

Autoregresné procesy

Beáta Stehlíková
FMFI UK Bratislava

Autokorelačná funkcia AR procesu

Príklad z prednášky

Pre AR(3) proces

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

sme odvodili diferenčnú rovnicu pre autokorelačnú funkciu a sústavu rovníc pre začiatkové podmienky (tzv. Yule-Wolkerove rovnice).

- ▶ Vyriešime v R túto sústavu rovníc - dostaneme ACF(k) pre $k=1,2,3$.
- ▶ V cykle vypočítame ďalšie hodnoty ACF:

```
rho <- .... # prve tri zložky ako riesenie sustavy  
for (i in 4:10) rho[i] <- ... # z diferencnej rovnice
```

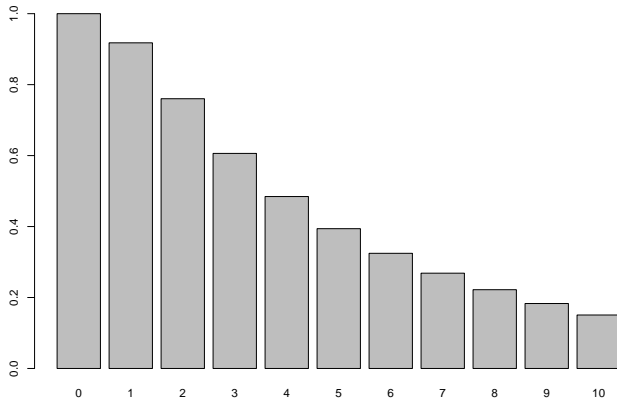
Matice a sústavy rovníc v R: `?matrix`, `?solve`

Výpočet ACF funkciou ARMAacf

```
r <- ARMAacf(ar=c(1.5, -0.8, 0.2), lag.max=10)
```

- ▶ Porovnajte r a rho
- ▶ Vykreslite priebeh (napríklad pomocou barplot)

Výsledok



Dáta Svetovej banky a grafy pomocou ggplot

Knižnice

WDI - World Development Indicators, prístup k dátam priamo z R
ggplot - pekné grafy

Nainštalujte a načítajte knižnicu WDI

```
library(WDI)  
library(ggplot2)
```

Vyhľadávanie dát vo WDI

Napríklad:

```
WDIsearch('gdp')
```

```
WDIsearch('gdp.*capita')
```

```
# pre tych co to poznaju:
```

```
# aj ine regularne vyrazy (regular expressions)
```

Ak je veľa výsledkov, môžeme vypísať niekoľko prvých:

```
WDIsearch('gdp.*capita')[1:5,]
```


Vyhľadávanie dát vo WDI

```
##      indicator
## [1,] "GDPPCKD"
## [2,] "GDPPCKN"
## [3,] "NV.AGR.PCAP.KD.ZG"
## [4,] "NY.GDP.PCAP.CD"
## [5,] "NY.GDP.PCAP.KD"
##      name
## [1,] "GDP per Capita, constant US$, millions"
## [2,] "Real GDP per Capita (real local currency units, va
## [3,] "Real agricultural GDP per capita growth rate (%)"
## [4,] "GDP per capita (current US$)"
## [5,] "GDP per capita (constant 2000 US$)"
```

Načítanie dát

```
data <- WDI(indicator='NY.GNP.PCAP.CD',  
            country=c('CA', 'US', 'FR', 'DE'),  
            start=1975)
```

- ▶ indicator pomocou WDIsearch
- ▶ country vo formáte iso2c
- ▶ defaultný start je 2005, ale mnohé dáta sú dostupné aj skôr (pozrieme si začiatok načítaných dát)

iso2c kódy sú napríklad tu:

https://github.com/vincentarelbundock/countrycode/blob/master/data/countrycode_data.csv

Načítanie dát

```
head(data)
```

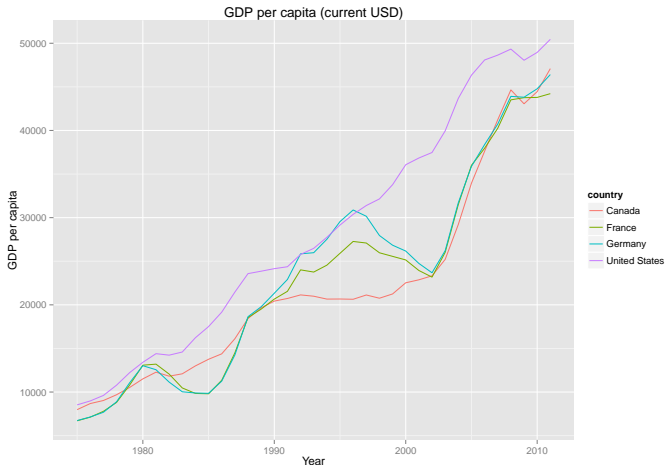
```
##   iso2c country NY.GNP.PCAP.CD year
## 1    CA  Canada      47090 2011
## 2    CA  Canada      44450 2010
## 3    CA  Canada      43060 2009
## 4    CA  Canada      44650 2008
## 5    CA  Canada      41230 2007
## 6    CA  Canada      37610 2006
```

Graf pomocou ggplot

```
ggplot(data, aes(year, NY.GNP.PCAP.CD, color=country))  
  + geom_line()  
  + xlab('Year') + ylab('GDP per capita')  
  + labs(title = 'GDP per capita (current USD)')
```

- ▶ celé je to jeden príkaz (v jednom riadku)
- ▶ data je *data frame*, v ktorom sú naše premenné
- ▶ year, NY.GNP.PCAP.CD - z dáta uložených v data bude na x-ovej osi premenná year a na y-ovej NY.GNP.PCAP.CD
- ▶ color=country - grafy budú odlišené farebne (preto color) a podľa premennej country

Graf pomocou ggplot



Cvičenie: Modelovanie logaritmu a rýchlosti rastu HDP

Dáta

```
# GDP per capita (constant 2000 US$)
data <- WDI(indicator='NY.GNP.PCAP.KD',
            country=c('US'),
            start=1965)

# zoradime rastuco podla rokov
data <- data[order(data$year),]

# nasa premenna na modelovanie
log.y <- log(data$NY.GNP.PCAP.KD)
# doplnime casovu strukturu
log.y <- ts(log.y, start=1965, frequency=1)
```

Zadanie

Z priebehu dát `plot(y)` vidíme, že dáta nie sú stacionárne, budeme pracovať s diferenciami (to je rýchlosť rastu HDP).

Nájdite AR model pre diferenciácie `log.y` a spravte predikcie `log.y` pre nasledujúce roky.

Cvičenie: Modelovanie dlhodobej nezamestnanosti

Dáta

```
WDIsearch('long.*term.*unemployment')
```

```
##      indicator
## [1,] "SL.UEM.LTRM.FE.ZS"
## [2,] "SL.UEM.LTRM.MA.ZS"
## [3,] "SL.UEM.LTRM.ZS"
##      name
## [1,] "Long-term unemployment, female (% of female unempl
## [2,] "Long-term unemployment, male (% of male unemployme
## [3,] "Long-term unemployment (% of total unemployment)"
```

Dáta

```
data <- WDI(indicator='SL.UEM.LTRM.ZS',  
            country=c('CA'),  
            start=1980)
```

```
data <- data[order(data$year),]  
y <- data$SL.UEM.LTRM.ZS  
y <- ts(y, start=1980, frequency=1)
```

Zadanie

Nájdite AR model pre y . Potvrdí model periodický charakter?

