ARMA modely
časť 4: Predikcie, odhadovanie parametrov

Beáta Stehlíková

Časové rady, FMFI UK, 2011/2012

IX.

Predikcie

Predikcie

- Sme v čase t, chceme predikciu hodnoty $x_{t+\tau}$, t. j. hodnotu procesu o τ období.
- Označme túto predikciu $\hat{x}_t(\tau)$, teda
  - index t označuje čas, v ktorom konštruujeme predikciu
  - argument τ označuje, na koľko období dopredu tá predikcia je
- Predikciou bude očakávaná hodnota procesu v tom čase, pri danej informácii, ktorú máme k dispozícii:

$$\hat{x}_t(\tau) = E_t[x_{t+\tau}]$$

(index t vo výraze E_t znamená, že strednú hodnotu počítame v čase t)

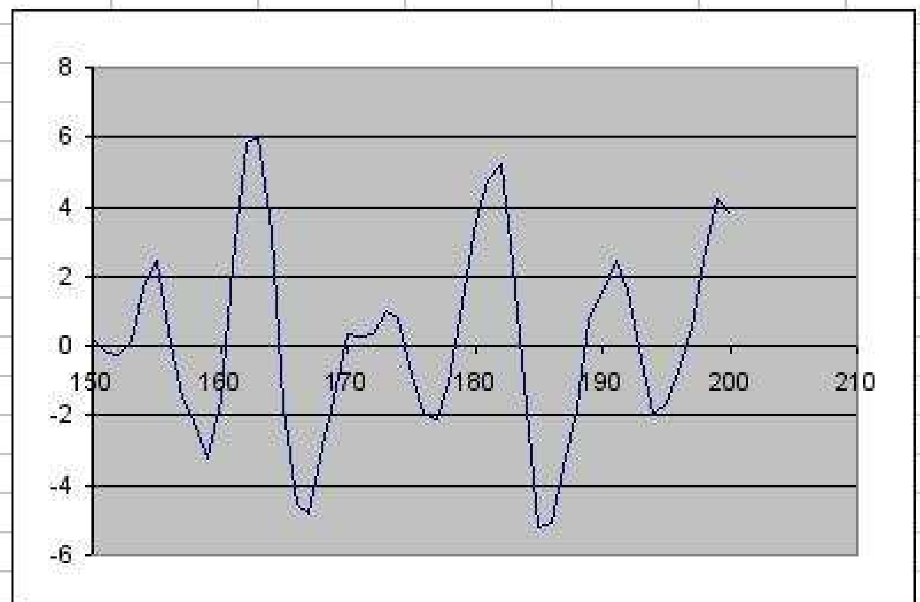
Predikcie pre AR proces

- Majme proces

$$x_t = 1.4x_{t-1} - 0.85x_{t-2} + u_t$$

a dáta:

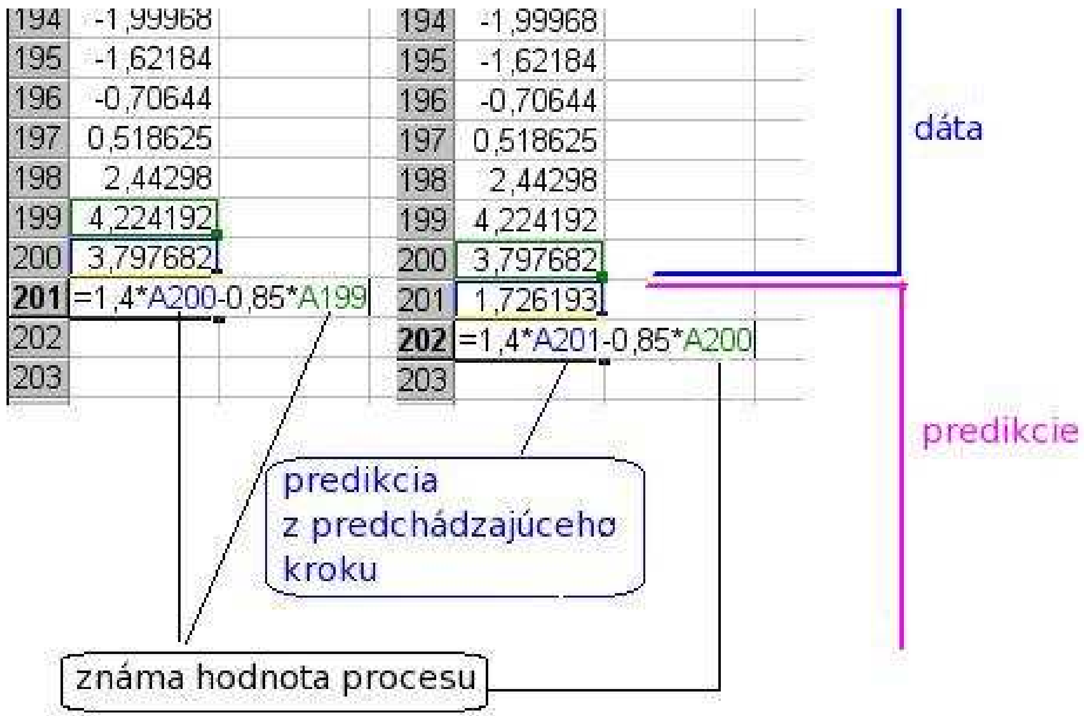
183	2,297308
184	-1,79321
185	-5,19365
186	-5,04636
187	-3,5133
188	-1,78753
189	0,706783
190	1,534983
191	2,447012
192	1,593168
193	-0,39105
194	-1,99968
195	-1,62184
196	-0,70644
197	0,518625
198	2,44298
199	4,224192
200	3,797682
201	
202	



- Chceme predikcie pre hodnoty v nasledujúcich časoch.

Predikcie pre AR proces

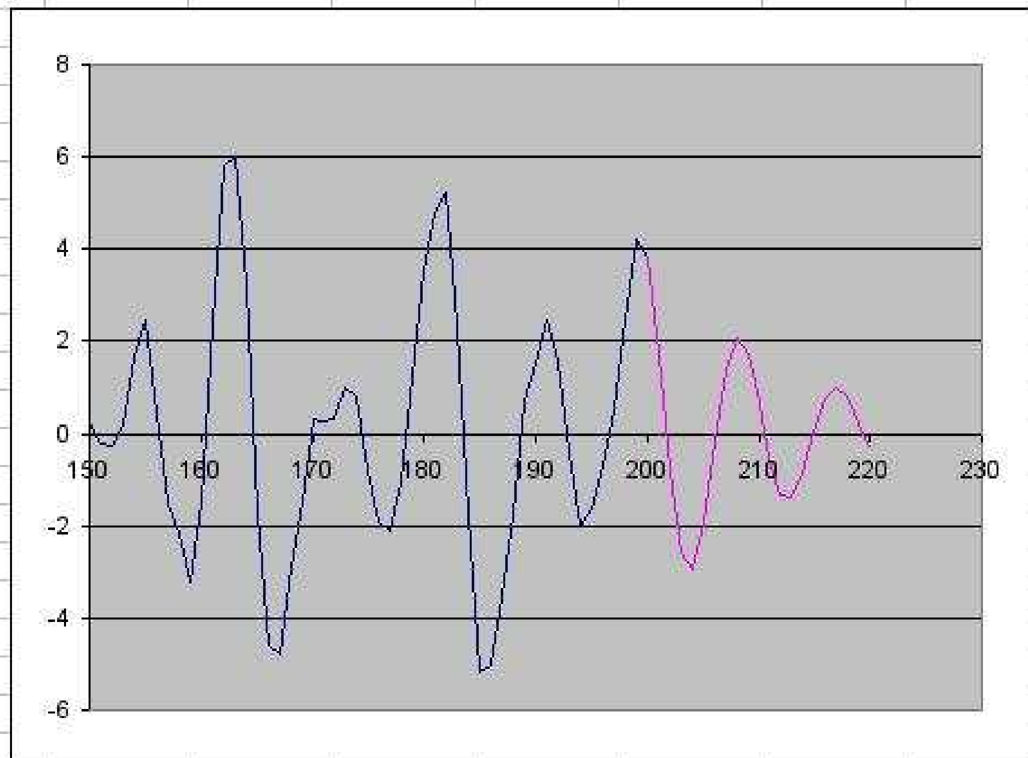
- Intuitívne:
  - ◇ do diferenčnej rovnice dosadzujeme známe hodnoty procesu a keď už k dispozícii nie sú, dosadzujeme predikcie
  - ◇ biely šum nahradíme nulou



Predikcie pre AR proces

- Výsledok:

189	0,706783
190	1,534983
191	2,447012
192	1,593168
193	-0,39105
194	-1,99968
195	-1,62184
196	-0,70644
197	0,518625
198	2,44298
199	4,224192
200	3,797682
201	1,726193
202	-0,81136
203	-2,60317
204	-2,95478
205	-1,924
206	-0,18203
207	1,38055
208	2,087499
209	1,749031
210	0,67427
211	-0,5427
212	-1,33291
213	-1,40478
214	[illegible]



Predikcie pre AR proces

- Vzťah $\hat{x}_t(\tau) = E_t[x_{t+\tau}]$ sa zhoduje s touto intuíciou.
- Majme AR(p) proces

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

Potom:

$$\begin{aligned} x_{t+\tau} &= \delta + \alpha_1 x_{t+\tau-1} + \ldots + \alpha_p x_{t+\tau-p} + u_{t+\tau}, \\ \underbrace{E_t[x_{t+\tau}]}_{\hat{x}_t(\tau)} &= \delta + \alpha_1 E_t[x_{t+\tau-1}] + \ldots + \alpha_p E_t[x_{t+\tau-p}] \\ &\quad + \underbrace{E_t[u_{t+\tau}]}_{0} \end{aligned}$$

pričom

$$E_t[x_{t+s}] = \begin{cases} \hat{x}_t(s) & \text{pre } s > 0 \\ x_{t+s} & \text{pre } s \leq 0 \end{cases}$$

Predikcie pre MA proces

- Majme MA(1) proces

$$x_t = \mu + u_t - \beta u_{t-1}$$

a počítajme predikcie $\hat{x}_t(\tau)$:

$$\begin{aligned} x_{t+s} &= \mu + u_{t+s} - \beta u_{t+s-1}, \\ \underbrace{E_t[x_{t+s}]}_{\hat{x}_t(s)} &= \mu + \underbrace{E_t[u_{t+s}]}_{0} - \beta \underbrace{E_t[u_{t+s-1}]}_{u_t \text{ pre } s=1, \text{ inak } 0} \end{aligned}$$

Teda:

$$\hat{x}_t(s) = \begin{cases} \mu - \beta u_t & \text{pre } s = 1 \\ \mu & \text{pre } s = 2, 3, \ldots \end{cases}$$

- CVIČENIE: Dokážte, že pre MA(q) proces platí $\hat{x}_t(s) = \mu$ pre $s > q$ a že pre $s \leq q$ predikcie obsahujú realizované hodnoty bieleho šumu u

Predikcie pre MA proces

- Predikcie pre MA(1) proces $x_t = \mu + u_t - \beta u_{t-1}$ sú

$$\hat{x}_t(s) = \begin{cases} \mu - \beta u_t & \text{pre } s = 1 \\ \mu & \text{pre } s = 2, 3, \ldots \end{cases}$$

  - obsahujú hodnotu bieleho šumu u_t. Táto hodnota už bola v čase konštrukcie predikcií realizovaná, nie sme však schopní ju pozorovať.
- Cvičenie z predch. slajdu ⇒ podobná situácia nastáva pre l'ubovol'ný MA(q) proces
- Ako prakticky počítat' predikcie?
  Myšlienka: vyjadríme u_t pomocou hodnôt procesu x
- Tento výpočet si ukážeme na MA(1) procese

Predikcie pre MA proces

- Vyjadrujeme teda u_t pomocou $x_1, x_2, \ldots, x_t$
- Využijeme pritom:
  - ◇ Pre MA(1) sme odvodili

$$\hat{x}_t(1) = \mu - \beta u_t \tag{1}$$

  - ◇ Pre predikčnú chybu platí

$$x_t - \hat{x}_{t-1}(1) = u_t \tag{2}$$

- Budeme postupne počítat' $\hat{x}_t(1)$ pre $t = 0, 1, 2, \ldots$

$$\hat{x}_0(1) \overset{(1)}{=} \mu - \beta u_0 \tag{3}$$

Predikcie pre MA proces

- Pre $t = 1$:

$$\begin{aligned}
\hat{x}_1(1) &\overset{(1)}{=} \mu - \beta u_1 \\
&\overset{(2)}{=} \mu - \beta[x_1 - \hat{x}_0(1)] \\
&\overset{(3)}{=} \mu - \beta[x_1 - (\mu - \beta u_0)] \\
&= \mu(1+\beta) - \beta x_1 - \beta^2 u_0 \qquad (4)
\end{aligned}$$

- Pre $t = 2$:

$$\begin{aligned}
\hat{x}_2(1) &\overset{(1)}{=} \mu - \beta u_2 \\
&\overset{(2)}{=} \mu - \beta[x_2 - \hat{x}_1(1)] \\
&\overset{(4)}{=} \mu - \beta[x_2 - (\mu(1+\beta) - \beta x_1 - \beta^2 u_0)] \\
&= \mu(1+\beta+\beta^2) - \beta x_2 - \beta^2 x_1 - \beta^3 u_0 \qquad (5)
\end{aligned}$$

Predikcie pre MA proces

- Pre všeobecné t:

$$\begin{aligned}\hat{x}_t(1) &= \mu(1+\beta+\beta^2+\ldots\beta^t) \\ &\quad -\beta x_t - \beta^2 x_{t-1} - \ldots - \beta^t x_1 - \beta^{t+1} u_0\end{aligned}$$

- Aj tu vystupuje nepozorovateľná hodnota u_0, ale vplyv člena $\beta^{t+1} u_0$ ide k nule pre $t \to \infty$ (kde t je počet dát) $\Rightarrow$ môžeme ho zanedbať

X.

Odhadovanie parametrov

Odhadovanie parametrov - EViews

- Viac možných prístupov k odhadovaniu parametrov, ukážeme si, čo robí softvér EViews, ktorý budeme používať na cvičeniach
- Ekonometria, 3. roč. - zápis odhadovanej rovnice napr.

```
y c x1 x2
```

  - model $y = c_0 + c_1 x_1 + c_2 x_2 + \varepsilon$, kde ε je biely šum
- Vo všeobecnosti ε nemusí byť biely šum. Ak je ε napr. AR(1) proces, zapíšeme to

```
y c x1 x2 ar(1)
```

- V modeloch, ktorými sa teraz zaoberáme, nie sú regresory, takže AR(1) model píšeme

```
y c ar(1)
```

  a znamená to $y = c + \varepsilon$, kde ε je AR(1) proces s $\delta = 0$

Odhadovanie parametrov - EViews

- Vo všeobecnosti ARMA(p,q) model:
  - ◇ Zápis ARMA(2,2) modelu:

    ```
    y c ar(1) ar(2) ma(1) ma(2)
    ```

  - ◇ Výstup:

Dependent Variable: Y
Method: Least Squares
Date: 10/02/09 Time: 18:30
Sample (adjusted): 3 216
Included observations: 214 after adjustments
Convergence achieved after 16 iterations
Backcast: 1 2

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.945871	0.481295	1.965263	0.0507
AR(1)	0.009916	0.117163	0.084634	0.9326
AR(2)	0.745537	0.110194	6.765656	0.0000
MA(1)	1.172727	0.118986	9.856026	0.0000
MA(2)	0.327399	0.069595	4.704361	0.0000

$y_t = \boxed{0.945871} + eps_t$

$eps_t = \boxed{0.009916\, eps_{t-1} + 0.745537\, eps_{t-2}} + u_t + \boxed{1.172727\, u_{t-1} + 0.327399\, u_{t-2}}$

u - biely šum

- Podrobnejšie na cvičení

Odhadovanie parametrov - AR

- Lineárny model s AR(1) chybami:

$$
\begin{aligned}
y_t &= \mathbf{x}_t^T \boldsymbol{\beta} + \varepsilon_t \\
\varepsilon_t &= \alpha \varepsilon_{t-1} + u_t
\end{aligned}
$$

  kde $\mathbf{x}_t$ je vektor regresorov, u_t je biely šum

- Model sa dá upraviť do tvaru

$$
y_t = \alpha y_{t-1} - (\mathbf{x}_t - \alpha \mathbf{x}_{t-1})^T \boldsymbol{\beta} + u_t \tag{6}
$$

- Parametre α, $\boldsymbol{\beta}$ sa odhadujú z transformovaného modelu (6) nelineárnou metódou najmenších štvorcov
- Analogicky AR(p) model

Odhadovanie parametrov - MA

- Lineárny model s MA(q) chybami:

$$\begin{aligned} y_t &= \mathbf{x}_t^T \boldsymbol{\beta} + \varepsilon_t \\ \varepsilon_t &= u_t + \theta_1 u_{t-1} + \ldots + \theta_q u_{t-q} \end{aligned}$$

  kde $\mathbf{x}_t$ je vektor regresorov, u_t je biely šum

- Dáta sú pozorované v časoch $t = 1, 2, \ldots, T$

Odhadovanie parametrov - MA

Postup I.

- Pre dané hodnoty $\boldsymbol{\theta}$, $\boldsymbol{\beta}$:
  1. vypočítame rezíduá $\hat{\varepsilon}_t = y_t - \mathbf{x}_t^T \hat{\boldsymbol{\beta}}$
  2. hodnoty bieleho šumu u pre $t = 0, -1, \ldots, -(q+1)$ (pred začiatkom pozorovania dát) položíme rovné nule: $\hat{u}_0 = 0, \hat{u}_{-1} = 0, \ldots, \hat{u}_{-(q-1)} = 0$
  3. nasledujúce hodnoty u počítame rekurzívne: $\hat{u}_t = \hat{\varepsilon}_t - \hat{\phi}_1 \hat{u}_{t-1} - \ldots \hat{\phi}_q \hat{u}_{t-q}$
  4. nakoniec minimalizujeme sumu štvorcov rezíduí
  $$\sum_{t=1}^{T} \left(y_t - \mathbf{x}_t^T \boldsymbol{\beta} - \phi_1 \hat{u}_{t-1} - \ldots - \phi_q \hat{u}_{t-q} \right)^2 \to \min_{\boldsymbol{\theta}, \boldsymbol{\beta}}$$
- S novými hodnotami $\boldsymbol{\theta}$, $\boldsymbol{\beta}$ cyklus opakujeme, kým neskonverguje

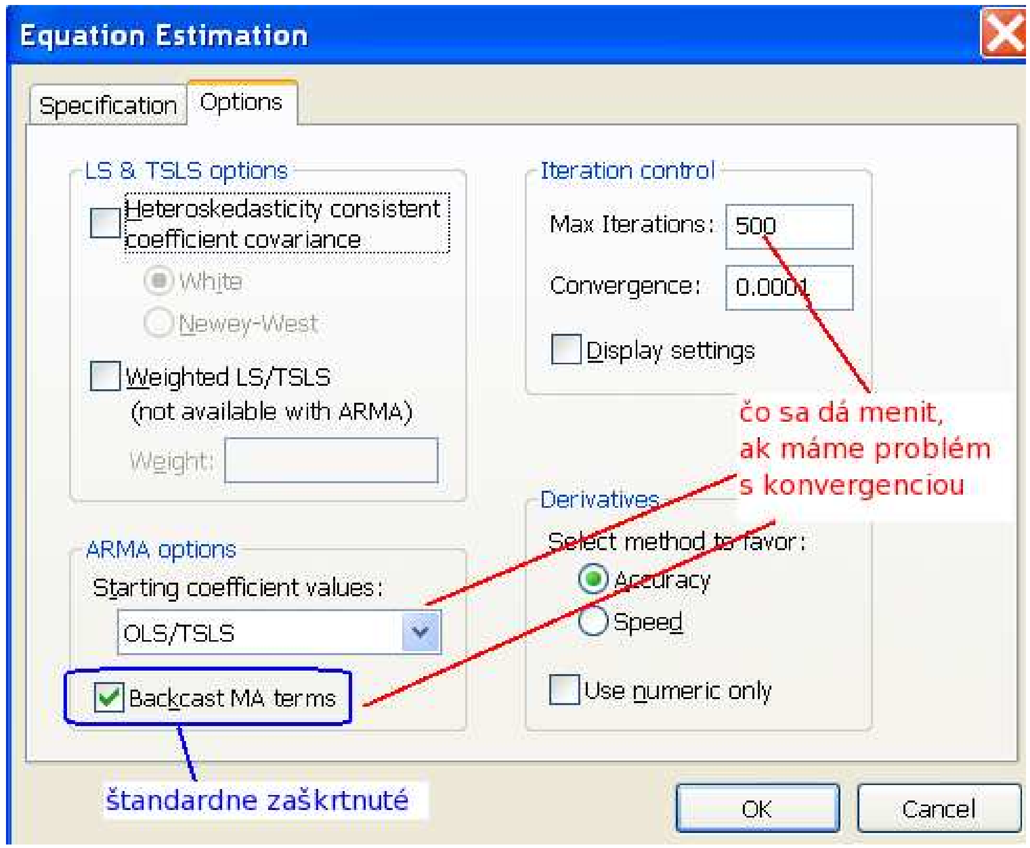
Odhadovanie parametrov - MA

Postup II. - backcasting

- Pre dané hodnoty $\boldsymbol{\theta}, \boldsymbol{\beta}$:
  1. vypočítame rezíduá - *rovnako ako v postupe I.*
  2. Backcast step - hodnoty bieleho šumu u pre $t = T+1, T+2, \dots T+q$ (po skončení pozorovania dát) položíme rovné nule: $\hat{u}_{T+1} = 0$, $\hat{u}_{T+2} = 0, \dots,$ $\hat{u}_{T+q} = 0$ a spätne dorátame u_t až po $\hat{u}_0, \hat{u}_{-1}, \dots,$ $\hat{u}_{-(q-1)}$
  3. nasledujúce hodnoty bieleho šumu u počítame rekurzívne - *rovnako ako v postupe I.*
  4. minimalizujeme sumu štvorcov rezíduí - *rovnako ako v postupe I.*
- S novými hodnotami $\boldsymbol{\theta}, \boldsymbol{\beta}$ cyklus opakujeme, kým neskonverguje

Odhadovanie parametrov - MA

- Backcasting v EViews:
  - ◇ V záložke *Options* pri odhadovaní modelu:



  - ◇ Štandardne je táto možnosť zaškrtnutá

Odhadovanie parametrov

- Ukážka výstupu - treba si skontrolovať konvergenciu:

Dependent Variable: D(LOGPCOCOA)
Method: Least Squares
Date: 10/18/11 Time: 12:29
Sample (adjusted): 1960M02 2002M09
Included observations: 512 after adjustments
Convergence achieved after 6 iterations
MA Backcast: 1960M01

	Coefficient	Std. Error	t-Statistic	Prob.
C	0.002454	0.003738	0.656614	0.5117
MA(1)	0.352651	0.041550	8.487292	0.0000

R-squared	0.112273	Mean dependent var	0.002398
Adjusted R-squared	0.110532	S.D. dependent var	0.066320
S.E. of regression	0.062548	Akaike info criterion	-2.701880
Sum squared resid	1.995221	Schwarz criterion	-2.685324
Log likelihood	693.6813	Hannan-Quinn criter.	-2.695390
F-statistic	64.50096	Durbin-Watson stat	1.987519
Prob(F-statistic)	0.000000		

Inverted MA Roots	-.35