ARMA modely
časť 1: autoregresné modely (AR)

Beáta Stehlíková

Časové rady, FMFI UK, 2011/2012

ARMA modely

- Terminológia:
  - AR - autoregresný model - niekoľko nasledujúcich hodín
  - MA - moving average, kĺzavé priemery
  - ARMA - ich kombinácia
- Najskôr: autoregresný proces prvého rádu - AR(1)
  - definícia
  - stacionarita, ohraničenie na parametre
  - výpočet momentov a ACF
  - simulované dáta
  - praktický príklad s reálnymi dátami
- Potom:
  - autoregresné procesy vyšších rádov
  - ako určiť vhodný rád procesu pre dané dáta

I.

Autoregresný proces prvého rádu - AR(1)

AR(1) - definícia

- AR(1) proces:

$$x_t = \delta + \alpha x_{t-1} + u_t,$$

kde δ a α sú konštanty a $\{u_t\}$ je biely šum

- Nech pre $t = t_0$ je daná hodnota x_{t_0} :

$$\begin{aligned} x_{t_0+1} &= \delta + \alpha x_{t_0} + u_{t_0+1}, \\ x_{t_0+2} &= \delta + \alpha x_{t_0+1} + u_{t_0+2} = \\ &\quad \delta(1+\alpha) + \alpha^2 x_{t_0} + (\alpha u_{t_0+1} + u_{t_0+2}), \\ x_{t_0+3} &= \ldots \end{aligned}$$

vo všeobecnosti:

$$x_{t_0+\tau} = \alpha^\tau x_{t_0} + \frac{1-\alpha^\tau}{1-\alpha}\delta + \sum_{j=0}^{\tau-1} \alpha^j u_{t_0+\tau-j} \qquad (1)$$

AR(1) - stacionarita

- Zo vzťahu (1):

$$x_t = \alpha^{t-t_0} x_{t_0} + \frac{1-\alpha^{t-t_0}}{1-\alpha}\delta + \sum_{j=0}^{t-t_0-1} \alpha^j u_{t-j}$$

- Deterministické začiatočné podmienky: hodnota procesu v čase t_0 je $x_0 \rightarrow$ proces nie je stacionárny
- Náhodné začiatočné podmienky:
  - ◇ Proces je generovaný pre $t \in \mathbb{R} \rightarrow$ hodnota x_{t_0} je náhodná.
  - ◇ Ak $-1 < \alpha < 1$, tak pre $t_0 \rightarrow -\infty$ dostaneme

$$x_t = \frac{1}{1-\alpha}\delta + \sum_{j=0}^{\infty} \alpha^j u_{t-j} \qquad (2)$$

  - ◇ Woldova reprezentácia: $\psi_j = \alpha^j$ pre $|\alpha| < 1 \rightarrow$ proces je slabo stacionárny.

AR(1) - výpočet pomocou operátora L

- Zápis procesu:

$$(1-\alpha L)x_t = \delta + u_t$$

- Vyjadríme x_t:

$$x_t = (1-\alpha L)^{-1}\delta + (1-\alpha L)^{-1}u_t$$

- Výpočet inverzného operátora - čo je $(1-\alpha L)^{-1}$ (píše sa aj $1/(1-\alpha L)$)?

$$(1-\alpha L)^{-1} = \phi_0 + \phi_1 L + \phi_2 L^2 + \ldots ...$$

$$(1-\alpha L)(\phi_0 + \phi_1 L + \phi_2 L^2 + \ldots ...) = 1$$

Roznásobíme a porovnáme koeficienty $\rightarrow$ $\phi_j = \alpha^j$, t. j.

$$(1-\alpha L)^{-1} = \frac{1}{1-\alpha L} = 1 + \alpha L + \alpha^2 L^2 + \alpha^3 L^3 + \ldots$$

AR(1) - výpočet pomocou operátora L

- Teda:

$$\begin{aligned} x_t &= (1+\alpha L+\alpha^2 L^2+\ldots)\delta+(1+\alpha L+\alpha^2 L^2+\ldots)u_t \\ &= (1+\alpha+\alpha^2+\ldots)\delta+(u_t+\alpha u_{t-1}+\alpha^2 u_{t-2}+\ldots) \\ &= \frac{\delta}{1-\alpha}+\sum_{j=0}^{\infty}\alpha^j u_{t-j}, \end{aligned}$$

  pričom na sčítanie nekonečného radu potrebujeme splnenie podmienky $|\alpha| < 1$.
- Z oboch postupov teda vyplýva podmienka stacionarity pre AR(1) proces: $|\alpha| < 1$
- V ďalších výpočtoch predpokladáme splnenie tejto podmienky

AR(1) - momenty, postup 1

- Pripomeňme si explicitné vyjadrenie procesu (2):

$$x_t = \frac{\delta}{1-\alpha} + \sum_{j=0}^{\infty} \alpha^j u_{t-j}$$

- Stredná hodnota:

$$\begin{aligned} E[x_t] &= E\left[\frac{\delta}{1-\alpha} + \sum_{j=0}^{\infty} \alpha^j u_{t-j}\right] \\ &= \frac{\delta}{1-\alpha} + \sum_{j=0}^{\infty} \alpha^j E[u_{t-j}] = \frac{\delta}{1-\alpha} \end{aligned}$$

  - ◇ $E[x_t] = 0$ práve vtedy, keď $\delta = 0$
  - ◇ vo všeobecnosti $E[x_t] \neq \delta$, ale majú rovnaké znamienko (lebo $|\alpha| < 1$)

AR(1) - momenty, postup 1

- Variancia:

$$
\begin{aligned}
Var[x_t] &= Var\left[\frac{\delta}{1-\alpha} + \sum_{j=0}^{\infty} \alpha^j u_{t-j}\right] \\
&= \sum_{j=0}^{\infty} Var[\alpha^j u_{t-j}] = \sum_{j=0}^{\infty} \alpha^{2j} Var[u_{t-j}] \\
&= \sigma^2 \sum_{j=0}^{\infty} \alpha^{2j} = \sigma^2 \frac{1}{1-\alpha^2}
\end{aligned}
$$

  kde
  - sme využili, že disperzia súčtu nekorelovaných premenných je súčet disperzií
  - σ^2 je variancia bieleho šumu $\{u_j\}$

AR(1) - momenty, postup 1

- Autokovariancie (využijeme, že $Cov[u_k, u_l] = \sigma^2$ pre $k = l$ a $Cov[u_k, u_l] = 0$ pre $k \neq l$):

$$
\begin{aligned}
Cov[x_t, x_{t-s}] &= E\left[\left(\sum_{i=0}^{\infty} \alpha^i u_{t-i}\right)\left(\sum_{j=0}^{\infty} \alpha^j u_{t-s-j}\right)\right] \\
&= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \alpha^{i+j} E[u_{t-i} u_{t-s-j}] \\
&= \sigma^2 \sum_{j=0}^{\infty} \alpha^{s+2j} = \alpha^s \frac{\sigma^2}{1-\alpha^2}
\end{aligned}
$$

- Autokorelácie :

$$
Cor[x_t, x_{t-s}] = \frac{Cor[x_t, x_{t-s}]}{Var[x_t]Var[x_{t-s}]} = \alpha^s
$$

AR(1) - momenty, postup 2

- Alternatívny postup, nevyužívajúci explicitné vyjadrenie x_t; užitočný pri procesoch vyššieho rádu
- Stredná hodnota:
  - vieme, že existuje konštanta μ, taká že $E[x_t] = \mu$ pre každé t (stacionarita)
  - spravíme strednú hodnotu l'avej aj pravej strany rovnosti z definície procesu:

$$
\begin{aligned}
x_t &= \delta + \alpha x_{t-1} + u_t \quad \Big/ \; E[.] \\
\mu &= \delta + \alpha\,\mu + 0 \\
\mu &= \frac{\delta}{1-\alpha}
\end{aligned}
$$

AR(1) - momenty, postup 2

- Disperzia a autokovariancie:
  - ak x_t je stacionárny proces (nielen AR(1)), tak procesy x_t a $x_t - \mu$ majú rovnaké disperzie a autokovariancie
  - pre AR(1) proces teda môžeme predpokladať, že $\delta = 0$
  - nech $s \geq 0$:

$$\begin{aligned} x_t &= \alpha x_{t-1} + u_t \qquad \Big/ \; \times x_{t-s},\; E[.] \\ E[x_t x_{t-s}] &= \alpha E[x_{t-1} x_{t-s}] + E[u_t x_{t-s}] \end{aligned}$$

z toho:

$$s = 0 \;\Rightarrow\; \gamma(0) = \alpha\gamma(1) + \sigma^2 \tag{3}$$

$$s = 1 \;\Rightarrow\; \gamma(1) = \alpha\gamma(0) \tag{4}$$

$$s \geq 2 \;\Rightarrow\; \gamma(s) = \alpha\gamma(s-1) \tag{5}$$

AR(1) - momenty, postup 2

- Disperzia a autokovariancie- pokračovanie:
  - (3) a (4) - dve lineárne rovnice s dvoma neznámymi $\rightarrow$ vyjadríme $\gamma(0), \gamma(1)$:

$$\gamma(0) = \frac{\sigma^2}{1-\alpha^2}, \ \gamma(1) = \frac{\alpha\sigma^2}{1-\alpha^2} \tag{6}$$

  - (5) je potom rekurentný predpis, pričom (6) dáva začiatočnú podmienku:

$$\gamma(1) = \frac{\alpha\sigma^2}{1-\alpha^2} \tag{7}$$

$$\gamma(s+1) = \alpha\gamma(s) \text{ pre } s \geq 1 \tag{8}$$

- Autokorelácie: rovnice (7) a (8) vydelíme $\gamma(0)$:

$$\rho(1) = \alpha, \quad \rho(s+1) = \alpha\rho(s) \text{ pre } s \geq 1$$

a teda $\rho(s) = \alpha^s$.

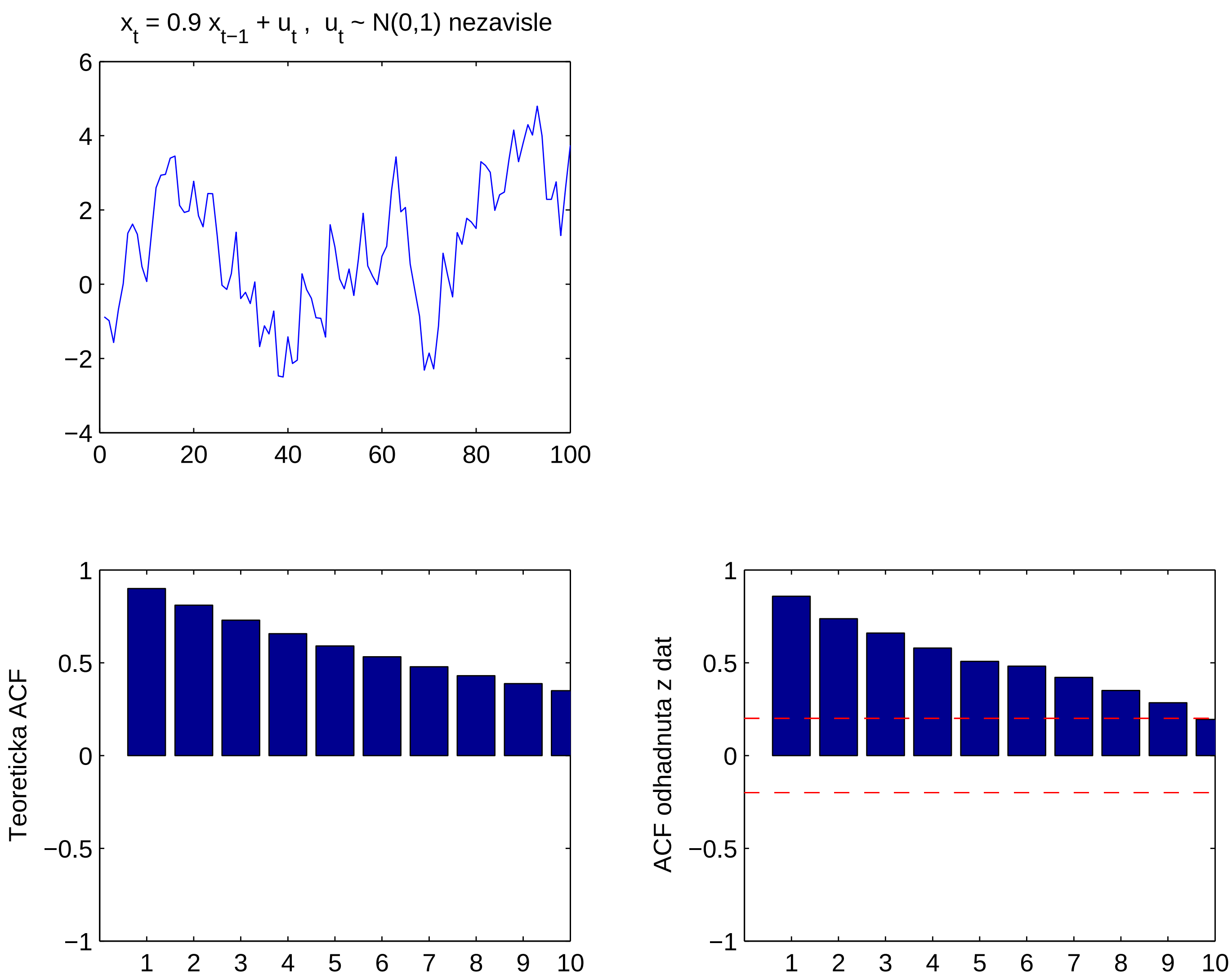
Príklad - simulované dáta

- AR(1) proces

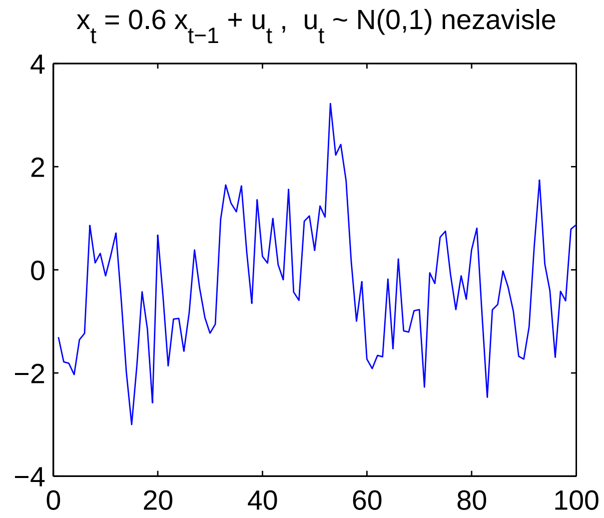
$$x_t = \delta + \alpha x_{t-1} + u_t,$$

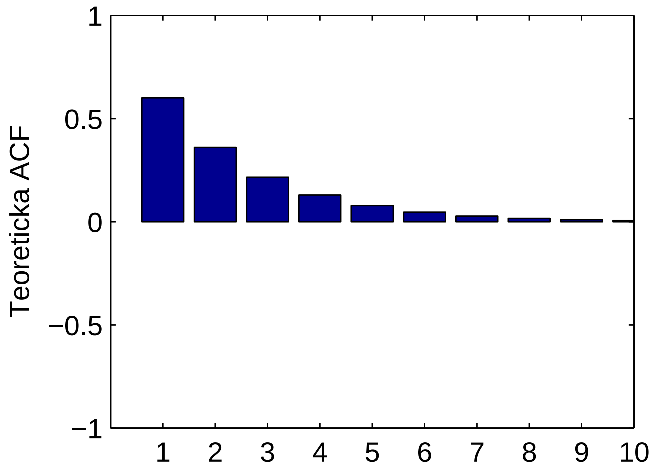
  kde biely šum u_t má normálne rozdelenie, $\delta = 0$, $\sigma^2 = 1$
- Postupne zoberieme $\alpha = \{0.9, 0.6, -0.9\}$
- Zobrazíme:
  - realizáciu procesu
  - teoretickú autokorelačnú funkciu
  - výberovú autokorelačnú funkciu odhadnutú z dát

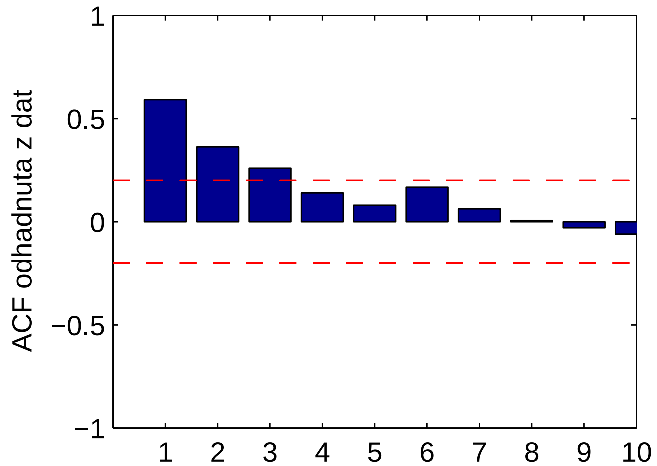
Príklad - simulované dáta, $\alpha = 0.9$



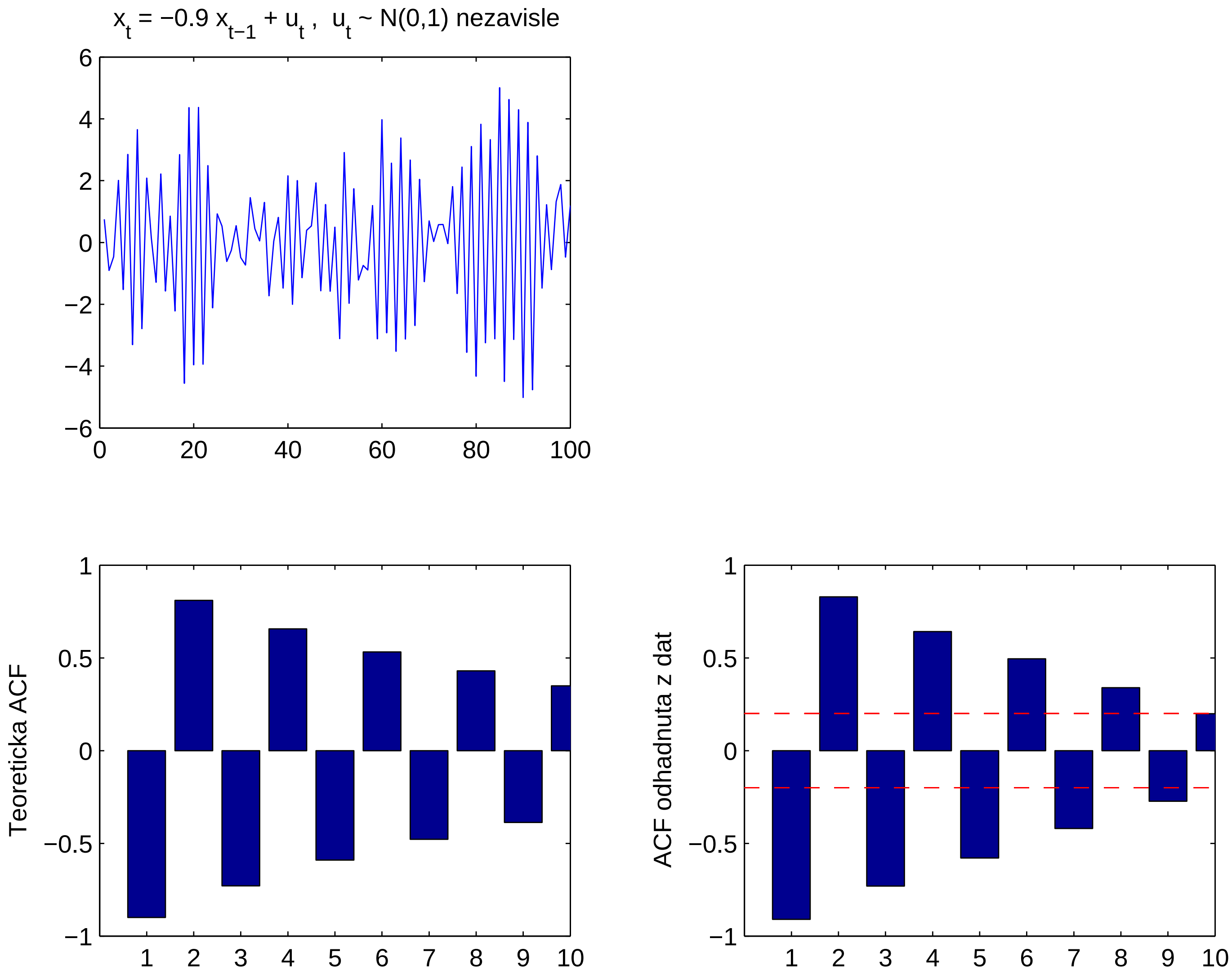
Príklad - simulované dáta, $\alpha = 0.6$







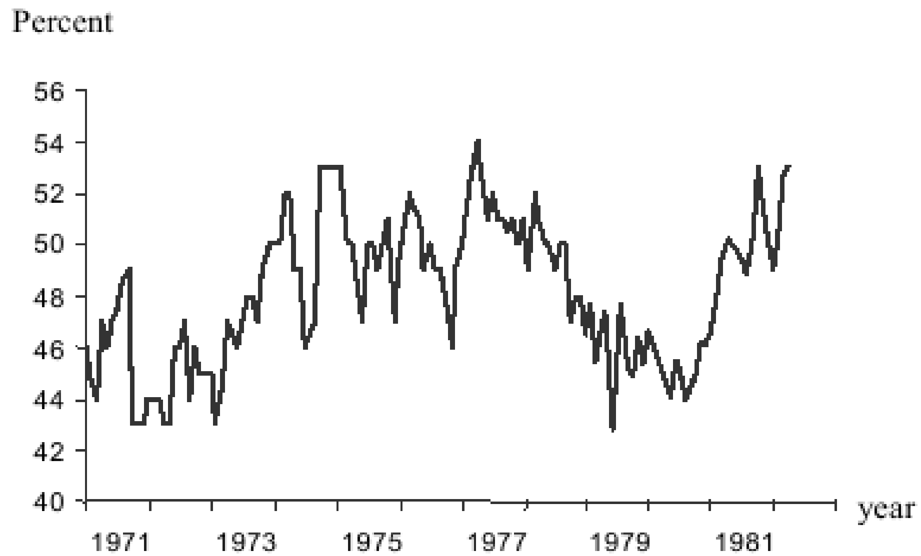
Príklad - simulované dáta, $\alpha = -0.9$



Príklad - reálne dáta

- G. Kirchgässner: **Causality Testing of the Popularity Function: An Empirical Investigation for the Federal Republic of Germany, 1971-1982**, Public Choice 45 (1985), p. 155-173.
- [Kirchgässner, Wolters], example 2.2

- Nemecko, január 1971 - apríl 1982
- CDU_t = volebné preferencie CDU/CSU



a) Popularity of the CDU/CSU, 1971 – 1982

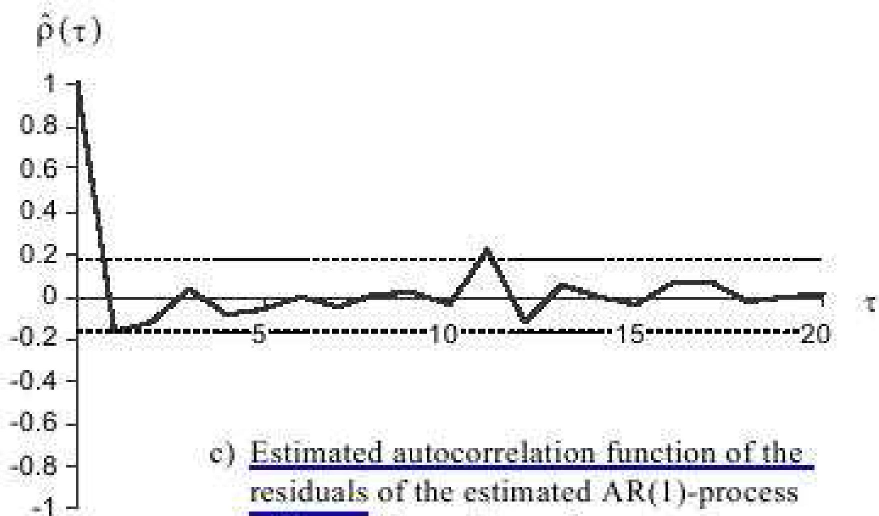
Príklad - reálne dáta

- Odhadnutý AR(1) model:

$$
\begin{aligned}
CDU_t &= \underset{(3.43)}{8.053} + \underset{(17.10)}{0.834}\, CDU_{t-1} + \hat{u}_t,
\end{aligned}
$$

$$\bar{R}^2 = 0.683, \quad SE = 1.586, \quad Q(11) = 12.516 \ (p = 0.326).$$

The estimated t values are given in parentheses. The autocorrelogram, which is also given in *Figure 2.4*, does not indicate any higher-order process. Moreover, the Box-Ljung Q Statistic with 12 correlation coefficients (i.e. with 11 degrees of freedom) gives no reason to reject this model.



c) Estimated autocorrelation function of the residuals of the estimated AR(1)-process with confidence intervals

Príklad - reálne dáta

- Otázky k odhadnutému modelu:
  - ◇ Je odhadnutý model stacionárny?
  - ◇ V texte sa spomínajú autokorelácie rezíduí a Ljung-Boxova Q štatistika - aké hypotézy sa testujú (a prečo), akým spôsobom a s akými závermi?
  - ◇ Na grafe sú pri autokoreláciách zostrojené intervaly. Vypočítajte pomocou známych údajov ich hranice.
  - ◇ Čomu sa rovná stredná hodnota premennej CDU_t?

II.

Autoregresný proces druhého rádu - AR(2)

AR(2) - definícia

- AR(2) proces:

$$x_t = \delta + \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + u_t$$

- Pomocou operátora posunu:

$$\begin{aligned} (1 - \alpha_1 L - \alpha_2 L^2) x_t &= \delta + u_t \\ \alpha(L) x_t &= \delta + u_t \end{aligned}$$

- Woldova reprezentácia a stacionarita:

$$x_t = \alpha^{-1}(L)\delta + \alpha^{-1}(L) u_t$$

$\rightarrow$ potrebujeme inverzný operátor $\alpha^{-1}(L)$; znovu ho nájdeme metódou neurčitých koeficientov:

$$\alpha^{-1}(L) = \psi_0 + \psi_1 L + \psi_2 L^2 + \dots$$

pričom

$$(9) \quad 1 = (1 - \alpha_1 L - \alpha_2 L^2)(\psi_0 + \psi_1 L + \psi_2 L^2 + \dots)$$

AR(2) - stacionarita

- Porovnáme koeficienty pri L^j na oboch stranách (9):
$$\psi_j - \alpha_1\psi_{j-1} - \alpha_2\psi_{j-2} = 0,$$
$$\psi_0 = 1,\ \psi_1 = \alpha_1$$
- Podmienka stacionarity: Kvôli splneniu podmienky $\sum \psi_j^2 < \infty$ musia byť korene charakteristickej rovnice
$$\lambda^2 - \alpha_1\lambda - \alpha_2 = 0$$
v absolútnej hodnote menšie ako 1
- Inak povedané: korene rovnice
$$\alpha(L) = 1 - \alpha_1 L - \alpha_2 L^2 = 0$$
musia byť v absolútnej hodnote väčšie ako 1, t. j. mimo jednotkového kruhu
- To isté vyšlo predtým pre AR(1): koreň $\alpha(L) = 0$ mimo jednotkového kruhu

AR(2) - momenty

- Slabo stacionárny AR(2) proces:

$$x_t = \delta + \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + u_t$$

- Stredná hodnota:
  - ◇ označme $\mu = E[x_i]$; potom

$$\begin{aligned} \mu &= \delta + \alpha_1 \mu + \alpha_2 \mu, \\ \mu &= \frac{\delta}{1 - \alpha_1 - \alpha_2} \end{aligned}$$

  - ◇ Podobne ako pre AR(1): stredná hodnota sa nerovná parametru δ (okrem prípadu $\delta = 0$), ale majú rovnaké znamienko

AR(2) - momenty

- Autokovariancie: znovu môžeme predpokladať nulovú strednú hodnotu, t. j.

$$\begin{aligned} x_t &= \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + u_t \quad \Big/ \quad \times x_{t-s},\ E[.] \\ E[x_{t-s}x_t] &= \alpha_1 E[x_{t-s}x_{t-1}] + \alpha_2 E[x_{t-s}x_{t-2}] + E[x_{t-s}u_t] \end{aligned}$$

- Pre $s = 0, 1, 2$ dostaneme:

$$\begin{aligned} \gamma(0) &= \alpha_1\gamma(1) + \alpha_2\gamma(2) + \sigma^2 \\ \gamma(1) &= \alpha_1\gamma(0) + \alpha_2\gamma(1) \\ \gamma(2) &= \alpha_1\gamma(1) + \alpha_2\gamma(0) \end{aligned}$$

  - sústava rovníc $\rightarrow \gamma(0) = Var[x_t]$, $\gamma(1)$, $\gamma(2)$

- Pre $s \geq 2$ - diferenčná rovnica:

  (10) $$\gamma(s) - \alpha_1\gamma(s-1) - \alpha_2\gamma(s-2) = 0,$$

  začiatočné podmienky z predchádzajúceho bodu

AR(2) - momenty

- Autokorelácie: diferenčnú rovnicu (10) a jej začiatočné podmienky vydelíme $\gamma(0)$:

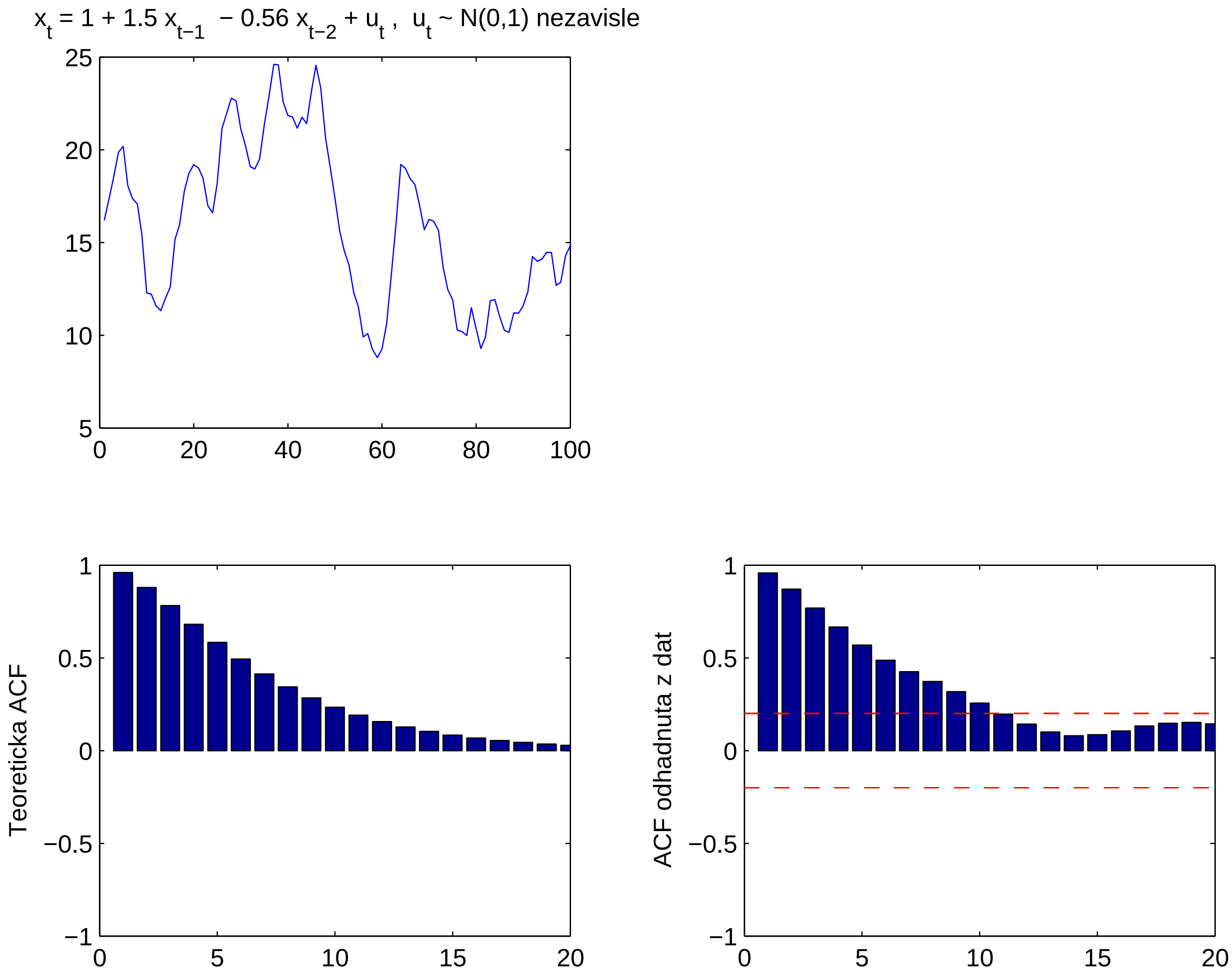
$$\rho(s) - \alpha_1 \rho(s-1) - \alpha_2 \rho(s-2) = 0$$

$$\rho(0) = 1, \rho(1) = \frac{\alpha_1}{1 - \alpha_2}$$

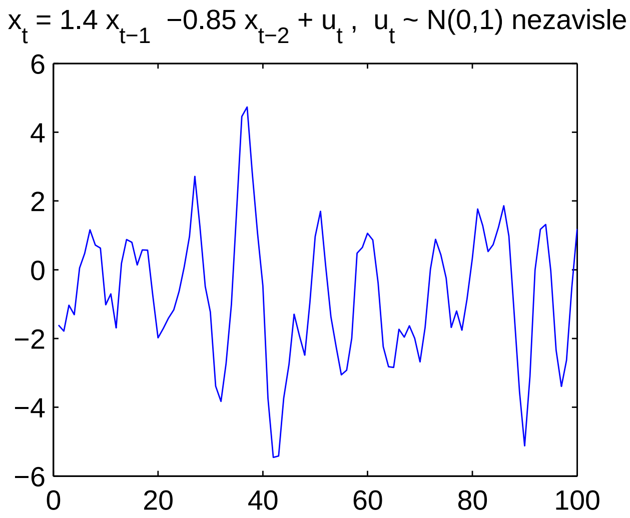
AR(2) - príklady

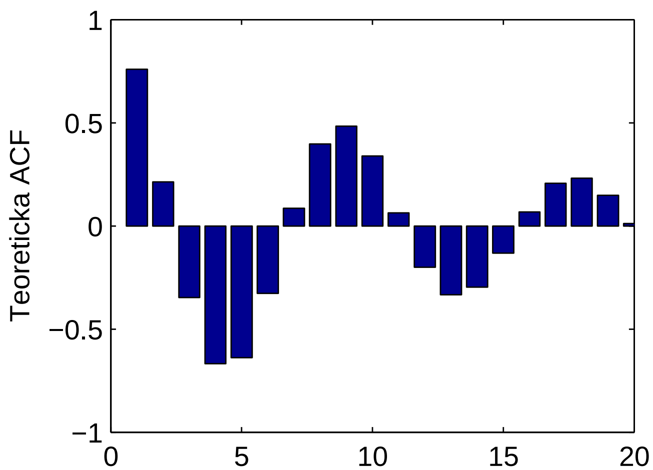
1. Proces: $x_t = 1 + 1.5x_{t-1} - 0.56x_{t-2} + u_t$
   - korelácie spĺňajú diferenčnú rovnicu
     $$\rho(t) - 1.5\rho(t-1) + 0.56\rho(t-2) = 0,$$
     ktorá má všeobecné riešenie
     $$\rho(t) = c_1(0.8)^t + c_2(0.7)^t$$
   - ACF monotónne klesá
2. Proces: $x_t = 1.4x_{t-1} - 0.85x_{t-2} + u_t$
   - korelácie spĺňajú diferenčnú rovnicu
     $$\rho(t) - 1.4\rho(t-1) + 0.85\rho(t-2) = 0,$$
     ktorá má všeobecné riešenie
     $$\rho(t) = 0.922^t(c_1\cos(0.709t) + c_2\sin(0.709t))$$
   - ACF osciluje

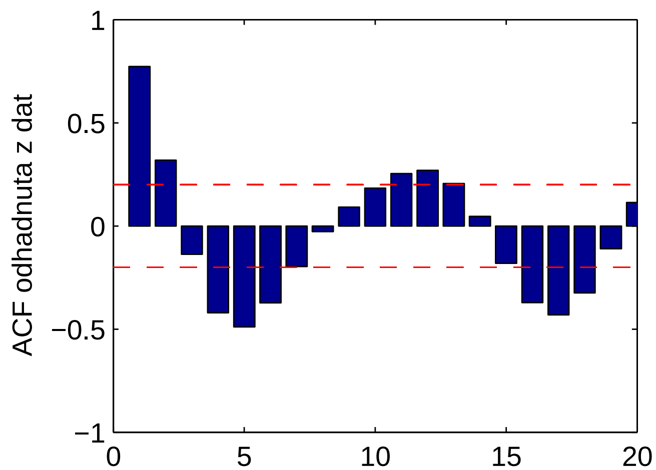
AR(2) - príklad 1



AR(2) - príklad 2





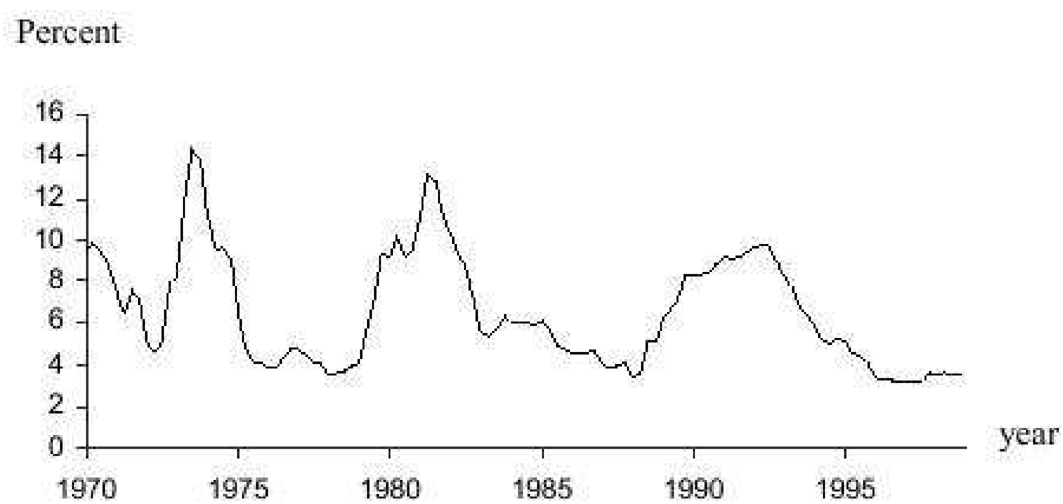


$\cos(kt), \sin(kt) \to$ perióda $\frac{2\pi}{k} = 8.862 \approx 9$

AR(2) - reálne dáta

[Kirchgässner, Wolters], example 2.6

- 3-mesačná úroková miera, Nemecko, 1970q1-1998q4

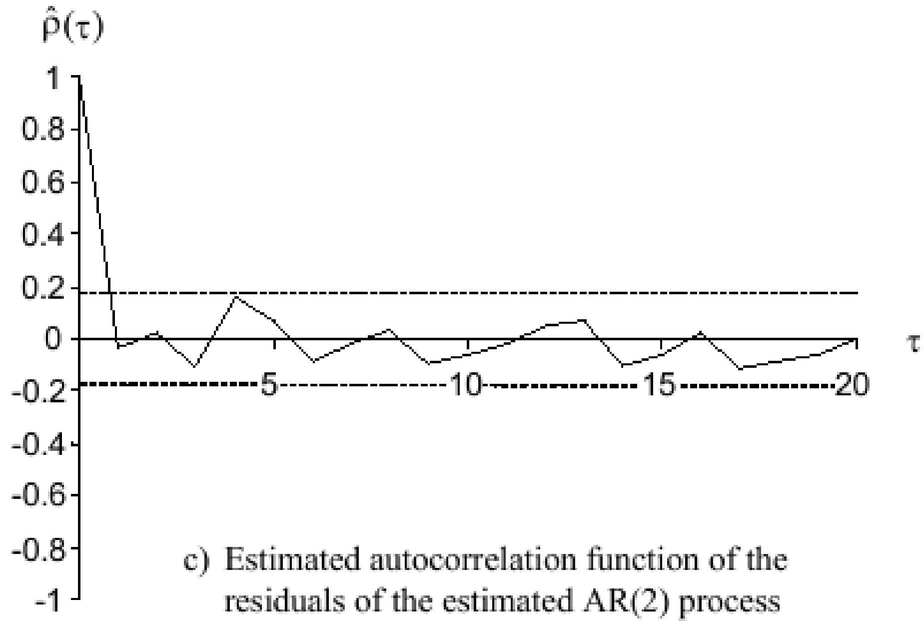


AR(2) - reálne dáta

- Odhadnutý AR(2) model:

$$GSR_t = \underset{(2.82)}{0.577} + \underset{(17.49)}{1.407}\, GSR_{t-1} - \underset{(-6.16)}{0.498}\, GSR_{t-2} + \hat{u}_t,$$

$$\bar{R}^2 = 0.910, \quad SE = 0.812, \quad Q(6) = 6.431 \quad (p = 0.377)$$



c) Estimated autocorrelation function of the residuals of the estimated AR(2) process with confidence intervals

AR(2) - reálne dáta

- Otázky k odhadnutému modelu:
  - Je stacionárny?
  - Analyzujte rezíduá - autokorelogram, Q-štatistika.
  - Aká je stredná hodnota odhadnutého procesu? Aký je priebeh jeho autokorelačnej funkcie?

III.

Autoregresný proces p-teho rádu - AR(p)

AR(p) proces - stacionarita

- AR(p) proces:

  (11) $x_t = \delta + \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \ldots + \alpha_p x_{t-p} + u_t,$

  t. j. $\alpha(L)x_t = \delta + u_t$, kde $\alpha(L) = 1 - \alpha_1 L - \ldots - \alpha_p L^p$

- Woldova reprezentácia a stacionarita:

  $$x_t = \alpha(L)^{-1}(\delta + u_t),$$

  inverzný operátor $\alpha(L)^{-1}$ hľadáme v tvare

  $$\alpha(L)^{-1} = 1 + \psi_1 L + \psi_2 L^2 + \ldots$$

- Pre koeficienty ψ_j dostaneme diferenčnú rovnicu

  $$\psi_k - \alpha_1 \psi_{k-1} - \ldots - \alpha_p \psi_{k-p} = 0$$

  $\Rightarrow$ kvôli konv. $\sum \psi_j^2$ musia byť korene charakt. rovnice $\lambda^k - \alpha_1 \lambda^{k-1} - \ldots - \alpha_p = 0$ vnútri jednotkového kruhu, t. j. korene $\alpha(L) = 0$ musia byť mimo jednotkového kruhu

AR(p) proces - momenty

- Stredná hodnota:
  označíme $\mu = E[x_t]$ a spravíme strednú hodnotu z ľavej aj pravej strany (11):

$$\mu = \delta + \alpha_1\mu + \ldots + \alpha_p\mu \;\Rightarrow\; \mu = \frac{\delta}{1 - \alpha_1 - \ldots - \alpha_p}$$

  aj teraz: stredná hodnota má rovnaké znamienko ako parameter δ

- Variancia, autokovariancie - nech $\delta = 0$

$$\begin{aligned} x_t &= \alpha_1 x_{t-1} + \ldots + \alpha_p x_{t-p} + u_t \;\Big/\; \times x_{t-s},\; E[.] \\ \gamma(s) &= \alpha_1\gamma(s-1) + \ldots \alpha_p\gamma(s-p) + E[u_t x_{t-s}] \end{aligned}$$

AR(p) proces - momenty

- Variancia, autokovariancie - pokračovanie:
  - $s = 0, 1, \ldots, p \rightarrow$ sústava $p+1$ rovníc s neznámymi $\gamma(0), \gamma(1), \ldots, \gamma(p)$:

$$\begin{aligned}
\gamma(0) &= \alpha_1\gamma(1) + \alpha_2\gamma(2) + \ldots + \alpha_p\gamma(p) + \sigma^2 \\
\gamma(1) &= \alpha_1\gamma(0) + \alpha_2\gamma(1) + \ldots + \alpha_p\gamma(p-1) \\
& \ldots \\
\gamma(p) &= \alpha_1\gamma(p-1) + \alpha_2\gamma(p-2) + \ldots + \alpha_p\gamma(0)
\end{aligned} \tag{12}$$

  - ostatné autokovariancie z diferenčnej rovnice

$$\gamma(t) - \alpha_1\gamma(t-1) - \ldots - \alpha_p\gamma(t-p) = 0 \tag{13}$$

AR(p) proces - momenty

- ACF :
  - diferenčná rovnica pre autokorelácie - rovnicu (13) vydelíme disperziou $\gamma(0)$:

$$\rho(t) - \alpha_1\rho(t-1) - \ldots - \alpha_p\rho(t-p) = 0$$

  - začiatočné podmienky - posledných p rovníc zo sústavy (12) vydelíme $\gamma(0)$:

$$\begin{aligned} \rho(1) &= \alpha_1 + \alpha_2\rho(1) + \ldots + \alpha_p\rho(p-1) \\ \rho(2) &= \alpha_1\rho(1) + \alpha_2 + \ldots + \alpha_p\rho(p-2) \\ &\ldots \\ \rho(p) &= \alpha_1\rho(p-1) + \alpha_2\rho(p-2) + \ldots + \alpha_p \end{aligned} \tag{14}$$

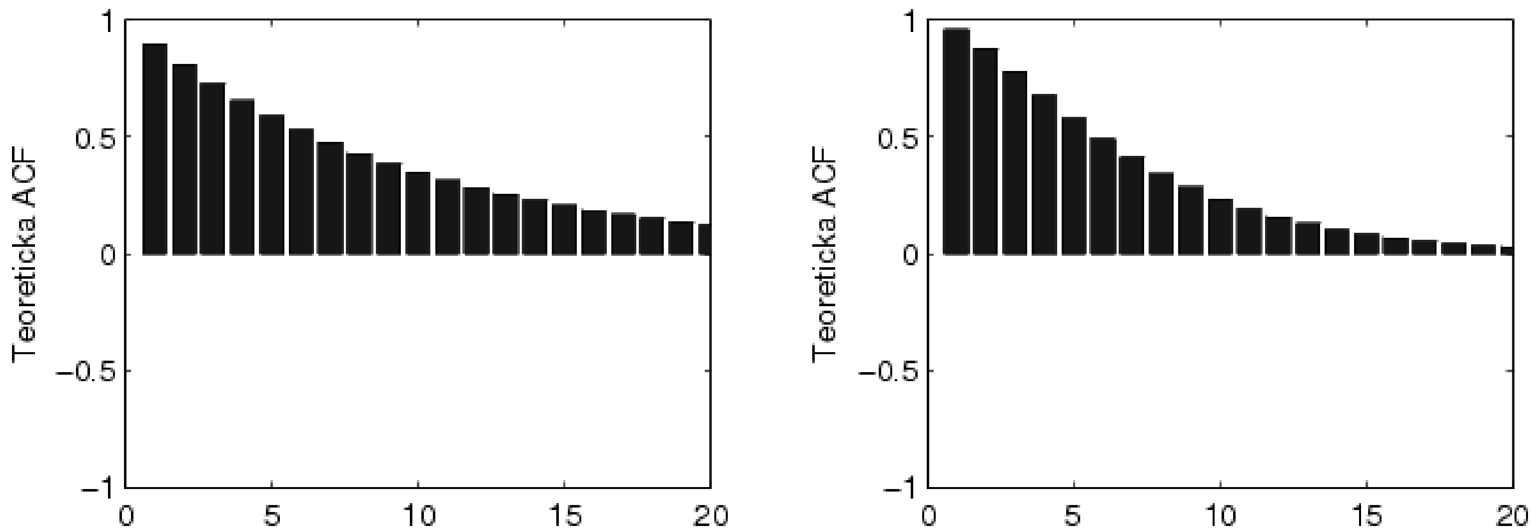
  - nazývajú sa Yule-Wolkerove rovnice

IV.

Parciálna autokorelačná funkcia - určovanie rádu AR procesu

PACF - motivácia

- Prečo sme na modelovanie preferencií použili AR(1) model a na modelovanie úrokovej miery AR(2)?
- Teoretická ACF pre AR(1) a AR(2) procesy z predchádzajúcich simulácií:



Ako ich rozlíšiť?

- Ako zistiť vhodnosť AR modelu určiť jeho rád?

PACF - motivácia

- Uvažujme nejaký náhodný proces x_t s nulovou strednou hodnotou a modelujme jeho hodnotu pomocou prechádzajúcich k hodnôt:

$$x_t = \beta_1 x_{t-1} + \beta_2 x_{t-2} + \ldots + \beta_k x_{t-k} + u_t$$

- Označme koeficienty Φ_{ki}, kde k je počet použitých starších hodnôt procesu x a i je koeficient pri x_{t-i}
- Teda:

$$\begin{aligned}
x_t &= \Phi_{11} x_{t-1} + u_t \\
x_t &= \Phi_{21} x_{t-1} + \Phi_{22} x_{t-2} + u_t \\
x_t &= \Phi_{31} x_{t-1} + \Phi_{32} x_{t-2} + \Phi_{33} x_{t-3} + u_t \\
&\ldots \\
x_t &= \Phi_{k1} x_{t-1} + \Phi_{k2} x_{t-2} + \Phi_{k3} x_{t-3} + \ldots + \Phi_{kk} x_{t-k} + u_t
\end{aligned}$$

- Ak x je AR(p) proces, tak $\Phi_{kk} = 0$ pre $k > p$.

PACF - definícia a výpočet

- Koeficient Φ_{kk} sa nazýva parciálna autokorelácia rádu k
- Postupnosť týchto koeficientov vytvára parciálnu autokorelačnú funkciu (PACF)
- Výpočet: vyjdeme zo vzťahu

  $$x_t = \Phi_{k1} x_{t-1} + \Phi_{k2} x_{t-2} + \Phi_{k3} x_{t-3} + \ldots + \Phi_{kk} x_{t-k} + u_t$$

  a rovnako ako pri odvodení Yule-Wolkerovych rovníc dostaneme

$$\begin{aligned}
\rho(1) &= \Phi_{k1} + \Phi_{k2}\,\rho(1) + \ldots + \Phi_{kk}\,\rho(k-1) \\
\rho(2) &= \Phi_{k1}\,\rho(1) + \Phi_{k2} + \ldots + \Phi_{kk}\,\rho(k-2) \\
&\ldots \\
\rho(k) &= \Phi_{k1}\,\rho(k-1) + \Phi_{k2}\,\rho(k-2) + \ldots + \Phi_{kk}
\end{aligned}$$

PACF - definícia a výpočet

- Maticový zápis:

$$\begin{bmatrix} 1 & \rho(1) & \dots & \rho(k-1) \\ \rho(1) & 1 & \dots & \rho(k-2) \\ & & \dots & \\ \rho(k-1) & \rho(k-2) & \dots & 1 \end{bmatrix} \begin{bmatrix} \Phi_{k1} \\ \Phi k2 \\ \dots \\ \Phi kk \end{bmatrix} = \begin{bmatrix} \rho(1) \\ \rho(2) \\ \dots \\ \rho(k) \end{bmatrix}$$

- Zaujíma nás iba Φ_{kk}, použijeme Cramerovo pravidlo:

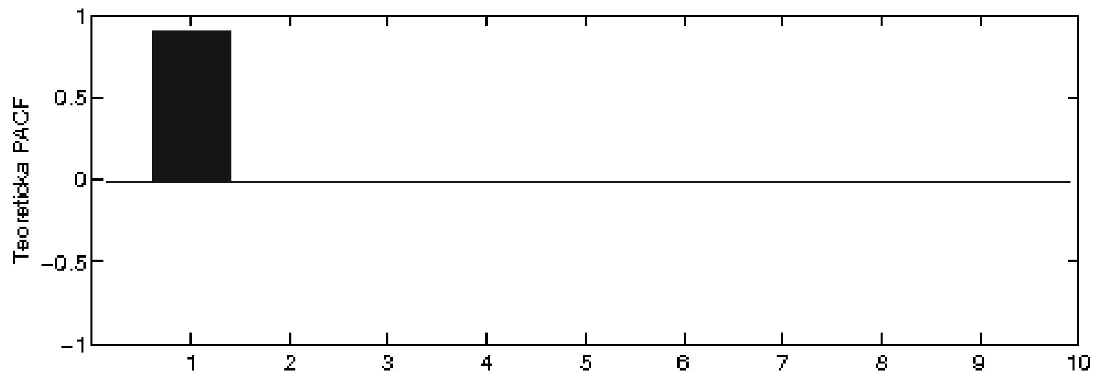
$$\Phi_{kk} = \frac{\det \begin{pmatrix} 1 & \rho(1) & \dots & \rho(1) \\ \rho(1) & 1 & \dots & \rho(2) \\ & \dots & \dots & \\ \rho(k-1) & \rho(k-2) & \dots & \rho(k) \end{pmatrix}}{\det \begin{pmatrix} 1 & \rho(1) & \dots & \rho(k-1) \\ \rho(1) & 1 & \dots & \rho(k-2) \\ & \dots & \dots & \\ \rho(k-1) & \rho(k-2) & \dots & 1 \end{pmatrix}} \tag{15}$$

PACF - príklad: AR(1) proces

- Postupne počítame:

$$
\begin{aligned}
\Phi_{11} &= \rho(1) \\
\Phi_{22} &= \frac{\det\begin{pmatrix} 1 & \rho(1) \\ \rho(1) & \rho(2) \end{pmatrix}}{\det\begin{pmatrix} 1 & \rho(1) \\ \rho(1) & 1 \end{pmatrix}} = \frac{\rho(2) - \rho(1)^2}{1 - \rho(1)^2} = 0 \\
&\dots
\end{aligned}
$$

- Pre $\alpha = 0.9$:

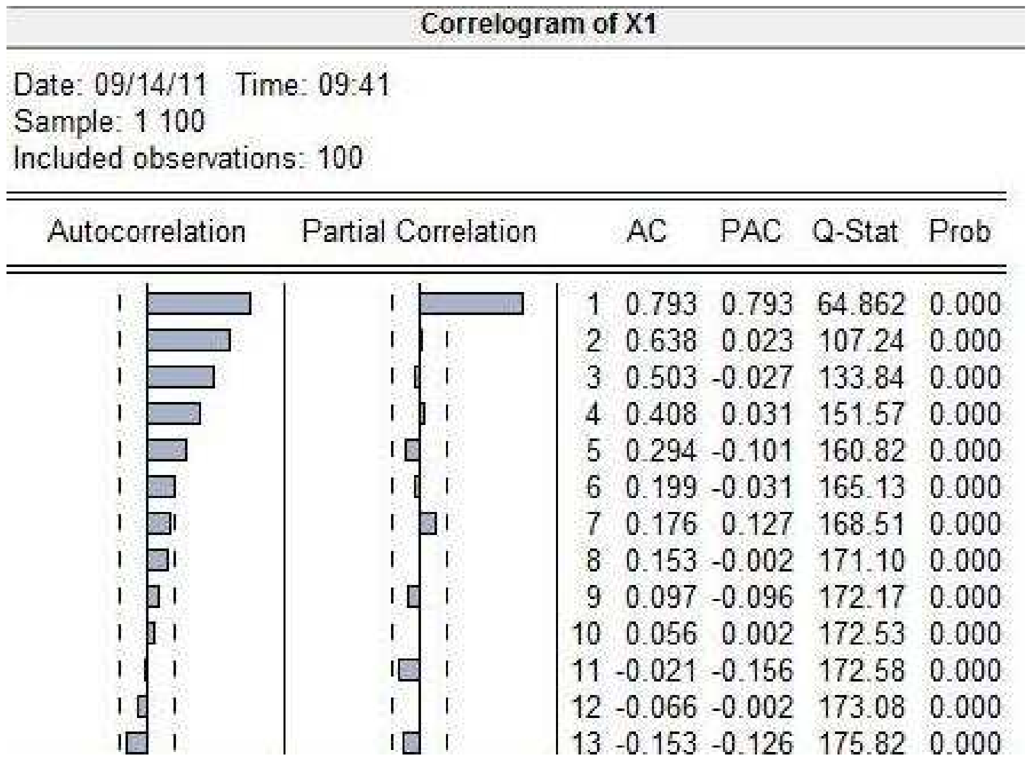


PACF - odhadovanie z dát

- Za teoretické autokorelácie vo vzťahu (15) dosadíme ich konzistentné odhady $\rightarrow$ dostaneme konzistentný odhad $\hat{\Phi}_{kk}$
- Pre AR(p) proces je $\Phi_{kk} = 0$ pre $k > p$, pre tieto k asymptoticky platí

$$Var[\hat{\Phi}_{kk}] \approx \frac{1}{T}$$

Príklad: simulované dáta (1)



Correlogram of X1

Date: 09/14/11 Time: 09:41
Sample: 1 100
Included observations: 100

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.793	0.793	64.862	0.000
		2	0.638	0.023	107.24	0.000
		3	0.503	-0.027	133.84	0.000
		4	0.408	0.031	151.57	0.000
		5	0.294	-0.101	160.82	0.000
		6	0.199	-0.031	165.13	0.000
		7	0.176	0.127	168.51	0.000
		8	0.153	-0.002	171.10	0.000
		9	0.097	-0.096	172.17	0.000
		10	0.056	0.002	172.53	0.000
		11	-0.021	-0.156	172.58	0.000
		12	-0.066	-0.002	173.08	0.000
		13	-0.153	-0.126	175.82	0.000

Z odhadnutej ACF a PACF usúdime, že dáta boli zrejme generované AR(1) procesom.

Príklad: simulované dáta (1)

Odhadneme AR(1) model:

Dependent Variable: X1
Method: Least Squares
Date: 09/14/11 Time: 09:44
Sample (adjusted): 2 100
Included observations: 99 after adjustments
Convergence achieved after 3 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	5.062255	0.481641	10.51044	0.0000
AR(1)	0.793600	0.060654	13.08403	0.0000

R-squared	0.638318	Mean dependent var	4.990404
Adjusted R-squared	0.634590	S.D. dependent var	1.633787
S.E. of regression	0.987611	Akaike info criterion	2.832939
Sum squared resid	94.61140	Schwarz criterion	2.885366
Log likelihood	-138.2305	F-statistic	171.1917
Durbin-Watson stat	2.076048	Prob(F-statistic)	0.000000

Inverted AR Roots	.79

Príklad: simulované dáta (1)

Získaný model

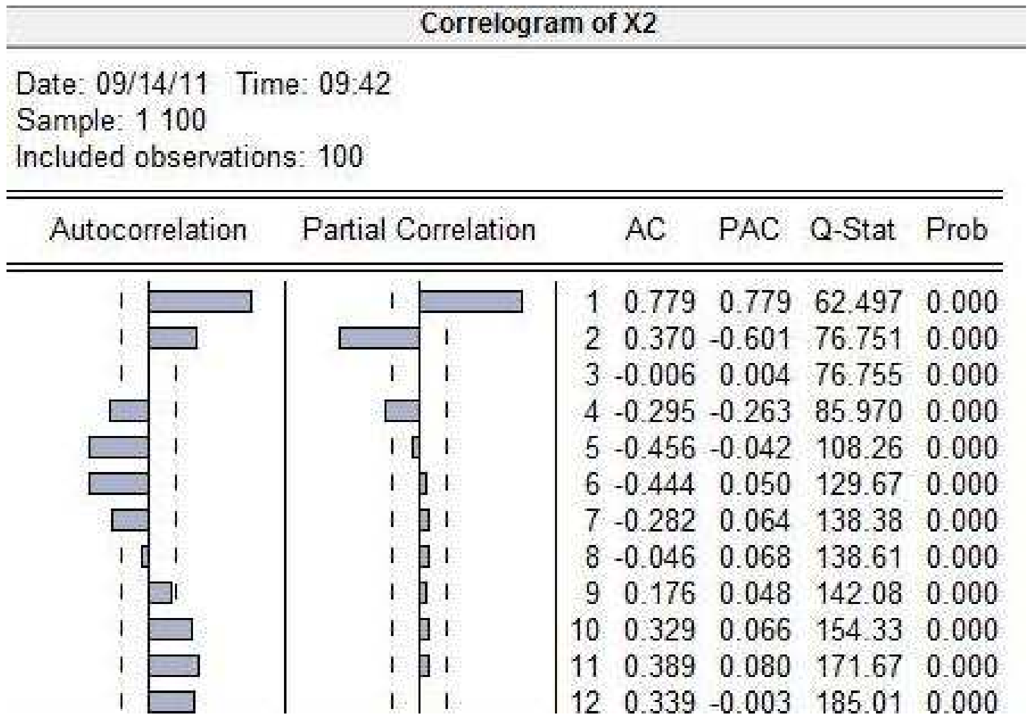
- je stacionárny
- má dobré rezíduá:

Correlogram of Residuals

Date: 09/14/11 Time: 09:47
Sample: 2 100
Included observations: 99
Q-statistic probabilities adjusted for 1 ARMA term(s)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.038	-0.038	0.1443	
		2	0.084	0.083	0.8703	0.351
		3	-0.020	-0.014	0.9107	0.634
		4	0.080	0.073	1.5889	0.662
		5	-0.008	-0.000	1.5959	0.810
		6	-0.121	-0.136	3.1791	0.672
		7	0.034	0.030	3.3043	0.770
		8	0.111	0.132	4.6631	0.701
		9	-0.079	-0.085	5.3556	0.719
		10	0.032	0.028	5.4693	0.792
		11	-0.071	-0.059	6.0422	0.812
		12	0.109	0.065	7.4115	0.765

Príklad: simulované dáta (2)


Correlogram of X2

Date: 09/14/11 Time: 09:42
Sample: 1 100
Included observations: 100

	AC	PAC	Q-Stat	Prob
1	0.779	0.779	62.497	0.000
2	0.370	-0.601	76.751	0.000
3	-0.006	0.004	76.755	0.000
4	-0.295	-0.263	85.970	0.000
5	-0.456	-0.042	108.26	0.000
6	-0.444	0.050	129.67	0.000
7	-0.282	0.064	138.38	0.000
8	-0.046	0.068	138.61	0.000
9	0.176	0.048	142.08	0.000
10	0.329	0.066	154.33	0.000
11	0.389	0.080	171.67	0.000
12	0.339	-0.003	185.01	0.000

$\rightarrow$ dáta vyzerajú na AR(2) proces

Príklad: simulované dáta (2)

Odhadneme AR(2) model:

Dependent Variable: X2
Method: Least Squares
Date: 09/14/11 Time: 09:44
Sample (adjusted): 3 100
Included observations: 98 after adjustments
Convergence achieved after 3 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.597631	0.285055	5.604647	0.0000
AR(1)	1.446250	0.067725	21.35463	0.0000
AR(2)	-0.764615	0.068891	-11.09894	0.0000

R-squared	0.858305	Mean dependent var	1.609184
Adjusted R-squared	0.855322	S.D. dependent var	2.360745
S.E. of regression	0.897947	Akaike info criterion	2.652722
Sum squared resid	76.59931	Schwarz criterion	2.731854
Log likelihood	-126.9834	F-statistic	287.7266
Durbin-Watson stat	1.999308	Prob(F-statistic)	0.000000

Inverted AR Roots	.72-.49i	.72+.49i

Príklad: simulované dáta (2)

Získaný model

- je stacionárny:

Inverse Roots of AR/MA Polynomial(s)
Specification: X2 C AR(1) AR(2)
Date: 09/14/11 Time: 09:53
Sample: 1 100
Included observations: 98

AR Root(s)	Modulus	Cycle
0.723125 ± 0.491635i	0.874423	10.52298

No root lies outside the unit circle.
ARMA model is stationary.

Príklad: simulované dáta (2)

- má dobré rezíduá:

Correlogram of Residuals

Date: 09/14/11 Time: 09:47
Sample: 3 100
Included observations: 98
Q-statistic probabilities adjusted for 2 ARMA term(s)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.003	-0.003	0.0012	
		2	-0.095	-0.095	0.9273	
		3	-0.045	-0.046	1.1386	0.286
		4	0.166	0.158	4.0041	0.135
		5	-0.009	-0.016	4.0118	0.260
		6	-0.073	-0.048	4.5762	0.334
		7	-0.030	-0.018	4.6744	0.457
		8	0.106	0.072	5.9023	0.434
		9	0.052	0.048	6.1963	0.517
		10	-0.052	-0.023	6.4958	0.592
		11	0.116	0.143	8.0085	0.533
		12	0.008	-0.025	8.0167	0.627

Príklad: simulované dáta (2)

Ak by sme odhadli AR(1), nedostali by sme dobrý model:

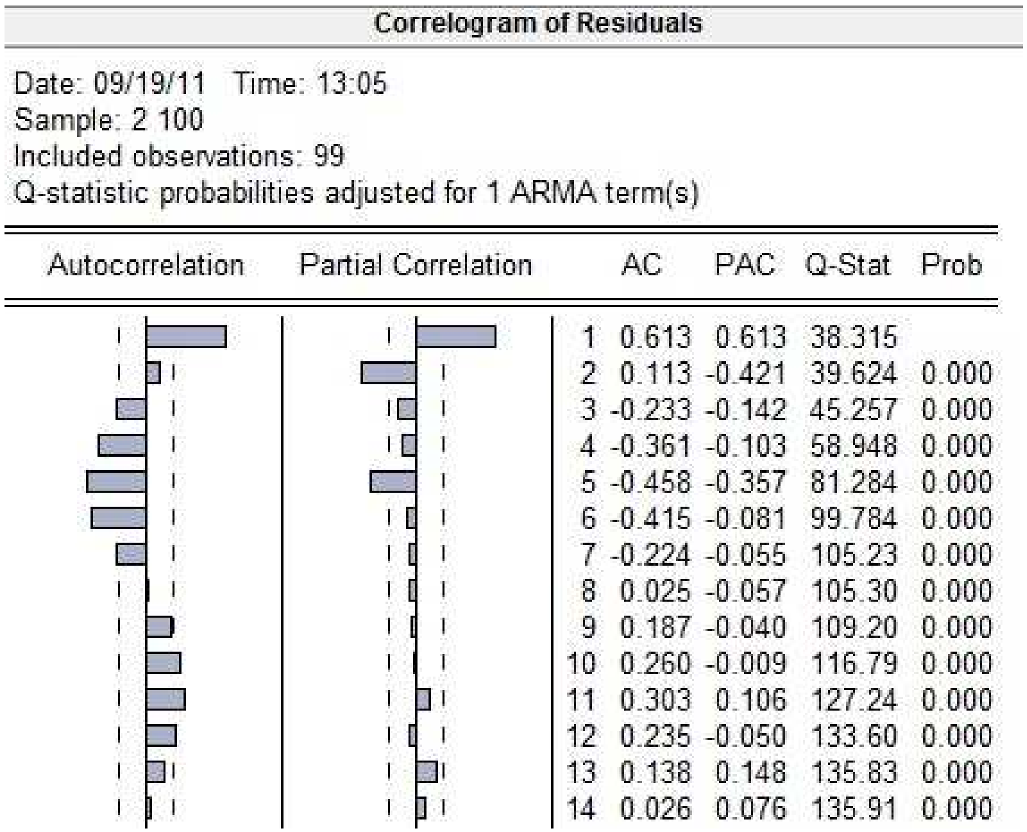
Dependent Variable: X2
Method: Least Squares
Date: 09/19/11 Time: 13:05
Sample (adjusted): 2 100
Included observations: 99 after adjustments
Convergence achieved after 3 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.772761	0.812831	2.180970	0.0316
AR(1)	0.832918	0.058714	14.18595	0.0000

R-squared	0.674760	Mean dependent var	1.615152
Adjusted R-squared	0.671407	S.D. dependent var	2.349420
S.E. of regression	1.346759	Akaike info criterion	3.453274
Sum squared resid	175.9347	Schwarz criterion	3.505701
Log likelihood	-168.9371	F-statistic	201.2411
Durbin-Watson stat	0.768890	Prob(F-statistic)	0.000000

Inverted AR Roots .83

Príklad: simulované dáta (2)



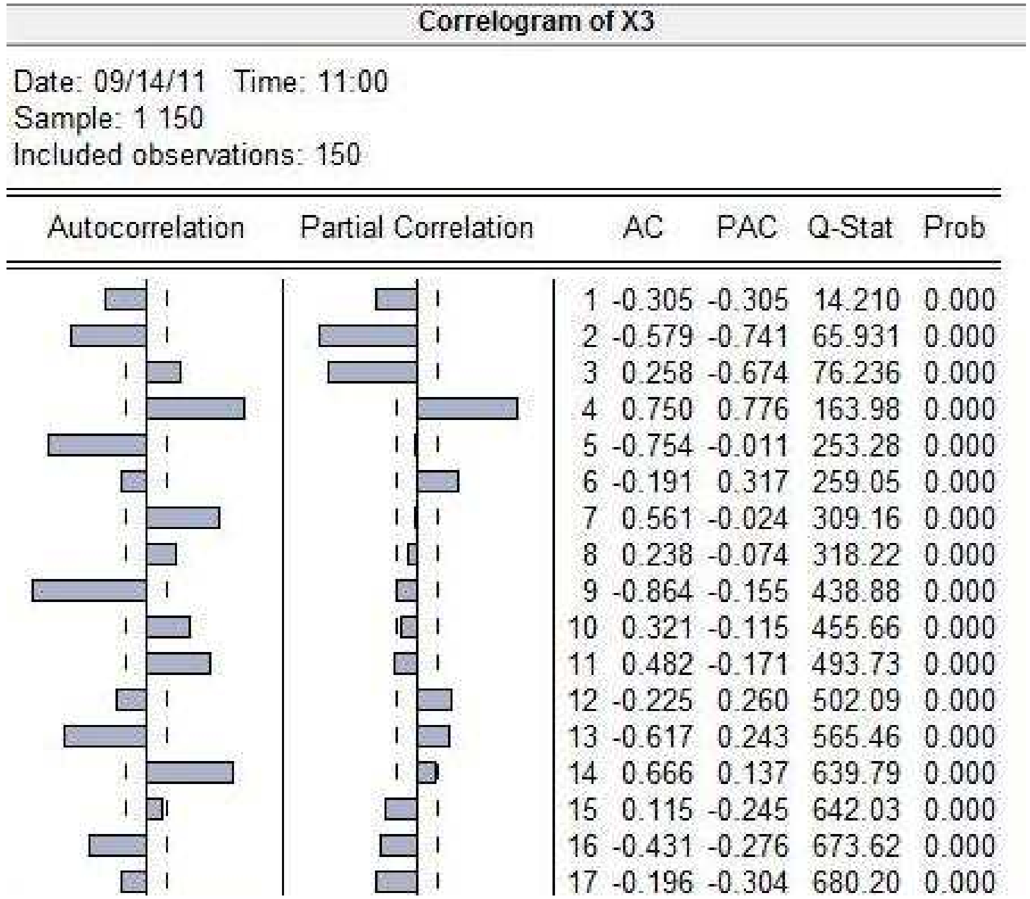
Correlogram of Residuals

Date: 09/19/11 Time: 13:05
Sample: 2 100
Included observations: 99
Q-statistic probabilities adjusted for 1 ARMA term(s)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.613	0.613	38.315	
		2	0.113	-0.421	39.624	0.000
		3	-0.233	-0.142	45.257	0.000
		4	-0.361	-0.103	58.948	0.000
		5	-0.458	-0.357	81.284	0.000
		6	-0.415	-0.081	99.784	0.000
		7	-0.224	-0.055	105.23	0.000
		8	0.025	-0.057	105.30	0.000
		9	0.187	-0.040	109.20	0.000
		10	0.260	-0.009	116.79	0.000
		11	0.303	0.106	127.24	0.000
		12	0.235	-0.050	133.60	0.000
		13	0.138	0.148	135.83	0.000
		14	0.026	0.076	135.91	0.000

- rezíduá nie sú biely šum

Príklad: simulované dáta (3)



Correlogram of X3

Date: 09/14/11 Time: 11:00
Sample: 1 150
Included observations: 150

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.305	-0.305	14.210	0.000
		2	-0.579	-0.741	65.931	0.000
		3	0.258	-0.674	76.236	0.000
		4	0.750	0.776	163.98	0.000
		5	-0.754	-0.011	253.28	0.000
		6	-0.191	0.317	259.05	0.000
		7	0.561	-0.024	309.16	0.000
		8	0.238	-0.074	318.22	0.000
		9	-0.864	-0.155	438.88	0.000
		10	0.321	-0.115	455.66	0.000
		11	0.482	-0.171	493.73	0.000
		12	-0.225	0.260	502.09	0.000
		13	-0.617	0.243	565.46	0.000
		14	0.666	0.137	639.79	0.000
		15	0.115	-0.245	642.03	0.000
		16	-0.431	-0.276	673.62	0.000
		17	-0.196	-0.304	680.20	0.000

$\rightarrow$ dáta vyzerajú na AR(4) proces

Príklad: simulované dáta (3)

Odhadneme AR(4) model:

Dependent Variable: X3
Method: Least Squares
Date: 09/14/11 Time: 11:00
Sample (adjusted): 5 150
Included observations: 146 after adjustments
Convergence achieved after 4 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.291702	0.164894	1.769026	0.0791
AR(1)	-0.446956	0.037374	-11.95901	0.0000
AR(2)	-0.142364	0.047729	-2.982727	0.0034
AR(3)	0.230836	0.047962	4.812920	0.0000
AR(4)	0.883163	0.037110	23.79828	0.0000

R-squared	0.964819	Mean dependent var	0.386781
Adjusted R-squared	0.963821	S.D. dependent var	4.926549
S.E. of regression	0.937062	Akaike info criterion	2.741512
Sum squared resid	123.8100	Schwarz criterion	2.843690
Log likelihood	-195.1304	F-statistic	966.7250
Durbin-Watson stat	1.788300	Prob(F-statistic)	0.000000

Inverted AR Roots	.90	-.18-.98i	-.18+.98i	-.99

Príklad: simulované dáta (3)

Získaný model

- je stacionárny:

Inverse Roots of AR/MA Polynomial(s)
Specification: X3 C AR(1) AR(2) AR(3) AR(4)
Date: 09/14/11 Time: 11:02
Sample: 1 150
Included observations: 146

AR Root(s)	Modulus	Cycle
-0.179619 ± 0.982122i	0.998412	3.586936
-0.986142	0.986142	
0.898425	0.898425	

No root lies outside the unit circle.
ARMA model is stationary.

Príklad: simulované dáta (3)

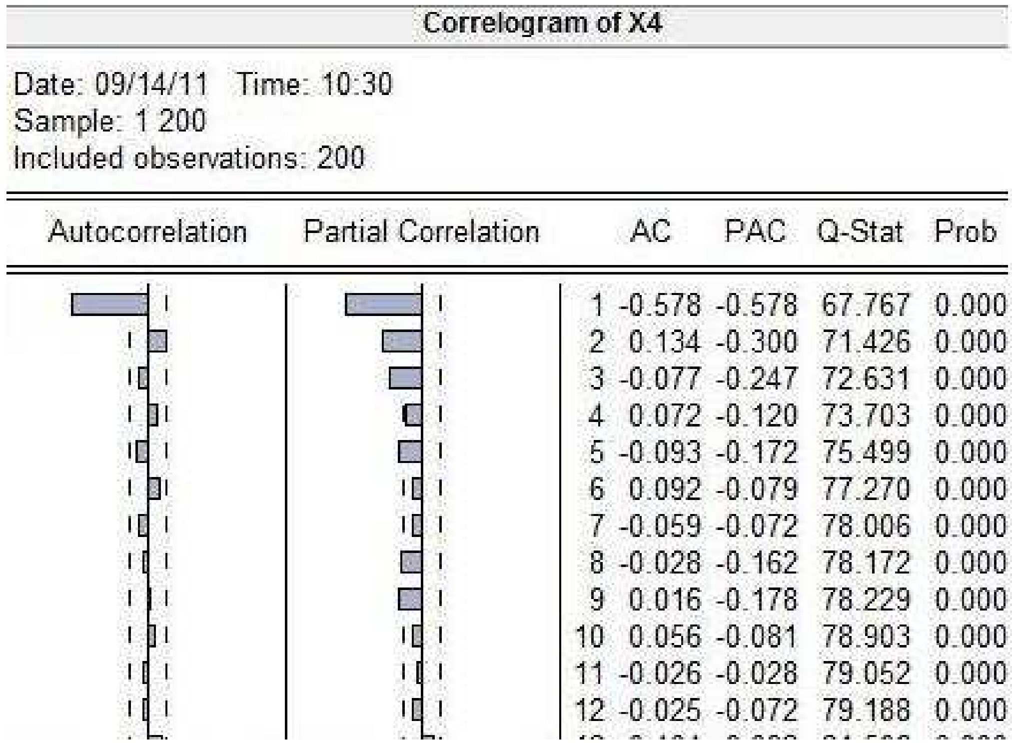
- má dobré rezíduá:

Correlogram of Residuals

Date: 09/14/11 Time: 11:01
Sample: 5 150
Included observations: 146
Q-statistic probabilities adjusted for 4 ARMA term(s)

	AC	PAC	Q-Stat	Prob
1	0.099	0.099	1.4701	
2	-0.049	-0.059	1.8249	
3	-0.060	-0.050	2.3672	
4	-0.025	-0.016	2.4594	
5	-0.035	-0.037	2.6425	0.104
6	-0.027	-0.025	2.7513	0.253
7	0.057	0.057	3.2548	0.354
8	-0.036	-0.056	3.4624	0.484
9	-0.024	-0.013	3.5525	0.615
10	-0.077	-0.075	4.4834	0.612
11	-0.015	-0.007	4.5213	0.718
12	-0.106	-0.115	6.3164	0.612
13	0.011	0.024	6.3353	0.706
14	0.062	0.036	6.9714	0.728
15	-0.109	-0.136	8.9264	0.629

Príklad: simulované dáta (4)



Correlogram of X4

Date: 09/14/11 Time: 10:30
Sample: 1 200
Included observations: 200

	AC	PAC	Q-Stat	Prob
1	-0.578	-0.578	67.767	0.000
2	0.134	-0.300	71.426	0.000
3	-0.077	-0.247	72.631	0.000
4	0.072	-0.120	73.703	0.000
5	-0.093	-0.172	75.499	0.000
6	0.092	-0.079	77.270	0.000
7	-0.059	-0.072	78.006	0.000
8	-0.028	-0.162	78.172	0.000
9	0.016	-0.178	78.229	0.000
10	0.056	-0.081	78.903	0.000
11	-0.026	-0.028	79.052	0.000
12	-0.025	-0.072	79.188	0.000

→ nebude to AR proces

Nabudúce: procesy s takouto ACF a PACF

Príklad: reálne dáta

- Z predchádzajúcich príkladov s reálnymi dátami:
  - ◇ volebné preferencie (vľavo) - AR(1)
  - ◇ úrokové miery (vpravo) - AR(2)