

CVIČENÍ 22.4.2013

AR MODELY ČASOVÝCH ŘAD

- Nasimulujte si časovou řadu o délce 200 pozorování, která se řídí AR(1) modelem s autoregresním parametrem $\varphi_1 = 0.7$ a s bílým šumem, který má směrodatnou odchylku 0.2.

```
x=arima.sim(n=n,list(ar=0.7),sd=0.2)
```

- Znázorněte si nasimulovanou řadu.
- Pro porovnání si nakreslete graf pro AR(1) modely s $\varphi_1 = 0.9$ a $\varphi_1 = -0.7$. Jaké vidíme rozdíly?
- Prozkoumejte, zda je řada x tvořena korelovanými pozorováními, pomocí metod, které zatím znáte. (Můžete využít funkci `lag.plot`).

- Necháme si vykreslit ACF a PACF nasimulované řady. Jaké grafy očekáváme?

```
acf(x)
pacf(x)
```

Chceme-li si nechat vypsat hodnoty a ne graf, pak stačí přidat volbu `plot=F`.

- Pomocí funkce `ARMAacf` si můžeme nechat spočítat teoretické ACF a PACF zadaného modelu a přidat si je do grafu

```
tacf=ARMAacf(ar=0.7,lag.max=20)
tpacf=ARMAacf(ar=0.7,lag.max=20,pacf=T)
acf(x)
lines(tacf,col="red")
pacf(x)
lines(tpacf,col="red")
```

- Čemu jsou rovny teoretické ACF? Uměli bychom je v tomto případě snadno spočítat i bez funkce `ARMAacf`?
 - Do jaké míry odpovídají výběrové ACF a PACF těm teoretickým? Zkuste zvýšit počet pozorování v řadě z 200 na 1000 a podívejte se, jak je to pro tento případ.
- Uvažujme opět řadu x výše uvedenou. Řekněme, že bychom na základě ACF a PACF identifikovali AR(1) model a chtěli ho odhadnout. V R lze využít několik různých funkcí (`ar`, `arima`, `arma`, `Arima (forecast)`, `auto.arima (forecast)`, `FitAR (FitAR)`, ...).

Funkce `ar()` vybírá optimálně řád autoregresního modelu na základě AIC kritéria. Lze volit různé metody odhadu.

```
ar(x)
ar(x,method="mle")
```

Obecnější je funkce `arima()`, která umožňuje odhadovat ARMA, ARIMA, SARIMA atd. a také regresní model s chyby řídicími se ARMA modelem. Řád modelu je nutné specifikovat

```
arima(x,order=c(1,0,0))
```

- Porovnejte odhady parametrů se zadanými parametry použitými v simulaci.
- Jaká je statistická významnost jednotlivých parametrů?
- Odhad modelu bez absolutního členu lze provést pomocí volby `include.mean=F`. Porovnejte tyto dva modely pomocí AIC. Srovnajte odhadnuté parametry s výsledkem funkce `ar`.
- Zkuste zvýšit řád autoregresního modelu na $AR(2)$. Posuďte statistickou významnost φ_2 .
- Výsledný model si nějak nazvěte, např. `mod1`.

5. Uvažujte rezidua z vybraného modelu. `mod1`.

- (a) Vykreslete si jejich graf a posuďte předpoklad homoskedasticity a nulovosti středních hodnot.
- (b) Zhodnoťte graficky (auto)korelovanost reziduí. Využijte, co znáte (graf, kumulativní periodogram, ACF, PACF).
- (c) Otestujeme nulovost autokorelace reziduí pomocí Ljungova-Boxova testu.

```
Box.test(r,lag=20,type="Ljung-Box")
```

 - Jaké by mělo být rozdělení testové statistiky za nulové hypotézy?
 - Ověřte, zda je p-hodnota testu spočtena správně dle toho, co znáte z přednášky.
- (d) Pro diagnostiku modelu lze také použít následující graf:

```
tsdiag(mod1)
```
- (e) Nakonec posuďte, zda jsou rezidua normálně rozdělená.

MA A ARMA MODELY

1. Nasimulujte si časovou řadu o délce 200 pozorování, která se řídí $MA(1)$ modelem s parametrem $\theta_1 = 0.4$ a s bílým šumem, který má směrodatnou odchylku 0.2.
 - (a) Znázorněte si graf nasimulované řady.
Podívejte se také na graf oproti zpožděným hodnotám a kumulativní periodogram.
 - (b) Znázorněte si graf ACF a PACF. Identifikovali byste $MA(1)$ model?
 - (c) Přidejte si do grafu teoretické ACF a PACF a podívejte se, jak přesný je odhad pro náš případ. Zvyšte počet pozorování a porovnejte.
 - (d) Ověřte pomocí Ljungova-Boxova testu, zda lze řadu považovat za bílý šum.
2. Nasimulujte si časovou řadu o délce 200 pozorování, která se řídí $ARMA(1,1)$ modelem s parametry $\varphi_1 = 0.6$ a $\theta_1 = 0.4$ a s bílým šumem, který má směrodatnou odchylku 0.2. Proveďte tytéž kroky jako výše s cílem
 - ověřit různými metodami, že je řada tvořena závislými pozorováními,
 - prozkoumat vlastnosti ACF a PACF pro $ARMA(p, q)$ modely a výpovědní hodnotu výběrových ACF a PACF.

Nakonec vhodný ARMA model odhadněte a zapiště.
3. Uvažujte řadu procentuálních čtvrtletních růstů GDP v USA od roku 1950 z `pGDP.csv`.
 - (a) Identifikujte na základě ACF a PACF vhodný model (modely) pro tato data.
 - (b) Proveďte odhad vybraných modelů a vyberte z nich jeden vhodný.
 - (c) Pro finální model proveďte reziduální diagnostiku.
 - (d) Zapiště odhadnutý model.