Autoregresné modely: AR(p) model, momenty a autokorelačná funkcia

Beáta Stehlíková

Časové rady

Fakulta matematiky, fyziky a informatiky, UK v Bratislave

Stredná hodnota

Zopakovanie príkladu

- Pripomeňme si príklad o úrokových mierach

```
library(astsa)
r20 <- read.table("data/R20Q.txt")$V1
rS <- read.table("data/RSQ.txt")$V1
spread <- r20 - rS
spread <- ts(spread, frequency = 4, start = c(1952, 1))
ar2 <- sarima(spread, 2, 0, 0, details = FALSE)
ar2$fit$coef
```

```
##         ar1        ar2      xmean
##   1.1808724 -0.2885893  1.0449355
```

- Už vieme, že tento model je stacionárny, takže naše výpočty so strednou hodnotou boli oprávnené.

- Máme parametre:

```
##       ar1        ar2      xmean
##  1.1808724 -0.2885893  1.0449355
```

- Pre $\alpha_1 =$ `ar1`, $\alpha_2 =$ `ar2`, $\mu = \mathbb{E}(x_s) =$ `xmean` platí:

$$x_t = \delta + \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + u_t \quad / \, \mathbb{E}()$$

$$\mu = \delta + \alpha_1 \mu + \alpha_2 \mu$$

$$\delta = \mu(1 - \alpha_1 - \alpha_2)$$

- Model teda je

$$x_t = 0.1125 + 1.1809 x_{t-1} - 0.2886 x_{t-2} + u_t$$

Zovšeobecnenie: stredná hodnota AR(p) procesu

- Majme stacionárny AR(p) proces

$$x_t = \delta + \alpha_1 x_{t-1} + \cdots + \alpha_p x_{t-p} + u_t$$

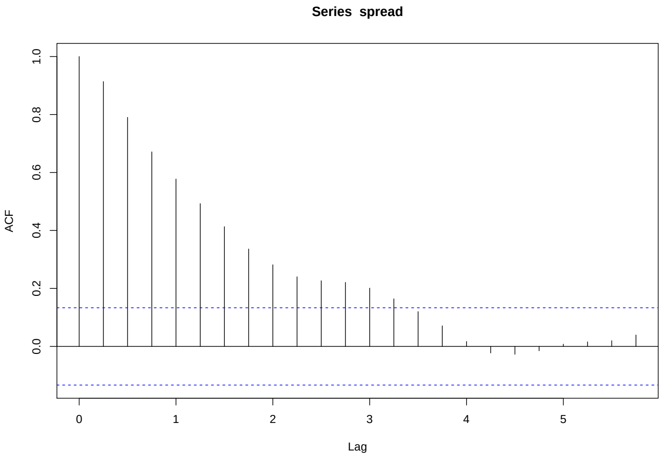
- Označme $\mu = \mathbb{E}(x_t)$ a spravme strednú hodnotu z obidvoch strán:

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

- Dá sa dokázať, že stredná hodnota procesu a parameter δ majú rovnaké znamienko.

Autokorelačná funkcia: motivácia

- Videli sme výberovú ACF pre spread:



Poznámky a otázky:

- Výberová ACF pre spread sa podobala na AR(1) proces
- Napriek tomu AR(1) nebol dobrý model kvôli rezíduám, ale AR(2) už áno
- Aký priebeh má ACF pre AR(2) proces?
- *Môže mať podobný priebeh ako pre AR(1)?* Zdá sa totiž, že áno.
- *Môže mať "úplne iný" priebeh ako pre AR(1)?* Teda, môžeme niekedy povedať, že *"toto určite nie je AR(1), ale AR(2) by to mohol byť"*?

AR(2): autokovariancie

- Môžeme uvažovať $\delta = 0$ (zodpovedá to procesu $x_t = y_t - \mathbb{E}(y_t)$, posun procesu o konštantu nezmení autokovariancie a autokorelácie - premyslite si, prečo):

$$\begin{aligned} x_t &= \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + u_t \ / \times x_{t-s}, \mathbb{E}(.) \\ \mathbb{E}(x_t x_{t-s}) &= \alpha_1 \mathbb{E}(x_{t-1} x_{t-s}) + \alpha_2 \mathbb{E}(x_{t-2} x_{t-s}) + \mathbb{E}(x_{t-s} u_t) \end{aligned}$$

- Potom $\mathbb{E}(x_t) = 0 \Rightarrow \gamma(s) = \mathrm{Cov}(x_t, x_{t-s}) = \mathbb{E}(x_t x_{t-s})$
- Pre $s = 0, 1, 2$ dostaneme sústavu rovníc:

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

$\rightarrow$ z nej $\gamma(0) = \mathbb{D}(x_t)$, $\gamma(1)$, $\gamma(2)$

- Pre $s \geq 2$ diferenčná rovnica

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

začiatočné podmienky z predchádzajúceho bodu

Cvičenie

Uvažujme AR(2) proces $x_t = 0,6x_{t-1} + 0,2x_{t-2} + u_t$, pričom disperzia bieleho šumu je 0,25.

- Ukážte, že je stacionárny.
- Odvoďte (t.j. spravte všetky odvodenia, nedosaďte do všeobecných vzťahov z prednášky) diferenčnú rovnicu pre autokovariancie a sústavu rovníc pre začiatočné podmienky.
- Vyriešte sústavu rovníc z predchádzajúceho bodu a rekurentne počítajte autokovariancie.
- Čomu sa rovná disperzia procesu?
- Vypočítajte z autokovariancií autokorelácie.

AR(2): autokorelácie

- Diferenčnú rovnicu (1) vydelíme $\gamma(0)$:

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

- Rovnice určujúce začiatočné podmienky pre autokovariancie boli tri (hľadali sme tri hodnoty - $\gamma(0)$, $\gamma(1)$, $\gamma(2)$), v prípade autokorelácií však $\rho(0)$ nehľadáme (je to 1)

Pre AR(2) môžeme spraviť:

- Druhú rovnicu vydelíme $\gamma(0)$a využijeme $\rho(0) = 1 \Rightarrow$ dostaneme $\rho(1)$ a máme začiatočné podmienky pre rovnicu (2), ktorá platí aj pre $s = 2$:

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

Univerzálny postup:

- Hodnotou $\gamma(0)$ vydelíme len posledné dve rovnice
- Ich riešením sú $\rho(1), \rho(2)$, čo sú začiatočné podmienky pre (2)

AR(2) model pre spread: ACF modelu

- Spread máme modelovaný AR(2) procesom

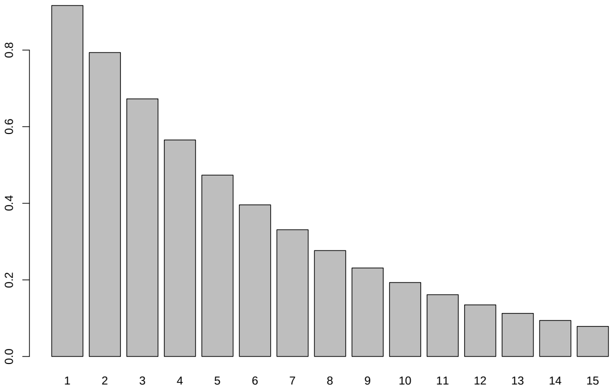
$$x_t = 0.1125 + 1.1809 x_{t-1} - 0.2886 x_{t-2} + u_t,$$

- Nájdeme ACF tohto procesu
- Diferenčná rovnica pre autokorelácie:

$$\rho(s) - 1.1809\rho(s-1) + 0.2886\rho(s-2) = 0$$

$$\rho(0) = 1, \rho(1) = \frac{1.1809}{1 - 0.2886}$$

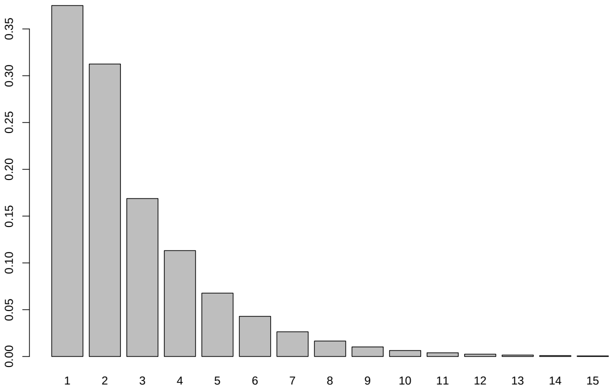
AR(2) model pre spread: ACF modelu



ACF pre AR(2) proces: rôzny charakter priebehu

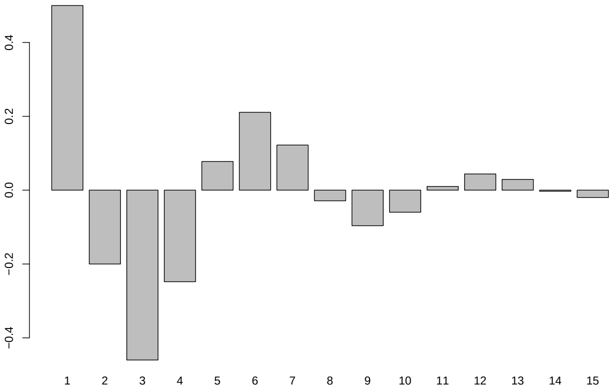
Príklad 1.

```
barplot(ARMAacf(ar=c(0.3, 0.2), lag.max = 15)[-1])
```



Príklad 2.

```
barplot(ARMAacf(ar=c(0.8, -0.6), lag.max = 15)[-1])
```



- ACF je riešením diferenčnej rovnice

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

$\Rightarrow$ priebeh závisí od koreňov charakteristickej rovnice

$$\lambda^2 - \alpha_1\lambda - \alpha_2 = 0$$

- λ_1, λ_2 reálne a rôzne: ACF má tvar

$$\rho(s) = c_1\lambda_1^s + c_2\lambda_2^s$$

zo stacionarity: $|\lambda_{1,2}| < 1$

- λ_1, λ_2 komplexné: ACF je tlmená kombinácia sínusu a kosínusu

$$\rho(s) = r^s(c_1\cos(ks) + c_2\sin(ks))$$

zo stacionarity: $|r| < 1$

Príklad s periodickým charakterom

- Proces: $x_t = 1.4x_{t-1} - 0.85x_{t-2} + u_t$
- Korelácie spĺňajú diferenčnú rovnicu

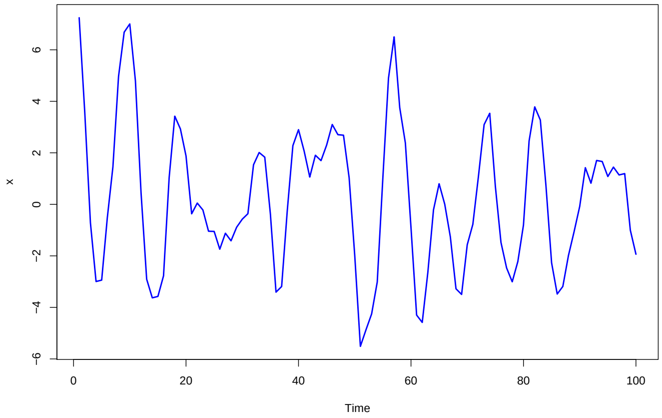
$$\rho(k) = 1.4\rho(k-1) - 0.85\rho(k-2)$$

- Jej všeobecné riešenie

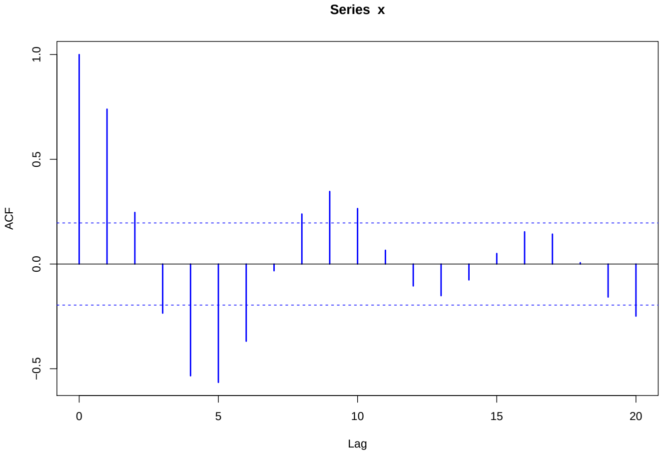
$$\rho(k) = 0.922^k(c_1\cos(0.709k) + c_2\sin(0.709k))$$

- Konštanty c_1, c_2 zo začiatočných podmienok $\rho(0), \rho(1)$
- $\cos(nt), \sin(nt)$ $\rightarrow$ perióda $\frac{2\pi}{n}$
- V našom prípade $\frac{2\pi}{n} = \frac{2\pi}{0.709} = 8.862 \approx 9 \Rightarrow$ v dátach generovaných týmto procesom sa dá čakať takáto perióda

```r
x <- arima.sim(model = list(ar = c(1.4, -0.85)), n = 100)
plot(x, lwd = 2, col = "blue")
```



```r
acf(x, lwd = 2, col = "blue")
```



ACF pre AR(p) proces

Disperzia, autokovariancie

- Znovu môžeme predpokladať nulovú strednú hodnotu, teda $x_t = \alpha_1 x_{t-1} + \cdots + \alpha_p x_{t-p} + u_t$
- Vynásobíme obidve strany členom x_{t-s} a spravíme strednú hodnotu:

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

- Pre $s = 0, 1, \ldots, p \to$ sústava $p + 1$ rovníc s neznámymi $\gamma(0), \gamma(1), \ldots, \gamma(p)$:

$$\begin{aligned} \gamma(0) &= \alpha_1\gamma(1) + \alpha_2\gamma(2) + \cdots + \alpha_p\gamma(p) + \sigma^2 \\ \gamma(1) &= \alpha_1\gamma(0) + \alpha_2\gamma(1) + \cdots + \alpha_p\gamma(p-1) \\ & \ldots \\ \gamma(p) &= \alpha_1\gamma(p-1) + \alpha_2\gamma(p-2) + \cdots + \alpha_p\gamma(0) \end{aligned} \tag{3}$$

- Ostatné autokovariancie z diferenčnej rovnice

$$\gamma(s) - \alpha_1\gamma(s-1) - \cdots - \alpha_p\gamma(s-p) = 0 \tag{4}$$

Autokorelácie

- Diferenčná rovnica pre autokorelácie - rovnicu (4) vydelíme disperziou $\gamma(0)$: $\rho(s) - \alpha_1 \rho_{s-1} - \cdots - \alpha_p \rho(s-p) = 0$
- Začiatočné podmienky - posledných p vydelíme $\gamma(0)$:

$$\begin{aligned} \rho(1) &= \alpha_1 \rho(0) + \alpha_2 \rho(1) + \cdots + \alpha_p \rho(p-1) \\ &\ldots \\ \rho(p) &= \alpha_1 \rho(p-1) + \alpha_2 \rho(p-2) + \cdots + \alpha_p \rho(0) \end{aligned}$$

Cvičenie

Uvažujme proces $x_t = 0.2x_{t-1} + 0.3x_{t-2} - 0.6x_{t-3} + u_t$ (overovali sme jeho stacionaritu). Odvoďte:

- jeho disperziu, ak je $\mathbb{D}(u_t) = 10$
- Yule-Wolkerove rovnice
- ACF pre lagy 1- 5

```
# pre kontrolu, funkcia ARMAacf
ARMAacf(ar = c(0.2, 0.3, -0.6), lag.max = 5)
```

```
##          0          1          2           3
##  1.00000000  0.04347826  0.28260870 -0.53043478 -0.04739
```

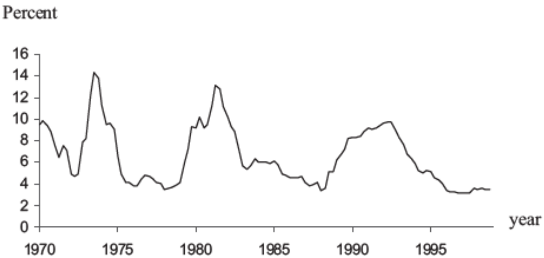
```r
# kvoli prehladnosti - hodnoty vypiseme pod seba
cat(ARMAacf(ar = c(0.2, 0.3, -0.6), lag.max = 5),
    sep = "\n")
```

```
## 1
## 0.04347826
## 0.2826087
## -0.5304348
## -0.0473913
## -0.3381739
```

AR(p) model použitý na reálne dáta

Dáta

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



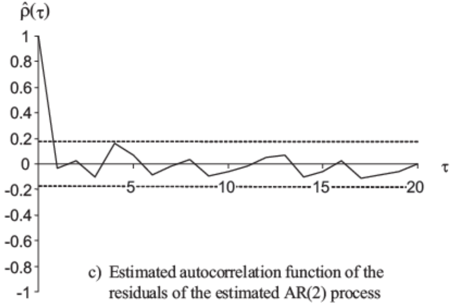
a) Three months money market rate in Frankfurt
1970 – 1998

Prebraté z učebnice *Kirchgässner & Wolters*, example 2.6

Odhadnutý AR(2) model

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

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



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

Otázky k odhadnutému modelu

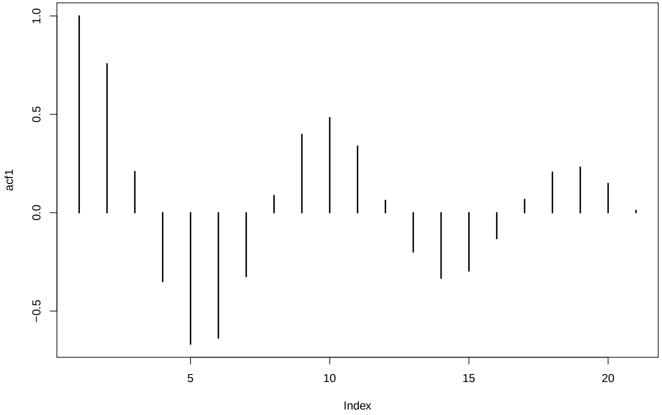
- Ukážte, že je stacionárny.
- Analyzujte rezíduá - autokorelogram a Q štatistiku. Aká hypotéza sa testuje, keď má Q štatistika rozdelenie so 6 stupňami voľnosti?
- Aká je stredná hodnota procesu?
- Aký je priebeh jeho autokorelačnej funkcie?
- Vysvetlite nasledovné tvrdenie z knihy: O aké korene ide? Odvoďte aj ostatné hodnoty uvedené v texte. Ako sa z nich vypočíta perióda?

The two roots of the process are 0.70 ± 0.06i, i.e. they indicate cycles which are strongly dampened. The modulus (dampening factor) is d = 0.706; the frequency f = 0.079 corresponds to a period of 79.7 quarters and therefore of nearly 20 years.

Záverečné poznámky k ACF, motivácia k PACF

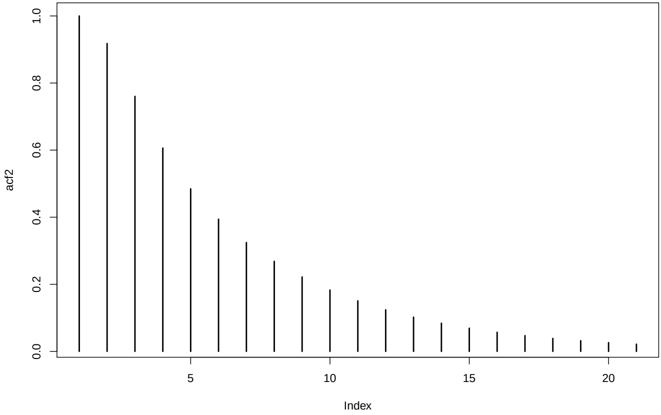
Autokorelácie - príklad 1

- AR(2) proces $x_t = 1.4x_{t-1} - 0.85x_{t-2} + u_t$



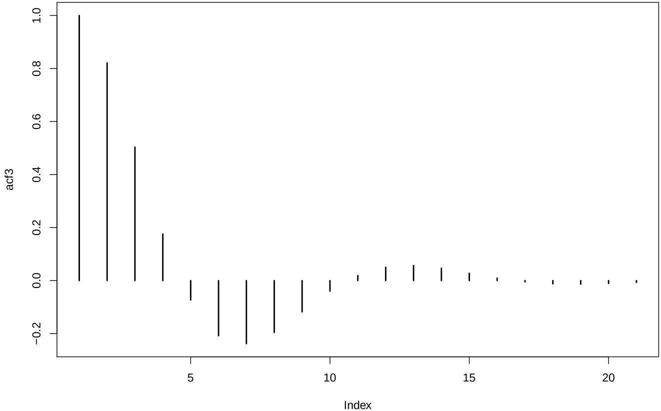
Autokorelácie - príklad 2

- AR(3) proces $x_t = 1.5x_{t-1} - 0.8x_{t-2} + 0.2x_{t-3} + u_t$

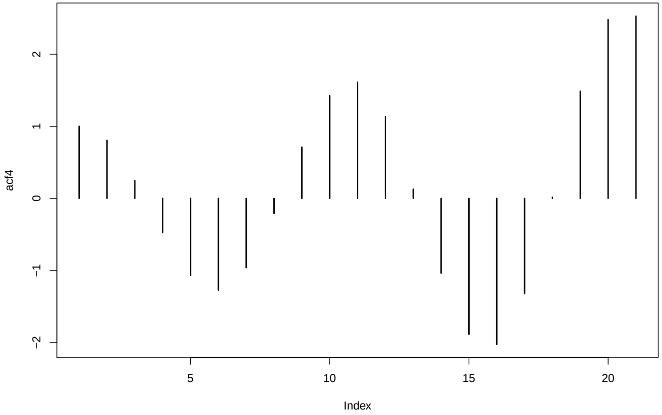


Autokorelácie - príklad 3

- AR(3) proces $x_t = 1.2x_{t-1} - 0.4x_{t-2} - 0.1x_{t-3} + u_t$
- Dajú sa očakávať komplexné korene



Autokorelácie - príklad 4



Autokorelácie - príklad 4

- Proces nie je stacionárny $\rightarrow$ predchádzajúci výpočet nemá zmysel

```
polyroot(c(1, -1.5, 0.8, 0.2))
```

```
## [1]  0.7610683+0.5711581i  0.7610683-0.5711581i -5.5221
```

```
abs(polyroot(c(1, -1.5, 0.8, 0.2)))
```

```
## [1] 0.9515496 0.9515496 5.5221367
```

Autokorelácie - príklad 5 - motivácia pre PACF

- ACF - jedna pre AR(2), druhá pre AR(3) - nevieme ich rozlíšiť
- Pri práci s dátami máme navyše len odhad ACF