Cvičení 18.3.2013

Data. Data kurzy.csv obsahují časovou řadu průměrných měsíčních kurzů eura vůči koruně od ledna 1999 do února 2013.

1. Načtěte si data kurzy.csv.Nejprve sdělíme R, že máme řadu měsíčních pozorování od ledna 1999 a necháme si nakreslit obrázek.

```
attach(kurzy)
EUR=ts(EUR,start=1999,freq=12)

plot(EUR)
```

Formát ts slouží pro pravidelné časové řady. Na objekt typu ts lze použít funkce jako např.

```
tsp(EUR)
time(EUR)
window(EUR,start=c(2012,1),end=c(2012,12))
plot(window(EUR,start=c(2008,1)))
```

Klouzavé průměry. Předpokládáme, že máme časovou řadu $\{y_t,\ t = 1, \dots .n\}$, kde $y_t = T_t + e_t$. Klouzavý průměr délky $2m + 1$ má obecný tvar

$$\widehat{T}_t = \sum_{j=-m}^{m} w_j y_{t+j},$$

kde $\{w_j\}$ jsou váhy takové, že $\sum_{j=-m}^{m} w_j = 1$. Jestliže $w_j = w_{-j}$, mluvíme o symetrickém klouzavém průměru.

2. Vyrovnejte časovou řadu pomocí jednoduchých klouzavých průměrů, tj. uvažujte $w_j = \frac{1}{2m+1}$.
   (a) Uvažujte různé hodnoty m a sledujte míru vyhlazení řady.
   Např. jednoduchý klouzavý průměr délky 3 spočítáme a zobrazíme následovně:
   ```
   f1=filter(EUR,filter=rep(1/3,3))
   lines(f1,col="red")
   ```
   Podobně proveďte pro další volby m (větší), dokud nedosáhnete nejlepšího vyrovnání.
   (b) Na základě vybraného modelu proveďte predikci o jedno období dopředu (použijte koeficienty známé z přednášky nebo knihy prof. Cipry).

3. Nyní uvažujte klouzavé průměry délky m, které zachovají polynom stupně r (viz přednáška).
   (a) Připomeňte si, o jaké klouzavé průměry se jedná.
   (b) Proveďte proložení např. pro $2m + 1 = 13$ a $r = 3$.
   (c) Zvolte optimálně parametr r dle kritéria V_k z přednášky, kde

$$V_k = \frac{\sum_{t=k+1}^{n} \left(\Delta^k y_t\right)^2}{\binom{2k}{k}(n-k)}.$$

   V R diference k-tého řádu získáte pomocí diff(EUR,k), kombinační číslo jako choose(2*k,k).

(d) Pro takto zvolené r najděte „ideální“ m a znázorněte si proložení.
(e) Zkonstruujte predikce na jedno období dopředu a srovnejte ji s predikcí získanou pomocí jednoduchých klouzavých průměrů.

4. Vyrovnejte řadu pomocí centrovaných klouzavých průměrů.

5. Dále lze uvažovat další vyhlazovací techniky

(a) Klouzavé mediány: Např. přes 5 měsíců získáme pomocí
```
library(zoo)
med=rollapply(EUR,5,median)
lines(med,col="red")
```
Podobně pro jiné délky období.
(b) Lokální vyhlazení pomocí polynomů (loess, lowess . . .).
(c) Vyhlazení pomocí splinů (smooth.spline).
(d) Hodrick-Prescott filter (např. hpfilter z balíku mFilter).

Data: Uvažujme roční úhrny srážek (v inch) v Londýně od roku 1813 do roku 1912 obsažené v souboru srazkyLondyn.txt. Načtěte si data a uvedenou časovou řadu nastavte do formátu ts.

Jednoduché exponenciální vyrovnávání používáme pro řady $\{y_t\}$, které lze lokálně považovat za konstantní. Pak

$$\widehat{y_t} = \alpha \sum_{j=0}^{\infty} (1-\alpha)^j y_{t-j},$$

kde $0 < \alpha < 1$ je vyrovnávací konstanta. V praxi se používá rekurentní vztah

$$\widehat{y_t} = \alpha y_t + (1-\alpha)\widehat{y_{t-1}}.$$

(a) Vyrovnejte zkoumanou řadu pomocí jednoduchého exponenciálního vyrovnávání pro $\alpha = 0.1$.
```
e1=HoltWinters(srazky,alpha=0.1,beta=FALSE,gamma=FALSE)
e1
e1$fitted
plot(e1)
```
Vyzkoušejte i jiné hodnoty pro α a sledujte, jak se mění míra vyhlazení.
(b) Nyní necháme parametr α nalézt „optimálně“ pomocí minimalizace SSE.
```
e2=HoltWinters(srazky,beta=FALSE,gamma=FALSE)
e2
plot(e2)
```
Jaká je hodnota optimálního parametru? Jak se liší SSE pro $\alpha = 0.1$ a optimální α?
(c) Spočtěte predikci na jedno období dopředu
```
predict(e2,ahead=1)
```
Vzpomeňte si, jak se predikce konstruuje (viz přednáška). Znázornění predikce v grafu:
```
points(predict(e2,ahead=1),col="red",pch=19)
```
(d) Jiné zajímavé grafické znázornění predikce (s využitím balíků forecast) a intervalová předpověď:

```
#install.packages("forecast")
# je-li potreba doinstalovat
library(forecast)
plot(forecast.HoltWinters(e2, h=1))
```

Oranžová oblast odpovídá 80% předpovědnímu intervalu a žlutá oblast 90% intervalu. Jak se změní obrázek, když budeme predikovat na 5 období dopředu?

HOLTOVA METODA se používá pro řady, u kterých lze trend považovat lokálně za lineární. Metoda používá dvě vyrovnávací konstanty α a β pro vyrovnání úrovně L_t a směrnice T_t:

$$L_t = \alpha y_t + (1-\alpha)(L_{t-1} + T_{t-1}),$$
$$T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1}.$$

Pak $\widehat{y}_t = L_t$ a $\widehat{y}_{t+j}(t) = L_t + T_t \cdot j$.

(a) Uvažujte opět časovou řady eurových kurzů a vyrovnejte ji pomocí Holtovy metody.

```
h1=HoltWinters(EUR,gamma=FALSE)
h1
plot(h1)
```

Jaká je hodnota optimálních parametrů? O čem to vypovídá?

(b) Použijte Holtovu metodu s jinou volbou α a β a sledujte, jak se mění vyhlazení řady.

(c) Proveďte predikci na jedno období dopředu a nechte si vykreslit graf.

(d) Jak se změní graf intervalové predikce, když budeme predikovat např. na pět období dopředu?

DVOJITÉ EXPONENCIÁLNÍ VYROVNÁVÁNÍ je speciálním případem Holtovy metody pro

$$\alpha = a(2-a), \quad \beta = \frac{a}{2-a},$$

kde a je vyhlazovací parametr dvojitého exponenciálního vyrovnání.

(a) Použijte dvojité exponenciální vyrovnávání na řadu eurových kurzů. Zvolte vyhlazovací parametr roven 0.1 a 0.2.

(b) Optimální vyhlazovací parametr lze nalézt následovně:

```
dvoj.exp=function(y,amin=0,amax=0.3){
  #y je zadana rada
  SEEdvoj.exp=function(a){
    HoltWinters(y,alpha=a*(2-a),beta=a/(2-a),gamma=FALSE)$SSE
  }
  minSSE=optimize(SEEdvoj.exp,c(amin,amax))
 # a je optimalni parametr z intervalu (amin, amax)
  a=minSSE$minimum
  print(a)
  dvoj.exp=HoltWinters(y,alpha=a*(2-a),beta=a/(2-a),gamma=FALSE)
}

d1=dvoj.exp(EUR)
d1
plot(d1)
```