Cvičení 25.3.2013

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

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

1. Načtěte si data kurzy.csv a stejným způsobem jako minule sdělte R, že máme řadu měsíčních pozorování od ledna 1999 a nechte si nakreslit obrázek.

2. Vyrovnejte řadu pomocí Holtovy metody.

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

   Jaká je hodnota optimálních parametrů? O čem to vypovídá? Dostali jsme očekávané vyhlazení?

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

4. Proveďte predikci na jedno období dopředu a nechte si vykreslit graf.

```
predict(h1,ahead=1)
points(predict(h1,ahead=1),col="red",pch=19)

install.packages("forecast")
library(forecast)
plot(forecast.HoltWinters(e2, h=1))
```

5. 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 vyrovnávání.

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

2. 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)
```

ČASOVÉ ŘADY SE SEZÓNNOSTÍ Data beerProduction.csv obsahují měsíční objem výroby piva (v megalitrech) v Austrálii v období leden 1956 až srpen 1995 (zahrnuje pouze piva s obsahem alkoholu více než 1.15 %).

1. Načtěte si data. V dalším budeme uvažovat časovou řadu logaritmů měsíční produkce. Vytvořte si ji a převeďte si ji do formátu ts, nazvěte ji např. beer.

2. Nechte si vykreslit obrázek celé řady a pro představu o sezónnosti i jeden vybraný rok.

3. Vyhlaďte řadu pomocí centrovaného klouzavého průměru.
   Pro srovnání proveďte také vyhlazení jednoduchým klouzavým průměrem kratší délky a podívejte se, že takto bychom nedostali požadované vyhlazení.

4. Chceme si rozložit řadu na trendovou, sezónní a náhodnou složku. Na základě grafu rozhodněte, zda je zde vhodný aditivní nebo multiplikativní model.
   Provedeme rozložení pomocí klouzavého průměru:

   ```
   d1=decompose(beer)
   plot(d1)
   ```

   Jednotlivé složky si můžeme vykreslit nebo s nimi pracovat pomocí

   ```
   plot(d1$trend)

   plot(d1$seasonal[1:12],type="h")
   points(d1$seasonal[1:12])
   abline(h=0)
   ```

   apod. Jak bychom popsali odhadnutou sezónnost?

   Poznámka: Sofistikovanější rozklad pomocí loess metody (lokální vážená regrese) lze dosáhnout pomocí stl(beer,s.window=13). Metoda iterativně hledá sezónní a trendovou složku pomocí loess vyhlazení pozorování v pohybujícím se okénku zadané délky.

5. Vytvořte sezónně očištěná data a znázorněte si jejich graf

```
beer.bezsez=beer -d1$seasonal
plot(beer.bezsez)
```

Modelování sezónnosti pomocí kvalitativní proměnné

1. Na základě výsledků dekompozice navrhněte parametrický model pro data:
   - Jak budeme modelovat trend?
   - Připomeňte si, jak lze modelovat sezónnost.
2. Nejprve odhadněte model bez sezónní složky a znázorněte si proložení (pro nalezení vhodného trendu). Do regresního modelu použijte jako čas proměnnou

```
t=c(time(beer))
```

3. Poté přidejte do modelu sezónní složku jako kvalitativní proměnnou udávající sezónní období (měsíc). Zavedeme si proto proměnnou označující měsíc a rovnou ji uložíme ve formátu factor, aby R vědělo, že s ní má pracovat jako s kvalitativní.

```
mesic=factor(c(rep(1:12,39),1:8))
m2=lm(beer~t+I(t^2)+mesic)
summary(m2)
```

   Jaká je interpretace parametrů takto odhadnutého modelu?
   Připomenutí ze statistiky (nebo regrese atd.): Statistickou významnost sezónnosti bychom zjistili pomocí anova(m2).

4. Odhadneme model trochu jinak:

```
m2.2=lm(beer~t+I(t^2)+mesic-1)
summary(m2.2)
```

   Jaká je nyní interpretace koeficientů?
   Znázornění odhadnutých absolutních členů:

```
plot(coef(m2.2)[3:14],type="h")
```

5. V časových řadách je ale výhodné pracovat s modelem, kde koeficienty sezónnosti udávají rozdíl proti celkovému průměru (tj. sečtou na nulu a model obsahuje absolutní člen) — viz aditivní normalizační pravidlo z přednášky. Takto přepočtené koeficienty získáme a znázorníme následovně:

```
sez1=coef(m2.2)[3:14]-mean(coef(m2.2)[3:14])
plot(sez1, type="h")
points(sez1,pch=19)
abline(h=0)

#nebo
barplot(sez1,names=1:12)
```

Poznámka: Téhož bychom dosáhli pomocí

```
m3=lm(beer~t+I(t^2)+mesic,contrasts=list(mesic=contr.sum))
b=c(coef(m3)[4:14],-sum(coef(m3)[4:14]))
barplot(b,names=1:12)
```