^{-1} \phi_n^{(h)}.$$

1.4 Rekurentní metody pro výpočet nejlepší lineární predikce

Věta 1.4.1 (Durbin-Levinsův algoritmus).

Nechť $\{Y_t, t \in \mathbb{Z}\} \in L^2(\Omega, \mathcal{A}, P)$ je centrováný stacionární proces s autokovarianční funkcí $\gamma(h)$ takovou, že $\gamma(0) > 0$ a $\gamma(h) \xrightarrow{h \rightarrow \infty} 0$. Jestliže

$\hat{Y}_{n+1} = P_{\overline{\text{sp}}\{Y_1, \dots, Y_n\}}(Y_{n+1}) = \phi_{n,1} Y_n + \dots + \phi_{n,n} Y_1$ je nejlepší lineární predikce, pak pro koeficienty $\phi_{n,j}$ ($j = 1, \dots, n$) a střední kvadratické chyby

$v_n = E \left(Y_{n+1} - \widehat{Y}_{n+1} \right)^2$ platí následující vztahy

$$\phi_{1,1} = \frac{\gamma(1)}{\gamma(0)} = \rho(1) \qquad v_0 = \gamma(0) \quad (1.3)$$

$$\phi_{n,n} = [\gamma(n) - \phi'_{n-1} \gamma_{n-1}] / v_{n-1} \quad (1.4)$$

$$\phi_n^{(1)} = \phi_{n-1} - \phi_{n,n} \phi_{n-1}^* \qquad v_n = v_{n-1} (1 - \phi_{n,n}^2) \quad (1.5)$$

kde

$$\phi_{n-1} = (\phi_{n-1,1}, \dots, \phi_{n-1,n-1})' \qquad \phi_{n-1}^* = (\phi_{n-1,n-1}, \dots, \phi_{n-1,1})'$$

$$\phi_n = (\phi_{n,1}, \dots, \phi_{n,n-1}, \phi_{n,n})' \qquad \phi_n^{(1)} = (\phi_{n,1}, \dots, \phi_{n,n-1})'$$

DŮKAZ. Mějme

$$\mathcal{M}_n = \overline{\text{sp}}\{Y_1, \dots, Y_n\}.$$

Označme

$$\mathcal{M}_{n-1} = \overline{\text{sp}}\{Y_2, \dots, Y_n\}$$

a

$$\mathcal{M}_{n-1}^\perp = \overline{\text{sp}}\{Y_1 - P_{\mathcal{M}_{n-1}}(Y_1)\}.$$

Vidíme, že \mathcal{M}_{n-1}^\perp je ortogonální komplement \mathcal{M}_{n-1} v \mathcal{M}_n . Takže pro libovolné $Y \in L^2(\Omega, \mathcal{A}, P)$, tedy i pro $Y_{n+1} \in L^2(\Omega, \mathcal{A}, P)$ musí platit

$$\widehat{Y}_{n+1} = P_{\mathcal{M}_n}(Y_{n+1}) = P_{\mathcal{M}_{n-1}}(Y_{n+1}) + P_{\mathcal{M}_{n-1}^\perp}(Y_{n+1}).$$

Označme

$$e_1 = \frac{Y_1 - P_{\mathcal{M}_{n-1}}(Y_1)}{\|Y_1 - P_{\mathcal{M}_{n-1}}(Y_1)\|} \quad \text{tj.} \quad \|e_1\| = 1.$$

Pak platí

$$\begin{aligned} \widehat{Y}_{n+1} &= P_{\mathcal{M}_{n-1}}(Y_{n+1}) + \overbrace{\langle Y_{n+1}, e_1 \rangle}_{\text{Fourier.koef.}} e_1 \\ &= P_{\mathcal{M}_{n-1}}(Y_{n+1}) + \left\langle Y_{n+1}, \frac{Y_1 - P_{\mathcal{M}_{n-1}}(Y_1)}{\|Y_1 - P_{\mathcal{M}_{n-1}}(Y_1)\|} \right\rangle \frac{Y_1 - P_{\mathcal{M}_{n-1}}(Y_1)}{\|Y_1 - P_{\mathcal{M}_{n-1}}(Y_1)\|} \\ &= P_{\mathcal{M}_{n-1}}(Y_{n+1}) + \frac{\langle Y_{n+1}, Y_1 - P_{\mathcal{M}_{n-1}}(Y_1) \rangle}{\|Y_1 - P_{\mathcal{M}_{n-1}}(Y_1)\|^2} (Y_1 - P_{\mathcal{M}_{n-1}}(Y_1)) \quad (1.6) \end{aligned}$$

Označme

$$a = \frac{\langle Y_{n+1}, Y_1 - P_{\mathcal{M}_{n-1}}(Y_1) \rangle}{\|Y_1 - P_{\mathcal{M}_{n-1}}(Y_1)\|^2}.$$

Počítejme postupně

$$P_{\mathcal{M}_{n-1}}(Y_1) = \phi_{n-1,1}Y_2 + \cdots + \phi_{n-1,n-1}Y_n$$

a označíme-li

$$\phi_{n-1} = (\phi_{n-1,1}, \dots, \phi_{n-1,n-1})',$$

pak složky vektoru ϕ_{n-1} jsou pro $k = 1, \dots, n-1$ řešením rovnic

$$EY_1Y_{1+k} = E(Y_1P_{\mathcal{M}_{n-1}}(Y_1)) = \sum_{j=1}^{n-1} \phi_{n-1,j}EY_{1+j}Y_{1+k},$$

což lze zapsat maticově

$$\begin{pmatrix} \gamma(1) \\ \gamma(2) \\ \vdots \\ \gamma(n-1) \end{pmatrix} = \begin{pmatrix} \gamma(0) & \gamma(1) & \cdots & \gamma(n-2) \\ \gamma(1) & \gamma(0) & \cdots & \gamma(n-3) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma(n-2) & \gamma(n-3) & \cdots & \gamma(0) \end{pmatrix} \begin{pmatrix} \phi_{n-1,1} \\ \phi_{n-1,2} \\ \vdots \\ \phi_{n-1,n-1} \end{pmatrix}$$

tj.

$$\gamma_{n-1} = \Gamma_{n-1} \phi_{n-1},$$

Jestliže platí $\gamma(0) > 0$ a $\gamma(h) \xrightarrow{h \rightarrow \infty} 0$, pak kovarianční matice Γ_{n-1} je regulární a platí

$$\phi_{n-1} = \Gamma_{n-1}^{-1} \gamma_{n-1}.$$

Obdobně

$$P_{\mathcal{M}_{n-1}}(Y_{n+1}) = \psi_{n-1,1}Y_2 + \cdots + \psi_{n-1,n-1}Y_n$$

a označíme-li

$$\psi_{n-1} = (\psi_{n-1,1}, \dots, \psi_{n-1,n-1})',$$

pak složky vektoru ψ_{n-1} jsou pro $k = 1, \dots, n-1$ řešením rovnic

$$EY_{n+1}Y_{n+1-k} = E(Y_{n+1}P_{\mathcal{M}_{n-1}}(Y_{n+1-k})) = \sum_{j=1}^{n-1} \psi_{n-1,j}EY_{n+1-j}Y_{n+1-k},$$

což maticově zapíšeme

$$\begin{pmatrix} \gamma(1) \\ \gamma(2) \\ \vdots \\ \gamma(n-1) \end{pmatrix} = \begin{pmatrix} \gamma(0) & \gamma(1) & \cdots & \gamma(n-2) \\ \gamma(1) & \gamma(0) & \cdots & \gamma(n-3) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma(n-2) & \gamma(n-3) & \cdots & \gamma(0) \end{pmatrix} \begin{pmatrix} \psi_{n-1,1} \\ \psi_{n-1,2} \\ \vdots \\ \psi_{n-1,n-1} \end{pmatrix}$$

tj.

$$\gamma_{n-1} = \Gamma_{n-1} \psi_{n-1},$$

tedy

$$\psi_{n-1} = \Gamma_{n-1}^{-1} \gamma_{n-1} = \phi_{n-1}.$$

Celkově tedy, označíme-li

$$\begin{aligned}\mathbf{Y}_{n-1}^* &= (Y_2, \dots, Y_n)' \\ \mathbf{Y}_{n-1} &= (Y_n, \dots, Y_2)'\end{aligned}$$

tak dostaneme

$$\begin{aligned}P_{\mathcal{M}_{n-1}}(Y_1) &= \phi'_{n-1} \mathbf{Y}_{n-1}^* \\ P_{\mathcal{M}_{n-1}}(Y_{n+1}) &= \phi'_{n-1} \mathbf{Y}_{n-1}\end{aligned}$$

Využijme vzorce (1.2) z důsledku 1.3.3 a počítejme střední kvadratické chyby obou predikcí

$$\begin{aligned}\|Y_1 - P_{\mathcal{M}_{n-1}}(Y_1)\|^2 &= \gamma(0) - \gamma'_{n-1} \mathbf{\Gamma}_{n-1}^{-1} \gamma_{n-1} \\ \|Y_{n+1} - P_{\mathcal{M}_{n-1}}(Y_{n+1})\|^2 &= \gamma(0) - \gamma'_{n-1} \mathbf{\Gamma}_{n-1}^{-1} \gamma_{n-1}\end{aligned}$$

Tedy

$$v_{n-1} = \|Y_1 - P_{\mathcal{M}_{n-1}}(Y_1)\|^2 = \|Y_{n+1} - P_{\mathcal{M}_{n-1}}(Y_{n+1})\|^2.$$

Vraťme se k rovnici (1.6), pak

$$\begin{aligned}\widehat{Y}_{n+1} &= P_{\mathcal{M}_n}(Y_{n+1}) = P_{\mathcal{M}_{n-1}}(Y_{n+1}) + a(Y_1 - P_{\mathcal{M}_{n-1}}(Y_1)) \\ &= aY_1 + \phi_{n-1,1}Y_n + \dots + \phi_{n-1,n-1}Y_2 - a(\phi_{n-1,1}Y_2 + \dots + \phi_{n-1,n-1}Y_n) \\ &= aY_1 + \sum_{j=1}^{n-1} (\phi_{n-1,j} - a\phi_{n-1,n-j})Y_{n+1-j} \\ &= \phi_{n,n}Y_1 + \phi_{n,n-1}Y_2 + \dots + \phi_{n,1}Y_n\end{aligned}$$

a odtud dostaneme

$$\begin{aligned}\phi_{n,n} &= a \\ \phi_{n,j} &= \phi_{n-1,j} - a\phi_{n-1,n-j} \quad \text{pro } j = 1, \dots, n-1.\end{aligned}$$

Nyní se vrátíme ke konstantě a . Protože

$$\begin{aligned}\langle Y_{n+1}, Y_1 - P_{\mathcal{M}_{n-1}}(Y_1) \rangle &= \langle Y_{n+1}, Y_1 \rangle - \langle Y_{n+1}, P_{\mathcal{M}_{n-1}}(Y_1) \rangle \\ &= EY_1Y_{n+1} - E \left(Y_{n+1} \sum_{j=1}^{n-1} \phi_{n-1,j} Y_{1+j} \right) \\ &= \gamma(n) - \sum_{j=1}^{n-1} \phi_{n-1,j} EY_{1+j}Y_{n+1} \\ &= \gamma(n) - \sum_{j=1}^{n-1} \phi_{n-1,j} \gamma(n-j),\end{aligned}$$

pak

$$a = \phi_{n,n} = \frac{\langle Y_{n+1}, Y_1 - P_{\mathcal{M}_{n-1}}(Y_1) \rangle}{\|Y_1 - P_{\mathcal{M}_{n-1}}(Y_1)\|^2} = \frac{\gamma(n) - \sum_{j=1}^{n-1} \phi_{n-1,j} \gamma(n-j)}{v_{n-1}}.$$

Zbývá dokázat, že platí

$$v_n = v_{n-1}(1 - \phi_{n,n}^2).$$

Počítejme proto střední kvadratickou chybu predikce $P_{\mathcal{M}_n}(Y_{n+1})$:

$$\begin{aligned} v_n &= \|Y_{n+1} - P_{\mathcal{M}_n}(Y_{n+1})\|^2 = \left\| \underbrace{Y_{n+1} - P_{\mathcal{M}_{n-1}}(Y_{n+1})}_{\text{orthogonal to } P_{\mathcal{M}_{n-1}}^{\perp}(Y_{n+1})} - P_{\mathcal{M}_{n-1}}^{\perp}(Y_{n+1}) \right\|^2 \\ &= \left\langle \underbrace{Y_{n+1} - P_{\mathcal{M}_{n-1}}(Y_{n+1})}_{\text{orthogonal to } P_{\mathcal{M}_{n-1}}^{\perp}(Y_{n+1})} - P_{\mathcal{M}_{n-1}}^{\perp}(Y_{n+1}), \underbrace{Y_{n+1} - P_{\mathcal{M}_{n-1}}(Y_{n+1})}_{\text{orthogonal to } P_{\mathcal{M}_{n-1}}^{\perp}(Y_{n+1})} - P_{\mathcal{M}_{n-1}}^{\perp}(Y_{n+1}) \right\rangle \\ &= \|Y_{n+1} - P_{\mathcal{M}_{n-1}}(Y_{n+1})\|^2 - 2 \left\langle Y_{n+1} - P_{\mathcal{M}_{n-1}}(Y_{n+1}), P_{\mathcal{M}_{n-1}}^{\perp}(Y_{n+1}) \right\rangle \\ &\quad + \|P_{\mathcal{M}_{n-1}}^{\perp}(Y_{n+1})\|^2 \end{aligned}$$

Protože

$$P_{\mathcal{M}_{n-1}}^{\perp}(Y_{n+1}) = a(Y_1 - P_{\mathcal{M}_{n-1}}(Y_1)),$$

tedy

$$\|P_{\mathcal{M}_{n-1}}^{\perp}(Y_{n+1})\|^2 = a^2 \|Y_1 - P_{\mathcal{M}_{n-1}}(Y_1)\|^2 = a^2 v_{n-1}.$$

Zbývá dopočítat

$$\begin{aligned} \langle Y_{n+1} - P_{\mathcal{M}_{n-1}}(Y_{n+1}), P_{\mathcal{M}_{n-1}}^{\perp}(Y_{n+1}) \rangle &= \langle Y_{n+1} - P_{\mathcal{M}_{n-1}}(Y_{n+1}), a(Y_1 - P_{\mathcal{M}_{n-1}}(Y_1)) \rangle \\ &= a \langle Y_{n+1}, Y_1 - P_{\mathcal{M}_{n-1}}(Y_1) \rangle - a \underbrace{\langle P_{\mathcal{M}_{n-1}}(Y_{n+1}), Y_1 - P_{\mathcal{M}_{n-1}}(Y_1) \rangle}_{=0(\text{ortogonal})} \end{aligned}$$

Dále víme, že

$$a = \frac{\langle Y_{n+1}, Y_1 - P_{\mathcal{M}_{n-1}}(Y_1) \rangle}{v_{n-1}} = \phi_{n,n},$$

tedy

$$a \langle Y_{n+1}, Y_1 - P_{\mathcal{M}_{n-1}}(Y_1) \rangle = a^2 v_{n-1}$$

a celkově

$$v_n = v_{n-1} - 2a^2 v_{n-1} + a^2 v_{n-1} = v_{n-1}(1 - a^2) = v_{n-1}(1 - \phi_{n,n}^2).$$

■

1.4.1 Důsledek Durbin-Levinsonova algoritmu

Důsledek 1.4.2.

Nechť $\{Y_t, t \in \mathbb{Z}\} \in L^2(\Omega, \mathcal{A}, P)$ je centrováný stacionární proces s autokovarianční funkcí $\gamma(h)$, pro kterou platí $\gamma(0) > 0$ a $\gamma(h) \xrightarrow{h \rightarrow \infty} 0$. Označme

$$\begin{aligned}\mathcal{M}_n &= \overline{\text{sp}}\{Y_1, \dots, Y_n\} \\ \mathcal{M}_{n-1} &= \overline{\text{sp}}\{Y_2, \dots, Y_n\}\end{aligned}$$

a nejlepší lineární predikci

$$\widehat{Y}_{n+1} = P_{\mathcal{M}_n}(Y_{n+1}) = \phi_{n,n}Y_1 + \phi_{n,n-1}Y_2 + \dots + \phi_{n,1}Y_n,$$

pak platí

$$\boxed{\phi_{n,n} = R(Y_{n+1} - P_{\mathcal{M}_{n-1}}(Y_{n+1}), Y_1 - P_{\mathcal{M}_{n-1}}(Y_1))}. \quad (1.7)$$

DŮKAZ. Označme

$$\mathcal{M}_{n-1}^\perp = \overline{\text{sp}}\{Y_1 - P_{\mathcal{M}_{n-1}}(Y_1)\}.$$

Tedy \mathcal{M}_{n-1}^\perp je ortogonální komplement \mathcal{M}_{n-1} v \mathcal{M}_n . Takže lze psát

$$\widehat{Y}_{n+1} = P_{\mathcal{M}_n}(Y_{n+1}) = P_{\mathcal{M}_{n-1}}(Y_{n+1}) + P_{\mathcal{M}_{n-1}^\perp}(Y_{n+1}).$$

Protože $P_{\mathcal{M}_{n-1}}(Y_{n+1})$ a $Y_1 - P_{\mathcal{M}_{n-1}}(Y_1)$ jsou ortogonální tj.

$$\langle P_{\mathcal{M}_{n-1}}(Y_{n+1}), Y_1 - P_{\mathcal{M}_{n-1}}(Y_1) \rangle = 0.$$

Tedy

$$\langle Y_{n+1}, Y_1 - P_{\mathcal{M}_{n-1}}(Y_1) \rangle = \langle Y_{n+1} - P_{\mathcal{M}_{n-1}}(Y_{n+1}), Y_1 - P_{\mathcal{M}_{n-1}}(Y_1) \rangle.$$

Navíc platí (viz důkaz Durbin-Levinsonova algoritmu)

$$\|Y_1 - P_{\mathcal{M}_{n-1}}(Y_1)\|^2 = \|Y_{n+1} - P_{\mathcal{M}_{n-1}}(Y_{n+1})\|^2 = v_{n-1},$$

takže lze psát

$$\begin{aligned}
\boxed{\phi_{n,n}} &= a = \frac{\langle Y_{n+1}, Y_1 - P_{\mathcal{M}_{n-1}}(Y_1) \rangle}{\|Y_1 - P_{\mathcal{M}_{n-1}}(Y_1)\|^2} \\
&= \frac{\langle Y_{n+1} - P_{\mathcal{M}_{n-1}}(Y_{n+1}), Y_1 - P_{\mathcal{M}_{n-1}}(Y_1) \rangle}{\|Y_1 - P_{\mathcal{M}_{n-1}}(Y_1)\| \|Y_{n+1} - P_{\mathcal{M}_{n-1}}(Y_{n+1})\|} \\
&= \frac{E(Y_{n+1} - P_{\mathcal{M}_{n-1}}(Y_{n+1})) (Y_1 - P_{\mathcal{M}_{n-1}}(Y_1))}{\sqrt{E(Y_{n+1} - P_{\mathcal{M}_{n-1}}(Y_{n+1}))^2} \sqrt{E(Y_1 - P_{\mathcal{M}_{n-1}}(Y_1))^2}} \\
&= \frac{C(Y_{n+1} - P_{\mathcal{M}_{n-1}}(Y_{n+1}), Y_1 - P_{\mathcal{M}_{n-1}}(Y_1))}{\sqrt{D(Y_{n+1} - P_{\mathcal{M}_{n-1}}(Y_{n+1}))} \sqrt{D(Y_1 - P_{\mathcal{M}_{n-1}}(Y_1))}} \\
&= \boxed{R(Y_{n+1} - P_{\mathcal{M}_{n-1}}(Y_{n+1}), Y_1 - P_{\mathcal{M}_{n-1}}(Y_1))}. \quad \blacksquare
\end{aligned}$$

1.5 Parciální autokorelační funkce (PACF)

Definice 1.5.1.

Nechť $\{Y_t, t \in \mathbb{Z}\} \in L^2(\Omega, \mathcal{A}, P)$ je stacionární proces. Pak **parciální autokorelační funkce** je definována vztahem

$$\begin{aligned}
\alpha(1) &= R(Y_t, Y_{t+1}) \\
\alpha(k) &= R(Y_t - \widehat{Y}_t, Y_{t-k} - \widehat{Y}_{t-k}) \quad \text{pro } |k| > 1
\end{aligned}$$

kde \widehat{Y}_t , resp. \widehat{Y}_{t-k} jsou nejlepší lineární predikce Y_t (resp. Y_{t-k}) pomocí $Y_{t-k+1}, \dots, Y_{t-1}$.

Nejlepší lineární predikce \widehat{Y}_t a \widehat{Y}_{t-k} jsou projekce $\boxed{\widehat{Y}_t = P_{\mathcal{M}_{k-1}}(Y_t)}$ a $\boxed{\widehat{Y}_{t-k} = P_{\mathcal{M}_{k-1}}(Y_{t-k})}$, kde $\mathcal{M}_{k-1} = \overline{\text{span}}\{Y_{t-k+1}, \dots, Y_{t-1}\}$. Přitom existují taková $\phi_{k-1} = (\phi_{k-1,1}, \dots, \phi_{k-1,k-1})'$, že platí

$$\widehat{Y}_t = \phi_{k-1,1} Y_{t-1} + \dots + \phi_{k-1,k-1} Y_{t-k+1}$$

a také taková $\psi_{k-1} = (\psi_{k-1,1}, \dots, \psi_{k-1,k-1})'$, že platí

$$\widehat{Y}_{t-k} = \psi_{k-1,1} Y_{t-k+1} + \dots + \psi_{k-1,k-1} Y_{t-1},$$

která minimalizují

$$E(Y_t - \widehat{Y}_t)^2 \quad \text{resp.} \quad E(Y_{t-k} - \widehat{Y}_{t-k})^2,$$

přičemž (jak již víme z důkazu Durbin-Levinsonova algoritmu) platí

$$\phi_{k-1,1} = \psi_{k-1,1}, \dots, \phi_{k-1,k-1} = \psi_{k-1,k-1} \quad \text{tj.} \quad \boldsymbol{\phi}_{k-1} = \boldsymbol{\psi}_{k-1}.$$

Celkově tedy, označíme-li

$$\begin{aligned} \mathbf{Y}_{k-1}^* &= (Y_{t-k+1}, \dots, Y_{t-1})' \\ \mathbf{Y}_{k-1} &= (Y_{t-1}, \dots, Y_{t-k+1})' \end{aligned}$$

tak dostaneme

$$\begin{aligned} P_{\mathcal{M}_{n-1}}(Y_{t-k}) &= \boldsymbol{\phi}_{k-1}' \mathbf{Y}_{k-1}^* \\ P_{\mathcal{M}_{n-1}}(Y_t) &= \boldsymbol{\phi}_{k-1}' \mathbf{Y}_{k-1} \end{aligned}$$

Víme, že pokud pro autokovarianční funkci $\gamma(h)$ platí $\gamma(0) > 0$ a $\gamma(h) \xrightarrow{h \rightarrow \infty} 0$, pak matice $\boldsymbol{\Gamma}_{k-1}$ je regulární a neznámé složky vektoru $\boldsymbol{\phi}_{k-1}$ jsou rovny

$$\boldsymbol{\phi}_{k-1} = \boldsymbol{\Gamma}_{k-1}^{-1} \boldsymbol{\gamma}_{k-1}.$$

Avšak podle důsledku 1.4.2 Durbin-Levinsonova algoritmu není třeba počítat inverzní matici $\boldsymbol{\Gamma}_{k-1}^{-1}$, odtud $\hat{Y}_{t-k} = P_{\mathcal{M}_{n-1}}(Y_{t-k})$ a $\hat{Y}_t = P_{\mathcal{M}_{n-1}}(Y_t)$ a nakonec korelační koeficient $\alpha(k) = R(Y_t - \hat{Y}_t, Y_{t-k} - \hat{Y}_{t-k})$, neboť platí

$$\alpha(k) = \phi_{k,k} = R(Y_t - P_{\mathcal{M}_{k-1}}(Y_t), Y_{t-k} - P_{\mathcal{M}_{k-1}}(Y_{t-k})).$$

1.6 Inovační algoritmus

Základní myšlenkou Durbin-Levinsonova algoritmu je rozdělení

$$\mathcal{M}_n = \overline{\text{sp}}\{Y_n, \dots, Y_1\}$$

na dva ortogonální podprostory

$$\mathcal{M}_{n-1} = \overline{\text{sp}}\{Y_n, \dots, Y_2\} \quad \text{a} \quad \mathcal{M}_{n-1}^\perp = \overline{\text{sp}}\{Y_1 - P_{\mathcal{M}_{n-1}}(Y_1)\}.$$

Následující rekurentní algoritmus spočívá v dekompozici \mathcal{M}_n na n ortogonálních Hilbertových podprostorů pomocí Gram-Schmidtova algoritmu.

Rekurentní algoritmus lze aplikovat nejen na stacionární procesy, ale obecně na procesy s **konečnými druhými momenty**. Pro jednoduchost předpokládejme, že jsou centrované. Nejprve zavedme následující značení:

$$\gamma(i, j) = EX_i X_j.$$

Stejně označme

$$\begin{aligned} \mathcal{M}_n &= \overline{\text{sp}}\{Y_n, \dots, Y_1\} \\ v_n &= \left\| Y_{n+1} - \hat{Y}_{n+1} \right\|^2. \end{aligned}$$

Pokud označíme

$$\widehat{Y}_n = \begin{cases} 0 (= \mu_Y) & \text{pro } n = 1 \\ P_{\mathcal{M}_{n-1}}(Y_n) & \text{pro } n = 2, 3, \dots \end{cases}$$

pak zřejmě

$$\mathcal{M}_n = \overline{\text{sp}}\{Y_n - \widehat{Y}_n, \dots, Y_1 - \widehat{Y}_1\} \quad n \geq 1.$$

Definujme tzv. **inovaci** vztahem

$$U_{n+1} = Y_{n+1} - \widehat{Y}_{n+1} = Y_{n+1} - \sum_{j=1}^n \phi_{n,j} Y_{n+1-j}.$$

Označme

$$\begin{aligned} \mathbf{U}_n &= (U_1, \dots, U_n)' \\ \mathbf{Y}_n &= (Y_1, \dots, Y_n)' \\ \widehat{\mathbf{Y}}_n &= (\widehat{Y}_1, \dots, \widehat{Y}_n)'. \end{aligned}$$

Pak lze psát

$$\mathbf{U}_n = \mathbf{A}_n \mathbf{Y}_n,$$

kde matice \mathbf{A}_n je dolní trojúhelníkovou maticí

$$\mathbf{A}_n = \begin{pmatrix} 1 & 0 & \dots & \dots & \dots & 0 \\ -\phi_{1,1} & 1 & 0 & \dots & \dots & 0 \\ -\phi_{2,2} & -\phi_{2,1} & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ -\phi_{n-2,n-2} & -\phi_{n-2,n-3} & \dots & -\phi_{n-2,1} & 1 & 0 \\ -\phi_{n-1,n-1} & -\phi_{n-1,n-2} & \dots & \dots & -\phi_{n-1,1} & 1 \end{pmatrix}.$$

Všimněme si, že determinant matice je roven 1, takže existuje inverzní matice

$$\mathbf{C}_n = \mathbf{A}_n^{-1},$$

která je také dolní trojúhelníkovou maticí. Upravujme postupně

$$\widehat{\mathbf{Y}}_n = \mathbf{Y}_n - \mathbf{U}_n = \mathbf{A}_n^{-1} \mathbf{U}_n - \mathbf{U}_n = (\mathbf{A}_n^{-1} \mathbf{U}_n - \mathbf{I}_n) \mathbf{U}_n = \boldsymbol{\theta}_n \mathbf{U}_n,$$

kde

$$\boldsymbol{\theta}_n = \mathbf{C}_n - \mathbf{I}_n = \begin{pmatrix} 0 & 0 & \dots & \dots & \dots & 0 \\ \theta_{1,1} & 0 & 0 & \dots & \dots & 0 \\ \theta_{2,2} & \theta_{2,1} & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \theta_{n-2,n-2} & \theta_{n-2,n-3} & \dots & -\phi_{n-2,1} & 0 & 0 \\ \theta_{n-1,n-1} & \theta_{n-1,n-2} & \dots & \dots & \theta_{n-1,1} & 0 \end{pmatrix}.$$

Protože

$$\widehat{\mathbf{Y}}_n = \boldsymbol{\theta}_n \mathbf{U}_n = \boldsymbol{\theta}_n (\mathbf{Y}_n - \widehat{\mathbf{Y}}_n)$$

a protože $\boldsymbol{\theta}_n$ je dolní trojúhelníkovou maticí, můžeme psát

$$\widehat{Y}_{n+1} = \begin{cases} 0 & n = 0 \\ \sum_{j=1}^n \theta_{n,j} (Y_{n+1-j} - \widehat{Y}_{n+1-j}) & n = 1, 2, \dots \end{cases}$$

Věta 1.6.1.

Nechť $\{Y_t, t \in \mathbb{Z}\}$ je centrováný náhodný proces s konečnými druhými momenty, přičemž kovarianční matice $(E Y_i Y_j)_{i,j=1}^n = (\gamma(i, j))_{i,j=1}^n$ je regulární pro každé $n \in \mathbb{N}$. Pak pro jedнокrokovou predikci platí následující rekurentní vztahy

$$\widehat{Y}_{n+1} = \begin{cases} 0 & n = 0 \\ \sum_{j=1}^n \theta_{n,j} (Y_{n+1-j} - \widehat{Y}_{n+1-j}) & n = 1, 2, \dots \end{cases} \quad (1.8)$$

$$v_0 = \gamma(1, 1) \quad (1.9)$$

$$\theta_{n,n-k} = v_k^{-1} \left[\gamma(n+1, k+1) - \sum_{j=0}^{k-1} \theta_{k,k-j} \theta_{n,n-j} v_j \right] \quad k = 0, \dots, n-1 \quad (1.10)$$

$$v_n = \gamma(n+1, n+1) - \sum_{j=0}^{n-1} \theta_{n,n-j}^2 v_j \quad (1.11)$$

DŮKAZ. Tvrzení (1.6.4) jsem dokázali již v předchozím textu. Pro $n \geq 1$ provedme následující přeindexování:

$$\begin{aligned} \boxed{\widehat{Y}_{n+1}} &= \sum_{j=1}^n \theta_{n,j} (Y_{n+1-j} - \widehat{Y}_{n+1-j}) \\ &= \theta_{n,1} (Y_n - \widehat{Y}_n) + \dots + \theta_{n,n} (Y_1 - \widehat{Y}_1) \\ &= \boxed{\sum_{k=0}^{n-1} \theta_{n,n-k} (Y_{k+1} - \widehat{Y}_{k+1})} \end{aligned}$$

Obě strany předchozí rovnice vynásobme výrazem $(Y_{k+1} - \widehat{Y}_{k+1})$ a vypočítejme střední hodnoty. Dostaneme:

$$E \left[\widehat{Y}_{n+1} (Y_{k+1} - \widehat{Y}_{k+1}) \right] = \sum_{j=0}^{n-1} \theta_{n,n-j} E (Y_{j+1} - \widehat{Y}_{j+1}) (Y_{k+1} - \widehat{Y}_{k+1})$$

nebo ekvivalentně pomocí skalárních součinů:

$$\begin{aligned}
\langle \widehat{Y}_{n+1}, Y_{k+1} - \widehat{Y}_{k+1} \rangle &= \sum_{j=0}^{n-1} \theta_{n,n-j} \langle Y_{j+1} - \widehat{Y}_{j+1}, Y_{k+1} - \widehat{Y}_{k+1} \rangle \\
&= \theta_{n,n-k} \langle Y_{k+1} - \widehat{Y}_{k+1}, Y_{k+1} - \widehat{Y}_{k+1} \rangle \\
&= \theta_{n,n-k} \left\| Y_{k+1} - \widehat{Y}_{k+1} \right\|^2 \\
&= \theta_{n,n-k} v_k,
\end{aligned}$$

s využitím vztahů

$$\langle Y_{j+1} - \widehat{Y}_{j+1}, Y_{k+1} - \widehat{Y}_{k+1} \rangle = 0 \quad \text{pro} \quad j \neq k.$$

Dále díky tomu, že pro $n > s$

$$\langle Y_{n+1} - \widehat{Y}_{n+1}, Y_{k+1} - \widehat{Y}_{k+1} \rangle = 0,$$

dostáváme

$$\langle \widehat{Y}_{n+1}, Y_{k+1} - \widehat{Y}_{k+1} \rangle = \langle Y_{n+1}, Y_{k+1} - \widehat{Y}_{k+1} \rangle = \theta_{n,n-k} v_k.$$

Dále upravujeme

$$\begin{aligned}
\theta_{n,n-k} v_k &= \langle Y_{n+1}, Y_{k+1} - \widehat{Y}_{k+1} \rangle = \underbrace{\langle Y_{n+1}, Y_{k+1} \rangle}_{=\gamma(n+1,k+1)} - \langle Y_{n+1}, \widehat{Y}_{k+1} \rangle \\
&= \gamma(n+1, k+1) - \left\langle Y_{n+1}, \sum_{j=0}^{k-1} \theta_{k,k-j} (Y_{j+1} - \widehat{Y}_{j+1}) \right\rangle \\
&= \gamma(n+1, k+1) - \sum_{j=0}^{k-1} \theta_{k,k-j} \underbrace{\langle Y_{n+1}, Y_{j+1} - \widehat{Y}_{j+1} \rangle}_{=\theta_{n,n-j} v_j} \\
&= \gamma(n+1, k+1) - \sum_{j=0}^{k-1} \theta_{k,k-j} \theta_{n,n-j} v_j
\end{aligned}$$

Odtud jednoduchou úpravou dostaneme tvrzení (1.10):

$$\boxed{\theta_{n,n-k} = v_k^{-1} \left[\gamma(n+1, k+1) - \sum_{j=0}^{k-1} \theta_{k,k-j} \theta_{n,n-j} v_j \right]}.$$

Nakonec díky tomu, že

$$\langle \widehat{Y}_{n+1}, Y_{n+1} - \widehat{Y}_{n+1} \rangle = 0,$$

pak s využitím Pythagorovy věty dostaneme

$$\underbrace{\|Y_{n+1}\|^2}_{=\gamma(n+1,n+1)} = \left\| (Y_{n+1} - \widehat{Y}_{n+1}) + \widehat{Y}_{n+1} \right\|^2 = \underbrace{\|Y_{n+1} - \widehat{Y}_{n+1}\|^2}_{=v_n} + \|\widehat{Y}_{n+1}\|^2.$$

Tedy

$$\begin{aligned} \boxed{v_n} &= \gamma(n+1, n+1) - \langle \widehat{Y}_{n+1}, \widehat{Y}_{n+1} \rangle \\ &= \gamma(n+1, n+1) - \left\langle \widehat{Y}_{n+1}, \sum_{j=0}^{n-1} \theta_{n,n-j} (Y_{j+1} - \widehat{Y}_{j+1}) \right\rangle \\ &= \gamma(n+1, n+1) - \sum_{j=0}^{n-1} \theta_{n,n-j} \underbrace{\langle \widehat{Y}_{n+1}, Y_{j+1} - \widehat{Y}_{j+1} \rangle}_{=\theta_{n,n-j}v_j} \\ &= \boxed{\gamma(n+1, n+1) - \sum_{j=0}^{n-1} \theta_{n,n-j}^2 v_j} \end{aligned}$$

čímž jsme dokázali poslední tvrzení (1.11). ■

POZNÁMKA 1.6.2.

Zatímco **Durbin-Levinsův algoritmus** dává koeficienty $\phi_{n,j}$ v reprezentaci

$$\widehat{Y}_{n+1} = \sum_{j=1}^n \phi_{n,j} Y_{n+1-j} = \sum_{j=0}^{n-1} \phi_{n,n-j} Y_{j+1},$$

inovační algoritmus dává koeficienty $\theta_{n,j}$ v ortogonálním rozvoji

$$\widehat{Y}_{n+1} = \sum_{j=1}^n \theta_{n,j} (Y_{n+1-j} - \widehat{Y}_{n+1-j}) = \sum_{j=0}^{n-1} \theta_{n,n-j} (Y_{j+1} - \widehat{Y}_{j+1}).$$

POZNÁMKA 1.6.3.

Inovační algoritmus dává „inovační reprezentaci“ samotných Y_{n+1} , neboť platí

$$\boxed{\mathbf{Y}_n} = \mathbf{Y}_n - \widehat{\mathbf{Y}}_n + \widehat{\mathbf{Y}}_n = (\mathbf{Y}_n - \widehat{\mathbf{Y}}_n) + (\mathbf{C}_n - \mathbf{I}_n)(\mathbf{Y}_n - \widehat{\mathbf{Y}}_n) = \boxed{\mathbf{C}_n(\mathbf{Y}_n - \widehat{\mathbf{Y}}_n)}.$$

a položíme-li $\theta_{n,0} = 1$, můžeme psát

$$Y_{n+1} = \sum_{j=0}^n \theta_{n,j} (Y_{n+1-j} - \widehat{Y}_{n+1-j}) = \sum_{j=0}^n \theta_{n,n-j} (Y_{j+1} - \widehat{Y}_{j+1}).$$

Tyto vztahy využijeme později při odvozování maximálně věrohodných odhadů neznámých parametrů $\theta_{n,j}$.

1.6.1 Jednokroková nejlepší lineární predikce v $AR(p)$

Nejprve si všimněme, jaké vlastnosti má predikce v případě autoregresních procesů řádu p .

Věta 1.6.4.

Nechť $\{Y_t, t \in \mathbb{Z}\} \in L^2(\Omega, \mathcal{A}, P)$ je centrovaný nedegenerovaný kauzální $AR(p)$ proces $Y_t = \varphi_1 Y_{t-1} + \dots + \varphi_p Y_{t-p} + \varepsilon_t$. Pak pro nejlepší lineární predikci platí

$$\widehat{Y}_{n+1} = \begin{cases} 0 & n = 0 \\ \sum_{j=1}^{\min(n,p)} \varphi_j Y_{n+1-j} & n = 1, 2, \dots \end{cases}$$

DŮKAZ. Vzhledem k tomu, že autokovarianční funkce $\gamma(h)$ exponenciálně klesá k nule, stačí předpokládat, že proces není degenerovaný, tj. rozptyl $\gamma(0) > 0$. Nejlepší lineární predikce podle definice je rovna

$$\widehat{Y}_{n+1} = \widehat{Y}_{n+1|n} = \begin{cases} 0 (= \mu_Y) & n = 0, \\ P_{\overline{sp}\{Y_1, \dots, Y_n\}}(Y_{n+1}) = P_{\mathcal{M}_n}(Y_{n+1}) & n \geq 1. \end{cases}$$

Předpokládejme tedy, že $n \geq p$ a postupně upravujeme

$$\begin{aligned} \widehat{Y}_{n+1} &= P_{\overline{sp}\{Y_1, \dots, Y_n\}}(Y_{n+1}) = P_{\overline{sp}\{Y_1, \dots, Y_n\}}(\varphi_1 Y_n + \dots + \varphi_p Y_{n+1-p} + \varepsilon_{n+1}) \\ &= \sum_{j=1}^p \varphi_j P_{\overline{sp}\{Y_1, \dots, Y_n\}}(Y_{n+1-j}) + P_{\overline{sp}\{Y_1, \dots, Y_n\}}(\varepsilon_{n+1}). \end{aligned}$$

Připomeňme, že pro projekci v případě $j = 1, \dots, n$ platí

$$P_{\overline{sp}\{Y_1, \dots, Y_n\}}(Y_j) = Y_j, \quad \text{neboť} \quad Y_j \in \overline{sp}\{Y_1, \dots, Y_n\}.$$

Dále počítejme pro $j = 1, \dots, n$ skalární součin

$$\langle \varepsilon_{n+1}, Y_j \rangle \stackrel{\text{kauzal.}}{=} \langle \varepsilon_{n+1}, \sum_{k=1}^{\infty} \psi_j \varepsilon_{j-k} \rangle = \sum_{k=1}^{\infty} \psi_j \underbrace{E(\varepsilon_{n+1} \varepsilon_{j-k})}_{=0} = 0,$$

tj. $\varepsilon_{n+1} \perp Y_1, \dots, Y_n$, a $P_{\overline{sp}\{Y_1, \dots, Y_n\}}(\varepsilon_{n+1}) = 0$, takže

$$\widehat{Y}_{n+1} = \varphi_1 Y_n + \dots + \varphi_p Y_{n+1-p} \quad \text{jestliže} \quad n \geq p,$$

čímž dostáváme tvrzení věty. ■

Tedy v případě $AR(p)$ procesu

$$Y_t = \varphi_1 Y_{t-1} + \dots + \varphi_p Y_{t-p} + \varepsilon_t$$

jsou koeficienty $\phi_{n,1}, \dots, \phi_{n,1}$ nejlepší lineární predikce pro $n \geq p$ rovny

$$\begin{aligned}\phi_{n,1} &= \varphi_1 \\ &\vdots \\ \phi_{n,p} &= \varphi_p \\ \phi_{n,p+1} &= 0 \\ &\vdots \\ \phi_{n,n} &= 0\end{aligned}$$

1.6.2 Vícekroková nejlepší lineární predikce v $AR(p)$

Podle definice h -kroková predikce je definována vztahem

$$\widehat{Y}_{n+h} = \widehat{Y}_{n+h|n} = \begin{cases} 0 (= \mu_Y) & n, h = 0, \\ P_{\overline{sp}\{Y_1, \dots, Y_n\}}(Y_{n+h}) = P_{\mathcal{M}_n}(Y_{n+h}) & n, h \geq 1. \end{cases}$$

Počítejme postupně

$$\begin{aligned}\boxed{\widehat{Y}_{n+2|n}} &= P_{\overline{sp}\{Y_1, \dots, Y_n\}}(Y_{n+2}) \\ &= P_{\overline{sp}\{Y_1, \dots, Y_n\}}(\varphi_1 Y_{n+1} + \dots + \varphi_p Y_{n+2-p} + \varepsilon_{n+2}) \\ &= \sum_{j=1}^p \varphi_j P_{\overline{sp}\{Y_1, \dots, Y_n\}}(Y_{n+2-j}) + \underbrace{P_{\overline{sp}\{Y_1, \dots, Y_n\}}(\varepsilon_{n+2})}_{=0(\text{viz předchozí důkaz})} \\ &= \varphi_1 P_{\overline{sp}\{Y_1, \dots, Y_n\}}(Y_{n+1}) + \sum_{j=2}^p \varphi_j \underbrace{P_{\overline{sp}\{Y_1, \dots, Y_n\}}(Y_{n+2-j})}_{=Y_{n+2-j}} \\ &= \boxed{\varphi_1 \widehat{Y}_{n+1|n} + \varphi_2 Y_n + \dots + \varphi_p Y_{n+2-p}}\end{aligned}$$

\vdots

$$\begin{aligned}\boxed{\widehat{Y}_{n+p|n}} &= P_{\overline{sp}\{Y_1, \dots, Y_n\}}(Y_{n+p}) \\ &= P_{\overline{sp}\{Y_1, \dots, Y_n\}}(\varphi_1 Y_{n+p-1} + \dots + \varphi_p Y_n + \varepsilon_{n+p}) \\ &= \boxed{\varphi_1 \widehat{Y}_{n+p-1|n} + \dots + \varphi_{p-1} \widehat{Y}_{n+1|n} + \varphi_p Y_n}\end{aligned}$$

pro $s > p$

$$\begin{aligned}\boxed{\widehat{Y}_{n+p|n}} &= P_{\overline{sp}\{Y_1, \dots, Y_n\}}(Y_{n+p}) \\ &= P_{\overline{sp}\{Y_1, \dots, Y_n\}}(\varphi_1 Y_{n+p-1} + \dots + \varphi_p Y_n + \varepsilon_{n+p}) \\ &= \boxed{\varphi_1 \widehat{Y}_{n+s-1|n} + \dots + \varphi_p \widehat{Y}_{n+s-p|n}}\end{aligned}$$

1.6.3 PACF pro $AR(p)$, $MA(q)$ a $ARMA(p, q)$

Věta 1.6.5.

Nechť $\{Y_t, t \in \mathbb{Z}\} \in L^2(\Omega, \mathcal{A}, P)$ je centrováný nedegenerovaný kauzální $AR(p)$ proces $Y_t = \varphi_1 Y_{t-1} + \dots + \varphi_p Y_{t-p} + \varepsilon_t$. Pak platí

(1) $\alpha(p) = \varphi_p$

(2) $\alpha(k) = 0$ pro $k > p$.

DŮKAZ. Již dříve jsme ukázali, že v případě $AR(p)$ procesu

$$Y_t = \varphi_1 Y_{t-1} + \dots + \varphi_p Y_{t-p} + \varepsilon_t$$

jsou koeficienty $\phi_{n,1}, \dots, \phi_{n,n}$ nejlepší lineární predikce pro $n \geq p$ rovny

$$\begin{aligned} \phi_{n,1} &= \varphi_1 \\ &\vdots \\ \phi_{n,p} &= \varphi_p \\ \phi_{n,p+1} &= 0 \\ &\vdots \\ \phi_{n,n} &= 0 \end{aligned}$$

Tedy pokud přímo $n = p$, tak podle důsledku Durbin–Lewinsonova algoritmu platí

$$\alpha(p) = \phi_{p,p} = \varphi_p.$$

Jestliže $k > p$, pak je parciální autokorelační funkce nulová $\alpha(k) = 0$, což je velmi důležitá identifikační vlastnost $AR(p)$ procesů. ■

POZNÁMKA 1.6.6.

Parciální autokorelační koeficienty $\alpha(1), \dots, \alpha(p-1)$ lze určit jako $\phi_{1,1}, \dots, \phi_{p-1,p-1}$ z Durbin–Levinsonova algoritmu.

Důsledek 1.6.7.

Nechť $\{Y_t, t \in \mathbb{Z}\} \in L^2(\Omega, \mathcal{A}, P)$ je centrováný nedegenerovaný invertibilní $MA(q)$ (resp. $ARMA(p, q)$) proces. Pak **neexistuje** takové $k_0 \in \mathbb{N}$, že pro $k > k_0$ platí $\alpha(k) = 0$.

DŮKAZ. Využijeme toho, že proces $MA(q)$ (resp. $ARMA(p, q)$) je invertibilní. Pak existuje absolutně konvergentní posloupnost reálných čísel $\pi = \{\pi_j\}_{j=0}^{\infty}$ (tj. $\sum_{j=0}^{\infty} |\pi_j| < \infty$) taková, že

$$\varepsilon_t = \sum_{j=0}^{\infty} \pi_j Y_{t-j}, \quad \text{tj. zkráceně} \quad Y_t \sim AR(\infty) : \varepsilon_t = \pi(B)Y_t,$$

tj. $p = \infty$, takže podle předchozí věty nenajdeme $k_0 \in \mathbb{N}$ takové, že pro $k > k_0$ platí $\alpha(k) = 0$. ■

1.6.4 Jednokroková nejlepší lineární predikce v $MA(q)$

Pro jednokrokovou predikci v případě $MA(q)$ procesů je velmi užitečný inovační algoritmus. Nejprve uvedeme podrobně rekurentní vzorce pro stacionární procesy, pro které platí

$$\gamma(i, j) = \gamma(i - j).$$

Predikci pomocí inovací lze vypočítat pomocí následujícího vzorce

$$\widehat{Y}_{n+1} = \begin{cases} 0 & n = 0 \\ \theta_{n,1} (Y_n - \widehat{Y}_n) \cdots + \cdots + \theta_{n,n} (Y_1 - \widehat{Y}_1) & n = 1, 2, \dots \end{cases},$$

přičemž pro

$$n = 0 \quad v_0 = \gamma(0),$$

dále

$$\begin{aligned} \theta_{n,n} &= \frac{\gamma(n)}{v_0} \\ \theta_{n,n-1} &= \frac{\gamma(n-1)}{v_1} - \theta_{1,1} \theta_{n,n} \frac{v_0}{v_1} \\ \theta_{n,n-2} &= \frac{\gamma(n-2)}{v_2} - \theta_{2,2} \theta_{n,n} \frac{v_0}{v_2} - \theta_{2,1} \theta_{n,n-1} \frac{v_1}{v_2} \\ &\vdots \\ \theta_{n,2} &= \frac{\gamma(2)}{v_{n-2}} - \underbrace{\theta_{n-2,n-2} \theta_{n,n} \frac{v_0}{v_{n-2}} - \cdots - \theta_{n-2,1} \theta_{n,3} \frac{v_{n-3}}{v_{n-2}}}_{(n-2) \text{ členů}} \\ \theta_{n,1} &= \frac{\gamma(1)}{v_{n-1}} - \underbrace{\theta_{n-1,n-1} \theta_{n,n} \frac{v_0}{v_{n-1}} - \theta_{n-1,n-2} \theta_{n,n-1} \frac{v_1}{v_{n-1}} - \cdots - \theta_{n-1,1} \theta_{n,2} \frac{v_{n-2}}{v_{n-1}}}_{(n-1) \text{ členů}} \end{aligned}$$

a nakonec

$$v_n = \gamma(0) - \theta_{n,n}^2 v_0 - \cdots - \theta_{n,1}^2 v_{n-1}.$$

Vzhledem k tomu, že $MA(q)$ proces má autokovarianční funkci $\boxed{\gamma(k) = 0}$ pro $k > q$, pak v případě, že $n > q$ jsou koeficienty $\theta_{n,n} = 0, \dots, \theta_{n,q+1} = 0$ a teprve $\theta_{n,q} \neq 0, \dots, \theta_{n,1} \neq 0$, takže nejlepší lineární predikce je tvaru

$$\widehat{Y}_{n+1} = \begin{cases} 0 & n = 0 \\ \sum_{j=1}^{\min(n,q)} \theta_{n,j} (Y_{n+1-j} - \widehat{Y}_{n+1-j}) & n = 1, 2, \dots \end{cases}$$

1.6.5 Nejlepší lineární predikce v $ARMA(p, q)$

Nechť $\{Y_t, t \in \mathbb{Z}\}$ je kauzální a invertibilní $ARMA$ proces

$$\{Y_t, t \in \mathbb{Z}\} \sim ARMA(p, q) : \Phi(B)Y_t = \Theta(B)\varepsilon_t \quad \varepsilon_t \sim WN(0, \sigma_\varepsilon^2).$$

Z kauzality vyplývá, že existuje posloupnost $\{\psi_j\}_{j=0}^\infty$ taková, že $\sum_{j=0}^\infty |\psi_j| < \infty$ a platí

$$Y_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}, \quad \text{tj. } Y_t \sim MA(\infty),$$

takže pro $|z| \leq 1$ dostáváme

$$\Psi(z) = \frac{\Theta(z)}{\Phi(z)} \Rightarrow \Phi(z)\Psi(z) = \Theta(z).$$

Koeficienty $\{\psi_j\}_{j=0}^\infty$ se určí ze vztahu

$$(1 - \varphi_1 z - \varphi_2 z^2 - \dots - \varphi_p z^p)(\psi_0 + \psi_1 z + \psi_2 z^2 + \dots) = (1 + \theta_1 z + \theta_2 z^2 + \dots + \theta_q z^q)$$

porovnáním koeficientů u mocnin proměnné z , tj.

$$\begin{aligned} z^0 : \psi_0 &= 1 & \Rightarrow \psi_0 &= 1 \\ z^1 : \psi_1 - \varphi_1 &= \theta_1 & \Rightarrow \psi_1 &= \theta_1 + \varphi_1 \\ z^2 : \psi_2 - \varphi_1 \psi_1 - \varphi_2 &= \theta_2 & \Rightarrow \psi_2 &= \theta_2 + \varphi_1 \psi_1 + \varphi_2 \\ z^3 : \psi_3 - \varphi_2 \psi_1 - \varphi_1 \psi_2 - \varphi_3 &= \theta_3 & \Rightarrow \psi_3 &= \theta_3 + \varphi_2 \psi_1 + \varphi_1 \psi_2 + \varphi_3 \\ & & & \vdots \end{aligned}$$

Obecně, položíme-li $\theta_j = 0$ pro $j > q$ a označíme-li $m = \max(p, q)$, dostaneme

$$\begin{aligned} \psi_0 &= 1 \\ \psi_j &= \theta_j + \sum_{i=1}^{\min(j,p)} \varphi_i \psi_{j-i} \quad \text{pro } 1 \leq j \leq m \\ \psi_j &= \sum_{i=1}^p \varphi_i \psi_{j-i} \quad \text{pro } j > m \end{aligned}$$

a vidíme, že pro $j > m$ se koeficienty θ_k neprosadí. Pokud bychom použili predikci pomocí inovací, bude vždy pro $n > m$ platit

$$\hat{Y}_{n+1} = \sum_{j=1}^n \theta_{n,j} (Y_{n+1-j} - \hat{Y}_{n+1-j})$$

takže použijeme vždy n předchozích inovací a ztrácíme tak výhodu, která byla u MA procesu konečného řádu.

Nechceme-li o tuto výhodu přijít, ukázalo se, že díky jednoduché transformaci využijeme jednak možnosti použít maximálně q předchozích inovací a také toho, že díky AR části je proces lineární kombinací předchozích p hodnot.

Položme nejprve

$$m = \max(p, q)$$

a definujeme

$$W_t = \begin{cases} \sigma_\varepsilon^{-1} Y_t & 1 \leq t \leq m \\ \sigma_\varepsilon^{-1} \Phi(B) Y_t = \sigma_\varepsilon^{-1} (Y_t - \varphi_1 Y_{t-1} - \dots - \varphi_p Y_{t-p}) & t > m \end{cases}$$

tedy pro $1 \leq t \leq m$ jde o $ARMA(p, q)$ proces s jednotkovým rozptylem a pro $t > m$ jde o $MA(q)$ proces (opět s jednotkovým rozptylem).

Zkoumejme jedнокrokovou predikci tohoto transformovaného procesu. Zřejmě

$$\mathcal{M}_n = \overline{sp}\{Y_1, \dots, Y_n\} = \overline{sp}\{W_1, \dots, W_n\},$$

takže položíme-li

$$\widehat{W}_1 = 0 = \mu_Y = \mu_W \quad \text{pro} \quad n = 1,$$

pak pro $1 \leq t \leq m$

$$\widehat{W}_t = P_{\overline{sp}\{Y_1, \dots, Y_{t-1}\}}(\sigma_\varepsilon^{-1} Y_t) = \sigma_\varepsilon^{-1} \widehat{Y}_t$$

a pro $t > m$

$$\begin{aligned} \widehat{W}_t &= P_{\overline{sp}\{Y_1, \dots, Y_{t-1}\}}(\sigma_\varepsilon^{-1} \Phi(B) Y_t) \\ &= \sigma_\varepsilon^{-1} P_{\overline{sp}\{Y_1, \dots, Y_{t-1}\}}(Y_t - \varphi_1 Y_{t-1} - \dots - \varphi_p Y_{t-p}) \\ &= \sigma_\varepsilon^{-1} \left(\widehat{Y}_t - \varphi_1 \widehat{Y}_{t-1} - \dots - \varphi_p \widehat{Y}_{t-p} \right). \end{aligned}$$

Z předchozího také plyne, že

$$W_t - \widehat{W}_t = \sigma_\varepsilon^{-1} (Y_t - \widehat{Y}_t).$$

Použijeme-li inovační algoritmus na proces W_t , dostaneme

$$\widehat{W}_{n+1} = \begin{cases} \sum_{j=1}^n \theta_{n,j} (W_{n+1-j} - \widehat{W}_{n+1-j}) & 1 \leq n \leq m-1, \\ \sum_{j=1}^q \theta_{n,j} (W_{n+1-j} - \widehat{W}_{n+1-j}) & n \geq m. \end{cases}$$

Koeficienty $\theta_{n,j}$ se určí pomocí autokovarianční funkce procesu W_t (viz inovační algoritmus). Zpětnou transformací k původnímu procesu bude **nejlepší lineární**

predikce o jeden krok dopředu rovna

$$\widehat{Y}_{n+1} = \begin{cases} 0 & n = 1 \\ \sum_{j=1}^n \theta_{n,j} (Y_{n+1-j} - \widehat{Y}_{n+1-j}) & 1 \leq n \leq m-1, \\ \sum_{j=1}^p \varphi_j Y_{n+1-j} + \sum_{j=1}^q \theta_{n,j} (Y_{n+1-j} - \widehat{Y}_{n+1-j}) & n \geq m. \end{cases}$$

Při odvození predikce o $h > 1$ kroků dopředu opět vyjdeme z transformovaného procesu W_t

$$\widehat{W}_{n+h|n} = \begin{cases} P_{\overline{sp}\{Y_1, \dots, Y_n\}}(\sigma_\varepsilon^{-1} Y_{n+h}) & n+h \leq m, \\ P_{\overline{sp}\{Y_1, \dots, Y_n\}}(\sigma_\varepsilon^{-1} (Y_{n+h} - \sum_{j=1}^p \varphi_j Y_{n+h-j})) & n+h > m. \end{cases}$$

1.7 Výstavba modelů v B-J metodologii

1.7.1 Odhady v ARMA procesech

Určení vhodného $ARMA(p, q)$ modelu pro danou realizaci stacionárního procesu $\{Y_t, t \in \mathbb{Z}\}$ v sobě zahrnuje celou řadu problémů

1. výběr řádu modelu p a q , tj. provést identifikaci modelu;
2. odhad parametrů $\varphi_1, \dots, \varphi_p, \theta_1, \dots, \theta_q$ a σ_ε^2 ;
3. ověření vhodnosti modelu.

Identifikace modelu je první fází výstavby modelu a jejím úkolem je rozhodnout, jaký typ modelu vybrat (tj. zda AR, MA nebo ARMA) a explicitně určit řád modelu. Před vlastní identifikací se doporučuje provést některé z následujících přípravných operací:

1. Pořídit grafický záznam řady.
2. Pokud střední hodnota je nenulová, provést odhad střední hodnoty a následně provést centrování.

VLASTNÍ IDENTIFIKACE PROCESU

Identifikace procesu se provádí na základě zkoumání průběhu

- odhadnuté autokorelační funkce ACF, tj. $r(k) = \hat{\rho}(k)$
- a parciální autokorelační funkce PACF, tj. $a(k) = \hat{\alpha}(k)$.

Snažíme se především zjistit existenci případného identifikačního bodu k_0 . Je však nutné mít na paměti, že pracujeme pouze s odhadnutými hodnotami $r(k)$ a $a(k)$, takže naše závěry mohou být někdy dost zkreslené. Doporučuje se proto netrvat na jednoznačné identifikaci určitého modelu, ale přijmout a přezkoušet několik alternativ.

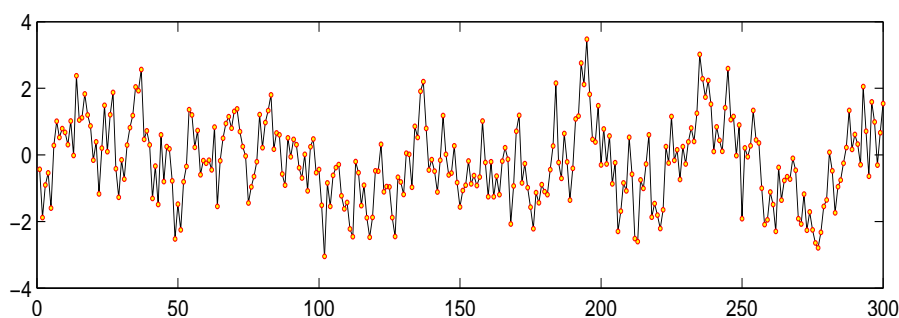
	$AR(p)$	$MA(q)$	$ARMA(p, q)$
ACF $\rho(k)$	neexistuje k_0 ve tvaru „U“	$k_0 = q$	neexistuje k_0 po $q - p$ hodnotách je ve tvaru „U“
PACF $\alpha(k)$	$k_0 = p$	neexistuje k_0 omezena křivkou „U“	neexistuje k_0 po $p - q$ hodnotách je omezena křivkou „U“

Identifikační bod k_0 je index, pro nějž platí: $\forall k > k_0$ je $\rho(k) = 0$ (resp. $\alpha(k) = 0$).

Křivka „U“ je křivka ve tvaru lineární kombinace klesajících geometrických posloupností a sinusoid s geometricky klesající amplitudou.

Na následujícím obrázku vidíme simulovaná data kauzálního $AR(2)$ procesu

$$Y_t = 0.5Y_{t-1} + 0.2Y_{t-2} + \epsilon_t, \quad \epsilon_t \sim N(0, 1),$$



Obrázek 1.8: Simulovaná data pro kauzální $AR(2)$ proces.

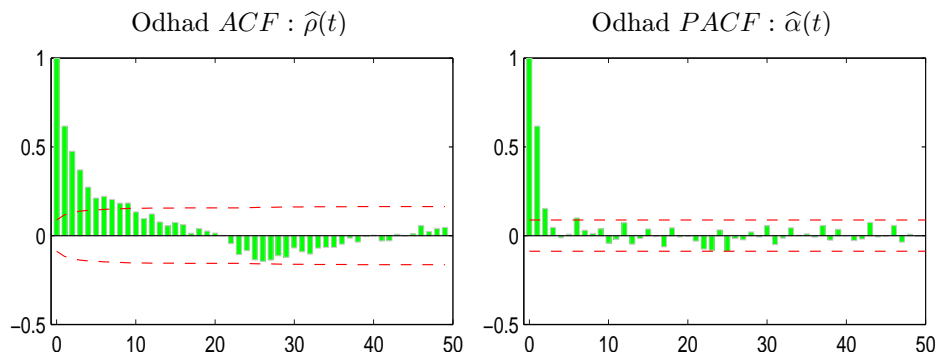
jehož kořeny polynomu

$$\Phi(z) = 1 - \varphi_1 z - \varphi_2 z^2 = 1 - 0.5z - 0.2z^2$$

jsou rovny

$$z_{01} = 3.81 \quad \text{a} \quad z_{02} = 1.31.$$

Na následujících dvou grafech jsou vykresleny odhady autokorelační ACF a parciální autokorelační funkce $PACF$ této konkrétní realizace.

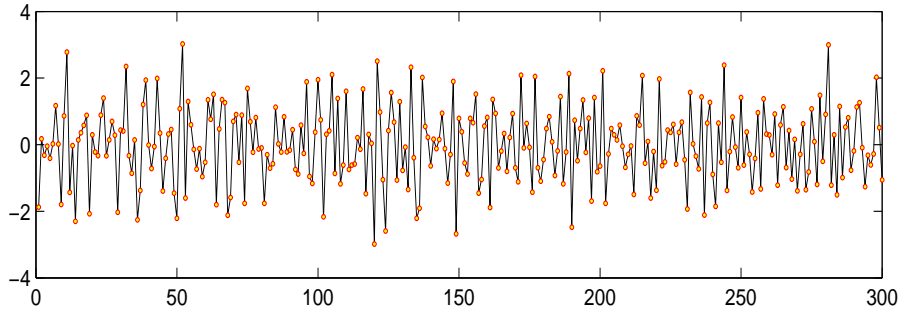


Obrázek 1.9: Odhady ACF a $PACF$ pro simulovaná data z $AR(2)$ procesu.

Vidíme, že odhad autokorelační funkce $\hat{\rho}(t)$ exponenciálně klesá k nule pro $t \rightarrow \infty$ a odhady parciální autokorelační funkce $\hat{\alpha}(t)$ pro $t > k_0 = 2$ se výrazně neliší od nuly.

Naproti tomu, pro simulovaná data z invertibilního $MA(2)$ procesu

$$Y_t = \epsilon_t - 0.5\epsilon_{t-1} - 0.2\epsilon_{t-2}, \quad \epsilon_t \sim N(0, 1),$$



Obrázek 1.10: Simulovaná data pro invertibilní $MA(2)$ proces.

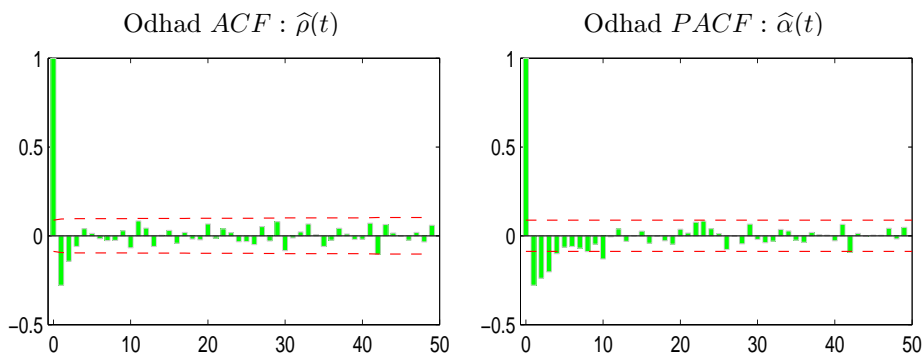
s kořeny polynomu

$$\Theta(z) = 1 + \theta_1 z + \theta_2 z^2 = 1 - 0.5z - 0.2z^2$$

rovnými hodnotám

$$z_{01} = 3.81 \quad \text{a} \quad z_{02} = 1.31,$$

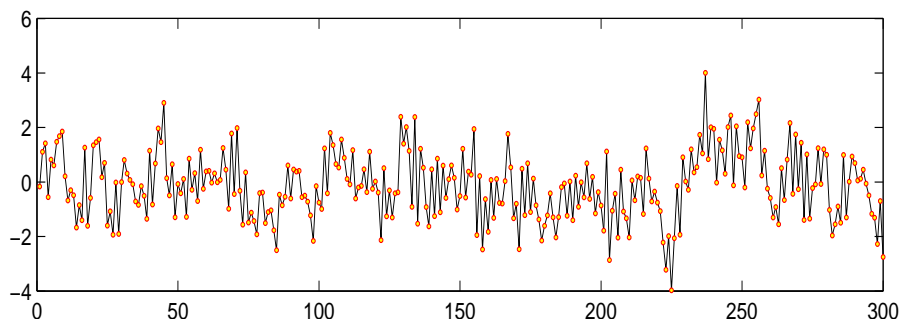
vidíme na odhadech ACF a PACF zobrazených na následujících dvou obrázcích, že naopak odhad parciální autokorelační funkce $\hat{\alpha}(t)$ exponenciálně klesá k nule pro $t \rightarrow \infty$ a odhady autokorelační funkce $\hat{\rho}(t)$ pro $t > k_0 = 2$ se výrazně neliší od nuly.



Obrázek 1.11: Odhady ACF a PACF pro simulovaná data z $MA(2)$ procesu.

Na závěr si všimneme simulovaných dat z kauzálního a invertibilního $ARMA(2, 2)$ procesu

$$Y_t = 0.5Y_{t-1} + 0.2Y_{t-2} + \epsilon_t - 0.4\epsilon_{t-1} + 0.3\epsilon_{t-2}, \quad \epsilon_t \sim N(0, 1),$$



Obrázek 1.12: Simulovaná data pro invertibilní $ARMA(2, 2)$ proces.

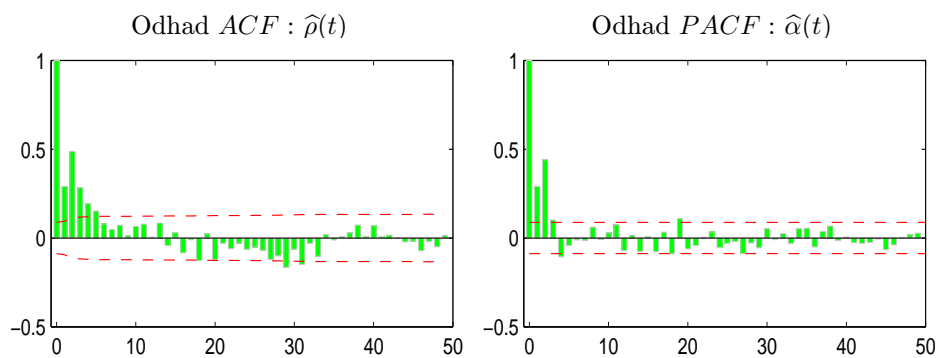
jehož všechny kořeny polynomů

$$\Phi(z) = 1 - \varphi_1 z - \varphi_2 z^2 = 1 - 0.5z - 0.2z^2 \quad \text{a} \quad \Theta(z) = 1 + \theta_1 z + \theta_2 z^2 = 1 - 0.4z + 0.3z^2$$

leží vně jednotkové kružnice, neboť

$$AR: z_{01} = 3.81 \quad \text{a} \quad z_{02} = 1.31 \quad \quad MA: z_{01} = 1.83 \quad \text{a} \quad z_{02} = 1.83.$$

Podíváme-li se na grafy odhadů autokorelační a parciální autokorelační funkce, vidíme že obě funkce exponenciálně klesají k nule pro $t \rightarrow \infty$. V tomto případě nelze určit identifikační bod k_0 , od kterého by se některá z korelačních funkcí již výrazně nelišila od nuly.



Obrázek 1.13: Odhady ACF a PACF pro simulovaná data z $ARMA(2, 2)$ procesu.

1.7.2 Yuleovy-Walkerovy rovnice a odhad parametrů v $AR(p)$

Nechť $\{Y_t, t \in \mathbb{Z}\}$ je centrováný kauzální autoregresní proces

$$\boxed{AR(p) : \Phi(B)Y_t = \varepsilon_t} \quad \varepsilon_t \sim WN(0, \sigma_\varepsilon^2).$$

Vraťme se k Yuleovým-Walkerovým rovnicím

$$\underbrace{\rho(0)}_{=1} = \varphi_1\rho(1) + \dots + \varphi_p\rho(p) + \frac{\sigma_\varepsilon^2}{\gamma(0)} \Rightarrow \sigma_\varepsilon^2 = \gamma(0) [1 - \varphi_1\rho(1) - \dots - \varphi_p\rho(p)]$$

$$\rho(k) = \varphi_1\rho(k-1) + \dots + \varphi_p\rho(k-p) \quad k \geq 1$$

Označíme-li

$$\begin{aligned} \widehat{\mathbf{R}}_p &= (\widehat{\rho}(i-j))_{i,j=1}^p & \widehat{\boldsymbol{\rho}}_p &= (\widehat{\rho}(1), \dots, \widehat{\rho}(p))' \\ \boldsymbol{\phi}_p &= (\varphi_1, \dots, \varphi_p)' & \widehat{\boldsymbol{\phi}}_p &= (\widehat{\varphi}_1, \dots, \widehat{\varphi}_p)' \end{aligned}$$

a v Yuleových-Walkerových rovnicích nahradíme $\rho(k)$ odpovídajícími odhady $\widehat{\rho}(k)$, pak (pokud $\widehat{\gamma}(0) > 0$) dostaneme tzv. **Yuleovy-Walkerovy odhady**:

$$\boxed{\begin{aligned} \widehat{\boldsymbol{\phi}}_p &= \widehat{\mathbf{R}}_p^{-1} \widehat{\boldsymbol{\rho}}_p \\ \widehat{\sigma}_\varepsilon^2 &= \widehat{\gamma}(0) \left(1 - \widehat{\boldsymbol{\rho}}_p' \widehat{\mathbf{R}}_p^{-1} \widehat{\boldsymbol{\rho}}_p\right) \end{aligned}}$$

Věta 1.7.1.

Nechť $\{Y_t, t \in \mathbb{Z}\}$ je centrováný kauzální autoregresní proces $AR(p) : \Phi(B)Y_t = \varepsilon_t$, kde $\varepsilon_t \sim IID(0, \sigma_\varepsilon^2)$ a $\widehat{\boldsymbol{\phi}}_p$ je Yuleovův-Walkerův odhad $\boldsymbol{\phi}_p = (\varphi_1, \dots, \varphi_p)'$, pak platí

$$\sqrt{n} \left(\widehat{\boldsymbol{\phi}}_p - \boldsymbol{\phi}_p \right) \overset{A}{\rightsquigarrow} N_p(\mathbf{0}, \sigma_\varepsilon^2 \boldsymbol{\Gamma}_p^{-1}),$$

kde $\boldsymbol{\Gamma}_p = (\gamma(i-j))_{i,j=1}^p$. Kromě toho platí

$$\widehat{\sigma}_\varepsilon^2 \xrightarrow{P} \sigma_\varepsilon^2.$$

DŮKAZ. viz Brockwell, Davis (1991), str. 255-257. ■

Z předchozích tvrzení plyne, že odhady získané řešením Yuleových-Walkerových rovnic jsou asymptoticky nestranné a lze pro ně konstruovat asymptotické intervaly spolehlivosti.

V praktických situacích však skutečný řád p autoregresního procesu neznáme. V tom případě se využijí tvrzení následující věty.

Věta 1.7.2.

Nechť $\{Y_t, t \in \mathbb{Z}\}$ je centrovaný kauzální autoregresní proces $AR(p) : \Phi(B)Y_t = \varepsilon_t$, kde $\varepsilon_t \sim IID(0, \sigma_\varepsilon^2)$ a

$$\hat{\phi}_m = \left(\hat{\phi}_{m,1}, \dots, \hat{\phi}_{m,m} \right)' = \hat{\mathbf{R}}_m^{-1} \hat{\rho}_m, \quad m > p,$$

pak platí

$$\sqrt{n} \left(\hat{\phi}_m - \phi_m \right) \stackrel{A}{\sim} N_m \left(\mathbf{0}, \sigma_\varepsilon^2 \mathbf{\Gamma}_m^{-1} \right),$$

kde $\mathbf{\Gamma}_m = (\gamma(i-j))_{i,j=1}^m$, ϕ_m jsou koeficienty nejlepší lineární predikce $\phi_m \mathbf{Y}_m = \mathcal{P}_{sp\{Y_m, \dots, Y_1\}} Y_{m+1}$, přičemž $\mathbf{Y}_m = (Y_m, \dots, Y_1)'$, tj. $\phi_m = \mathbf{R}_m^{-1} \rho_m$, přičemž $\mathbf{R}_m = (\rho(i-j))_{ij=1}^m$. Speciálně pro $m > p$ platí

$$\sqrt{n} \phi_{m,m} \stackrel{A}{\sim} N(0, 1).$$

DŮKAZ. viz Brockwell, Davis (1991), str. 255-257. ■

1.7.3 Předběžné odhady v $AR(p)$ a Durbin-Levinsův algoritmus.

Předpokládejme, že máme k dispozici pozorování

$$y_1, \dots, y_n$$

centrované stacionární posloupnosti

$$\{Y_t, t \in \mathbb{Z}\} \sim AR(m) : \Phi(B)Y_t = \varepsilon_t \quad \varepsilon_t \sim WN(0, \sigma_\varepsilon^2).$$

Za předpokladu, že $\hat{\gamma}(0) > 0$, pak můžeme odhadnout neznámé parametry autoregresního modelu řádu $m < n$ pomocí Yuleových-Walkerových rovnic. Odhadnutý $AR(m)$ proces je tvaru

$$Y_t - \hat{\phi}_{m,1}Y_{t-1} - \dots - \hat{\phi}_{m,m}Y_{t-m} = \varepsilon_t \quad \varepsilon_t \sim WN(0, \hat{v}_m),$$

kde

$$\begin{aligned} \hat{\phi}_m &= \left(\hat{\phi}_{m,1}, \dots, \hat{\phi}_{m,m} \right)' = \hat{\mathbf{R}}_m^{-1} \hat{\rho}_m \\ \hat{v}_m &= \hat{\gamma}(0) \left(1 - \hat{\rho}_m' \hat{\mathbf{R}}_m^{-1} \hat{\rho}_m \right). \end{aligned}$$

Jestliže $\hat{\gamma}(0) > 0$, pak $\mathbf{R}_1, \mathbf{R}_2, \dots$ nejsou singulární a můžeme využít Durbin-Levinsův algoritmus pro postupné odhady autoregresních koeficientů $\hat{\phi}_1, \hat{\phi}_2$ a odhady variability bílého šumu $\hat{v}_1, \hat{v}_2, \dots$.

Věta 1.7.3 (Durbin-Levinsův algoritmus).

Jestliže $\hat{\gamma}(0) > 0$, pak parametry $\hat{\phi}_{m,1}, \dots, \hat{\phi}_{m,m}$ a \hat{v}_m autoregresního modelu

$$Y_t - \hat{\phi}_{m,1}Y_{t-1} - \dots - \hat{\phi}_{m,m}Y_{t-m} = \varepsilon_t \quad \varepsilon_t \sim WN(0, \hat{v}_m),$$

pro $m = 1, \dots, n-1$ lze získat rekurzivně ze vztahů

$$\hat{\phi}_{1,1} = \frac{\hat{\gamma}(1)}{\hat{\gamma}(0)} = \hat{\rho}(1) \quad v_0 = \hat{\gamma}(0) \quad (1.12)$$

$$\hat{\phi}_{m,m} = \left[\hat{\gamma}(m) - \hat{\phi}'_{m-1} \hat{\gamma}_{m-1} \right] / \hat{v}_{m-1} \quad (1.13)$$

$$\hat{\phi}_m^{(1)} = \hat{\phi}_{m-1} - \hat{\phi}_{m,m} \hat{\phi}_{m-1}^* \quad \hat{v}_m = \hat{v}_{m-1} \left(1 - \hat{\phi}_{m,m}^2 \right) \quad (1.14)$$

kde

$$\begin{aligned} \hat{\phi}_{m-1} &= (\hat{\phi}_{m-1,1}, \dots, \hat{\phi}_{m-1,m-1})' & \hat{\phi}_{m-1}^* &= (\hat{\phi}_{m-1,m-1}, \dots, \hat{\phi}_{m-1,1})' \\ \hat{\phi}_m &= (\hat{\phi}_{m,1}, \dots, \hat{\phi}_{m,m-1}, \hat{\phi}_{m,m})' & \hat{\phi}_m^{(1)} &= (\hat{\phi}_{m,1}, \dots, \hat{\phi}_{m,m-1})' \end{aligned}$$

1.7.4 Předběžné odhady v $MA(q)$ a inovační algoritmus.

Jestliže chceme na základě pozorování y_1, \dots, y_n centrované stacionární posloupnosti provést odhad $MA(m)$ ($m = 1, 2, \dots, n-1$) ve tvaru

$$Y_t = \varepsilon_t + \hat{\theta}_{m,1}\varepsilon_{t-1} + \dots + \hat{\theta}_{m,m}\varepsilon_{t-m} \quad \varepsilon_t \sim WN(0, \hat{v}_m),$$

můžeme využít inovační algoritmus.

Věta 1.7.4.

Jestliže $\hat{\gamma}(0) > 0$, pak odhady parametrů MA procesů lze provést pomocí následujících rekurentních vztahů

$$\begin{aligned} \hat{v}_0 &= \hat{\gamma}(0) \\ \hat{\theta}_{m,m-k} &= \hat{v}_k^{-1} \left[\hat{\gamma}(m-k) - \sum_{j=0}^{k-1} \hat{\theta}_{m,k-j} \hat{\theta}_{m,m-j} \hat{v}_j \right] \quad k = 0, \dots, m-1 \\ \hat{v}_m &= \hat{\gamma}(0) - \sum_{j=0}^{m-1} \hat{\theta}_{m,m-j}^2 \hat{v}_j \end{aligned}$$

Označme

$$\widehat{\boldsymbol{\theta}}_m = \left(\widehat{\theta}_{m,1}, \dots, \widehat{\theta}_{m,m} \right)'$$

Pak platí věta

Věta 1.7.5.

Nechť $\{Y_t, t \in \mathbb{Z}\}$ je kauzální a invertibilní ARMA proces $\Phi(B)Y_t = \Theta(B)\varepsilon_t$, $\varepsilon_t \sim IID(0, \sigma_\varepsilon^2)$, $E\varepsilon_t^4 < \infty$ a $\psi(z) = \sum_{j=0}^{\infty} \Psi_j z^j = \frac{\Theta(z)}{\Phi(z)}$, $|z| \leq 1$. Pak pro libovolnou posloupnost kladných celých čísel $\{m(n)\}_{n=1}^{\infty}$ takovou, že $m < n$, $m \rightarrow \infty$ a $m = o\left(n^{\frac{1}{3}}\right)$, když $n \rightarrow \infty$, pro každé k platí

$$\sqrt{n} \left(\widehat{\theta}_{m,1} - \psi_1, \widehat{\theta}_{m,2} - \psi_2, \dots, \widehat{\theta}_{m,k} - \psi_k \right)' \stackrel{A}{\sim} N_k(\mathbf{0}, \mathbf{A}),$$

kde $\mathbf{A} = (a_{ij})_{i,j=1}^k$ a

$$a_{ij} = \sum_{r=1}^{\min(i,j)} \psi_{i-r} \psi_{j-r}.$$

Kromě toho platí

$$\widehat{v}_m \xrightarrow{P} \sigma_\varepsilon^2.$$

DŮKAZ. viz Brockwell, Davis (1987), str. 239. ■

I když rekurentní odhady koeficientů MA procesů pomocí inovačního algoritmu jsou analogické jako rekurentní odhady koeficientů AR procesů pomocí Durbin-Levinsonova algoritmu, je mezi nimi přece jen jistý rozdíl. Pro odhady $\widehat{\boldsymbol{\phi}}_p = (\widehat{\phi}_{p,1}, \dots, \widehat{\phi}_{p,p})'$ pomocí Durbin-Levinsonova algoritmu platí

$$\widehat{\boldsymbol{\phi}}_p \xrightarrow{P} \boldsymbol{\phi}_p,$$

avšak odhady $\widehat{\boldsymbol{\theta}}_q = (\widehat{\theta}_{q,1}, \dots, \widehat{\theta}_{q,q})'$ nekonvergují podle pravděpodobnosti k $\boldsymbol{\theta}_q$. Ke konvergenci podle pravděpodobnosti je třeba odhad $(\widehat{\theta}_{m,1}, \dots, \widehat{\theta}_{m,q})'$, kde posloupnost $\{m(n)\}_{n=1}^{\infty}$ splňuje podmínky předchozí věty. Výběr m (maximálně až do $\frac{n}{4}$) pro výběr pevné délky se volí tak, aby odhady $(\widehat{\theta}_{m,1}, \dots, \widehat{\theta}_{m,q})'$ se stabilizovaly.

1.7.5 Předběžné odhady v ARMA(p, q) procesu.

Nechť $\{Y_t, t \in \mathbb{Z}\}$ je kauzální a invertibilní ARMA proces

$$\{Y_t, t \in \mathbb{Z}\} \sim ARMA(p, q) : \Phi(B)Y_t = \Theta(B)\varepsilon_t \quad \varepsilon_t \sim WN(0, \sigma_\varepsilon^2).$$

Z kauzality vyplývá, že existuje posloupnost $\{\psi_j\}_{j=0}^{\infty}$ taková, že $\sum_{j=0}^{\infty} |\psi_j| < \infty$ a platí $Y_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$, tj. pro $|z| \leq 1$ dostáváme

$$\Psi(z) = \frac{\Theta(z)}{\Phi(z)} \Rightarrow \Phi(z)\Psi(z) = \Theta(z).$$

Koeficienty $\{\psi_j\}_{j=0}^{\infty}$ se určí ze vztahu

$$(1 - \varphi_1 z - \varphi_1 z^2 - \dots - \varphi_p z^p)(\psi_0 + \psi_1 z + \psi_2 z^2 + \dots) = (1 + \theta_1 z + \theta_2 z^2 + \dots + \theta_q z^q)$$

porovnáním koeficientů u mocnin proměnné z , tj.

$$\begin{aligned} z^0: \quad \psi_0 &= 1 & \Rightarrow \quad \psi_0 &= 1 \\ z^1: \quad \psi_1 - \varphi_1 &= \theta_1 & \Rightarrow \quad \psi_1 &= \theta_1 + \varphi_1 \\ z^2: \quad \psi_2 - \varphi_1 \psi_1 - \varphi_2 &= \theta_2 & \Rightarrow \quad \psi_2 &= \theta_2 + \varphi_1 \psi_1 + \varphi_2 \\ z^3: \quad \psi_3 - \varphi_2 \psi_1 - \varphi_1 \psi_2 - \varphi_3 &= \theta_3 & \Rightarrow \quad \psi_3 &= \theta_3 + \varphi_2 \psi_1 + \varphi_1 \psi_2 + \varphi_3 \\ & & & \vdots \end{aligned}$$

Obecně, položíme-li $\theta_j = 0$ pro $j > q$ dostaneme
 $\varphi_j = 0$ $j > p$

$$\boxed{\begin{aligned} \psi_0 &= 1 \\ \psi_j &= \theta_j + \sum_{i=1}^{\min(j,p)} \varphi_i \psi_{j-i} \end{aligned}} \quad j = 1, 2, \dots$$

Za předběžné odhady $\psi_1, \psi_2, \dots, \psi_{p+q}$ použijeme inovační odhady $\hat{\theta}_{m,1}, \dots, \hat{\theta}_{m,p+q}$, jejichž asymptotické vlastnosti dává předchozí věta. Takže dostáváme

$$\boxed{\hat{\sigma}_\varepsilon^2 = \hat{v}_m}$$

a

$$\boxed{\hat{\theta}_{m,j} = \theta_j + \sum_{i=1}^{\min(j,p)} \varphi_i \hat{\theta}_{m,j-i}} \quad j = 1, 2, \dots, p+q.$$

Nejprve uvažujeme rovnice pro $j = q+1, \dots, p+q$

$$\begin{pmatrix} \hat{\theta}_{m,q+1} \\ \hat{\theta}_{m,q+2} \\ \vdots \\ \hat{\theta}_{m,q+p-1} \\ \hat{\theta}_{m,q+p} \end{pmatrix} = \begin{pmatrix} \hat{\theta}_{m,q} & \hat{\theta}_{m,q-1} & \cdots & \cdots & \hat{\theta}_{m,q+1-p} \\ \hat{\theta}_{m,q+1} & \hat{\theta}_{m,q} & \hat{\theta}_{m,q-1} & \cdots & \hat{\theta}_{m,q+2-p} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \hat{\theta}_{m,q+p-2} & \cdots & \hat{\theta}_{m,q+1} & \hat{\theta}_{m,q} & \hat{\theta}_{m,q-1} \\ \hat{\theta}_{m,q+p-1} & \cdots & \cdots & \hat{\theta}_{m,q+1} & \hat{\theta}_{m,q} \end{pmatrix} \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_{p-1} \\ \varphi_p \end{pmatrix}.$$

Řešením těchto rovnic dostaneme odhady $\hat{\varphi}_1, \dots, \hat{\varphi}_p$. Nakonec získáme odhady $\hat{\theta}_1, \dots, \hat{\theta}_q$ ze vztahů

$$\boxed{\hat{\theta}_j = \hat{\theta}_{m,j} - \sum_{i=1}^{\min(j,p)} \hat{\varphi}_i \hat{\theta}_{m,j-i}} \quad j = 1, \dots, q.$$

POZNÁMKA 1.7.6

Pro $MA(q)$ platí $\boxed{\hat{\theta}_j = \hat{\theta}_{m,j}}$, neboť $p = 0$.

1.7.6 Maximálně věrohodné odhady.

Předpokládejme, že $\{Y_t, t \in \mathbb{Z}\}$ je **Gaussovský proces** s nulovou střední hodnotou a kovarianční funkcí

$$\gamma(i, j) = EX_i X_j.$$

Označme

$$\mathbf{Y}_n = (Y_1, \dots, Y_n)' \quad \text{a} \quad \mathbf{\Gamma}_n = (\gamma(i, j))_{i, j=1}^n.$$

Věrohodnostní funkce náhodného vektoru \mathbf{Y}_n je tvaru

$$L(\mathbf{\Gamma}_n) = (2\pi)^{-\frac{n}{2}} |\mathbf{\Gamma}_n|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2} \mathbf{Y}_n' \mathbf{\Gamma}_n^{-1} \mathbf{Y}_n\right\}.$$

Dále označme

$$\mathcal{M}_n = \overline{sp}\{Y_n, \dots, Y_1\}$$

a

$$\widehat{Y}_n = \begin{cases} 0 (= \mu_Y) & \text{pro } n = 1 \\ P_{\mathcal{M}_{n-1}}(Y_n) & \text{pro } n = 2, 3, \dots \end{cases}$$

pak zřejmě

$$\mathcal{M}_n = \overline{sp}\{Y_n - \widehat{Y}_n, \dots, Y_1 - \widehat{Y}_1\} \quad n \geq 1.$$

Pro nejlepší lineární predikce použijme inovační algoritmus, podle kterého

$$\widehat{Y}_{n+1} = \begin{cases} 0 & n = 0 \\ \sum_{j=1}^n \theta_{n,j} (Y_{n+1-j} - \widehat{Y}_{n+1-j}) & n = 1, 2, \dots \end{cases}$$

příčemž střední kvadratickou chybu označme

$$v_n = \left\| Y_{n+1} - \widehat{Y}_{n+1} \right\|^2.$$

Označíme-li

$$\widehat{\mathbf{Y}}_n = (\widehat{Y}_1, \dots, \widehat{Y}_n)'$$

a

$$\mathbf{C}_n = \begin{pmatrix} 1 & 0 & \cdots & \cdots & \cdots & 0 \\ \theta_{1,1} & 1 & 0 & \cdots & \cdots & 0 \\ \theta_{2,2} & \theta_{2,1} & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \theta_{n-2,n-2} & \theta_{n-2,n-3} & \cdots & -\theta_{n-2,1} & 1 & 0 \\ \theta_{n-1,n-1} & \theta_{n-1,n-2} & \cdots & \cdots & \theta_{n-1,1} & 1 \end{pmatrix},$$

pak můžeme psát

$$\boxed{\widehat{\mathbf{Y}}_n = (\mathbf{C}_n - \mathbf{I}_n)(\mathbf{Y}_n - \widehat{\mathbf{Y}}_n)}.$$

Postupně upravujeme

$$\boxed{\mathbf{Y}_n} = \mathbf{Y}_n - \widehat{\mathbf{Y}}_n + \widehat{\mathbf{Y}}_n = (\mathbf{Y}_n - \widehat{\mathbf{Y}}_n) + (\mathbf{C}_n - \mathbf{I}_n)(\mathbf{Y}_n - \widehat{\mathbf{Y}}_n) = \boxed{\mathbf{C}_n(\mathbf{Y}_n - \widehat{\mathbf{Y}}_n)}.$$

Tento výsledek použijeme při vyjádření varianční matice

$$\begin{aligned} \Gamma_n &= E\mathbf{Y}_n\mathbf{Y}_n' = E\left[\mathbf{C}_n(\mathbf{Y}_n - \widehat{\mathbf{Y}}_n)(\mathbf{Y}_n - \widehat{\mathbf{Y}}_n)'\mathbf{C}_n'\right] \\ &= \mathbf{C}_n E\left[(\mathbf{Y}_n - \widehat{\mathbf{Y}}_n)(\mathbf{Y}_n - \widehat{\mathbf{Y}}_n)'\right]\mathbf{C}_n' \end{aligned}$$

Nyní počítejme

$$\begin{aligned} &E\left[(\mathbf{Y}_n - \widehat{\mathbf{Y}}_n)(\mathbf{Y}_n - \widehat{\mathbf{Y}}_n)'\right] \\ &= \begin{pmatrix} \underbrace{E(Y_1 - \widehat{Y}_1)^2}_{v_0} & \underbrace{E(Y_1 - \widehat{Y}_1)(Y_2 - \widehat{Y}_2)}_{=0} & \cdots & \underbrace{E(Y_1 - \widehat{Y}_1)(Y_n - \widehat{Y}_n)}_{=0} \\ \underbrace{E(Y_2 - \widehat{Y}_2)(Y_1 - \widehat{Y}_1)}_{=0} & \underbrace{E(Y_2 - \widehat{Y}_2)^2}_{v_1} & \cdots & \underbrace{E(Y_2 - \widehat{Y}_2)(Y_n - \widehat{Y}_n)}_{=0} \\ \vdots & \vdots & \ddots & \vdots \\ \underbrace{E(Y_n - \widehat{Y}_n)(Y_1 - \widehat{Y}_1)}_{=0} & \cdots & \underbrace{E(Y_n - \widehat{Y}_n)(Y_{n-1} - \widehat{Y}_{n-1})}_{=0} & \underbrace{E(Y_n - \widehat{Y}_n)^2}_{v_{n-1}} \end{pmatrix} \\ &= \text{diag}\{v_0, \dots, v_{n-1}\} = \mathbf{D}_n. \end{aligned}$$

Takže

$$\boxed{\Gamma_n = \mathbf{C}_n \mathbf{D}_n \mathbf{C}_n'}.$$

Počítejme dále

$$\begin{aligned} \boxed{\mathbf{Y}_n' \Gamma_n^{-1} \mathbf{Y}_n} &= (\mathbf{Y}_n - \widehat{\mathbf{Y}}_n)' \mathbf{C}_n' [\mathbf{C}_n \mathbf{D}_n \mathbf{C}_n']^{-1} \mathbf{C}_n (\mathbf{Y}_n - \widehat{\mathbf{Y}}_n) \\ &= (\mathbf{Y}_n - \widehat{\mathbf{Y}}_n)' \mathbf{C}_n' (\mathbf{C}_n')^{-1} \mathbf{D}_n^{-1} \mathbf{C}_n^{-1} \mathbf{C}_n (\mathbf{Y}_n - \widehat{\mathbf{Y}}_n) \\ &= (\mathbf{Y}_n - \widehat{\mathbf{Y}}_n)' \mathbf{D}_n^{-1} (\mathbf{Y}_n - \widehat{\mathbf{Y}}_n) \\ &= \boxed{\sum_{j=1}^n \frac{(Y_j - \widehat{Y}_j)^2}{v_{j-1}}}. \end{aligned}$$

Dále zřejmě platí

$$\boxed{|\Gamma_n|} = |\mathbf{C}_n \mathbf{D}_n \mathbf{C}_n'| = \underbrace{|\mathbf{C}_n|}_{=1} |\mathbf{D}_n| \underbrace{|\mathbf{C}_n'|}_{=1} = \boxed{v_0 v_1 \cdots v_{n-1}}.$$

Takže věrohodnostní funkce náhodného vektoru \mathbf{Y}_n je tvaru

$$\boxed{L(\Gamma_n) = (2\pi)^{-\frac{n}{2}} (v_0 v_1 \cdots v_{n-1})^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \sum_{j=1}^n \frac{(Y_j - \widehat{Y}_j)^2}{v_{j-1}} \right\}}.$$

Nechť $\{Y_t, t \in \mathbb{Z}\}$ je kauzální a invertibilní *ARMA* proces

$$\{Y_t, t \in \mathbb{Z}\} \sim ARMA(p, q) : \Phi(B)Y_t = \Theta(B)\varepsilon_t \quad \varepsilon_t \sim N(0, \sigma_\varepsilon^2).$$

Ukazuje se (viz Brockwell, Davis (1991), str. 168/170), že k velkému zjednodušení jednokrokové predikce dojde, pokud inovační algoritmus aplikujeme ne na Y_t , ale na následující transformovaný proces

$$W_t = \begin{cases} \sigma_\varepsilon^{-1}Y_t & t = 1, \dots, m; \quad m = \max(p, q) \\ \Phi(B)Y_t & t > m. \end{cases}$$

Poznamenejme nejprve, že zřejmě

$$\overline{sp}\{Y_1, \dots, Y_n\} = \overline{sp}\{W_1, \dots, W_n\} \quad n \geq 1.$$

Označme

$$\widehat{W}_{j+1} = \begin{cases} 0 = \widehat{Y}_1 & j = 0, \\ P_{\overline{sp}\{W_1, \dots, W_j\}} W_{j+1} & j \geq 1. \end{cases}$$

Pak platí

$$\widehat{W}_t = \begin{cases} \sigma_\varepsilon^{-1}\widehat{Y}_t & t = 1, \dots, m; \quad m = \max(p, q), \\ \sigma_\varepsilon^{-1}[\widehat{Y}_t - \varphi_1 Y_{t-1} - \dots - \varphi_p Y_{t-p}] & t > m, \end{cases}$$

takže

$$Y_t - \widehat{Y}_t = \sigma_\varepsilon(W_t - \widehat{W}_t).$$

Při aplikaci inovačního algoritmu na W_t dostaneme $\theta_{n,j}$ a střední kvadratické chyby, které označme r_j .

Pak z předchozích vztahů vyplývá, že platí

$$\widehat{Y}_{n+1} = \begin{cases} \sum_{j=1}^n \theta_{n,j}(Y_{n+1-j} - \widehat{Y}_{n+1-j}) & 1 \leq n < m, \\ \varphi_1 Y_n + \dots + \varphi_p Y_{n+1-p} + \sum_{j=1}^q \theta_{n,j}(Y_{n+1-j} - \widehat{Y}_{n+1-j}) & n \geq m. \end{cases}$$

a

$$\boxed{v_n} = E(Y_{n+1} - \widehat{Y}_{n+1})^2 = \sigma_\varepsilon^2 E(W_{n+1} - \widehat{W}_{n+1})^2 = \boxed{\sigma_\varepsilon^2 r_n}.$$

Takže věrohodnostní funkce náhodného vektoru \mathbf{Y}_n je tvaru

$$\boxed{L(\phi, \theta, \sigma_\varepsilon^2) = (2\pi\sigma_\varepsilon^2)^{-\frac{n}{2}} (r_0 r_1 \dots r_{n-1})^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2\sigma_\varepsilon^2} \sum_{j=1}^n \frac{(Y_j - \widehat{Y}_j)^2}{r_{j-1}} \right\}}.$$

Pokud položíme

$$\frac{\partial \ln L}{\partial \sigma_\varepsilon^2} = 0,$$

a budeme předpokládat, že \hat{Y}_j a r_j jsou nezávislé na σ_ε^2 , dostaneme

$$\boxed{\hat{\sigma}_\varepsilon^2 = \frac{1}{n} S(\hat{\phi}, \hat{\theta})},$$

kde

$$S(\hat{\phi}, \hat{\theta}) = \sum_{j=1}^n \frac{(Y_j - \hat{Y}_j)^2}{r_{j-1}}$$

a $\hat{\phi}$ a $\hat{\theta}$ jsou hodnoty, které minimalizují tzv. **redukovaný logaritmus věrohodnostní funkce**

$$\boxed{l(\hat{\phi}, \hat{\theta}) = \ln \left(\frac{1}{n} S(\hat{\phi}, \hat{\theta}) \right) + \frac{1}{n} \sum_{j=1}^n \ln r_{j-1}}.$$

POZNÁMKA 1.7.7.

Alternativou k maximalizaci $L(\phi, \theta, \sigma_\varepsilon^2)$ je **minimalizace váženého součtu čtverců**

$$S(\phi, \theta) = \sum_{j=1}^n \frac{(Y_j - \hat{Y}_j)^2}{r_{j-1}},$$

přičemž

$$\tilde{\sigma}_\varepsilon^2 = \frac{1}{n - p - q} S(\tilde{\phi}, \tilde{\theta})$$

a platí

$$\frac{S(\tilde{\phi}, \tilde{\theta})}{\tilde{\sigma}_\varepsilon^2} \stackrel{A}{\sim} \chi^2(n - p - q).$$

(viz Brockwell, Davis (1991), §8.9).

Takto získané odhady se nazývají **odhady metodou nejmenších čtverců**, což vede k systému nelineárních rovnic.

Pokud chceme zkoumat **asymptotické vlastnosti maximálně věrohodných odhadů**, musíme zesílit předpoklady: necht' $\{Y_t, t \in \mathbb{Z}\}$ je kauzální a invertibilní *ARMA* proces

$$\{Y_t, t \in \mathbb{Z}\} \sim ARMA(p, q) : \Phi(B)Y_t = \Theta(B)\varepsilon_t \quad \varepsilon_t \sim IID(0, \sigma_\varepsilon^2)$$

a necht' $\Phi(z)$ a $\Theta(z)$ nemají společné kořeny.

Pak, označíme-li maximálně věrohodný odhad neznámých parametrů

$$\beta = (\phi', \theta', \sigma_\varepsilon^2)'$$

symbolem

$$\hat{\beta}_{MLE} = (\hat{\phi}', \hat{\theta}', \hat{\sigma}_\varepsilon^2)',$$

platí

$$\sqrt{n} \left(\widehat{\boldsymbol{\beta}}_{MLE} - \boldsymbol{\beta} \right) \overset{A}{\approx} N_{n+p+1}(\mathbf{0}, V(\boldsymbol{\beta})),$$

kde

$$V(\boldsymbol{\beta}) = \sigma_\varepsilon^2 \begin{pmatrix} E\mathbf{U}_t \mathbf{U}_t' & E\mathbf{U}_t \mathbf{V}_t' \\ E\mathbf{V}_t \mathbf{U}_t' & E\mathbf{V}_t \mathbf{V}_t' \end{pmatrix}^{-1},$$

přičemž

$$\begin{aligned} \mathbf{U}_t &= (U_t, \dots, U_{t+1-p})' \\ \mathbf{V}_t &= (V_t, \dots, V_{t+1-q})' \end{aligned}$$

a $\{U_t, t \in \mathbb{Z}\}$ i $\{V_t, t \in \mathbb{Z}\}$ jsou autoregresní procesy

$$\begin{aligned} \Phi(B)U_t &= \varepsilon_t \\ \Theta(B)V_t &= \varepsilon_t \end{aligned}$$

(viz Brockwell, Davis (1991), §8.9).

1.8 Výstavba modelů a predikce v ARIMA procesech

Až dosud jsme uvažovali pouze o procesech (*slabě*) *stacionárních*. V reálných situacích se však se stacionárními procesy setkáváme pouze zřídka.

Obecně rozlišujeme dva druhy **nestacionarity**:

- ve střední hodnotě
- v rozptylu.

1.8.1 Procesy nestacionární ve střední hodnotě

Nyní je třeba si vysvětlit a odlišit pojmy

Deterministický trend ,

kdy **nestacionaritu ve střední hodnotě** chápeme jako funkci času, tj. k jeho modelování použijeme například

$$\begin{aligned} \text{polynomický trend} & f(t) = \beta_0 + \beta_1 t + \dots + \beta_d t^d \\ \text{periodický trend} & f(t) = \mu + \sum_{j=1}^p (\alpha_j \cos \lambda_j t + \beta_j \sin \lambda_j t) \end{aligned}$$

Stochastický trend . U procesů ARMA jsem požadovali, aby všechny kořeny

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

ležely vně jednotkové kružnice, tj. proces bude kauzální. Pokud nějaký kořen leží

- **na jednotkové kružnici**, mluvíme o procesu **nestacionárním se stochastickým trendem**
- **vně jednotkové kružnice**, mluvíme o procesu **nestacionárním explozivního typu**

Nestacionární proces obsahující stochastický trend lze převést na stacionární **diferencováním**. Zavedme proto tzv. **diferenční operátor**:

$$\begin{aligned} \boxed{\Delta Y_t} &= Y_t - Y_{t-1} = \boxed{(1 - B)Y_t} \\ \boxed{\Delta^2 Y_t} &= \Delta(\Delta Y_t) = \Delta(Y_t - Y_{t-1}) \\ &= (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2}) \\ &= Y_t - 2Y_{t-1} + Y_{t-2} = \boxed{(1 - B)^2 Y_t} \\ &\vdots \\ \boxed{\Delta^d Y_t} &= \boxed{(1 - B)^d Y_t} \end{aligned}$$

Nestacionární proces se stochastickým trendem nazýváme **integrovaným smíšeným modelem** $ARIMA(p, d, q)$.

Formálně jej zapíšeme pomocí operátoru zpětného chodu takto:

$$ARIMA(p, d, q) : \Phi(B)(1 - B)^d Y_t = \Theta(B)\varepsilon_t$$

a položíme-li

$$W_t = (1 - B)^d Y_t,$$

pak W_t je **stacionární** $ARMA(p, q)$.

Zvláštní případy $ARIMA(p, d, q)$

p	d	q	Zkratka	Název
0			$IMA(d, q)$	Integrovaný proces klouzavých součtů
0	0		$MA(q)$	Proces klouzavých součtů
		0	$ARI(p, d)$	Integrovaný autoregresní proces
	0	0	$AR(p)$	Autoregresní proces
0		0	$I(d)$	Integrovaný proces
0	1	0	$I(1)$	Náhodná procházka (<i>random walk</i>)

NÁHODNÁ PROCHÁZKA: $Y_t = Y_{t-1} + \varepsilon_t$ $\varepsilon_t \sim VW(0, \sigma_\varepsilon^2)$

Proces náhodné procházky je limitním případem procesu $AR(1)$, kde $\varphi_1 = 1$, takže

- (a) hodnoty $ACF = \rho(k)$ budou klesat velmi pomalu (lineárně),
- (b) hodnoty $PACF = \alpha(k)$ jsou logicky velmi podobné procesu $AR(1)$.

V praxi se používá následující **modifikace**:

$$Y_t = \beta + Y_{t-1} + \varepsilon_t \quad \beta \in \mathbb{R}.$$

Potom, pokud budeme postupně upravovat, dostaneme

$$\boxed{Y_t} = \beta + Y_{t-1} + \varepsilon_t = Y_{t-2} + 2\beta + \varepsilon_t + \varepsilon_{t-1} = \dots = \underbrace{Y_0 + \beta \cdot t}_{\text{deterministický lineární trend}} + \sum_{j=1}^t \varepsilon_j.$$

POZNÁMKA 1.8.1.

Tvary $ACF = \rho(k)$ a $PACF = \alpha(k)$ procesů $ARIMA(p, d, q)$ a náhodné procházky $I(1)$ jsou **prakticky totožné**. Přítomnost jednotkových kořenů způsobuje „zakrytí“ téměř všech identifikačních detailů těchto funkcí.

POZNÁMKA 1.8.2.

Operátor

$$\nu(B) = \Phi(B)(1 - B)^d$$

se někdy nazývá **zobecněný autoregresní operátor**.

Pokud $\nu(B)$ chápeme jako polynom v proměnné B , pak vzhledem ke kauzalitě modelu $(1 - B)^d W_t = \Theta(B)\varepsilon_t$ má $\nu(B)$ **právě p kořenů ležících vně jednotkového kruhu a d kořenů rovných 1**. V praxi se nejprve diferencováním časové řady získá stacionární řada W_t a pro ni se vybuduje proces $ARMA(p, q)$. Pokud jsme původně měli Y_1, \dots, Y_n , po diferencování zůstanou W_{d+1}, \dots, W_n .

POZNÁMKA 1.8.3.

$ARIMA(p, d, q)$ **nemá smysl centrovat**, neboť platí:

$$\Delta^d(Y_t - \bar{Y}) = \Delta^d Y_t.$$

POZNÁMKA 1.8.4.

Kromě trendů vyžadujících stochastické modelování mohou ARIMA modely zachytit i čistě deterministické trendy, pokud provedeme takovéto **zobecnění ARIMA(p, d, q) modelů**:

$$ARIMA(p, d, q) : \Phi(B)(1 - B)^d Y_t = \beta + \Theta(B)\varepsilon_t \quad \beta \in \mathbb{R};$$

Pak této definici vyhovují procesy tvaru:

$$\underbrace{\beta_0 + \beta_1 t + \dots + \beta_d t^d}_{\text{polynomický trend řádu } d} + Y_t.$$

s využitím poznatků o diferencování polynomů lze totiž psát:

$$\begin{aligned} \Phi(B)(1 - B)^d(\beta_0 + \beta_1 t + \dots + \beta_d t^d + Y_t) = \\ \underbrace{\Phi(B)(d! \beta_d)}_{\beta = (1 - \varphi_1 - \dots - \varphi_p)d! \beta_d} + \Phi(B)(1 - B)^d Y_t = \beta + \Theta(B)\varepsilon_t. \end{aligned}$$

1.8.2 Procesy nestacionární v rozptylu

Není-li splněna podmínka **neměnnosti rozptylu v čase**, je proces **nestacionární v rozptylu**. Takovýto proces je ovšem třeba nejprve vhodně transformovat. Vysvětleme si stručně pojem **transformace stabilizující rozptyl**.

Situace nestabilního rozptylu nastává především v případě, kdy náhodná veličina Y_t má rozdělení, které závisí na jediném parametru ϑ_t , který obecně nemusí mít pro všechna t stejnou hodnotu. Předpokládejme, že tento parametr je zvolen tak, aby platilo

$$E_{\mu_t} Y_t = \mu_t.$$

Ve většině případů (ne však u normálního rozdělení) na μ_t závisí i rozptyl veličiny Y_t , takže můžeme psát

$$D_{\mu_t} Y_t = \sigma^2(\mu_t).$$

Přitom $\sigma(\mu_t)$ bývá obvykle hladká funkce proměnné μ_t .

Protože μ_t může souviset s časem t , není splněna podmínka neměnnosti rozptylu v čase. Vzniká tedy otázka, zda lze najít netriviální funkci g tak, aby náhodná veličina

$$Z_t = g(Y_t)$$

měla rozptyl nezáviselý na μ_t . (Požadavkem netriviality se vylučují konstantní funkce g , které by vedly k veličinám s nulovým rozptylem).

Uvedená úloha v obecném případě nemá řešení. Používá se však určitých aproximací, které se ukázaly velmi užitečné.

Pokud se zabýváme jen dostatečně hladkými funkcemi g , z Taylorova rozvoje dostaneme aproximaci

$$g(Y_t) \approx g(\mu_t) + g'(\mu_t)(Y_t - \mu_t).$$

Potom střední hodnotu lze aproximovat takto

$$E_{\mu_t} g(Y_t) \approx E [g(\mu_t) + g'(\mu_t)(Y_t - \mu_t)] = g(\mu_t)$$

a rozptyl

$$D_{\mu_t} [g(Y_t)] \approx [g'(Y_t)]^2 D_{\mu_t} Y_t = [g'(\mu_t)]^2 \sigma^2(\mu_t).$$

Chceme, aby po transformaci byl rozptyl konstantní a nezávisel na střední hodnotě, tj.

$$c^2 = D_{\mu_t} [g(Y_t)] = [g'(\mu_t)]^2 \sigma^2(\mu_t) \quad \Rightarrow \quad g'(\mu_t) = \frac{c}{\sigma(\mu_t)},$$

kde c je nějaká konstanta. Odtud snadno dostaneme tvar transformace stabilizující rozptyl

$$g(\mu_t) = c \int \frac{1}{\sigma(\mu_t)} d\mu_t + K.$$

Konstanty c a K se volí tak, aby funkce g vypočtená podle předchozího vzorce měla výhodný tvar.

Ukázalo se, že funkce g vypočtená podle předchozího vzorce nejen výrazně stabilizuje rozptyl, takže rozptyl $D_{\mu_t} g(Y_t)$ závisí na μ_t jen velmi málo, ale zároveň také rozdělení náhodné veličiny $Z_t = g(Y_t)$ bývá již velmi blízké normálnímu, i když třeba samotné rozdělení veličiny Y_t je výrazně nenormální.

Mocninné transformace

Pro přehlednost vynechejme index t a uvažujme kladnou náhodnou veličinu X z rozdělení, které závisí na parametru μ se střední hodnotou

$$E_{\mu}X = \mu$$

(pokud tomu tak není, provede se vhodná reparametrizace) a rozptylem

$$D_{\mu}X = \sigma^2(\mu) = (\sigma\mu^{\vartheta})^2, \quad \sigma, \vartheta \in \mathbb{R},$$

tj.

$$X \sim \mathcal{L}(\mu, \sigma^2\mu^{2\vartheta}).$$

Podle **obecného vzorce** se transformace stabilizující rozptyl vypočítá takto:

$$g(\mu) = \int \frac{cd\mu}{\sigma(\mu)} + K = \frac{c}{\sigma} \int \frac{d\mu}{\mu^{\vartheta}} + K = \begin{cases} \frac{c}{\sigma} \ln |\mu| + K & \vartheta = 1, \\ \frac{c}{1-\vartheta} \mu^{1-\vartheta} + K & \vartheta \neq 1. \end{cases}$$

Položme v dalším

$$\lambda = 1 - \vartheta$$

a tento parametr nazvěme **transformačním parametrem** pro mocninnou transformaci.

Různou volbou c a K dostaneme následující často užívané transformace

- **Box-Coxova mocninná transformace** pro kladné náhodné veličiny při volbě

$$c = \sigma \quad \text{a} \quad K = \begin{cases} 0 & \lambda = 0 \Rightarrow \vartheta = 1, \\ -\frac{1}{\lambda} = -\frac{1}{1-\vartheta} & \lambda \neq 0 \Rightarrow \vartheta \neq 1, \end{cases}$$

a odtud

$$g(X) = X^{(\lambda)} = \begin{cases} \ln X & \lambda = 0 \ (\vartheta = 1), \\ \frac{X^{\lambda}-1}{\lambda} & \lambda \neq 0 \ (\vartheta \neq 1). \end{cases}$$

- **Box-Coxova mocninná transformace s posunutím** se použije v případě, že hodnoty náhodné veličiny nejsou kladné. Nalezneme proto takové reálné číslo a tak, aby pro všechny realizace platilo $x + a > 0$ a transformace bude mít tvar:

$$g(X + a) = (X + a)^{(\lambda)} = \begin{cases} \ln(X + a) & \lambda = 0 \ (\vartheta = 1), \\ \frac{(X+a)^{\lambda}-1}{\lambda} & \lambda \neq 0 \ (\vartheta \neq 1). \end{cases}$$

- **Mocninná transformace se znaménkem** lze opět použít v případě, že náhodné veličiny nejsou kladné:

$$g(X) = \text{sign}(X)|X|^{(\lambda)} = \begin{cases} \text{sign}(X) \ln |X| & \lambda = 0 \ (\vartheta = 1), \\ \text{sign}(X) \frac{|X|^{\lambda}-1}{\lambda} & \lambda \neq 0 \ (\vartheta \neq 1). \end{cases}$$

Odhad transformačního parametru mocninné transformace

- **Parametrický přístup pomocí metody maximální věrohodnosti.**

Mějme nezávislé realizace náhodné veličiny

$$X \sim \mathcal{L}(\mu_X, \sigma_X^2 \mu_X^{2\vartheta}).$$

Předpokládejme, že existuje takové $\lambda = 1 - \vartheta$, že transformovaný náhodný vektor $\mathbf{Y} = (Y_1 = g(X_1), \dots, Y_n = g(X_n))'$ je výběr z normálního rozdělení se střední hodnotou μ a rozptylem σ^2 . Označme $\mathbf{y} = (y_1, \dots, y_n)'$ realizaci náhodného výběru. Hledejme maximum **věrohodnostní funkce** pro $\boldsymbol{\theta} = (\mu, \sigma^2)'$, tj. pro funkci

$$\begin{aligned} L(\mu, \sigma^2) &= \prod_{i=1}^n \left[-\frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{1}{2} \left(\frac{y_i - \mu}{\sigma} \right)^2 \right\} \right] \\ &= (2\pi\sigma^2)^{-\frac{n}{2}} \exp \left\{ -\frac{1}{2} \sum_{i=1}^n \left(\frac{y_i - \mu}{\sigma} \right)^2 \right\}, \end{aligned}$$

což je stejná úloha jako hledat maximum **logaritmu věrohodnostní funkce**

$$l(\mu, \sigma^2) = \ln L(\mu, \sigma^2) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) - \frac{1}{2} \sum_{i=1}^n \left(\frac{y_i - \mu}{\sigma} \right)^2.$$

Maxima nalezneme, položíme-li $\frac{\partial l}{\partial \mu} = 0$ a $\frac{\partial l}{\partial \sigma^2} = 0$.

$$\begin{aligned} 0 &= \frac{\partial l}{\partial \mu} = \frac{2}{2\sigma^2} \sum_{i=1}^n (y_i - \mu) \\ 0 &= \frac{\partial l}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^n (y_i - \mu)^2 \end{aligned}$$

a odtud pak dostaneme

$$\begin{aligned} \hat{\mu} &= \frac{1}{n} \sum_{i=1}^n y_i = \bar{y}, \\ \hat{\sigma}^2 &= \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2 = s^2. \end{aligned}$$

Upravme nyní logaritmus věrohodnostní funkce takto:

$$\begin{aligned} l(\mu, \sigma^2) &= -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n [(y_i - \bar{y}) + (\bar{y} - \mu)]^2 \\ &= -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) - \frac{1}{2\sigma^2} \left\{ \sum_{i=1}^n (y_i - \bar{y})^2 + n(\bar{y} - \mu)^2 \right\} \\ &= -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) - \frac{1}{2\sigma^2} [ns^2 + n(\bar{y} - \mu)^2] \end{aligned}$$

Nyní dokažme, že funkce $l(\mu, \sigma^2)$ nabývá v bodě $(\hat{\mu}, \hat{\sigma}^2) = (\bar{y}, s^2)$ svého maxima. Platí

$$l(\bar{y}, s^2) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(s^2) - \frac{n}{2},$$

Ověřme, zda platí nerovnost

$$\begin{aligned} l(\mu, \sigma^2) &\stackrel{?}{\leq} l(\bar{y}, s^2) \\ -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) - \frac{ns^2 + n(\bar{y} - \mu)^2}{2\sigma^2} &\stackrel{?}{\leq} -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(s^2) - \frac{n}{2} \\ -\frac{1}{2} \ln(\sigma^2) - \frac{s^2}{2\sigma^2} - \frac{(\bar{y} - \mu)^2}{2\sigma^2} &\stackrel{?}{\leq} -\frac{n}{2} \ln(s^2) - \frac{1}{2} \\ 0 &\stackrel{?}{\leq} \left[\underbrace{\left(\frac{s^2}{2\sigma^2} - \frac{1}{2} \right) - \ln \frac{s}{\sigma}}_{\text{1. člen}} \right] + \underbrace{\frac{(\bar{y} - \mu)^2}{2\sigma^2}}_{\geq 0} \end{aligned}$$

Protože pro všechna kladná $x = \frac{s}{\sigma} > 0$ platí $\ln x < \frac{x^2 - 1}{2}$, je první i druhý člen nezáporný a nerovnost platí.

Celkově jsme tedy dostali, že

$$\max_{\mu, \sigma^2} l(\mu, \sigma^2) = l(\bar{y}, s^2) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(s^2) - \frac{n}{2}$$

a

$$\max_{\mu, \sigma^2} L(\mu, \sigma^2) = L(\bar{y}, s^2) = (2\pi s^2)^{-\frac{n}{2}} e^{-\frac{n}{2}}.$$

Nyní toto maximum vyjádříme v původních proměnných x_i , kdy

$$y_i = g(x_i) = \begin{cases} \ln x_i & \lambda = 0, \\ \frac{x_i^\lambda - 1}{\lambda} & \lambda \neq 0. \end{cases}$$

Nejprve vypočítáme jakobián této transformace:

$$|J| = \prod_{i=1}^n \left| \frac{dy_i}{dx_i} \right| = \prod_{i=1}^n \frac{\lambda x_i^{\lambda-1}}{\lambda} = \prod_{i=1}^n x_i^{\lambda-1}.$$

Pak

$$\begin{aligned} \max_{\mu, \sigma^2} L(\mu, \sigma^2, \lambda) &= (2\pi s^2(\lambda))^{-\frac{n}{2}} e^{-\frac{n}{2}} |J| \\ &= (2\pi s^2(\lambda))^{-\frac{n}{2}} e^{-\frac{n}{2}} \prod_{i=1}^n x_i^{\lambda-1} \\ &= (2\pi s^2(\lambda))^{-\frac{n}{2}} e^{-\frac{n}{2}} \prod_{i=1}^n e^{(\lambda-1) \ln x_i} \\ \max_{\mu, \sigma^2} l(\mu, \sigma^2, \lambda) &= -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(s^2(\lambda)) - \frac{n}{2} + (\lambda - 1) \sum_{i=1}^n \ln x_i. \end{aligned}$$

Nyní hledíme maximum funkce $l(\hat{\mu}, \hat{\sigma}^2, \lambda) = l(\bar{y}, s^2, \lambda)$ pro parametr λ . Protože maximum vzhledem k λ nezávisí na konstantách, budeme maximalizovat funkci

$$l^*(\lambda) = -\frac{n}{2} \ln(s^2(\lambda)) + (\lambda - 1) \sum_{i=1}^n \ln x_i.$$

Teoretickým odvozením maximálně věrohodného odhadu parametru λ se zde již dále nebudeme zabývat, ale ukážeme si **jednodušší přístup**, který pro ekvidistantní hodnoty $\lambda_1 < \lambda_2 < \dots < \lambda_m$ (pro dostatečně velké m) ze vhodně zvoleného intervalu $(\lambda_1^*, \lambda_2^*)$, (kde $\lambda_1^*, \lambda_2^* \in \mathbb{R}$, $\lambda_1^* < \lambda_2^*$) vypočítá hodnoty $l^*(\lambda)$ a hledá argument $\hat{\lambda}$ maxima těchto hodnot.

Ve své práci Box a Cox (1964) odvodili asymptotické rozdělení statistiky

$$K = -2 \left[l^*(\lambda) - l^*(\hat{\lambda}) \right] \overset{A}{\sim} \chi^2(1),$$

takže můžeme zkonstruovat jednostranný asymptotický **interval spolehlivosti pro parametr λ**

$$\begin{aligned} 1 - \alpha &= P(K < \chi_{1-\alpha}^2(1)) = P\left(-2 \left[l^*(\lambda) - l^*(\hat{\lambda}) \right] < \chi_{1-\alpha}^2(1)\right) \\ &= P\left(\underbrace{l^*(\hat{\lambda}) - \frac{1}{2} \chi_{1-\alpha}^2(1)}_{=D_\alpha} \leq l^*(\lambda)\right), \end{aligned}$$

tj. všechna λ splňující nerovnost

$$l^*(\lambda) \geq D_\alpha$$

leží v intervalu spolehlivosti a jsou tedy přijatelná.

Testování hypotéz typu $H_0 : \lambda = \lambda_0$ proti alternativě $H_1 : \lambda > \lambda_0$:

1. Budeme testovat hypotézu

$$H_0^1 : \lambda = 1.$$

Pokud hypotézu **nezamítáme**, tj. $l^*(1) \geq D_\alpha$, **nemusíme data transformovat**.

2. Pokud předchozí hypotézu **zamítáme**, můžeme testovat další hypotézu

$$H_0^2 : \lambda = 0.$$

Pokud H_0^2 **nezamítáme**, tj. $l^*(0) \geq D_\alpha \wedge l^*(1) < D_\alpha$, transformace bude tvaru

$$y_i = \ln x_i.$$

Pokud však se $l^*(0) < D_\alpha \wedge l^*(1) < D_\alpha$, provedeme transformaci

$$y_i = \frac{x_i^{\hat{\lambda}} - 1}{\hat{\lambda}}.$$

• **Jednoduchý algoritmus v praktických úlohách**

1. Algoritmus nejprve zkontroluje vstupní data tak, aby byla **nezáporná**, tj. případně přičte kladnou konstantu.
2. Upravený vektor dat rozdělí na krátké úseky o délce 4 až 12 údajů.
3. V každém úseku dat se provede pokud možno robustní odhad střední hodnoty $\hat{\mu}_i$ (průměr, medián) a robustní odhad variability $\hat{\sigma}_i^2$ (např. max-min, interkvartilové rozpětí).
4. Protože předpokládáme, že platí

$$\sigma(\mu) = \sigma\mu^\vartheta$$

pak logaritmováním dostaneme vztah

$$\ln(\sigma(\mu)) = \underbrace{\ln \sigma}_a + \vartheta \ln(\mu),$$

takže neznámé ϑ můžeme odhadnout pomocí **metody nejmenších čtverců** díky hodnotám

$$z_i = \ln(\hat{\sigma}_i) \quad \text{a} \quad u_i = \ln \hat{\mu}_i$$

v regresním modelu

$$z_i = a + \vartheta u_i + \varepsilon_i \quad \varepsilon_i \sim WN(0, \sigma_\varepsilon^2).$$

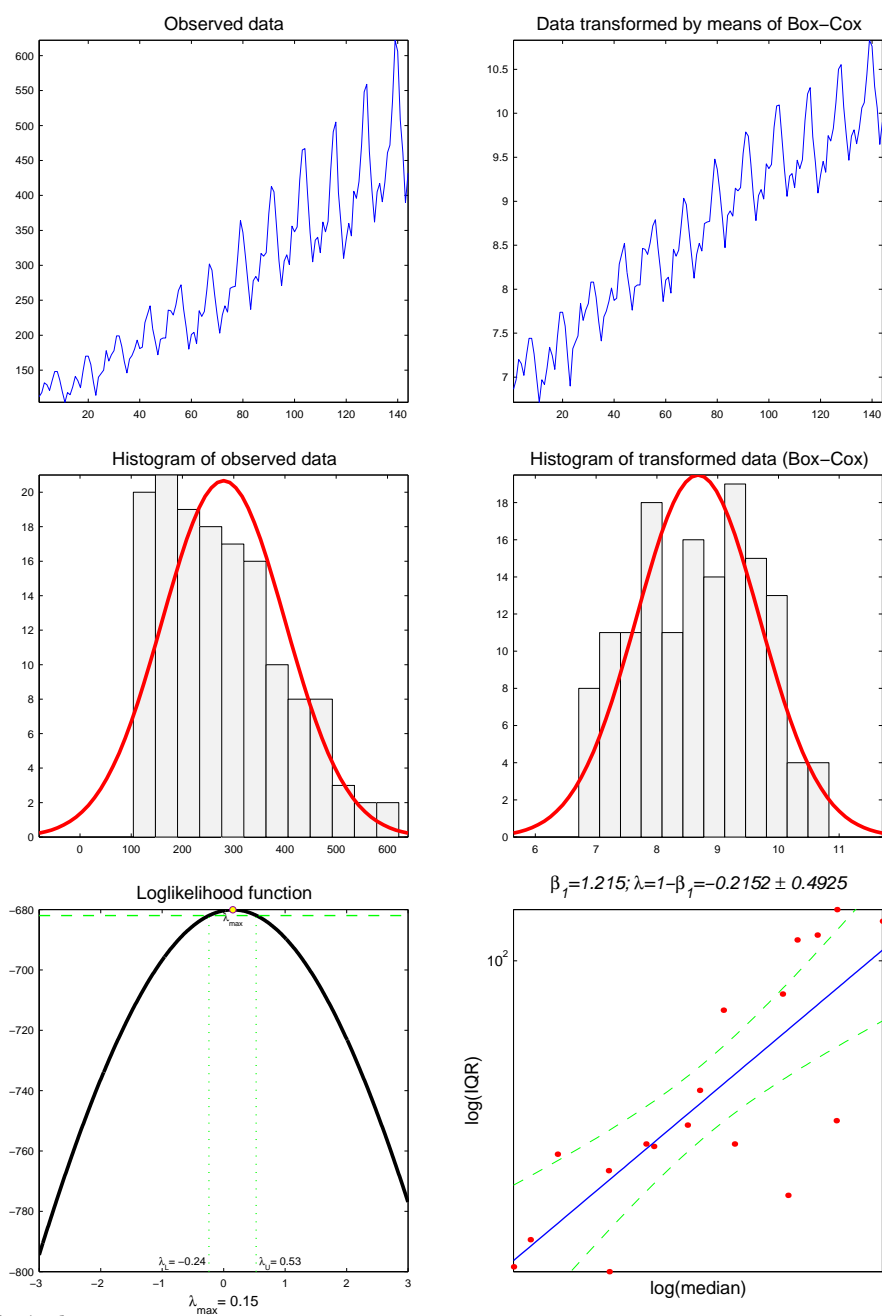
5. Pro odhad $\hat{\vartheta} = 1 - \hat{\lambda}$ pomocí t-statistiky zkonstruujeme interval spolehlivosti $I(\hat{\vartheta})$.
 - Pokud tento interval bude obsahovat **nulu**, tj. $0 \in I(\hat{\vartheta})$ data se **nebudou transformovat**.
 - Pokud $0 \notin I(\hat{\vartheta}) \wedge 1 \in I(\hat{\vartheta})$, volí se logaritmická transformace

$$y_i = \ln x_i.$$

- Jinak se volí mocninná transformace

$$y_i = \frac{x_i^{\hat{\lambda}} - 1}{\hat{\lambda}}.$$

International airline passenger monthly totals (in thousands), Jan.49–Dec.60



Obrázek 1.14: Ukázka mocninné transformace. Protože 0 leží uvnitř intervalu spolehlivosti pro parametr λ , lze doporučit logaritmicou transformaci.

1.8.3 Volba řádu modelu

Vhodná volba řádu modelu je důležitou součástí výstavby modelů v Box-Jenkinsonově metodologii.

K identifikaci typu modelu nám poslouží především průběh odhadnuté autokorelační funkce $ACF = \rho(k)$ a parciální autokorelační funkce $PACF = \alpha(k)$.

Formálnější přístup při rozhodování o výběru řádů p a q je však založen na jistých kritériálních funkcích.

Všimněme si nejprve kritériální funkce odvozené pro $AR(p)$ modely.

Odvození kritéria FPE (Final prediction error)

Mějme dvě nezávislé realizace

$$(X_1, \dots, X_n) \quad \text{a} \quad (Y_1, \dots, Y_n)$$

procesu $AR(p)$ s koeficienty

$$\boldsymbol{\phi}_p = (\varphi_1, \dots, \varphi_p)'$$

Pomocí prvních realizací odvodíme maximálně věrohodné odhady

$$\widehat{\boldsymbol{\phi}}_p = (\widehat{\varphi}_1, \dots, \widehat{\varphi}_p)'$$

Definujeme

$$\mathbf{Y}_p = (Y_n, Y_{n-1}, \dots, Y_{n+1-p})'$$

a uvažujeme jedнокrokovou predikci

$$\widehat{Y}_{n+1} = \widehat{\varphi}_1 Y_n + \dots + \widehat{\varphi}_p Y_{n+1-p} = \widehat{\boldsymbol{\phi}}_p' \mathbf{Y}_p.$$

Počítejme střední kvadratickou chybu této predikce:

$$\begin{aligned} \widehat{v}_n &= E(Y_{n+1} - \widehat{Y}_{n+1})^2 = E(Y_{n+1} - \widehat{\varphi}_1 Y_n - \dots - \widehat{\varphi}_p Y_{n+1-p})^2 \\ &= E \left[\widehat{Y}_{n+1} - \varphi_1 Y_n - \dots - \varphi_p Y_{n+1-p} - (\widehat{\varphi}_1 - \varphi_1) Y_n - \dots - (\widehat{\varphi}_p - \varphi_p) Y_{n+1-p} \right]^2 \\ &= E \left[\varepsilon_{n+1} - (\widehat{\boldsymbol{\phi}}_p - \boldsymbol{\phi}_p)' \mathbf{Y}_p \right]^2 \\ &= \underbrace{E \varepsilon_{n+1}^2}_{=\sigma_\varepsilon^2} + (\widehat{\boldsymbol{\phi}}_p - \boldsymbol{\phi}_p)' \underbrace{E \mathbf{Y}_p \mathbf{Y}_p'}_{\boldsymbol{\Gamma}_p} (\widehat{\boldsymbol{\phi}}_p - \boldsymbol{\phi}_p) \end{aligned}$$

neboť

$$E(\varepsilon_{n+1} Y_{n-j}) = 0 \quad \text{pro} \quad j = 0, \dots, p-1.$$

Protože

$$\sqrt{n}(\widehat{\boldsymbol{\phi}}_p - \boldsymbol{\phi}_p) \overset{A}{\rightsquigarrow} N_p(\mathbf{0}, \sigma_\varepsilon^2 \boldsymbol{\Gamma}_p^{-1})$$

nebo-li

$$(\widehat{\boldsymbol{\phi}}_p - \boldsymbol{\phi}_p) \overset{A}{\rightsquigarrow} N_p\left(\mathbf{0}, \frac{\sigma_\varepsilon^2}{n} \boldsymbol{\Gamma}_p^{-1}\right).$$

Odtud dostaneme, že

$$\frac{n}{\sigma_\varepsilon^2}(\widehat{\phi}_p - \phi_p)' \mathbf{T}_p (\widehat{\phi}_p - \phi_p) = \frac{n}{\sigma_\varepsilon^2} Z_p \stackrel{A}{\sim} \chi^2(p).$$

Protože platí

$$E\left(\frac{n}{\sigma_\varepsilon^2} Z_p\right) = p,$$

ihned dostaneme

$$EZ_p = \sigma_\varepsilon^2 \frac{p}{n}$$

a odtud

$$\boxed{\widehat{v}_n \approx \sigma_\varepsilon^2 \left(1 + \frac{p}{n}\right)}.$$

Jestliže $\widehat{\sigma}_\varepsilon^2$ je maximálně věrohodným odhadem neznámého parametru σ_ε^2 , pak pro $n \rightarrow \infty$ platí

$$\frac{n\widehat{\sigma}_\varepsilon^2}{\sigma_\varepsilon^2} \stackrel{A}{\sim} \chi^2(n-p)$$

a protože

$$E\left(\frac{n\widehat{\sigma}_\varepsilon^2}{\sigma_\varepsilon^2}\right) = n-p,$$

pak

$$E\widehat{\sigma}_\varepsilon^2 = \frac{n-p}{n} \sigma_\varepsilon^2 \quad \Rightarrow \quad \sigma_\varepsilon^2 = \frac{n}{n-p} E\widehat{\sigma}_\varepsilon^2.$$

Pokud ve výpočtu \widehat{v}_n nahradíme σ_ε^2 jeho nestranným odhadem $\widehat{\sigma}_\varepsilon^2$, dostaneme následující kritérium:

$$\boxed{FPE} = \widehat{\sigma}_\varepsilon^2 \frac{n}{n-p} \left(1 + \frac{p}{n}\right) = \boxed{\widehat{\sigma}_\varepsilon^2 \frac{n+p}{n-p}}.$$

Mnohem obecnějším kritériem, které není odvozeno pro $AR(p)$ proces, je tzv. **AIC kritérium**:

$$AIC = nl(\widehat{\phi}_{pq}, \widehat{\theta}_{pq}) + 2(p+q),$$

nebo jeho korigovaná verze (neboť AIC nadhodnocuje řád modelu)

$$AICC = -2 \ln L(\widehat{\phi}_{pq}, \widehat{\theta}_{pq}, S(\widehat{\phi}_{pq}, \widehat{\theta}_{pq})) + \frac{2(p+q+1)n}{n-p-q-2}.$$

Dalším kritériem je tzv. **BIC kritérium**:

$$BIC = (n-p-q) \ln \left(\frac{n\widehat{\sigma}_\varepsilon^2}{n-p-q} \right) + (p+q) \ln \left[\frac{\sum Y_t^2 - n\widehat{\sigma}_\varepsilon^2}{p+q} \right].$$

1.8.4 Verifikace modelu – analýza reziduí

Po odhadu neznámých parametrů $ARMA(p, q)$ modelu metodou maximální věrohodnosti označme rezidua:

$$\widehat{W}_t = \frac{Y_t - \widehat{Y}_t(\widehat{\phi}, \widehat{\theta})}{r_{t-1}(\widehat{\phi}, \widehat{\theta})} \quad t = 1, \dots, n.$$

POZNÁMKA 1.8.5.

Existují i další rezidua, např.

$$\widehat{\varepsilon}_t = \widehat{\Theta}^{-1}(B)\widehat{\Phi}(B)Y_t \quad t = 1, \dots, n.$$

Velikost těchto reziduí však závisí na měrných jednotkách náhodné proměnné Y_t , proto se doporučuje používat raději \widehat{W}_t , která jsou standardizovaná a nezávislá na měřítku.

Protože rezidua \widehat{W}_t by měla být pro $n \rightarrow \infty$ (pokud byl model vhodně zvolen) přibližně bílým šumem $WN(0, \sigma_\varepsilon^2)$ (popř. $IID(0, \sigma_\varepsilon^2)$), ihned se nabízí myšlenka, analyzovat nejprve jeho výběrovou autokorelační funkci

$$\widehat{\rho}_W(h) = \frac{\sum_{t=1}^{n-h} (\widehat{W}_t - \bar{W})(\widehat{W}_{t+h} - \bar{W})}{\sum_{t=1}^n (W_t - \bar{W})^2},$$

kde

$$\bar{W} = \sum_{t=1}^n \widehat{W}_t.$$

Celá řada článků je věnována odvození asymptotických intervalů spolehlivosti pro tuto výběrovou autokorelační funkci.

Mnohem jednodušší je použití tzv. **Portmanteaovy statistiky**

$$Q_W = \sum_{j=1}^h \widehat{\rho}_W^2(j) \overset{A}{\sim} \chi^2(h - p - q) \quad \text{pro } h \in N.$$

Existují také jisté modifikace této statistiky, např. (Ljung, Box, 1978)

$$\tilde{Q}_W = n(n+2) \sum_{j=1}^h \frac{\widehat{\rho}_W^2(j)}{n-j} \overset{A}{\sim} \chi^2(h - p - q) \quad \text{pro } h \in N,$$

nebo (Granger, Anderson, 1978)

$$\tilde{Q}_{WW} = n(n+2) \sum_{j=1}^h \frac{\widehat{\rho}_{WW}(j)}{n-j} \overset{A}{\sim} \chi^2(h) \quad \text{pro } h \in N.$$

1.9 Modelování sezónnosti pomocí SARIMA modelů

Sezónnost je v Box-Jenkinsonově metodologii stejně jako trend modelována **stochasticky**.

Nejprve zavedme **sezónní diferenciální operátor o délce $L > 0$** :

$$\begin{aligned} \boxed{\Delta_L Y_t} &= Y_t - Y_{t-L} = (1 - B^L)Y_t \\ \boxed{\Delta_L^2 Y_t} &= \Delta_L(\Delta_L Y_t) = \Delta_L(Y_t - Y_{t-L}) \\ &= (Y_t - Y_{t-L}) - (Y_{t-L} - Y_{t-2L}) \\ &= Y_t - 2Y_{t-L} + Y_{t-2L} = \boxed{(1 - B^L)^2 Y_t} \\ &\vdots \\ \boxed{\Delta_L^D Y_t} &= \boxed{(1 - B^L)^D Y_t} \end{aligned}$$

Při konstrukci se uvažuje způsobem, který budeme demonstrovat pomocí následujícího příkladu:

Nechť časová řada $\{Y_t\}$ vykazuje sezónnost o délce $L = 12$.

1. Zkonstruuje se nejprve $ARIMA(P_1, D_1, Q_1)$ model pro řadu **lednových měření**, tj. pro $\{S_t^1 = B^{12}Y_t\}$

$$\pi_1(B^{12})\Delta_{12}^{D_1}Y_t = \Psi_1(B^{12})\eta_t^{(1)} \sim ARIMA(P_1, D_1, Q_1)$$

kde časový index t odpovídá **lednovým** obdobím a o η_t se budeme zajímat později.

Přitom

$$\pi_1(B^{12}) = 1 - \pi_{1,1}B^{12} - \dots - \pi_{1,P_1}B^{12 \cdot P_1}$$

je tzv. **sezónní autoregresní operátor $SAR(P_1)$**

$$\Psi_1(B^{12}) = 1 + \psi_{1,1}B^{12} + \dots + \psi_{1,Q_1}B^{12 \cdot Q_1}$$

je tzv. **sezónní operátor klouzavých součtů $SMA(Q_1)$**

$$\Delta_{12}^{D_1} = (1 - B^L)^{D_1} \quad \text{je tzv. sezónní diferenciální operátor } SI(D_1)$$

2. Podobné modely zkonstruuje se pro ostatní měsíce:

$$\pi_2(B^{12})\Delta_{12}^{D_2}Y_t = \Psi_2(B^{12})\eta_t^{(2)} \sim ARIMA(P_2, D_2, Q_2)$$

\vdots

$$\pi_{12}(B^{12})\Delta_{12}^{D_{12}}Y_t = \Psi_{12}(B^{12})\eta_t^{(12)} \sim ARIMA(P_{12}, D_{12}, Q_{12})$$

3. Předpokládejme přitom, že tyto **modely jsou pro jednotlivé měsíce přibližně stejné**, tj.

$$\begin{aligned} P_1 &\approx \dots \approx P_{12} \approx P \\ Q_1 &\approx \dots \approx Q_{12} \approx Q \\ D_1 &\approx \dots \approx D_{12} \approx D \\ \pi_1(B^{12}) &\approx \dots \approx \pi_{12}(B^{12}) \approx \pi(B^{12}) \\ \Psi_1(B^{12}) &\approx \dots \approx \Psi_{12}(B^{12}) \approx \Psi(B^{12}) \end{aligned}$$

4. **Náhodné veličiny** $\eta_t^{(j)}$ ($j = 1, \dots, 12$) by však v těchto modelech měly být pro různé měsíce mezi sebou **korelované**, neboť by měl existovat např. vztah mezi lednovými a únorovými hodnotami. Předpokládejme proto, že také řada η_t je popsána modelem $ARIMA(p, d, q)$ tvaru

$$\Phi(B)\Delta^d\eta_t = \Theta(B)\varepsilon_t \sim ARIMA(p, d, q)$$

kde $\varepsilon_t \sim WN(0, \sigma^2)$ je bílý šum.

5. **Spojme** předchozí dva modely do jediného tzv. **multiplikativního sezónního modelu řádu** $(p, d, q) \times (P, D, Q)_L$

$$\boxed{\Phi(B)\pi(B^L)\Delta^d\Delta_L^D Y_t = \Theta(B)\Psi(B^L)\varepsilon_t \sim SARIMA(p, d, q) \times (P, D, Q)_L}$$

$L = 12.$

PŘÍKLAD 1.9.1.

Model $SARIMA(0, 1, 1) \times (0, 1, 1)_{12}$ má tvar:

$$\Delta\Delta_{12}Y_t = (1 - B)(1 - B^{12})Y_t = (1 + \theta_1 B)(1 + \psi_1 B^{12})\varepsilon_t,$$

nebo ekvivalentně

$$Y_t - Y_{t-1} - Y_{t-12} + Y_{t-13} = \varepsilon_t + \theta_1\varepsilon_{t-1} + \psi_1\varepsilon_{t-12} + \theta_1\psi_1\varepsilon_{t-13}.$$

POZNÁMKA 1.9.2.

Existují také **aditivní sezónní modely**, které se však používají jen zřídka. Jako příklad lze uvést model

$$Y_t = \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_{12}\varepsilon_{t-12} + \theta_{13}\varepsilon_{t-13}.$$

Výstavba sezónních modelů

Označme řád běžného diferencování na odstranění trendu jako d a D jako řád diferencování na odstranění sezónnosti. V praktických situacích většinou $d = 0, 1, 2$ a $D = 0, 1$. Dále nechť L je délka sezóny.

Výstavba sezónních modelů probíhá ve třech stejných fázích jako pro modely $ARIMA$. Všimněme si pouze **FÁZE IDENTIFIKACE MODELŮ**, neboť ostatní dvě fáze jsou totožné.

1. Odhad parametrů d, D :

(a) Proveďte se studium odhadnuté **autokorelační funkce** $ACF = \hat{\gamma}(k)$, neboť identifikuje přítomnost trendu. Doporučuje se prozkoumat $4L$ hodnot $\hat{\gamma}(k)$.

- **Určení D** : Má-li funkce $\hat{\rho}(k)$ v bodech $L, 2L, 3L \dots$ **lokální maxima**, pak (bez ohledu na její průběh mezi těmito časovými body) je nutné položit $D = 1$.

To plyne z toho, že hodnoty $\hat{\rho}(L), \hat{\rho}(2L), \hat{\rho}(3L), \dots$ představují odhadnuté hodnoty autokorelační funkce pro řady $\{S_t^j = B^L Y_t\}$, $j = 1, \dots, L$ modelu

$$\pi(B^L)\Delta_L^D Y_t = \Psi(B^L)\eta_t \sim ARIMA(P, D, Q),$$

přičemž **nestacionaritě** tohoto ARIMA modelu **odpovídá pomalý pokles autokorelační funkce** $\hat{\rho}(L), \hat{\rho}(2L), \hat{\rho}(3L), \dots$, tj. tuto řadu je nutno diferencovat (s krokem L) a pokládáme $D = 1$.

- **Určení d** : Jestliže funkce r_k klesá **mezi body jL a $(j+1)L$ pouze přibližně lineárně**, je třeba provést také běžné diferencování.
- (b) Čísla d, D se také někdy určují tak, že se hledá nejmenší číslo mezi odhadnutými hodnotami

$$\hat{\sigma}_Y^2, \hat{\sigma}_{\Delta Y_t}^2, \hat{\sigma}_{\Delta_L Y_t}, \hat{\sigma}_{\Delta_L^2 Y_t}, \dots$$

rozptylů dané řady a jejich diferencí.

2. Odhad parametrů p, P, q, Q :

Po určení řádu d a D zkonstruujeme řadu $W_t = \Delta^d \Delta_L^D Y_t$,

pro kterou je nutné identifikovat model tvaru

$$\Phi(B)\pi(B^L)W_t = \Theta(B)\Psi(B^L)\varepsilon_t \sim SARIMA(p, 0, q) \times (P, 0, Q)_L.$$

Pro tento účel se použije odhadnutá $ACF = \hat{\rho}(k)$ a $PACF = \hat{\alpha}(k)$ řady W_t .

(a) MA-HOMOGENNÍ MODELY

- Jestliže ACF funkce $\hat{\rho}(k)$ je zhruba **významně nenulová** v bodech

$$\begin{array}{c} 1, \dots, q \\ L - q, \dots, L + q \\ 2L - q, \dots, 2L + q \\ \vdots \\ QL - q, \dots, QL + q \end{array}$$

přičemž **mezi těmito body se neodlišují významně od nuly**

- a funkce $\hat{\alpha}(k)$ v jednotlivých úsecích mezi body jL a $(j+1)L$ vždy **v absolutní hodnotě klesá** (geometricky nebo po sinusoidě s geometricky klesající amplitudou) a **zároveň klesá**, když ji sledujeme v bodech $L, 2L, 3L, \dots$,

pak položíme

$$\boxed{p = 0 \text{ a } P = 0},$$

tj. budeme identifikovat odpovídající model pro řadu W_t jako

$$W_t = \Theta(B)\Psi(B^L)\varepsilon_t \sim SARIMA(0, 0, q) \times (0, 0, Q)_L$$

a tedy model pro řadu Y_t jako

$$\Delta^d \Delta_L^D Y_t = \Theta(B)\Psi(B^L)\varepsilon_t \sim SARIMA(0, d, q) \times (0, D, Q)_L.$$

(b) AR-HOMOGENNÍ MODELY

- Jestliže naopak funkce $\hat{\rho}(k)$ **klesá** v absolutní hodnotě (geometricky nebo po sinusoidě s geometricky klesající amplitudou) **v úsecích mezi body** jL a $(j+1)L$ a **zároveň klesá**, když ji sledujeme v bodech $L, 2L, 3L, \dots$

- a funkce $\hat{\alpha}(k)$ je zhruba **významně nenulová** v bodech

$$\begin{array}{c} 1, \dots, p \\ L, \dots, L+p \\ 2L, \dots, 2L+p \\ \vdots \\ PL, \dots, PL+p \end{array}$$

přičemž **mezi těmito body se neodlišují významně od nuly**,

pak položíme

$$\boxed{q = 0 \text{ a } Q = 0},$$

tj. budeme identifikovat odpovídající model pro řadu W_t jako

$$\Phi(B)\pi(B^L)W_t = \varepsilon_t \sim SARIMA(p, 0, 0) \times (P, 0, 0)_L$$

a tedy model pro řadu Y_t jako

$$\Phi(B)\pi(B^L)\Delta^d \Delta_L^D Y_t = \varepsilon_t \sim SARIMA(p, d, 0) \times (P, D, 0)_L.$$

(c) NEHOMOGENNÍ MODELY typu

$$SARIMA(p, d, 0) \times (0, D, Q)_L \text{ nebo } SARIMA(0, d, q) \times (P, D, 0)_L$$

se většinou **nepoužívají**, neboť obvykle vedou při srovnání s předchozími tzv. **homogenní modely**

$$SARIMA(0, d, q) \times (0, D, Q)_L \text{ nebo } SARIMA(p, d, 0) \times (P, D, 0)_L$$

k odhadu neúnosně velkého počtu parametrů.

(d) Identifikace obecných modelů

$$SARIMA(p, d, q) \times (P, D, Q)_L,$$

v nichž čísla p, q, P a Q mohou být vesměs nenulová, je již dosti komplikovanou záležitostí a obvykle zde hodně záleží na zkušenostech statistika, který analýzu provádí.

Kapitola 2

Dynamické lineární modely

2.1 Motivační příklad

- (1) Představme si, že jsme se ztratili na moři, popřípadě v podmínkách vnitrozemského státu v hlubokém lese. Nacházíme se v nějaké pozici x , která objektivně existuje, ale kterou neznáme a snažíme se ji určit. Skutečná pozice x je neznámým **stavem**.
- (2) Pomocí hvězd odhadneme naši pozici jako y_0 s určitou nepřesností, neurčitostí $\sigma_{y_0}^2$. Na základě dosavadních znalostí známe hustotu pozice x , kterou označíme $p_{1|0}(x) = p(x|y_0)$, se střední hodnotou $\hat{x}_{1|0} = y_0$ a rozptylem $\widehat{\sigma_{1|0}^2} = \sigma_{y_0}^2$.
- (3) Mraky se rozestoupí a jasněji vidíme hvězdy. Proto upřesníme náš odhad jako y_1 s větší jistotou $\sigma_{y_1}^2 < \sigma_{y_0}^2$. Na základě širších znalostí dostaneme novou hustotu pozice x jako $p_{1|1}(x) = p(x|y_0, y_1)$ se střední hodnotou $\hat{x}_{1|1}$ a rozptylem $\widehat{\sigma_{1|1}^2}$. **Střední hodnotu** vypočteme jako vážený průměr pozorování y_0 a y_1 , kde váha je o to větší, o č je pozorování přesnější, tj. má menší rozptyl (přitom součet vah je roven 1)

$$\hat{x}_{1|1} = \frac{\frac{1}{\sigma_{y_0}^2}}{\frac{1}{\sigma_{y_0}^2} + \frac{1}{\sigma_{y_1}^2}} y_0 + \frac{\frac{1}{\sigma_{y_1}^2}}{\frac{1}{\sigma_{y_0}^2} + \frac{1}{\sigma_{y_1}^2}} y_1 = \underbrace{\frac{\sigma_{y_1}^2}{\sigma_{y_0}^2 + \sigma_{y_1}^2}}_{=(1-K)} y_0 + \underbrace{\frac{\sigma_{y_0}^2}{\sigma_{y_0}^2 + \sigma_{y_1}^2}}_{=K} y_1 = (1-K)y_0 + Ky_1$$

Za předpokladu, že y_0 a y_1 jsou nezávislá pozorování, pak **rozptyl** váženého

průměru je roven

$$\widehat{\sigma_{1|1}^2} = \left(\frac{\sigma_{y_1}^2}{\sigma_{y_0}^2 + \sigma_{y_1}^2} \right)^2 \sigma_{y_0}^2 + \left(\frac{\sigma_{y_0}^2}{\sigma_{y_0}^2 + \sigma_{y_1}^2} \right)^2 \sigma_{y_1}^2 = \frac{\sigma_{y_0}^2}{\sigma_{y_0}^2 + \sigma_{y_1}^2} \sigma_{y_1}^2 = K \sigma_{y_1}^2 = (1 - K) \sigma_{y_0}^2$$

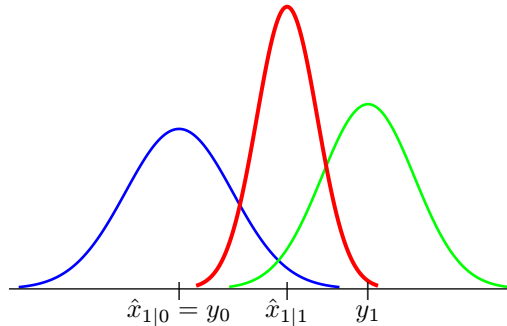
tj. $\widehat{\sigma_{1|1}^2} < \sigma_{y_1}^2$

Celkově dostaneme tyto rekurentní vztahy

$$\widehat{x_{1|1}} = \widehat{x_{1|0}} + K(y_1 - \widehat{x_{1|0}})$$

$$\widehat{\sigma_{1|1}^2} = (1 - K) \widehat{\sigma_{1|0}^2},$$

tj. využili jsme informací z prvního i druhého měření.



(4) Nyní zdynamizujeme (rozpohybujeme) náš příklad

(a) Zatímco děláme svoje měření, nestojíme na místě, ale pohybujeme se, tj. existuje jednoduchý dynamický model

$$\frac{dx}{dt} = \underbrace{v}_{\text{konstantní posun}} + \underbrace{w}_{\text{náhodná složka}}.$$

Předtím, než provedeme další měření, uděláme predikci $\widehat{x_{2|1}} = F_1 \widehat{x_{1|1}}$ na základě předchozího stavu $\widehat{x_{1|1}}$ a dynamického modelu (nějaké funkce přechodu F_1) s určitou dávkou neurčitosti $\widehat{\sigma_{2|1}^2}$

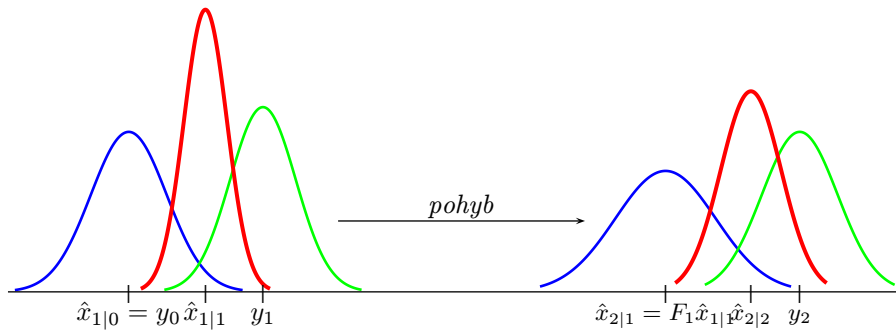
(b) Provedeme pomocí hvězd další měření polohy y_2 s neurčitostí, nepřesností $\sigma_{y_2}^2$.

(5) Všechny předchozí informace shrneme do odhadu polohy

$$\widehat{x_{2|2}} = \widehat{x_{2|1}} + K(y_2 - \widehat{x_{2|1}})$$

s vahou $K = \frac{\widehat{\sigma_{2|1}^2}}{\widehat{\sigma_{2|1}^2} + \sigma_{y_2}^2}$ (též tzv. Kalmanův zisk)

$$\widehat{\sigma_{2|2}^2} = (1 - K) \widehat{\sigma_{2|1}^2}.$$



2.2 Stavově-prostorové modely

Místo jednorozměrné náhodné posloupnosti $\{Y_t, t \in \mathbb{Z}\}$ uvažujme posloupnost w -rozměrných náhodných vektorů $\{\mathbf{Y}_t, t \in \mathbb{Z}\}$, $\mathbf{Y}_t \in \mathbb{R}^w$, které splňují tzv. DATOVÉ A STAVOVÉ ROVNICE

DATOVÁ ROVNICE

$$\mathbf{Y}_t = \mathbf{G}_t \mathbf{X}_t + \mathbf{W}_t \quad t = 1, 2, 3, \dots$$

STAVOVÁ ROVNICE

$$\mathbf{X}_{t+1} = \mathbf{F}_t \mathbf{X}_t + \mathbf{V}_t \quad t = 1, 2, 3, \dots$$

příčemž

- \mathbf{X}_t ... je tzv. *stavový* v -rozměrný náhodný vektor
- \mathbf{W}_t ... je *šum měření*
- \mathbf{V}_t ... je *šum procesu*
- \mathbf{G}_t ... je posloupnost matic typu $w \times v$
(popisují vztah pozorování ke stavu)
- \mathbf{F}_t ... je posloupnost matic typu $v \times v$
(modelují dynamiku - tzv. *matice přechodu*)

Dále platí

$$E\mathbf{V}_t = \mathbf{0}$$

$$E\mathbf{W}_t = \mathbf{0}$$

$$D \begin{pmatrix} \mathbf{W}_t \\ \mathbf{V}_t \end{pmatrix} = \begin{pmatrix} \mathbf{R}_t & \mathbf{S}_t \\ \mathbf{S}'_t & \mathbf{Q}_t \end{pmatrix} \quad \text{tj.} \quad \begin{aligned} E\mathbf{W}_t \mathbf{W}'_t &= \mathbf{R}_t \\ E\mathbf{V}_t \mathbf{V}'_t &= \mathbf{Q}_t \\ E\mathbf{W}_t \mathbf{V}'_t &= \mathbf{S}_t \end{aligned}$$

$C(\mathbf{X}_t, (\mathbf{W}'_t, \mathbf{V}'_t)') = \mathbf{0}$, tj. jsou nekorelované

Všechny náhodné vektory mají konečné druhé momenty.

PŘÍKLAD 2.2.1.**NÁHODNÁ PROCHÁZKA S TRENDEM**

Mějme $\beta \in \mathbb{R}$,
šum procesu $V_t \sim WN(0, \sigma_v^2)$,
náhodné veličiny Tr_t , přičemž $Tr_0 = \mu_0 = 0$.

Dále necht' pro $t = 1, 2, \dots$ platí

$$C(Tr_t, V_t) = 0 \quad \text{tj. } Tr_t \text{ a } V_t \text{ jsou nekorelované, což značíme } Tr_t \perp V_t.$$

Definujeme

$$\begin{aligned} \boxed{Tr_{t+1}} &= \boxed{Tr_t + \beta + V_t} = Tr_{t-1} + \beta + V_{t-1} + \beta + V_t \\ &= Tr_{t-1} + 2\beta + V_t + V_{t-1} = \dots \\ \text{po } t \text{ krocích} &= \underbrace{Tr_0}_{=\mu_0=0} + \beta t + \sum_{j=1}^t V_j \end{aligned}$$

Položme

$$\mathbf{X}_t = \begin{pmatrix} Tr_t \\ \beta \end{pmatrix} \quad \mathbf{V}_t = \begin{pmatrix} V_t \\ 0 \end{pmatrix} \quad \mathbf{F}_t = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Pak

$$\boxed{\mathbf{X}_{t+1}} = \begin{pmatrix} Tr_{t+1} \\ \beta \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} Tr_t \\ \beta \end{pmatrix} + \begin{pmatrix} V_t \\ 0 \end{pmatrix} = \boxed{\mathbf{F}_t \mathbf{X}_t + \mathbf{V}_t} \quad t = 1, 2, \dots$$

Označme šum měření $W_t \sim WN(0, \sigma_w^2)$ a položme

$$\boxed{Y_t} = \underbrace{\begin{pmatrix} 1 & 0 \end{pmatrix}}_{=\mathbf{G}_t} \begin{pmatrix} Tr_t \\ \beta \end{pmatrix} + W_t = \boxed{\mathbf{G}_t \mathbf{X}_t + \mathbf{W}_t} \quad t = 1, 2, \dots$$

Jestliže $\mathbf{X}_1 = \begin{pmatrix} Tr_1 \\ \beta \end{pmatrix}$, $\underbrace{V_1, W_1}, \underbrace{V_2, W_2}, \dots$ jsou nekorelované, dostáváme *stavově-prostorovou reprezentaci náhodné procházky*, pro kterou platí

$$\begin{aligned} E\mathbf{V}_t &= \mathbf{0} & D\mathbf{V}_t &= E\mathbf{V}_t \mathbf{V}_t' = \begin{pmatrix} \sigma_v^2 & 0 \\ 0 & 0 \end{pmatrix} = \mathbf{Q}_t = \mathbf{Q} \\ E\mathbf{W}_t &= E\mathbf{W}_t = \mathbf{0} & D\mathbf{W}_t &= E\mathbf{W}_t \mathbf{W}_t' = E\mathbf{W}_t^2 = \sigma_w^2 = R_t = R \\ E\mathbf{V}_t \mathbf{W}_t' &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \mathbf{S}_t = \mathbf{S}. \end{aligned}$$

PŘÍKLAD 2.2.2.
SEZÓNŇÍ ŘADA SE ŠUMEM

Uvažujme sezónu délky d
 a sezónní komponenty s_1, \dots, s_d

přičemž platí

$$s_{t+d} = s_t \quad \text{a} \quad s_1 + \dots + s_d = 0.$$

Takže dostaneme

$$\begin{aligned} s_{t+1} &= s_{t+1-d} \\ &\vdots \\ s_{t+1} + s_t + s_{t-1} + \dots + s_{t+1-d+1} &= 0 \end{aligned}$$

Odtud získáme *deterministickou rovnici*

$$s_{t+1} = -s_t - s_{t-1} - \dots - s_{t+2-d}$$

Přidejme šum procesu $V_t \sim WN(0, \sigma_v^2)$ a dostaneme po přeznačení *stochastickou rovnici*

$$Y_{t+1} = -Y_t - Y_{t-1} - \dots - Y_{t+2-d} + V_t.$$

Položme

$$\mathbf{X}_{t+1} = \begin{pmatrix} Y_{t+1} \\ Y_t \\ Y_{t-1} \\ \vdots \\ Y_{t+3-d} \end{pmatrix} = \underbrace{\begin{pmatrix} -1 & -1 & \dots & -1 & -1 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & 0 \\ 0 & \dots & 0 & 1 & 0 \end{pmatrix}}_{=\mathbf{F}_t} \underbrace{\begin{pmatrix} Y_t \\ Y_{t-1} \\ Y_{t-2} \\ \vdots \\ Y_{t+2-d} \end{pmatrix}}_{=\mathbf{X}_t} + \underbrace{\begin{pmatrix} V_t \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}}_{=\mathbf{V}_t}$$

tj. stavově-prostorový model sezónní řady se šumem je roven

$$\begin{aligned} \mathbf{X}_{t+1} &= \mathbf{F}\mathbf{X}_t + \mathbf{V}_t \\ Y_t &= \underbrace{(1 \quad 0 \quad \dots \quad 0)}_{=\mathbf{G}_t} \mathbf{X}_t \end{aligned}$$

2.3 Stacionární stavově-prostorové modely

DATOVÁ ROVNICE

$$\mathbf{Y}_t = \mathbf{G}\mathbf{X}_t + \mathbf{W}_t \quad t = 1, 2, 3, \dots$$

STAVOVÁ ROVNICE

$$\mathbf{X}_{t+1} = \mathbf{F}\mathbf{X}_t + \mathbf{V}_t \quad t = 1, 2, 3, \dots$$

příčemž

- \mathbf{X}_t ... je tzv. *stavový* v -rozměrný náhodný vektor
- \mathbf{W}_t ... je *šum měření*
- \mathbf{V}_t ... je *šum procesu*
- \mathbf{G} ... je matice typu $w \times v$ (popisují vztah pozorování ke stavu)
- \mathbf{F} ... je matice typu $v \times v$ tzv. *matice přechodu*

Dále platí

$$\begin{aligned} E\mathbf{V}_t &= \mathbf{0} \\ E\mathbf{W}_t &= \mathbf{0} \\ D \begin{pmatrix} \mathbf{W}_t \\ \mathbf{V}_t \end{pmatrix} &= \begin{pmatrix} \mathbf{R}_t & \mathbf{S}_t \\ \mathbf{S}'_t & \mathbf{Q}_t \end{pmatrix} \quad \text{tj.} \quad \begin{aligned} E\mathbf{W}_t\mathbf{W}'_t &= \mathbf{R}_t \\ E\mathbf{V}_t\mathbf{V}'_t &= \mathbf{Q}_t \\ E\mathbf{W}_t\mathbf{V}'_t &= \mathbf{S}_t \end{aligned} \\ C(\mathbf{X}_t, (\mathbf{W}'_t, \mathbf{V}'_t)') &= \mathbf{0}, \quad \text{tj. jsou nekorelované} \end{aligned}$$

Všechny náhodné vektory mají konečné druhé momenty.

Stavová rovnice se nazývá *stabilní* (také *kauzální*), právě když všechna vlastní čísla matice \mathbf{F} leží uvnitř jednotkové kružnice, tj.

$$\det(\mathbf{I} - \mathbf{F}\mathbf{z}) \neq \mathbf{0} \quad \text{pro } \forall |\mathbf{z}| < 1.$$

Pokud je systém *stabilní* (*kauzální*), pak

$$\begin{aligned} \mathbf{X}_{t+1} &= \sum_{j=0}^{\infty} \mathbf{F}^j \mathbf{V}_{t-j} \\ \mathbf{Y}_t &= \mathbf{W}_t + \mathbf{G} \sum_{j=0}^{\infty} \mathbf{F}^j \mathbf{V}_{t-1-j} \end{aligned}$$

PŘÍKLAD 2.3.1.
AUTOREGRESNÍ MODEL ŘÁDU p

$$AR(p) : \boxed{Y_t = \varphi_1 Y_{t-1} \cdots + \varphi_p Y_{t-p} + \varepsilon_t}$$

kde

$$\varepsilon_t \sim WN(0, \sigma_\varepsilon^2).$$

$$\underbrace{\begin{pmatrix} Y_{t+1} \\ Y_t \\ Y_{t-1} \\ \vdots \\ Y_{t+1-p} \end{pmatrix}}_{=\mathbf{X}_{t+1}} = \underbrace{\begin{pmatrix} \varphi_1 & \varphi_2 & \cdots & \varphi_{p-1} & \varphi_p \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & 0 \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix}}_{=\mathbf{F}} \underbrace{\begin{pmatrix} Y_t \\ Y_{t-1} \\ Y_{t-2} \\ \vdots \\ Y_{t-p} \end{pmatrix}}_{=\mathbf{X}_t} + \underbrace{\begin{pmatrix} \varepsilon_t \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}}_{=\mathbf{V}_t}$$

$$Y_t = \underbrace{(1 \quad 0 \quad \cdots \quad 0)}_{=\mathbf{G}} \mathbf{X}_t$$

PŘÍKLAD 2.3.2.
MA PROCES ŘÁDU q

$$MA(q) : \boxed{Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}}$$

kde

$$\varepsilon_t \sim WN(0, \sigma_\varepsilon^2).$$

$$\underbrace{\begin{pmatrix} \varepsilon_t \\ \varepsilon_{t-1} \\ \varepsilon_{t-2} \\ \vdots \\ \varepsilon_{t+1-q} \end{pmatrix}}_{=\mathbf{X}_{t+1}} = \underbrace{\begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & 0 \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix}}_{=\mathbf{F}} \underbrace{\begin{pmatrix} \varepsilon_{t-1} \\ \varepsilon_{t-2} \\ \varepsilon_{t-3} \\ \vdots \\ \varepsilon_{t-q} \end{pmatrix}}_{=\mathbf{X}_t} + \underbrace{\begin{pmatrix} \varepsilon_t \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}}_{=\mathbf{V}_t}$$

$$Y_t = \underbrace{(\theta_1 \quad \cdots \quad \theta_q)}_{=\mathbf{G}} \mathbf{X}_t + \underbrace{\varepsilon_t}_{=\mathbf{W}_t}$$

PŘÍKLAD 2.3.3.ARMA PROCES ŘÁDU p, q

$$ARMA(p, q) : \boxed{Y_t = \varphi_1 Y_{t-1} + \dots + \varphi_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}},$$

kde

$$\varepsilon_t \sim WN(0, \sigma_\varepsilon^2).$$

$$\underbrace{\begin{pmatrix} Y_t \\ Y_{t-1} \\ Y_{t-2} \\ \vdots \\ \vdots \\ Y_{t+1-p} \\ \varepsilon_{t+1} \\ \varepsilon_t \\ \vdots \\ \varepsilon_{t+1-q} \end{pmatrix}}_{=\mathbf{X}_{t+1}} = \underbrace{\begin{pmatrix} \varphi_1 & \cdots & \varphi_{p-1} & \varphi_p & 1 & \theta_1 & \cdots & \theta_{q-1} & \theta_q \\ 1 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & & & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & & & & \vdots \\ \vdots & & \ddots & 1 & \ddots & & & & \vdots \\ \vdots & & & \ddots & 0 & \ddots & & & \vdots \\ \vdots & & & & \ddots & 1 & \ddots & & \vdots \\ \vdots & & & & & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & 1 & 0 \end{pmatrix}}_{=\mathbf{F}} \underbrace{\begin{pmatrix} Y_{t-1} \\ Y_{t-2} \\ Y_{t-3} \\ \vdots \\ \vdots \\ Y_{t-p} \\ \varepsilon_t \\ \varepsilon_{t-1} \\ \vdots \\ \varepsilon_{t-q} \end{pmatrix}}_{=\mathbf{X}_t} + \underbrace{\begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ \vdots \\ 0 \\ \varepsilon_{t+1} \\ 0 \\ \vdots \\ 0 \end{pmatrix}}_{=\mathbf{V}_t}$$

$$Y_t = \underbrace{(\varphi_1 \quad \cdots \quad \varphi_p \quad 1 \quad \theta_1 \quad \cdots \quad \theta_q)}_{=\mathbf{G}} \mathbf{X}_t$$

2.4 Nejlepší lineární predikce pomocí projekce náhodných vektorů druhého řádu

Mějme pravděpodobnostní prostor (Ω, \mathcal{A}, P) . Pro pevně zvolené $v \in \mathbb{N}$ označme

$$L_2^v = \{\mathbf{X} = (X_1, \dots, X_v)' : X_1 \in L^2(\Omega, \mathcal{A}, P), \dots, X_v \in L^2(\Omega, \mathcal{A}, P)\}$$

a označme

$$L_2^\infty = \bigcup_{v=1}^{\infty} L_2^v.$$

Pak lze nad tímto prostorem definovat *skalární součin* pro $\mathbf{X} \in L_2^v$ a $\mathbf{Y} \in L_2^w$ ($v, w \in \mathbb{N}$) předpisem

$$\langle \mathbf{X}, \mathbf{Y} \rangle = E\mathbf{X}\mathbf{Y}'$$

za předpokladu, že existuje sdružené rozdělení náhodného vektoru $\mathbf{Z} = \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix}$ a $\mathbf{Z} \in L_2^{v+w}$.

Označme pro $\mathbf{Y}_0, \dots, \mathbf{Y}_t \in L_2^w$

$$\mathcal{M}_t = \overline{\text{sp}}\{\mathbf{Y}_0, \dots, \mathbf{Y}_t\}$$

uzavřený podprostor generovaný všemi možnými lineárními kombinacemi typu

$$\mathbf{C}_0 \mathbf{Y}_0 + \dots + \mathbf{C}_t \mathbf{Y}_t,$$

kde $\mathbf{C}_0, \dots, \mathbf{C}_t$ jsou reálné matice.

Pak uvažujme nad L_2^∞ **projekci** $\mathbf{X} \in L_2^v$ do \mathcal{M}_t

$$P_{\mathcal{M}_t}(\mathbf{X}) = (P_{\mathcal{M}_t}(X_1), \dots, P_{\mathcal{M}_t}(X_v))',$$

kteřou budeme značit různými způsoby, a to

$$P_{\mathcal{M}_t}(\mathbf{X}) = \hat{\mathbf{X}} = P_t(\mathbf{X}) = P_t(\mathbf{X}|\mathbf{Y}_0, \dots, \mathbf{Y}_t).$$

Připomeňme vlastnosti predikce, které v následujících důkazech využijeme

- (a) vždy existuje jediný vektor $P_t(\mathbf{X})$ takový, že pro $\forall \mathbf{Y} \in \mathcal{M}_t = \overline{\text{sp}}\{\mathbf{Y}_0, \dots, \mathbf{Y}_t\}$ platí

$$\langle \mathbf{X} - \hat{\mathbf{X}}, \mathbf{Y} \rangle = \mathbf{0} \Leftrightarrow \langle \mathbf{X}, \mathbf{Y} \rangle = \langle \hat{\mathbf{X}}, \mathbf{Y} \rangle \Leftrightarrow E\mathbf{X}\mathbf{Y}' = E\hat{\mathbf{X}}\mathbf{Y}'.$$

Protože $\hat{\mathbf{X}} \in \mathcal{M}_t$, pak

$$E\mathbf{X}\hat{\mathbf{X}}' = E\hat{\mathbf{X}}\hat{\mathbf{X}}'.$$

- (b) Jestliže $\mathbf{X}, \mathbf{Y}_1, \dots, \mathbf{Y}_t$ mají sdružené normální rozdělení, pak (pokud $\mathbf{Y}_0 = \mathbf{1} = (1, \dots, 1)'$) platí

$$\hat{\mathbf{X}} = P_t(\mathbf{X}) = E(\mathbf{X}|\mathbf{Y}_1, \dots, \mathbf{Y}_t) \quad t \geq 1.$$

- (c) Predikce $\hat{\mathbf{X}} = P_t(\mathbf{X})$ je **lineární** v tom smyslu, že pro libovolnou matici $\mathbf{A} \in \mathbb{R}^{k+v}$ a $\mathbf{X}, \mathbf{Z} \in L_2^v$ platí:

$$\begin{aligned} P_t(\mathbf{A}\mathbf{X}) &= \mathbf{A}P_t(\mathbf{X}) \\ P_t(\mathbf{X} + \mathbf{Z}) &= P_t(\mathbf{X}) + P_t(\mathbf{Z}) \end{aligned}$$

- (d) Pokud $\mathbf{X} \in L_2^v$ a $\mathbf{Y} \in L_2^w$, pak

$$P_t(\mathbf{X}|\mathbf{Y}) = \mathbf{M}\mathbf{Y},$$

kde $\mathbf{M} \in \mathbb{R}^{v+w}$, pro níž platí

$$\mathbf{M} = E\mathbf{X}\mathbf{Y}' [E\mathbf{Y}\mathbf{Y}']^{-}$$

a \mathbf{A}^{-} značí *pseudoinverzní matici* k matici \mathbf{A} , což je taková matice, pro níž

$$\mathbf{A}^{-} = \mathbf{A}\mathbf{A}^{-}\mathbf{A}$$

Každá matice má alespoň jednu pseudoinverzní matici. Pokud matice \mathbf{A} je regulární, pak

$$\mathbf{A}^{-} = \mathbf{A}^{-1}.$$

Připomeňme opět definici stavového modelu

DATOVÁ ROVNICE

$$\mathbf{Y}_t = \mathbf{G}_t \mathbf{X}_t + \mathbf{W}_t \quad t = 1, 2, 3, \dots$$

STAVOVÁ ROVNICE

$$\mathbf{X}_{t+1} = \mathbf{F}_t \mathbf{X}_t + \mathbf{V}_t \quad t = 1, 2, 3, \dots$$

příčemž

- \mathbf{X}_t ... je tzv. *stavový* v -rozměrný náhodný vektor
- \mathbf{W}_t ... je *šum měření*
- \mathbf{V}_t ... je *šum procesu*
- \mathbf{G}_t ... je posloupnost matic typu $w \times v$
- \mathbf{F}_t ... je posloupnost matic typu $v \times v$

Dále platí

$$\begin{aligned} E\mathbf{V}_t &= \mathbf{0} \\ E\mathbf{W}_t &= \mathbf{0} \\ D \begin{pmatrix} \mathbf{W}_t \\ \mathbf{V}_t \end{pmatrix} &= E \begin{pmatrix} \mathbf{W}_t \\ \mathbf{V}_t \end{pmatrix} (\mathbf{W}_t' \quad \mathbf{V}_t') = \begin{pmatrix} E\mathbf{W}_t \mathbf{W}_t' & E\mathbf{W}_t \mathbf{V}_t' \\ E\mathbf{V}_t \mathbf{W}_t' & E\mathbf{V}_t \mathbf{V}_t' \end{pmatrix} = \begin{pmatrix} \mathbf{R}_t & \mathbf{S}_t \\ \mathbf{S}_t' & \mathbf{Q}_t \end{pmatrix} \end{aligned}$$

$C(\mathbf{X}_t, (\mathbf{W}_t', \mathbf{V}_t)') = \mathbf{0}$, tj. jsou nekorelované: $\mathbf{X}_t \perp (\mathbf{W}_t', \mathbf{V}_t)'$

Označme

$$\begin{aligned} \hat{\mathbf{X}}_{t|k} &= P_{\overline{sp}\{\mathbf{Y}_0, \dots, \mathbf{Y}_k\}}(\mathbf{X}_t) = P(\mathbf{X}_t | \mathbf{Y}_0, \dots, \mathbf{Y}_k) \\ \mathbf{\Omega}_{t|k} &= E(\mathbf{X}_t - \hat{\mathbf{X}}_{t|k})(\mathbf{X}_t - \hat{\mathbf{X}}_{t|k})' \end{aligned}$$

Pokud $k = t - 1$... jde o tzv. problém (jednokrokové) *predikce*
 $k = t$... jde o tzv. problém *filtrace*
 $k = n > t$... jde o tzv. problém *vyhlazení*

Přidejme předpoklady pro $\forall t$

- $\mathbf{W}_t \perp \{\mathbf{Y}_0, \dots, \mathbf{Y}_{t-1}\}$, tj. jsou nekorelované
- $\mathbf{V}_t \perp \{\mathbf{Y}_0, \dots, \mathbf{Y}_{t-1}\}$
- $\mathbf{S}_t = \mathbf{0}$ (tj. šumy procesu \mathbf{V}_t a měření \mathbf{W}_t jsou nekorelované)

Věta 2.4.1.

JEDNOKROKOVÁ KALMANOVA PREDIKCE

$$\boxed{\widehat{\mathbf{X}}_t} = \widehat{\mathbf{X}}_{t|t-1} = P_{t-1}(\mathbf{X}_t) = P(\mathbf{X}_t | \mathbf{Y}_0, \dots, \mathbf{Y}_{t-1}) = P_{\overline{sp}\{\mathbf{Y}_0, \dots, \mathbf{Y}_{t-1}\}}(\mathbf{X}_t)$$

a CHYBOVÁ PREDIKČNÍ KOVARIANČNÍ MATICE

$$\boxed{\Omega_{t|t-1}} = E(\mathbf{X}_t - \widehat{\mathbf{X}}_t)(\mathbf{X}_t - \widehat{\mathbf{X}}_t)' = E(\mathbf{X}_t - \widehat{\mathbf{X}}_{t|t-1})(\mathbf{X}_t - \widehat{\mathbf{X}}_{t|t-1})'$$

jsou jednoznačně určeny

(1) počátečními podmínkami:

$$\begin{aligned} \boxed{\widehat{\mathbf{X}}_1} &= \widehat{\mathbf{X}}_{1|0} = P(\mathbf{X}_1 | \mathbf{Y}_0) = P_{\overline{sp}\{\mathbf{Y}_0\}}(\mathbf{X}_1) \\ \boxed{\Omega_{1|0}} &= \Omega_1 = \Sigma_{\mathbf{X}_1 \mathbf{X}_1} - \Sigma_{\widehat{\mathbf{X}}_1 \widehat{\mathbf{X}}_1} \end{aligned}$$

kde

$$\begin{aligned} \Sigma_{\mathbf{X}_1 \mathbf{X}_1} &= E\mathbf{X}_1 \mathbf{X}_1' \\ \Sigma_{\widehat{\mathbf{X}}_1 \widehat{\mathbf{X}}_1} &= E\widehat{\mathbf{X}}_1 \widehat{\mathbf{X}}_1' \end{aligned}$$

(2) a platí pro ně následující rekurentní vztahy:

$$\boxed{\widehat{\mathbf{X}}_{t+1}} = \widehat{\mathbf{X}}_{t+1|t} = \mathbf{F}_t \widehat{\mathbf{X}}_{t|t-1} + \mathbf{K}_{t+1|t} (\mathbf{Y}_t - \mathbf{G}_t \widehat{\mathbf{X}}_{t|t-1}),$$

kde $\mathbf{K}_{t+1|t}$ je tzv. PREDIKČNÍ KALMANŮV ZISK, pro nějž platí:

$$\mathbf{K}_{t+1|t} = \Sigma_{\mathbf{X}_{t+1} \mathbf{I}_t} \Sigma_{\mathbf{I}_t \mathbf{I}_t}^{-1}$$

přičemž

$$\begin{aligned} \Sigma_{\mathbf{X}_{t+1} \mathbf{I}_t} &= \mathbf{F}_t \Omega_{t|t-1} \mathbf{G}_t' \\ \Sigma_{\mathbf{I}_t \mathbf{I}_t} &= \mathbf{G}_t \Omega_{t|t-1} \mathbf{G}_t' + \mathbf{R}_t \end{aligned}$$

a

$$\boxed{\Omega_{t+1|t}} = \Sigma_{\mathbf{X}_{t+1} \mathbf{X}_{t+1}} - \Sigma_{\widehat{\mathbf{X}}_{t+1} \widehat{\mathbf{X}}_{t+1}}$$

přičemž

$$\begin{aligned} \Sigma_{\mathbf{X}_{t+1} \mathbf{X}_{t+1}} &= \mathbf{F}_t \Sigma_{\mathbf{X}_t \mathbf{X}_t} \mathbf{F}_t' + \mathbf{Q}_t \\ \Sigma_{\widehat{\mathbf{X}}_{t+1} \widehat{\mathbf{X}}_{t+1}} &= \mathbf{F}_t \Sigma_{\widehat{\mathbf{X}}_t \widehat{\mathbf{X}}_t} \mathbf{F}_t' + \mathbf{K}_{t+1|t} \Sigma_{\mathbf{I}_t \mathbf{I}_t} \mathbf{K}_{t+1|t}' \end{aligned}$$

a kde \mathbf{I}_t jsou INOVACE pro \mathbf{Y}_t , tj.

$$\mathbf{I}_t = \mathbf{Y}_t - \widehat{\mathbf{Y}}_t = \mathbf{Y}_t - P_{\overline{sp}\{\mathbf{Y}_0, \dots, \mathbf{Y}_{t-1}\}}(\mathbf{Y}_t)$$

DŮKAZ. Nejprve definujme inovaci pro \mathbf{Y}_t

$$\begin{aligned}
 \mathbf{I}_0 &= \mathbf{Y}_0 \\
 \mathbf{I}_t &= \mathbf{Y}_t - \widehat{\mathbf{Y}}_t \\
 &= \mathbf{Y}_t - P_{\overline{sp}\{\mathbf{Y}_0, \dots, \mathbf{Y}_{t-1}\}}(\mathbf{Y}_t) \\
 &= \mathbf{Y}_t - P_{t-1}(\mathbf{Y}_t) \\
 &= \mathbf{Y}_t - P_{t-1}(\mathbf{G}_t \mathbf{X}_t + \mathbf{W}_t) \\
 &= \mathbf{Y}_t - \mathbf{G}_t P_{t-1}(\mathbf{X}_t) - P_{t-1}(\mathbf{W}_t)
 \end{aligned}$$

Díky nezávislosti náhodných vektorů

$$\mathbf{W}_t \perp \{\mathbf{Y}_0, \dots, \mathbf{Y}_{t-1}\}$$

platí

$$P_{t-1}(\mathbf{W}_t) = P_{\overline{sp}\{\mathbf{Y}_0, \dots, \mathbf{Y}_{t-1}\}}(\mathbf{W}_t) = 0,$$

takže dostaneme

$$\boxed{\mathbf{I}_t} = \boxed{\mathbf{Y}_t - \mathbf{G}_t \widehat{\mathbf{X}}_t} = \mathbf{G}_t \mathbf{X}_t + \mathbf{W}_t - \mathbf{G}_t \widehat{\mathbf{X}}_t = \boxed{\mathbf{G}_t (\mathbf{X}_t - \widehat{\mathbf{X}}_t) + \mathbf{W}_t}.$$

Je třeba si uvědomit, že inovace jsou ortogonální (tj. nekorelované)

$$\mathbf{I}_0 \perp \mathbf{I}_1 \perp \mathbf{I}_2 \perp \dots \perp \mathbf{I}_t,$$

takže pro libovolné \mathbf{X} platí

$$\begin{aligned}
 P_t(\mathbf{X}) &= P(\mathbf{X} | \mathbf{Y}_0, \dots, \mathbf{Y}_t) \\
 &= P(\mathbf{X} | \mathbf{I}_0, \dots, \mathbf{I}_t) \\
 &= P(\mathbf{X} | \mathbf{I}_0, \dots, \mathbf{I}_{t-1}) + P(\mathbf{X} | \mathbf{I}_t) = P_{t-1}(\mathbf{X}) + P(\mathbf{X} | \mathbf{I}_t) \\
 &= P_{t-1}(\mathbf{X}) + \mathbf{M} \mathbf{I}_t,
 \end{aligned}$$

kde

$$\mathbf{M} = EX \mathbf{I}'_t [E \mathbf{I}_t \mathbf{I}'_t]^{-1}.$$

Takže

$$\begin{aligned}
 \widehat{\mathbf{X}}_{t+1} &= \widehat{\mathbf{X}}_{t+1|t} = P_t(\mathbf{X}_{t+1}) \\
 &= P_{t-1}(\mathbf{X}_{t+1}) + P(\mathbf{X}_{t+1} | \mathbf{I}_t) \\
 &= P_{t-1}(\mathbf{F}_t \mathbf{X}_t + \mathbf{V}_t) + EX_{t+1} \mathbf{I}'_t [E \mathbf{I}_t \mathbf{I}'_t]^{-1} \mathbf{I}_t
 \end{aligned}$$

a označíme-li

$$\begin{aligned}
 \Sigma_{\mathbf{X}_{t+1} \mathbf{I}_t} &= EX_{t+1} \mathbf{I}'_t \\
 \Sigma_{\mathbf{I}_t \mathbf{I}_t} &= E \mathbf{I}_t \mathbf{I}'_t,
 \end{aligned}$$

pak

$$\widehat{\mathbf{X}}_{t+1} = \mathbf{F}_t \underbrace{P_{t-1}(\mathbf{X}_t)}_{=\widehat{\mathbf{X}}_{t|t-1}} + \underbrace{P_{t-1}(\mathbf{V}_t)}_{=0} + \Sigma_{\mathbf{X}_{t+1}\mathbf{I}_t} \Sigma_{\mathbf{I}_t\mathbf{I}_t}^{-1} (\mathbf{Y}_t - \mathbf{G}_t \widehat{\mathbf{X}}_t)$$

Vyjádřeme nyní

$$\begin{aligned} \boxed{\Sigma_{\mathbf{X}_{t+1}\mathbf{I}_t}} &= E\mathbf{X}_{t+1}\mathbf{I}_t' \\ &= E(\mathbf{F}_t\mathbf{X}_t + \mathbf{V}_t) \left[\mathbf{G}_t(\mathbf{X}_t - \widehat{\mathbf{X}}_t) + \mathbf{W}_t \right]' \\ &= E \left[\mathbf{F}_t(\mathbf{X}_t - \widehat{\mathbf{X}}_t) + \mathbf{F}_t\widehat{\mathbf{X}}_t + \mathbf{V}_t \right] \left[\mathbf{G}_t(\mathbf{X}_t - \widehat{\mathbf{X}}_t) + \mathbf{W}_t \right]' \\ &= \mathbf{F}_t \underbrace{E(\mathbf{X}_t - \widehat{\mathbf{X}}_t)(\mathbf{X}_t - \widehat{\mathbf{X}}_t)'}_{=\Omega_{t|t-1}} \mathbf{G}_t' + \mathbf{F}_t \underbrace{E(\mathbf{X}_t - \widehat{\mathbf{X}}_t)\mathbf{W}_t'}_{=0(\text{nekorel.})} + \mathbf{F}_t \underbrace{E\widehat{\mathbf{X}}_t(\mathbf{X}_t - \widehat{\mathbf{X}}_t)'}_{=0(\text{nekorel.})} \\ &\quad + \mathbf{F}_t \underbrace{E\widehat{\mathbf{X}}_t\mathbf{W}_t'}_{=0(\text{nekorel.})} + \underbrace{E\mathbf{V}_t(\mathbf{X}_t - \widehat{\mathbf{X}}_t)'}_{=0(\text{nekorel.})} \mathbf{G}_t' + \underbrace{E\mathbf{V}_t\mathbf{W}_t'}_{=\mathbf{S}_t=0} \\ &= \boxed{\mathbf{F}_t\Omega_{t|t-1}\mathbf{G}_t'} \end{aligned}$$

Dále počítejme

$$\begin{aligned} \boxed{\Sigma_{\mathbf{I}_t\mathbf{I}_t}} &= E\mathbf{I}_t\mathbf{I}_t' = E \left[\mathbf{G}_t(\mathbf{X}_t - \widehat{\mathbf{X}}_t) + \mathbf{W}_t \right] \left[\mathbf{G}_t(\mathbf{X}_t - \widehat{\mathbf{X}}_t) + \mathbf{W}_t \right]' \\ &= \mathbf{G}_t E(\mathbf{X}_t - \widehat{\mathbf{X}}_t)(\mathbf{X}_t - \widehat{\mathbf{X}}_t)' \mathbf{G}_t' \\ &\quad + \mathbf{G}_t \underbrace{E(\mathbf{X}_t - \widehat{\mathbf{X}}_t)\mathbf{W}_t'}_{=0(\text{nekorel.})} + \underbrace{E\mathbf{W}_t(\mathbf{X}_t - \widehat{\mathbf{X}}_t)'}_{=0(\text{nekorel.})} \mathbf{G}_t' + E\mathbf{W}_t\mathbf{W}_t' \\ &= \boxed{\mathbf{G}_t\Omega_{t|t-1}\mathbf{G}_t' + \mathbf{R}_t} \end{aligned}$$

Tedy celkově máme

$$\widehat{\mathbf{X}}_{t+1} = \widehat{\mathbf{X}}_{t+1|t} = \mathbf{F}_t\widehat{\mathbf{X}}_{t|t-1} + \Sigma_{\mathbf{X}_{t+1}\mathbf{I}_t} \Sigma_{\mathbf{I}_t\mathbf{I}_t}^{-1} \underbrace{(\mathbf{Y}_t - \mathbf{G}_t\widehat{\mathbf{X}}_t)}_{=\mathbf{I}_t}$$

a

$$\mathbf{K}_{t+1|t} = \Sigma_{\mathbf{X}_{t+1}\mathbf{I}_t} \Sigma_{\mathbf{I}_t\mathbf{I}_t}^{-1}$$

je tzv. KALMANŮV PREDIKČNÍ ZISK a můžeme tedy psát

$$\boxed{\widehat{\mathbf{X}}_{t+1} = \mathbf{F}_t\widehat{\mathbf{X}}_{t|t-1} + \mathbf{K}_{t+1|t}(\mathbf{Y}_t - \mathbf{G}_t\widehat{\mathbf{X}}_{t|t-1})}$$

Zbývá najít rekurentní vztah pro $\Omega_{t+1|t}$. Přitom využijeme důležitý vztah, který vychází z vlastností ortogonální projekce, tj. že pro $\forall \mathbf{Y} \in P_{\overline{sp}\{\mathbf{Y}_0, \dots, \mathbf{Y}_{t-1}\}}$ platí

$$\begin{aligned} \langle \mathbf{X}_t - \widehat{\mathbf{X}}_t, \mathbf{Y} \rangle &= \mathbf{0} \\ \langle \mathbf{X}_t, \mathbf{Y} \rangle &= \langle \widehat{\mathbf{X}}_t, \mathbf{Y} \rangle \\ E\mathbf{X}_t\mathbf{Y} &= E\widehat{\mathbf{X}}_t\mathbf{Y} \end{aligned}$$

a protože $\widehat{\mathbf{X}}_t \in P_{sp\{\mathbf{Y}_0, \dots, \mathbf{Y}_{t-1}\}}$, dostaneme

$$\boxed{E\mathbf{X}_t\widehat{\mathbf{X}}_t = E\widehat{\mathbf{X}}_t\widehat{\mathbf{X}}_t'}$$

Proto počítejme

$$\begin{aligned}\Omega_{1|0} &= \Omega_1 = E(\mathbf{X}_1 - \widehat{\mathbf{X}}_1)(\mathbf{X}_1 - \widehat{\mathbf{X}}_1)' \\ &= E\mathbf{X}_1\mathbf{X}_1' - \underbrace{E\widehat{\mathbf{X}}_1\mathbf{X}_1'}_{=E\widehat{\mathbf{X}}_1\widehat{\mathbf{X}}_1'} - \underbrace{E\mathbf{X}_1\widehat{\mathbf{X}}_1}_{=E\widehat{\mathbf{X}}_1\widehat{\mathbf{X}}_1'} + E\widehat{\mathbf{X}}_1\widehat{\mathbf{X}}_1' \\ &= E\mathbf{X}_1\mathbf{X}_1' - E\widehat{\mathbf{X}}_1\widehat{\mathbf{X}}_1' = \Sigma_{\mathbf{X}_1\mathbf{X}_1} - \Sigma_{\widehat{\mathbf{X}}_1\widehat{\mathbf{X}}_1}\end{aligned}$$

Úplnou matematickou indukcí obdobně dokážeme, že pokud budeme předpokládat, že platí

$$\Omega_{t|t-1} = \Sigma_{\mathbf{X}_t\mathbf{X}_t} - \Sigma_{\widehat{\mathbf{X}}_t\widehat{\mathbf{X}}_t},$$

pak

$$\begin{aligned}\Omega_{t+1|t} &= E(\mathbf{X}_{t+1} - \widehat{\mathbf{X}}_{t+1})(\mathbf{X}_{t+1} - \widehat{\mathbf{X}}_{t+1})' \\ &= E\mathbf{X}_{t+1}\mathbf{X}_{t+1}' - \underbrace{E\widehat{\mathbf{X}}_{t+1}\mathbf{X}_{t+1}'}_{=E\widehat{\mathbf{X}}_{t+1}\widehat{\mathbf{X}}_{t+1}'} - \underbrace{E\mathbf{X}_{t+1}\widehat{\mathbf{X}}_{t+1}}_{=E\widehat{\mathbf{X}}_{t+1}\widehat{\mathbf{X}}_{t+1}'} + E\widehat{\mathbf{X}}_{t+1}\widehat{\mathbf{X}}_{t+1}' \\ &= E\mathbf{X}_{t+1}\mathbf{X}_{t+1}' - E\widehat{\mathbf{X}}_{t+1}\widehat{\mathbf{X}}_{t+1}' = \Sigma_{\mathbf{X}_{t+1}\mathbf{X}_{t+1}} - \Sigma_{\widehat{\mathbf{X}}_{t+1}\widehat{\mathbf{X}}_{t+1}}\end{aligned}$$

$$\begin{aligned}\Sigma_{\mathbf{X}_{t+1}\mathbf{X}_{t+1}} &= E\mathbf{X}_{t+1}\mathbf{X}_{t+1}' = E(\mathbf{F}_t\mathbf{X}_t + \mathbf{W}_t)(\mathbf{F}_t\mathbf{X}_t + \mathbf{W}_t)' \\ &= \mathbf{F}_t \underbrace{E\mathbf{X}_t\mathbf{X}_t'}_{=\Sigma_{\mathbf{X}_t\mathbf{X}_t}} \mathbf{F}_t' + \mathbf{F}_t \underbrace{E\mathbf{X}_t\mathbf{W}_t'}_{=\mathbf{0}(\text{nekorel.})} + \underbrace{E\mathbf{W}_t\mathbf{X}_t'}_{=\mathbf{0}(\text{nekorel.})} \mathbf{F}_t' + E\mathbf{W}_t\mathbf{W}_t' \\ &= \mathbf{F}_t\Sigma_{\mathbf{X}_t\mathbf{X}_t}\mathbf{F}_t' + \mathbf{Q}_t\end{aligned}$$

$$\begin{aligned}\Sigma_{\widehat{\mathbf{X}}_{t+1}\widehat{\mathbf{X}}_{t+1}} &= E\widehat{\mathbf{X}}_{t+1}\widehat{\mathbf{X}}_{t+1}' \\ &= E \left[\mathbf{F}_t\widehat{\mathbf{X}}_t + \mathbf{K}_{t+1|t}(\mathbf{Y}_t - \mathbf{G}_t\widehat{\mathbf{X}}_t) \right] \left[\mathbf{F}_t\widehat{\mathbf{X}}_t + \mathbf{K}_{t+1|t}(\mathbf{Y}_t - \mathbf{G}_t\widehat{\mathbf{X}}_t) \right]' \\ &= \mathbf{F}_t \underbrace{E\widehat{\mathbf{X}}_t\widehat{\mathbf{X}}_t'}_{\Sigma_{\widehat{\mathbf{X}}_t\widehat{\mathbf{X}}_t}} \mathbf{F}_t' + \mathbf{F}_t \underbrace{E\widehat{\mathbf{X}}_t(\mathbf{Y}_t - \mathbf{G}_t\widehat{\mathbf{X}}_t)'}_{=\mathbf{0}(\text{nekorel.})} \mathbf{K}_{t+1|t}' \\ &\quad + \mathbf{K}_{t+1|t} \underbrace{E(\mathbf{Y}_t - \mathbf{G}_t\widehat{\mathbf{X}}_t)\widehat{\mathbf{X}}_t'}_{=\mathbf{0}(\text{nekorel.})} \\ &\quad + \mathbf{K}_{t+1|t} \underbrace{E(\mathbf{Y}_t - \mathbf{G}_t\widehat{\mathbf{X}}_t)(\mathbf{Y}_t - \mathbf{G}_t\widehat{\mathbf{X}}_t)'}_{=\Sigma_{\mathbf{I}_t\mathbf{I}_t}} \mathbf{K}_{t+1|t}' \\ &= \mathbf{F}_t\Sigma_{\widehat{\mathbf{X}}_t\widehat{\mathbf{X}}_t}\mathbf{F}_t' + \mathbf{K}_{t+1|t}\Sigma_{\mathbf{I}_t\mathbf{I}_t}\mathbf{K}_{t+1|t}'\end{aligned}$$

■

Věta 2.4.2.

Pro KALMANOVU FILTRACI

$$\boxed{\widehat{\mathbf{X}}_{t|t}} = P(\mathbf{X}_t | \mathbf{Y}_0, \dots, \mathbf{Y}_t)$$

a FILTROVACÍ CHYBOVOU KOVARIANČNÍ MATICÍ

$$\boxed{\boldsymbol{\Omega}_{t|t}} = E(\mathbf{X}_t - \widehat{\mathbf{X}}_{t|t})(\mathbf{X}_t - \widehat{\mathbf{X}}_{t|t})'$$

pro $\forall t \geq 1$ platí následující rekurentní vztahy

$$\boxed{\widehat{\mathbf{X}}_{t|t} = \widehat{\mathbf{X}}_{t|t-1} + \mathbf{K}_{t|t}(\mathbf{Y}_t - \mathbf{G}_t \widehat{\mathbf{X}}_{t|t-1})},$$

kde $\mathbf{K}_{t|t}$ je tzv. FILTRAČNÍ KALMANŮV ZISK, pro nějž platí

$$\boxed{\mathbf{K}_{t|t} = \boldsymbol{\Sigma}_{\mathbf{X}_t \mathbf{I}_t} \boldsymbol{\Sigma}_{\mathbf{I}_t \mathbf{I}_t}^{-1}},$$

přičemž

$$\begin{aligned} \boldsymbol{\Sigma}_{\mathbf{X}_t \mathbf{I}_t} &= \boldsymbol{\Omega}_{t|t-1} \mathbf{G}_t' \\ \boldsymbol{\Sigma}_{\mathbf{I}_t \mathbf{I}_t} &= \mathbf{G}_t \boldsymbol{\Omega}_{t|t-1} \mathbf{G}_t' + \mathbf{R}_t \end{aligned}$$

a

$$\boxed{\boldsymbol{\Omega}_{t|t} = (\mathbf{I} - \mathbf{K}_{t|t} \mathbf{G}_t) \boldsymbol{\Omega}_{t|t-1}},$$

kde \mathbf{I} je jednotková matice řádu $v \times v$.

DŮKAZ. Využijme opět inovací

$$\begin{aligned} \mathbf{I}_0 &= \mathbf{Y}_0 \\ \mathbf{I}_t &= \mathbf{Y}_t - \widehat{\mathbf{Y}}_t \\ &= \mathbf{Y}_t - P_{\overline{\text{sp}}\{\mathbf{Y}_0, \dots, \mathbf{Y}_{t-1}\}}(\mathbf{Y}_t) \\ &= \mathbf{Y}_t - P_{t-1}(\mathbf{Y}_t) \\ &= \mathbf{Y}_t - P_{t-1}(\mathbf{G}_t \mathbf{X}_t + \mathbf{W}_t) \\ &= \mathbf{Y}_t - \mathbf{G}_t P_{t-1}(\mathbf{X}_t) - P_{t-1}(\mathbf{W}_t) \\ &= \mathbf{Y}_t - \mathbf{G}_t \widehat{\mathbf{X}}_t \\ &= \mathbf{G}_t(\mathbf{X}_t - \widehat{\mathbf{X}}_t) + \mathbf{W}_t \end{aligned}$$

které jsou navzájem kolmé (tj. nekolerované).

Počítejme

$$\begin{aligned}
\widehat{\mathbf{X}}_{t|t} &= P(\mathbf{X}_t | \mathbf{Y}_0, \dots, \mathbf{Y}_t) \\
&= P(\mathbf{X}_t | \mathbf{I}_0, \dots, \mathbf{I}_t) \\
&= P(\mathbf{X}_t | \mathbf{I}_0, \dots, \mathbf{I}_{t-1}) + P(\mathbf{X}_t | \mathbf{I}_t) \\
&= \widehat{\mathbf{X}}_t + \mathbf{M}\mathbf{I}_t \\
&= \widehat{\mathbf{X}}_t + E\mathbf{X}_t\mathbf{I}_t' [E\mathbf{X}_t\mathbf{I}_t']^{-1} \mathbf{I}_t \\
&= \widehat{\mathbf{X}}_t + \Sigma_{\mathbf{X}_t\mathbf{I}_t} \Sigma_{\mathbf{I}_t\mathbf{I}_t}^{-1} (\mathbf{Y}_t - \mathbf{G}_t\widehat{\mathbf{X}}_t)
\end{aligned}$$

přičemž

$$\begin{aligned}
\Sigma_{\mathbf{X}_t\mathbf{I}_t} &= E\mathbf{X}_t\mathbf{I}_t' \\
&= E\mathbf{X}_t \left[\mathbf{G}_t(\mathbf{X}_t - \widehat{\mathbf{X}}_t) + \mathbf{W}_t \right]' \\
&= E \left[(\mathbf{X}_t - \widehat{\mathbf{X}}_t) + \widehat{\mathbf{X}}_t \right] \left[\mathbf{G}_t(\mathbf{X}_t - \widehat{\mathbf{X}}_t) + \mathbf{W}_t \right]' \\
&= E(\mathbf{X}_t - \widehat{\mathbf{X}}_t)(\mathbf{X}_t - \widehat{\mathbf{X}}_t)' \mathbf{G}_t + \underbrace{E(\mathbf{X}_t - \widehat{\mathbf{X}}_t)\mathbf{W}_t'}_{=\mathbf{0}(\text{nekorel.})} \\
&\quad + \underbrace{E\widehat{\mathbf{X}}_t(\mathbf{X}_t - \widehat{\mathbf{X}}_t)'}_{=\mathbf{0}(\text{nekorel.})} \mathbf{G}_t' + \underbrace{E\widehat{\mathbf{X}}_t\mathbf{W}_t'}_{=\mathbf{0}(\text{nekorel.})} = \Omega_{t|t-1} \mathbf{G}_t'
\end{aligned}$$

Takže celkově dostaneme

$$\widehat{\mathbf{X}}_{t|t} = \widehat{\mathbf{X}}_t + \underbrace{\Sigma_{\mathbf{X}_t\mathbf{I}_t} \Sigma_{\mathbf{I}_t\mathbf{I}_t}^{-1}}_{\text{ozn. } \mathbf{K}_{t|t}} \underbrace{(\mathbf{Y}_t - \mathbf{G}_t\widehat{\mathbf{X}}_t)}_{=\mathbf{I}_t},$$

odtud

$$\widehat{\mathbf{X}}_{t|t} - \widehat{\mathbf{X}}_t = \mathbf{K}_{t|t}\mathbf{I}_t.$$

Zbývá dopočítat $\Omega_{t|t}$. Víme, že

$$\begin{aligned}
\Omega_{t|t-1} &= E(\mathbf{X}_t - \widehat{\mathbf{X}}_t)(\mathbf{X}_t - \widehat{\mathbf{X}}_t)' \\
&= E \left[(\mathbf{X}_t - \widehat{\mathbf{X}}_{t|t}) + \underbrace{(\widehat{\mathbf{X}}_{t|t} - \widehat{\mathbf{X}}_{t|t-1})}_{=\mathbf{K}_{t|t}\mathbf{I}_t} \right] \left[(\mathbf{X}_t - \widehat{\mathbf{X}}_{t|t}) + \underbrace{(\widehat{\mathbf{X}}_{t|t} - \widehat{\mathbf{X}}_{t|t-1})}_{=\mathbf{K}_{t|t}\mathbf{I}_t} \right]' \\
&= \underbrace{E(\mathbf{X}_t - \widehat{\mathbf{X}}_{t|t})(\mathbf{X}_t - \widehat{\mathbf{X}}_{t|t})'}_{\Omega_{t|t}} + \underbrace{E(\mathbf{X}_t - \widehat{\mathbf{X}}_{t|t})\mathbf{I}_t' \mathbf{K}_{t|t}'}_{=\mathbf{0}(\text{nekorel.})} \\
&\quad + \mathbf{K}_{t|t} \underbrace{E\mathbf{I}_t(\mathbf{X}_t - \widehat{\mathbf{X}}_{t|t})'}_{=\mathbf{0}(\text{nekorel.})} + \mathbf{K}_{t|t} E\mathbf{I}_t\mathbf{I}_t' \mathbf{K}_{t|t}' = \Omega_{t|t} + \mathbf{K}_{t|t} \Sigma_{\mathbf{I}_t\mathbf{I}_t} \mathbf{K}_{t|t}'
\end{aligned}$$

Protože

$$\mathbf{K}_{t|t} = \Sigma_{\mathbf{X}_t \mathbf{I}_t} \Sigma_{\mathbf{I}_t \mathbf{I}_t}^{-1} = \Omega_{t|t-1} \mathbf{G}'_t \Sigma_{\mathbf{I}_t \mathbf{I}_t}^{-1}$$

dostáváme

$$\begin{aligned} \Omega_{t|t-1} &= \Omega_{t|t} + \Sigma_{\mathbf{X}_t \mathbf{I}_t} \Sigma_{\mathbf{I}_t \mathbf{I}_t}^{-1} \Sigma_{\mathbf{I}_t \mathbf{I}_t} \Sigma_{\mathbf{I}_t \mathbf{I}_t}^{-1} \Sigma'_{\mathbf{X}_t \mathbf{I}_t} \\ &= \Omega_{t|t} + \Sigma_{\mathbf{X}_t \mathbf{I}_t} \Sigma_{\mathbf{I}_t \mathbf{I}_t}^{-1} \Sigma'_{\mathbf{X}_t \mathbf{I}_t} = \Omega_{t|t} + \underbrace{\Omega_{t|t-1} \mathbf{G}'_t \Sigma_{\mathbf{I}_t \mathbf{I}_t}^{-1} \mathbf{G}_t}_{=\mathbf{K}_{t|t}} \Omega_{t|t-1} \end{aligned}$$

Odtud

$$\Omega_{t|t} = \Omega_{t|t-1} - \mathbf{K}_{t|t} \mathbf{G}_t \Omega_{t|t-1} = (\mathbf{I} - \mathbf{K}_{t|t} \mathbf{G}_t) \Omega_{t|t-1}.$$

■

2.5 Kalmanův iterační proces

Shrňme předchozí výsledky Kalmanovy predikce a filtrace takto

1) Protože

$$\left. \begin{aligned} \mathbf{K}_{t+1|t} &= \Sigma_{\mathbf{X}_{t+1} \mathbf{I}_t} \Sigma_{\mathbf{I}_t \mathbf{I}_t}^{-1} = \mathbf{F}_t \Omega_{t|t-1} \mathbf{G}'_t \Sigma_{\mathbf{I}_t \mathbf{I}_t}^{-1} \\ \mathbf{K}_{t|t} &= \Sigma_{\mathbf{X}_t \mathbf{I}_t} \Sigma_{\mathbf{I}_t \mathbf{I}_t}^{-1} = \Omega_{t|t-1} \mathbf{G}'_t \Sigma_{\mathbf{I}_t \mathbf{I}_t}^{-1} \end{aligned} \right\} \Rightarrow \boxed{\mathbf{K}_{t+1|t} = \mathbf{F}_t \mathbf{K}_{t|t}}.$$

2) Dále platí

$$\begin{aligned} \boxed{\mathbf{F}_t \widehat{\mathbf{X}}_{t|t}} &= \mathbf{F}_t \widehat{\mathbf{X}}_{t|t-1} + \mathbf{F}_t \mathbf{K}_{t|t} (\mathbf{Y}_t - \mathbf{G}_t \widehat{\mathbf{X}}_t) \\ &= \mathbf{F}_t \widehat{\mathbf{X}}_{t|t-1} + \mathbf{K}_{t+1|t} (\mathbf{Y}_t - \mathbf{G}_t \widehat{\mathbf{X}}_t) \\ &= \boxed{\widehat{\mathbf{X}}_{t+1|t}}. \end{aligned}$$

3) Budeme se snažit nově vyjádřit $\Omega_{t+1|t}$. Protože

$$\begin{aligned} \boxed{\mathbf{X}_{t+1} - \widehat{\mathbf{X}}_{t+1|t}} &= \mathbf{F}_t \mathbf{X}_t + \mathbf{V}_t - \mathbf{F}_t \widehat{\mathbf{X}}_{t|t} \\ &= \boxed{\mathbf{F}_t (\mathbf{X}_t - \widehat{\mathbf{X}}_{t|t}) + \mathbf{V}_t} \end{aligned}$$

a vzhledem k tomu, že $\mathbf{V}_t \perp (\mathbf{X}_t - \widehat{\mathbf{X}}_{t|t})$, dostáváme

$$\begin{aligned} \boxed{\Omega_{t+1|t}} &= E(\mathbf{X}_{t+1} - \widehat{\mathbf{X}}_{t+1|t})(\mathbf{X}_{t+1} - \widehat{\mathbf{X}}_{t+1|t})' \\ &= E \left[\mathbf{F}_t (\mathbf{X}_t - \widehat{\mathbf{X}}_{t|t}) + \mathbf{V}_t \right] \left[\mathbf{F}_t (\mathbf{X}_t - \widehat{\mathbf{X}}_{t|t}) + \mathbf{V}_t \right]' \\ &= \mathbf{F}_t E(\mathbf{X}_t - \widehat{\mathbf{X}}_{t|t})(\mathbf{X}_t - \widehat{\mathbf{X}}_{t|t})' \mathbf{F}'_t + E \mathbf{V}_t \mathbf{V}'_t = \boxed{\mathbf{F}_t \Omega_{t|t} \mathbf{F}'_t + \mathbf{Q}_t}. \end{aligned}$$

Všechny předchozí mezivýsledky použijeme pro odvození velmi jednoduchého Kalmanova iteračního procesu, který je spojením filtrace a predikce.

KALMANŮV ITERAČNÍ PROCES

(I) POČÁTEČNÍ PODMÍNKY

$$\begin{aligned}\widehat{\mathbf{X}}_{1|0} &= \widehat{\mathbf{X}}_1 = E\mathbf{X}_1 \quad \text{při} \quad \mathbf{Y}_0 = \mathbf{1} \\ \boldsymbol{\Omega}_{1|0} &= E(\mathbf{X}_1 - E\mathbf{X}_1)(\mathbf{X}_1 - E\mathbf{X}_1)' = D\mathbf{X}_1\end{aligned}$$

(II) DATOVÝ (FILTRAČNÍ) KROK KALMANOVA FILTRU

Nejprve se spočítá tzv. KALMANŮV ZISK (nebo též *Kalmanovo zesílení*)

$$\mathbf{K}_{t|t} = \boldsymbol{\Omega}_{t|t-1} \mathbf{G}'_t (\mathbf{G}_t \boldsymbol{\Omega}_{t|t-1} \mathbf{G}'_t + \mathbf{R}_t)^{-},$$

pak

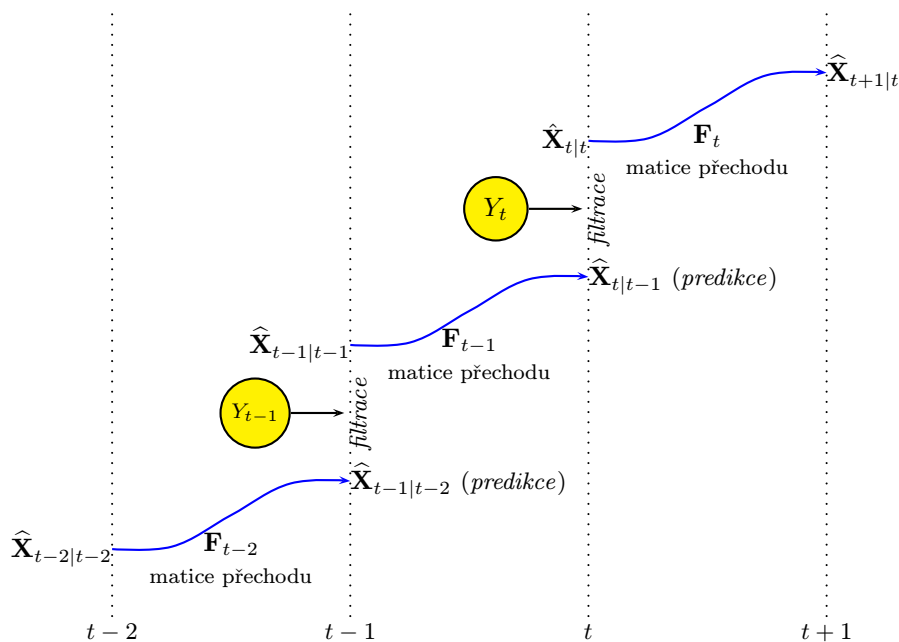
$$\widehat{\mathbf{X}}_{t|t} = \widehat{\mathbf{X}}_{t|t-1} + \mathbf{K}_{t|t} (\mathbf{Y}_t - \mathbf{G}_t \widehat{\mathbf{X}}_{t|t-1})$$

a FILTRAČNÍ CHYBOVOU KOVARIANČNÍ MATICI

$$\boldsymbol{\Omega}_{t|t} = (\mathbf{I} - \mathbf{K}_{t|t} \mathbf{G}_t) \boldsymbol{\Omega}_{t|t-1}.$$

(III) ČASOVÝ (PREDIKČNÍ) KROK KALMANOVA FILTRU

$$\begin{aligned}\widehat{\mathbf{X}}_{t+1|t} &= \mathbf{F}_t \widehat{\mathbf{X}}_{t|t} \\ \boldsymbol{\Omega}_{t+1|t} &= \mathbf{F}_t \boldsymbol{\Omega}_{t|t} \mathbf{F}'_t + \mathbf{Q}_t.\end{aligned}$$



Literatura

- [1] ANDĚL, J. *Matematická statistika*. SNTL/ALFA Praha, 1978.
- [2] ANDĚL, J. *Statistické metody*. Matfyzpress Praha, 1993.
- [3] ANDĚL, J. *Statistická analýza časových řad*. Praha. SNTL 1976.
- [4] BOX, G.E.P, COX, D.R.: Analysis of Transformations. *Journals of the Royal Statistical Society*, Biometrika 26, 1964, 211-252.
- [5] BOX, G., JENKINS, G. *Time series analysis - forecasting and control*. Holden-Day 1976.
- [6] BROCKWELL, P.J., DAVIS, R.A. *Time Series: Theory and Methods*. Springer-Verlag, New York, 1991.
- [7] BROCKWELL, P.J., DAVIS, R.A. *Introduction to time series and forecasting*. Springer-Verlag, New York, 2002.
- [8] CIPRA, T. *Analýza časových řad s aplikacemi v ekonomii*. SNTL, Praha, 1986.
- [9] GICHMAN, I.I., SKOROCHOD, A.V. *Teoriya slučajnych processov*. Moskva. Nauka 1971.
- [10] GRANGER, C.W.J., ANDERSEN, A. *An introduction to bilinear time series models*. Vandenhoeck and Ruprecht, Göttingen 1978.
- [11] HAMILTON, J.D. *Time Series Analysis*. Princeton University Press. 1994.
- [12] KALMAN, R. *A new approach to linear filtering and prediction problems*. Trans. ASME J. Basic Eng. D 82 (1960), 34-45.
- [13] KUBÁČKOVÁ, L., KUBÁČEK, L., KUKUČA, J. *Pravdepodobnosť a štatistika v geodézii a geofyzike*. Veda, Bratislava (1982).
- [14] LJUNG, G. M., BOX, G. E. P. On a measure of lack of fit in time series models. *Biometrika* 65. 1978, 553-564.
- [15] NEUBRUNN, T., RIEČAN, B. *Miera a integrál*. Bratislava. Veda 1981.

- [16] PRIESTLEY, M. *Spectral analysis and time series*. Academic Press 1989.
- [17] RAO, R.C. *Lineární metody statistické indukce a jejich aplikace*. ACADEMIA Praha, 1978.