

Autoregresné procesy

Beáta Stehlíková
FMFI UK Bratislava

Autoregresné procesy, podmienky stacionarity

Autoregresné procesy

Všade u_t označuje biely šum.

AR(1) proces z prednášky

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

Na príklade sme videli, že tento model nie je vždy vyhovujúci (modelovanie spreadu úrokových mier, rezíduá neboli bielym šumom). Preto:

Zovšeobecnenie: AR(p) proces

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

Podmienky stacionarity

AR(1) proces z prednášky

Pre proces

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

sme odvodili, že musí platiť $|\alpha| < 1$.

Proces môžeme prepísať pomocou operátora posunu L :

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

Podmienku stacionarity potom môžeme formulovať tak, že koreň

$$1 - \alpha L = 0$$

je v absolútnej hodnote väčší ako 1.

Podmienky stacionarity

AR(p) proces

Analogicky, AR(p) proces môžeme prepísať pomocou operátora posunu L :

$$(1 - \alpha_1 L - \alpha_2 L^2 - \dots - \alpha_p L^p)x_t = \delta + u_t$$

Na prednáške odvodíme podmienku stacionarity: korene (môžu byť reálne alebo komplexné)

$$1 - \alpha_1 L - \alpha_2 L^2 - \dots - \alpha_p L^p = 0$$

sú v absolútnej hodnote väčšie ako 1.

Geometricky: **korene sú mimo jednotkového kruhu**

Overenie stacionarity v R

Budeme potrebovať knižnicu `fArma` a z nej funkciu `armaRoots`

```
library(fArma)
```

Korene

$$1 - \alpha_1 L - \alpha_2 L^2 - \dots - \alpha_p L^p = 0$$

nájdeme pomocou

```
armaRoots(c(alfa_1, alfa_2, ..., alfa_p))
```

Overenie stacionarity v R

Napríklad:

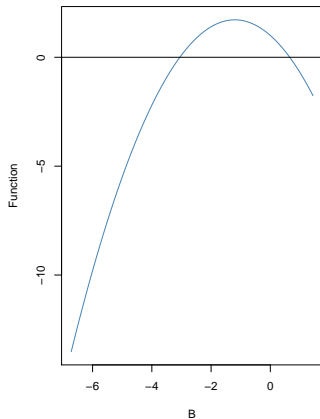
```
# ak chceme oba grafy vedľa seba  
# (inak budú na samostatných obrazkoch):  
par(mfrow = c(1, 2))  
  
# korene polynomu 1 - 1.2 L - 0.5 L^2 = 0  
armaRoots(c(1.2, 0.5))  
  
# povodne nastavenie  
par(mfrow = c(1, 1))
```

Výstup je grafický aj číselný v tabuľke (reálna a imaginárna časť, absolútna hodnota)

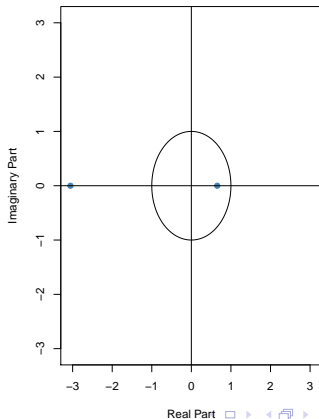
Overenie stacionarity v R

Graficky:

Polynomial Function vs. B



Roots and Unit Circle



Overenie stacionarity v R

Jeden z koreňov nie je mimo jednotkového kruhu, preto proces

$$(1 - 1.2L - 0.5L^2)x_t = \delta + u_t$$

nie je stacionárny.

Cvičenia:

- ▶ Nájdite príklad stacionárneho AR(2) procesu.
- ▶ Nájdite príklad stacionárneho a príklad nestacionárneho AR(3) procesu.
- ▶ Pre zadané procesy budeme overovať stacionaritu (súčasť kostry predmetu na skúške).

Odhadovanie parametrov AR modelu v R a konštrukcia predikcií

Odhadovanie parametrov AR modelu

Knižnica `astsa` (*applied statistical time series analysis*) a z nej funkcia `sarima` - aj pre zložitejšie modely:

```
library(astsa)
```

Príklad z prednášky - modelovanie spreadu

Príklad z prednášky - modelovanie spreadu, teda rozdielu medzi dlhodobou a krátkodobou úrokovou mierou.

Stiahnite si zo stránky dáta a načítajte ich:

```
rs <- read.table("RSQ.txt")      # short term rate  
r20 <- read.table("R20Q.txt")   # 20Y rate
```

Vytvoríme premennú spread:

```
spread <- r20 - rs
```

Príklad z prednášky - modelovanie spreadu

Z premennej `spread` **spravíme objekt typu time-series**, ktorý bude obsahovať time-series informáciu o časovej štruktúre dát.

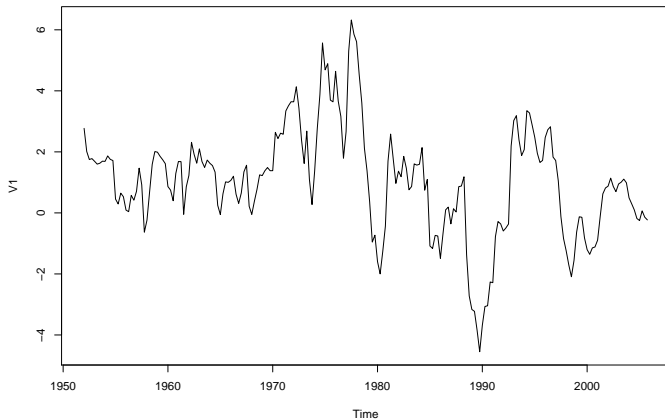
Použijeme funkciu `ts`:

```
spread <- ts(spread, frequency=4, start=c(1952,1))
```

Vykreslíme priebeh - na x-ovej osi budú správne časy:

```
plot(spread)
```

Príklad z prednášky - modelovanie spreadu



Príklad z prednášky - modelovanie spreadu

Úloha: Zobrazte výberovú autokorelačnú funkciu.

Na prednáške sme videli, že hoci sa podobá na ACF procesu AR(1), nie je to dobrý model pre dáta. Zopakujeme to - odhadneme model a skontrolujeme rezíduá.

Použitie funkcie `sarima`:

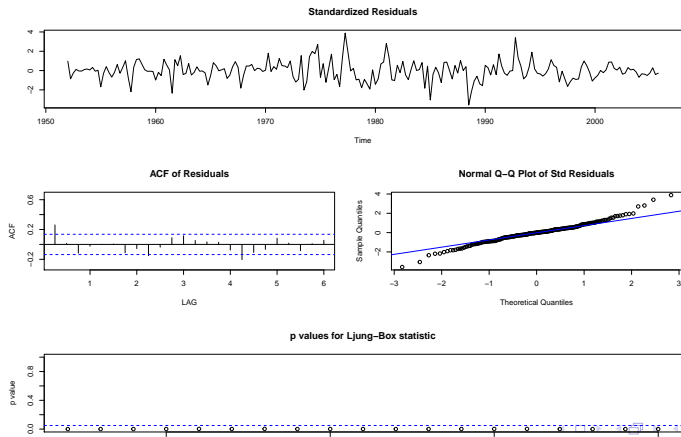
```
sarima(data,p,0,0) # AR(p) model pre data  
sarima(data,p,0,0, details=FALSE)
```

Dostaneme:

- ▶ odhady parametrov, štandardné odchýlky, informačné kritériá
- ▶ graficky rezíduá, ich ACF, p-hodnoty Ljung-Boxovej štatistiky

Príklad z prednášky - modelovanie spreadu

```
sarima(spread,1,0,0, details=FALSE)
```



Príklad z prednášky - modelovanie spreadu

Úlohy:

- ▶ Prečo nie sú tieto rezíduá vyhovujúce?
- ▶ Odhadnite pre spread AR(2) model.
- ▶ Má tento model dobré rezíduá? Je stacionárny?

Predikcie z AR modelu

Funkcia `sarima.for`:

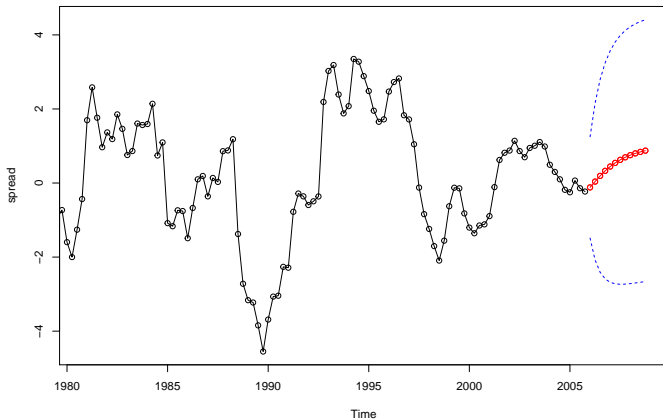
```
# predikcia z AR(p) modelu na n období dopredu:  
sarima.for(data,n,p,0,0)
```

Príklad z prednášky - modelovanie spreadu

Predikcie z AR(2) modelu na nasledujúce tri roky (teda 12 kvartálov):

```
sarima.for(spread, 12, 2, 0, 0)
```

Príklad z prednášky - modelovanie spreadu



AR model pre diferencie

AR model pre diferencie

AR model je modelom **pre stacionárne dáta**

Príkladom nestacionárnych dát sú napríklad dáta, ktoré majú trend (iný dôvod neskôr na prednáške aj cvičeniach, spolu so štatistickým testom)

Často pomôže **diferencovanie dát** a diferencie sú už stacionárne

AR model pre diferencie v R

Znovu funkcia `sarima`:

```
sarima(data,p,k,0) # AR(p) model pre k-te diferencie
```

Predikcie analogicky: `sarima(data,p,k,0)`, budú **pre pôvodnú premennú** (nie pre diferencie)

Príklad - teplota

Použijeme dáta zabudované v R-ku:

```
data(gtemp);
```

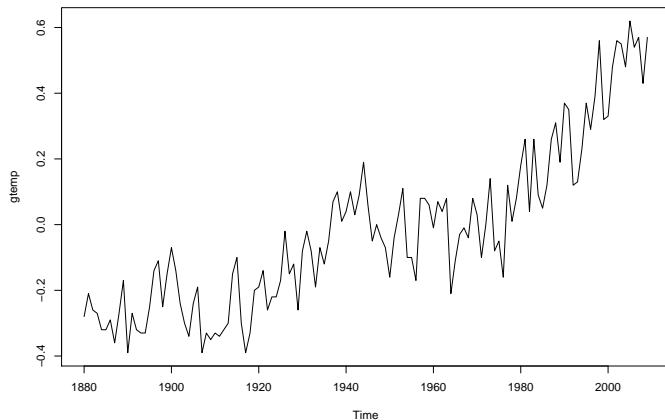
Dáta sú teraz v premennej `gtemp`.

Z popisu v helpe: *Global mean land-ocean temperature deviations (from 1951-1980 average) measured in degrees centigrade; annual data 1880-2009*

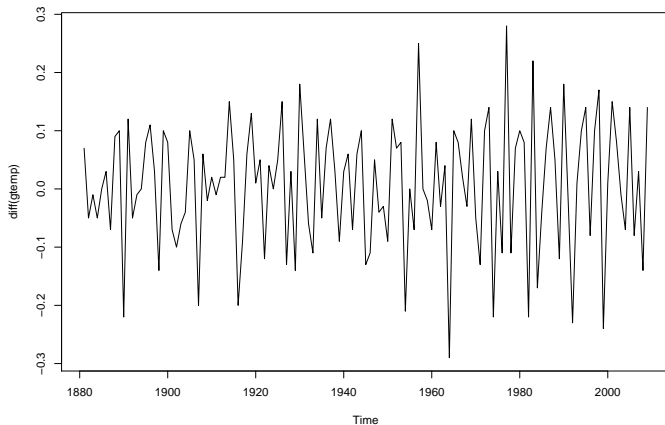
Úlohy:

- ▶ Zobrazte ich priebeh - uvidíme, že je v nich trend, a teda nie sú stacionárne
- ▶ Zopakujte pre diferencie `diff(gtemp)`

Príklad - teplota



Príklad - teplota



Príklad - teplota

Úlohy:

- ▶ Nájdite model pre diferencie premennej $gtemp$ - proces musí byť stacionárny a rezíduá musia byť bielym šumom
- ▶ Spravte predikcie pre nasledjúce roky.

Cvičenia

Cvičenia - pokyny

Špeciálny príklad ARIMA modelovania, ktoré je súčasťou kostry predmetu na skúške.

Úlohy:

- ▶ Nájdite model pre zadané dáta
- ▶ Spravte predikcie pre nasledujúce obdobie.

Cvičenie 1

Dáta, z popisu v helpe: *Quarterly U.S. GNP from 1947(1) to 2002(3)*:

```
data(gnp)
```

Nájdite model a spravte predikcie pre logaritmus premennej `gnp` tak, že budete jej diferencie (teda rýchlosť rastu hrubého národného produktu) modelovať AR procesom.

Cvičenie 2

Dáta z knižnice `astsa`, z popisu v helpe: *New York Harbor conventional regular gasoline weekly spot price FOB (in cents per gallon) from 2000 to mid-2010*. Zoberieme dáta od roku 2006.

```
data(gas)
gas <- window(gas, start=c(2006,1))
```

Nájdite model pre premennú `gas` tak, že budete jej diferencie modelovať AR procesom.

Príklad zo skúšky 2014, súčasťou bolo testovanie, či treba dáta diferencovať (budeme robiť neskôr počas semestra).

Cvičenie 3

Dáta z knižnice `astsa`, z popisu v helpe: *Leading indicator, 150 months; taken from Box and Jenkins (1970)*.

```
data(lead)
```

Nájdite model pre premennú `lead` tak, že budete jej diferencie modelovať AR procesom.

Príklad zo skúšky 2015, súčasťou bolo testovanie, či treba dáta diferencovať (budeme robiť neskôr počas semestra).