

---

## CVIČENÍ 29.4.2013

---

Data `pGDP.csv` obsahují čtvrtletní časovou řadu procentuálních přírůstků GDP v USA od Q2 roku 1950.

1. Načtete si danou řadu a nastavte si ji do formátu `ts`. Znázorněte si graf a posuďte, zda lze řadu považovat za stacionární.
2. Na základě ACF a PACF se pokuste identifikovat vhodný model (resp. kandidáty na vhodné modely) pro tato data.
3. Proveďte odhad vybraných modelů pomocí funkce `arima`.
  - (a) U každého modelu posuďte statistickou významnost všech jeho parametrů (můžete si napsat v R funkci, která to bude za Vás provádět).
  - (b) Proveďte odhad i společného nadmodelu Vašich kandidátních modelů. Opět posuďte statistickou významnost.
  - (c) U všech kandidátních modelů se stručně podívejte na reziduální diagnostiku.
  - (d) Nakonec porovnejte modely pomocí AIC kritéria.
  - (e) Jaký je ve skutečnosti rozdíl mezi uvažovanými modely? Pomocí funkce `ARMAtoMA` si nechte vypsat koeficienty ze zápisu pomocí lineárního procesu a porovnejte je.
  - (f) Který model bychom zvolili jako lepší?
4. Zavolejte si balík `forecast`. Funkce `auto.arima` hledá optimální model (dle zvoleného kritéria) pro zadaná data. Vyzkoušejte, jaký model zvolí pro naše data:

```
auto.arima(y,max.P=0,max.Q=0)
# nebo take auto.arima(y,d=0,D=0,max.P=0,max.Q=0)
```

Volbu kritéria, na jehož základě se vybírá optimální model, lze nastavit volbou `ic='bic'` (nebo `'aic'`, `'aicc'`).

- (a) Posuďte statistickou významnost koeficientů takto nalezeného optimálního modelu.
  - (b) Proveďte pro model reziduální diagnostiku.
5. Vyberte jeden finální model. Pro něj proveďte:
    - (a) Znázorněte si graficky vyrovnání řady.
    - (b) Srovnajte v jednom obrázku výběrovou a teoretickou ACF a PACF.
    - (c) Zkonstruuje předpověď na čtyři období dopředu (funkce `predict`) a znázorněte je do grafu původní řady. (Lze využít funkci `ts.plot`).
    - (d) Případně můžeme využít grafy z balíku `forecast`:

```
p2=forecast.Arima(mod.final,h=4,level=0.95)
p2
plot(p2)
```
    - (e) Znázorněte si graficky vyrovnání původní (nediferencované) řady.
    - (f) Zapište výsledný odhadnutý model. Jaký je rozdíl mezi střední hodnotou řady a absolutním členem z modelu?

ARIMA MODELY Data USAmacroG.txt obsahují čtvrtletní makroekonomická data o USA z let 1950 až 2000.

1. Uvažujte časovou řadu investic (v mld. USD) (*realinv*).
  - (a) Na základě grafu zhodnoťte, zda lze řadu považovat za stacionární.
  - (b) Vytvořte řadu prvních diferencí a zhodnoťte její graf. Podobně si prohlédněte časovou řadu logaritmů původní řady a časovou řadu prvních diferencí. Kterou bude lepší modelovat?
  - (c) Prohlédněte si ACF a PACF analyzované řady (nejprve nediferencované, pak diferencované).
  - (d) Navrhněte vhodný model (resp. kandidáty na modely) pro diferencovanou řadu a odhadněte je. Vyberte z nich jeden „nejlepší“.
  - (e) Porovnejte odhad ARIMA modelu pro původní (nediferencovanou) řadu a příslušný ARMA model pro diferencovanou řadu. V čem je rozdíl? Co je správně?
  - (f) Použijte funkci *Arima* z balíku *forecast* pro odhad nediferencované řady. Pro zapojení driftu zvolte `include.drift=T`.
  - (g) Zapište výsledný model jakožto model pro nediferencovanou řadu.
  - (h) Proveďte predikci o 8 období dopředu a znázorněte ji do grafu.
2. Podobným způsobem jako v předchozím bodě modelujte míru nezaměstnanosti (*unemp*). Nalezněte vhodný model a proveďte predikci na roky 2001 a 2002.
3. Nalezněte vhodný model pro časovou řadu míry inflace (*infl*) a proveďte predikci na roky 2001 a 2002.