M5201 – 9. CVIČENÍ: *Exponenciální vyrovnávání*

1 Exponenciální vyrovnávání

Hlavní myšlenka **lokální (vážené) metody nejmenších čtverců** spočívá v tom, že provedeme odhad trendu polynomem na **lokálním intervalu** $(t-s, t+s)$ na rozdíl od klasické (vážené) metody nejmenších čtverců, kdy trend odhadujeme polynomem na celém intervalu možných hodnot parametru t.

Exponenciální vyrovnávání aproximuje neznámý trend polynomem stupně m, který se odhaduje na základě pozorování z asymetrického vyhlazovacího okénka směrem do minulosti. Exponenciální vyrovnávání vychází z polynomiální lokální **vážené** metody nejmenších čtverců, kde váhy jednotlivých čtverců se směrem do minulosti exponenciálně snižují – odtud název metody.

Uvažujeme-li vyhlazovací okno pouze směrem do minulosti, pak pro každé $t, \tau = 0, 1, \ldots$ dostaneme regresní model tvaru

$$\boxed{Y_{t-\tau} = \sum_{j=0}^{m} (-\tau)^j \beta_j(t) + \varepsilon_{t-\tau}}, \; E\varepsilon_{t-\tau} = 0; \; E\varepsilon_q\varepsilon_s = 0, \; q \neq s; \; D\varepsilon_{t-\tau} = \alpha^{-\tau}\sigma^2; \; \alpha \in (0,1).$$

tj. matice vah je rovna

$$\mathbf{W} = diag\{w_1, \ldots, w_n, \ldots\} = diag\{\alpha^0, \alpha^1, \ldots, \alpha^\tau \ldots\}.$$

Odhad parametrů $\boldsymbol{\beta}$ provedeme metodou **nejmenších vážených čtverců** (nebot' rozptyly nejsou konstantní) která je dána vzorcem:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}\mathbf{Y}$$

kde

$$\mathbf{X}'\mathbf{W}\mathbf{X} = \begin{pmatrix} \sum_{\tau=0}^{\infty} \alpha^\tau & \sum_{\tau=0}^{\infty} (-\tau)^1\alpha^\tau & \cdots & \sum_{\tau=0}^{\infty} (-\tau)^m\alpha^\tau \\ \sum_{\tau=0}^{\infty} (-\tau)^1\alpha^\tau & \sum_{\tau=0}^{\infty} (-\tau)^2\alpha^\tau & \cdots & \sum_{\tau=0}^{\infty} (-\tau)^{m+1}\alpha^\tau \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{\tau=0}^{\infty} (-\tau)^m\alpha^\tau & \sum_{\tau=0}^{\infty} (-\tau)^{m+1}\alpha^\tau & \cdots & \sum_{\tau=0}^{\infty} (-\tau)^{2m}\alpha^\tau \end{pmatrix}, \; \mathbf{X}'\mathbf{W}\mathbf{Y} = \begin{pmatrix} \sum_{\tau=0}^{\infty} \alpha^\tau Y_{t-\tau} \\ \sum_{\tau=0}^{\infty} (-\tau)^1\alpha^\tau Y_{t-\tau} \\ \vdots \\ \sum_{\tau=0}^{\infty} (-\tau)^m\alpha^\tau Y_{t-\tau} \end{pmatrix}.$$

Tento přístup založený na vážené polynomiální metodě nejmenších čtverců se nazývá *Brownův přístup*.

<u>ZNAČENÍ</u>:

Pro dobrou srozumitelnost zavedeme následující značení. Nechť $\{Y_t, t \in \mathbb{Z}\}$ je náhodná posloupnost, její realizace v časových okamžicích $t_1, t_2, \ldots, t_n$ označme $y_1, y_2, ..., y_n$. Symbolem $\boxed{\hat{y}_{t|k}}$ označme odhad hodnoty Y_t v čase t na základě hodnot do časového okamžiku k včetně.

- Jestliže $k < t$, pak $\boxed{\hat{y}_{t|k}}$ nazýváme **predikcí**,

- pokud $k = t$, $\boxed{\hat{y}_{t|t}}$ nazýváme **filtrací**
- a je-li $k = n > t$, pak $\boxed{\hat{y}_{t|n}}$ nazýváme **vyrovnáním** (*smoothing*).

2 Jednoduché exponenciální vyrovnávání

Exponenciální vyrovnávání pro $m = 0$ se nazývá jednoduché exponenciální vyrovnávání. Použijeme–li označení $\hat{\beta}_0(t) = b_0(t)$ a uvážíme–li, že pro $\alpha \in (0,1)$ je $\sum\limits_{\tau=0}^{\infty} \alpha^\tau = \frac{1}{1-\alpha}$, dostaneme

$$b_o(t) \sum_{\tau=0}^{\infty} \alpha^\tau = \sum_{\tau=0}^{\infty} \alpha^\tau Y_{t-\tau} \quad \Rightarrow \quad \boxed{b_0(t) = \hat{Y}_t = (1-\alpha) \sum_{\tau=0}^{\infty} \alpha^\tau Y_{t-\tau}}$$

Abychom získali rekurentní vztah, upravujme

$$\begin{aligned} \hat{Y}_t &= (1-\alpha) \textstyle\sum_{\tau=0}^{\infty} \alpha^\tau Y_{t-\tau} = (1-\alpha) Y_t + (1-\alpha) \sum_{\tau=1}^{\infty} \alpha^\tau Y_{t-\tau} = \left| \begin{array}{c} \scriptstyle subst. \\ \scriptstyle k = \tau - 1 \end{array} \right| \\ &= (1-\alpha) Y_t + (1-\alpha) \textstyle\sum_{k=0}^{\infty} \alpha^{k+1} Y_{t-1-k} \\ &= (1-\alpha) Y_t + \alpha \underbrace{(1-\alpha) \sum_{k=0}^{\infty} \alpha^k Y_{t-1-k}}_{\hat{Y}_{t-1}} = (1-\alpha) Y_t + \alpha \hat{Y}_{t-1} \end{aligned}$$

Protože predikce o τ ($\tau > 0$) kroků dopředu pro jednoduché exponenciální vyrovnávání je rovna $\hat{Y}_{t+\tau|t} = \hat{Y}_t = b_0(t)$, můžeme předchozí rekurentní vztah přepsat pro realizace a dále upravovat

$$\begin{aligned} \hat{y}_{t+1|t} &= (1-\alpha) y_t + \alpha \hat{y}_{t|t-1} \\ &= (1-\alpha) y_t + \alpha \hat{y}_{t|t-1} + \hat{y}_{t|t-1} - \hat{y}_{t|t-1} \\ &= \hat{y}_{t|t-1} + (1-\alpha) \underbrace{(y_t - \hat{y}_{t|t-1})}_{\text{chyba predikce } \hat{\varepsilon}_{t|t-1}} \end{aligned}$$

a o rekurentním vzorci s chybou predikce $\hat{\varepsilon}_{t|t-1}$ se říká, že je ve formě korekce chyby předpovědi (*error correction form*).

3 Ad hoc přístupy Holta a Winterse

Pokud chceme na základě pozorování $y_1, \ldots, y_t$ sestrojit předpověď budoucí hodnoty y_{t+1} v čase $t+1$, označme ji $\widehat{y}_{t+1|t}$, pak nejjednodušším odhadem může být obyčejmý průměr. Tato předpověď je vhodná, pokud hodnoty časové řady náhodně kolísají kolem střední hodnoty, která se v čase nemění. Jako rozumější se však jeví použít pro predikci budoucí hodnoty ve větší míře pozorování, která jsou časově nejblíže. Pak se nabízejí vážené průměry

$$\widehat{y}_{t+1|t} = \sum_{j=0}^{t-1} w_{j,t} y_{t-j}, \tag{1}$$

kde součet vah je roven jedné, tj. $\sum_{j=0}^{n} w_{j,t} = 1$.

Exponenciální vyrovnávání je založeno na myšlence použití vah, které do minulosti klesají exponenciálně.

S využitím vztahu

$$\sum_{j=0}^{t-1} \alpha^j = \frac{1-\alpha^t}{1-\alpha}, \quad \text{pro } \alpha \in (0,1)\,, \tag{2}$$

chceme-li, aby součet vah, které exponenciálně klesají, byl roven jedné, položíme

$$w_{j,t} = \frac{1-\alpha}{1-\alpha^t}\alpha^j \;. \tag{3}$$

Protože pro $t \to \infty$ konvergují váhy $w_{j,t} \to w_j = (1-\alpha)\alpha^j$, můžeme uvažovat jednokrokovou předpověď ze všech minulých pozorování ve tvaru

$$\widehat{y}_{t+1|t} = (1-\alpha)\sum_{j=0}^{\infty} \alpha^j y_{t-j}, \quad \text{pro } \alpha \in (0,1) \;\;. \tag{4}$$

Analogicky jako u Brownova přístupu odvodíme rekurentní vztahy

$$\begin{aligned}
\widehat{y}_{t+1|t} &= (1-\alpha)y_t + (1-\alpha)\sum_{j=1}^{\infty} \alpha^j y_{t-j} \\
&= (1-\alpha)y_t + \alpha(1-\alpha)\sum_{k=0}^{\infty} \alpha^k y_{t-1-k} \\
&= (1-\alpha)y_t + \alpha\widehat{y}_{t|t-1}
\end{aligned}$$

Obdobně získáme i tvar využívající korekci chyby předpovědi

$$\begin{aligned}
\widehat{y}_{t+1|t} &= (1-\alpha)y_t + \alpha\widehat{y}_{t|t-1} + \widehat{y}_{t|t-1} - \widehat{y}_{t|t-1} \\
&= \widehat{y}_{t|t-1} + (1-\alpha)(y_t - \widehat{y}_{t|t-1}) \\
&= \widehat{y}_{t|t-1} + (1-\alpha)\widehat{\varepsilon}_{t|t-1}
\end{aligned}$$

Na tomto ad-hoc přístupu se nám podařilo ukázat, že se v podstatě jedná o **jednoduché exponenciální vyrovnávání**, které přepokládá model

$$Y_t = \beta_0(t) + \varepsilon_t$$

s lokální hladinou $\beta_0(t)$.

Použijeme-li značení obvyklá pro tento přístup, kdy váhy mají tvar

$$w_j = \beta(1-\beta)^j, \tag{5}$$

tj. $\alpha = 1-\beta$, místo $\widehat{\beta_0(t)}$, píšeme L_t (level). Odvozené vztahy v novém značení:

$$\widehat{y}_{t+1|t} = \beta y_t + (1-\beta)\widehat{y}_{t|t-1} = \widehat{y}_{t|t-1} + \beta\widehat{\varepsilon}_{t|t-1} \tag{6}$$

$$L_{t+1} = L_t + \beta\widehat{\varepsilon}_{t+1|t} \tag{7}$$

4 Holtovo exponenciální vyrovnávání

Oproti jednoduchému exponenciálnímu vyrovnávání Holtova metoda předpokládá lokálně lineární trend, jehož koeficienty $\beta_0(t)$ i $\beta_1(t)$ se v čase mění. Hodnota časové řady v okamžiku t je určena jednak její úrovní $\beta_0(t)$, jednak směrnicí $\beta_1(t)$. V Holtově metodě se úroveň v čase t značí symbolem L_t (zkratka pro *level*) a směrnice jako T_t (zkratka pro *trend*).

Úroveň L_t je zároveň vyrovnanou hodnotou realizace y_t v okamžiku t. Směrnice lokálně lineárního trendu T_t (někdy mluvíme krátce o trendu) vyjadřuje očekávanou změnu úrovně časové řady při jednotkové časové změně. Pokud chceme pomocí Holtovy metody přepovídat hodnotu časové řady o $h > 0$ jednotek dopředu, položíme

$$\widehat{y}_{t+h|t} = L_t + T_t h \ . \tag{8}$$

Takže, je-li $h = 1$, dostaneme jednokrokovou předpověď jako

$$\widehat{y}_{t+1|t} = L_t + T_t \ . \tag{9}$$

Protože by přibližně mělo platit, že realizace $y_{t+1} \approx L_{t+1}$, pak se jeví vhodné získat L_{t+1}, jako konvexní lineární kombinaci hodnot $(L_t + T_t)$ a y_{t+1}. V Holtově metodologii bývá zvykem místo $\alpha \in (0, 1)$ používat $\beta = 1 - \alpha$, takže konvexní lineární kombinace bude mít tvar

$$L_{t+1} = (1 - \beta)(L_t + T_t) + \beta y_{t+1} \ . \tag{10}$$

Hodnota β se nazývá *vyrovnávací konstanta pro úroveň řady.*

Analogickou úvahu použijeme i pro směrnici trendu T_t. Z přepokladu, že řada má lokálně lineární trend vyplývá, že by přibližně mělo platit

$$T_{t+1} \approx T_t,$$

ale zároveň má také smysl očekávat, že směrnice trendu je přibližně rozdílem sousedních úrovní, tj.

$$T_{t+1} \approx L_{t+1} - L_t \ .$$

Novou hodnotu směrnice T_{t+1} budeme uvažovat jako konvexní lineární kombinaci

$$T_{t+1} = (1 - \gamma)T_t + \gamma(L_{t+1} - L_t), \quad \text{kde } \gamma \in (0, 1) \tag{11}$$

γ je tzv. *vyrovnávací konstanta pro lineární růst (pro směrnici).*

Na závěr odstavce ještě ukážeme přepsání předchozích rekurentních vztahů do chybového tvaru.

$$\begin{aligned}
L_{t+1} &= (1-\beta)(L_t+T_t)+\beta y_{t+1} = (1-\beta)(L_t+T_t)+\beta y_{t+1}+\beta\widehat{y}_{t+1|t}-\beta\widehat{y}_{t+1|t}\\
&= \beta\underbrace{(y_{t+1}-\widehat{y}_{t+1|t})}_{\widehat{\varepsilon}_{t+1|t}}+(1-\beta)(L_t+T_t)+\beta\underbrace{\widehat{y}_{t+1|t}}_{L_t+T_t}\\
&= \beta\widehat{\varepsilon}_{t+1|t}+L_t+T_t
\end{aligned}$$

$$\begin{aligned}
T_{t+1} &= (1-\gamma)T_t+\gamma(L_{t+1}-L_t) = T_t-\gamma T_t+\gamma\underbrace{L_{t+1}}_{L_t+T_t+\beta\widehat{\varepsilon}_{t+1|t}}-\gamma L_t\\
&= T_t-\gamma T_t+\gamma(L_t+T_t+\beta\widehat{\varepsilon}_{t+1|t})-\gamma L_t\\
&= T_t+\gamma\beta\widehat{\varepsilon}_{t+1|t}\ .
\end{aligned}$$

5 Holtovo-Wintersovo exponenciální vyrovnávání

V případě, kdy časová řada má sezonní charakter, nevystačíme se žádnou z předchozích metod. Rozšíření Holtovy metody na sezónní časové řady je známo jako **Holtova–Wintersova metoda**. Autorem je Holtův student Peter R. Winters.

Holtova-Wintersova metoda je založena na třech vyrovnávacích konstantách. Jedna je pro hladinu, druhá pro trend a třetí pro sezónnost. Dle charakteru dat využívá aditivní nebo multiplikativní notaci.

Uvažujme časovou řadu s lokálně lineárním trendem a sezónností s periodou $p \geq 2$. Stejně jako u Holtovy metody označme symbolem L_t úroveň v čase t, symbolem T_t směrnici lokálně lineárního trendu a symbolem S_t sezónní výkyv čase t. Součet úrovně L_t s hodnotou sezónního výkyvu S_t představuje v okamžiku t vyrovnanou hodnotu realizace y_t. Předpověď hodnoty časové řady o $h > 0$ jednotek dopředu je pak dána vztahem

$$\widehat{y}_{t+h|t} = L_t + S_{t-p+h} + T_t h, \tag{12}$$

takže v případě jednokrokové predikce platí

$$\widehat{y}_{t+1|t} = L_t + S_{t+1-p} + T_t \tag{13}$$

Protože by mělo přibližně platit

$$y_{t+1} \approx L_{t+1} + S_{t+1-p}$$

a

$$L_{t+1} \approx L_t + T_t,$$

má smysl získat úroveň L_{t+1} jako konvexní lineární kombinaci hodnot $(L_t + S_t)$ a $(y_{t+1} - S_{t+1-p})$, tj.

$$L_{t+1} = (1-\beta)(L_t+T_t) + \beta(y_{t+1} - S_{t+1-p}). \tag{14}$$

Protože řada má lokálně lineární trend, mělo by přibližně platit

$$T_{t+1} \approx T_t,$$

ale zároveň lze směrnici lokálně lineárního trendu vyjádřit pomocí rozdílu sousedních hladin

$$T_{t+1} \approx L_{t+1} - L_t.$$

Oba předchozí vztahy využijeme při konstrukci směrnice lokálně linárního trendu díky konvexní linární kombinaci

$$T_{t+1} = (1-\gamma)T_t + \gamma(L_{t+1} - L_t),$$

kde $\gamma \in (0,1)$ se nazývá *vyrovnávací konstanta pro směrnici trendu*. Pro sezónní výkyvy musí platit vztah

$$S_{t+1} \approx S_{t+1-p}\ ,$$

a také

$$S_{t+1} \approx y_{t+1} - L_{t+1}$$

Tedy označíme-li symbolem $\delta \in (0,1)$ *vyrovnávací konstantu pro sezónní výkyvy*, pak

$$S_{t+1} = (1-\delta)S_{t+1-p} + \delta(y_{t+1} - L_{t+1})$$

Na závěr odstavce odvodíme rekuretní vztahy v chybové formě. Tedy upravujme

$$\begin{aligned}
L_{t+1} &= (1-\beta)(L_t + T_t) + \beta(y_{t+1} - S_{t+1-p}) \\
&= (1-\beta)(L_t + T_t) + \beta(y_{t+1} - S_{t+1-p}) + \beta\widehat{y}_{t+1|t} - \beta\widehat{y}_{t+1|t} \\
&= \beta(y_{t+1} - \widehat{y}_{t+1|t}) + L_t + T_t - \beta L_t - \beta T_t - \beta S_{t+1-p} + \beta \underbrace{\widehat{y}_{t+1|t}}_{L_t + S_{t+1-p} + T_t} \\
&= L_t + T_t + \beta\widehat{\varepsilon}_{t+1|t}
\end{aligned}$$

$$\begin{aligned}
T_{t+1} &= (1-\gamma)T_t + \gamma(L_{t+1} - L_t) = T_t - \gamma T_t + \gamma \underbrace{L_{t+1}}_{L_t + T_t + \beta\widehat{\varepsilon}_{t+1|t}} - L_t \\
&= T_t - \gamma T_t + \gamma(L_t + T_t + \beta\widehat{\varepsilon}_{t+1|t} - \gamma L_t) \\
&= T_t + \gamma\beta\widehat{\varepsilon}_{t+1|t}
\end{aligned}$$

$$\begin{aligned}
S_{t+1} &= (1-\delta)S_{t+1-p} + \delta(y_{t+1} - L_{t+1}) \\
&= S_{t+1-p} - \delta(S_{t+1-p} - y_{t+1}) - \delta(L_t + T_t + \beta\widehat{\varepsilon}_{t+1|t}) \\
&= S_{t+1-p} + \delta y_{t+1} - \delta(L_t + T_t + S_{t+1-p}) - \delta\beta\widehat{\varepsilon}_{t+1|t} \\
&= S_{t+1-p} + \delta(1-\beta)\widehat{\varepsilon}_{t+1|t}.
\end{aligned}$$

6 Exponenciální vyrovnávání v prostředí R

V prostředí R je pro exponenciální vyhlazování v balíčku `stats` k dispozici funkce `HoltWinters()`, která například v případě aditivního modelu uvažuje rekurentní vztahy

úroveň (*level*)	• $L_t = \alpha(y_t - S_{t-p}) + (1-\alpha)(L_{t-1} + T_{t-1})$
lineární růst (*growth*)	• $T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1}$
sezónní výkyvy (*seasonal*)	• $S_t = \gamma(y_t - L_{t-1} - T_{t-1}) + (1-\gamma)S_{t-p}$
předpověď (*forecast*)	• $\hat{y}_{t+h\|t} = L_t + T_t h + S_{t-p+h_p^+}$

kde $h_p^+ = [(h-1) \mod p] + 1$.

Počáteční stavy $L_0, T_0, S_{1-p}, \ldots, S_0$ a tzv. vyrovnávací konstanty α, β, γ jsou odhadnuty na základě dat. Na podrobný popis funkce se podívejte pomocí příkazu `?HoltWinters`.

Mnohem komplexnější možnosti nabízí funkce `ets()` z balíčku `forecast`, která navíc dovoluje tzv. tlumící faktor.

Tak například, označíme–li

$$\mu_t = \hat{y}_t,$$

pak model s **tlumeným lineárním trendem**, kdy

$$\mu_t = \hat{y}_t = L_{t-1} + T_{t-1}$$

a který je ve formě korekce chyby predikce (*error correction form*), bude definován rekuretními vztahy

$$\begin{aligned} y_t &= L_{t-1} + \phi T_{t-1} + \varepsilon_t \\ L_t &= L_{t-1} + \phi T_{t-1} + \alpha\varepsilon_t \\ T_t &= \phi T_{t-1} + \beta(L_t - L_{t-1} - \phi T_{t-1}) = \phi T_{t-1} + \alpha\beta\varepsilon_t \end{aligned}$$

Optimální model je vybrán na základě tzv. AIC kritéria

$$AIC = -2\log(Likelihood) + 2m,$$

kde m je počet parametrů.

Při výstupu procedura `ets()` používá následující notaci

Trendová komponenta	Sezónní komponenta		
	N (None)	A (Additive)	M (Multiplicative)
N (None)	N, N	N, A	N, M
A (Additive)	A, N	A, A	A, M
A_d (Additive damped)	A_d, N	A_d, A	A_d, M
M (Multiplicative)	M, N	M, A	M, M
M_d (Multiplicative damped)	M_d, N	M_d, A	M_d, M

Označení optimálního modelu tvoří trojici $ETS(E, T, S)$, kde

E	*error*	možné hodnoty	A, M
T	*trend*		N, A, A_d, M, M_d
S	*seasonal*		N, A, M

PŘÍKLAD 1

Metodu exponenciálního vyrovnávání aplikujeme na časovou řadu v souboru `wool.txt`. Jedná se o čtvrtletní údaje o množství vlny (v tunách) vyprodukované v Austrálii v letech 1965-1994. Datový soubor obsahuje pouze jeden sloupec s daty, proto bude nejjednodušší k načtení použít funkci `scan`:

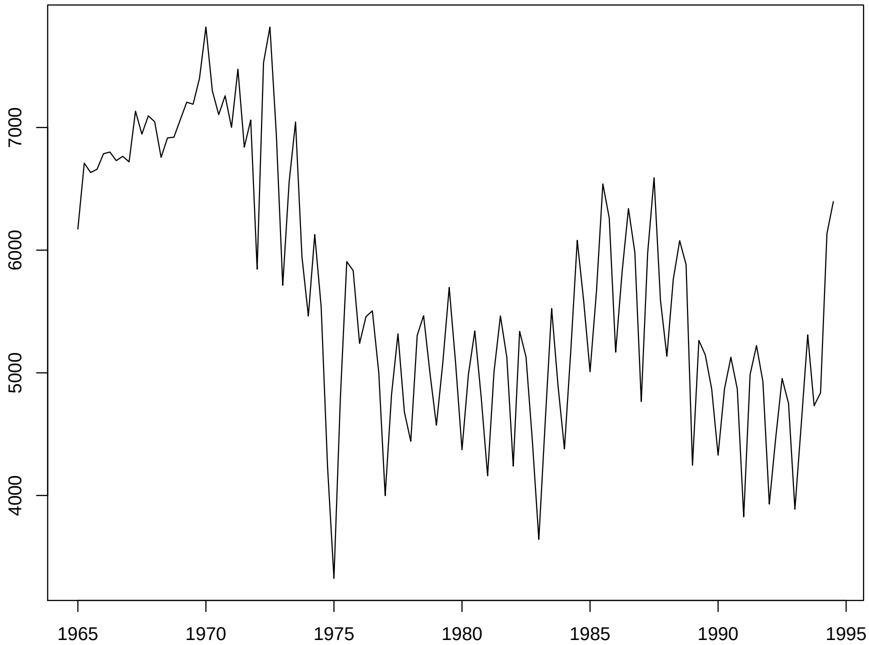
```
> fileDat <- paste(data.library, "wool.txt", sep = "")
> wool <- scan(fileDat)
```

Z načtených dat vytvoříme časovou řadu a vykreslíme ji.

```
> woolTS <- ts(wool, start = 1965, frequency = 4)
> str(woolTS)
```

```
 Time-Series [1:119] from 1965 to 1994: 6172 6709 6633 6660 6786 ...
```

```
> par(mar = c(2, 2, 0, 0) + 0.5)
> plot(woolTS)
```



Obrázek 1: *Produkce vlny v Austrálii v letech 1965-1994.*

Na načtená data nejprve použijeme jednoduché exponenciální vyrovnávání (lokálně konstantní trend). Potřebujeme určit jedinou vyrovnávací konstantu α. Model bude ve tvaru

$$\hat{Y}_{t+h} = L_t,$$

kde L_t je dáno vztahem

$$L_t = \alpha Y_t + (1 - \alpha) L_{t-1}$$

K tomu můžeme použít funkci `HoltWinters`, která je sice určená pro obecnější Holt–Wintersovo exponenciální vyrovnávání, kde počítáme nejen s úrovní (jako v případě jednoduchého exponenciálního vyrovnávání), ale i s lineárním trendem a sezónní složkou. Abychom odhadli pouze model jednoduchého exponenciálního vyrovnávání s jedinou vyrovnávací konstantou α, je třeba zadat, že nás nezajímají vyrovnávací konstanty β a γ, a toho dosáhneme zadáním argumentu `beta=FALSE, gamma=FALSE`.

Odhadnutou konstantu L_t pro úroveň v čase t odpovídající poslednímu pozorování zobrazíme pomocí funkce `coefficients`. V R je tato konstanta označena jako `a`.

```
> x <- woolTS
> model1 <- HoltWinters(x, beta = FALSE, gamma = FALSE)
> summary(model1)
```

```
             Length Class  Mode
fitted       236    mts    numeric
x            119    ts     numeric
alpha          1    -none- numeric
beta           1    -none- logical
gamma          1    -none- logical
coefficients   1    -none- numeric
seasonal       1    -none- character
SSE            1    -none- numeric
call           4    -none- call
```

```
> model1$alpha
```

```
[1] 0.376407
```

```
> coefficients(model1)
```

```
       a
5711.117
```

Na základě tohoto modelu provedeme predikci na dva roky dopředu. Ta by měla vypadat takto

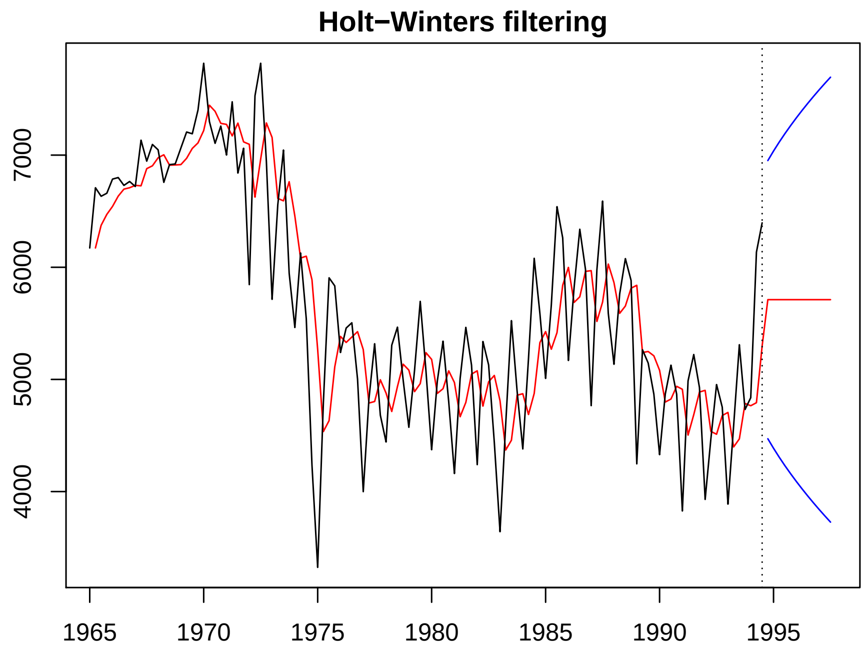
$$\hat{y}_{t+h|t} = L_t = \mathtt{a}$$

Výsledky jednoduchého exponenciálního vyrovnávání i predikci vykreslíme do grafu.

```
> pred1 <- predict(model1, n.ahead = 12, prediction.interval = TRUE)
> summary(pred1)
```

```
      fit            upr            lwr
 Min.   :5711   Min.   :6951   Min.   :3727
 1st Qu.:5711   1st Qu.:7173   1st Qu.:3885
 Median :5711   Median :7365   Median :4057
 Mean   :5711   Mean   :7349   Mean   :4073
 3rd Qu.:5711   3rd Qu.:7537   3rd Qu.:4250
 Max.   :5711   Max.   :7695   Max.   :4471
```

```
> par(mar = c(2, 2, 1, 0) + 0.5)
> plot(model1, predicted.values = pred1)
```



Obrázek 2: Jednoduché exponenciální vyrovnávání pomocí funkce `HoltWinters()` pro časovou řadu: *Produkce vlny v Austrálii v letech 1965-1994*

Z grafu je jasně patrné, že predikované hodnoty mají stejnou konstantní úroveň a ta se rovná odhadnuté úrovni pro poslední pozorování L_t – v R jí odpovídá konstanta `a`.

Zkusíme použít poněkud složitější model – Holtovo exponenciální vyrovnávání, které nevyužívá lokálně konstantní trend jako v předchozím případě, ale lokálně lineární trend. Tento typ exponenciálního vyrovnávání vyžaduje určení dvou vyrovnávacích konstant α a β.

Model exponenciálního vyrovnávání má tvar

$$\hat{Y}_{t+h} = L_t + T_t h,$$

kde L_t a T_t jsou dány

$$\begin{aligned} L_t &= \alpha Y_t + (1-\alpha)(L_{t-1} + T_{t-1}) \\ T_t &= \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1} \end{aligned}$$

Opět použijeme funkci `HoltWinters`, tentokrát použijeme argument `gamma=FALSE`, protože sezónní složka nás zatím nezajímá.

```
> model2 <- HoltWinters(x, gamma = FALSE)
> summary(model2)
```

```
             Length Class  Mode
fitted       351    mts    numeric
x            119    ts     numeric
alpha          1    -none- numeric
```

```
beta          1     -none- numeric
gamma         1     -none- logical
coefficients  2     -none- numeric
seasonal      1     -none- character
SSE           1     -none- numeric
call          3     -none- call
```

```
> model2$alpha
```

```
    alpha
0.4180789
```

```
> model2$beta
```

```
     beta
0.1901893
```

```
> coefficients(model2)
```

```
        a          b
5888.1247   205.5479
```

Na základě druhého modelu znovu provedeme predikci na dva roky dopředu. Ta bude ve tvaru

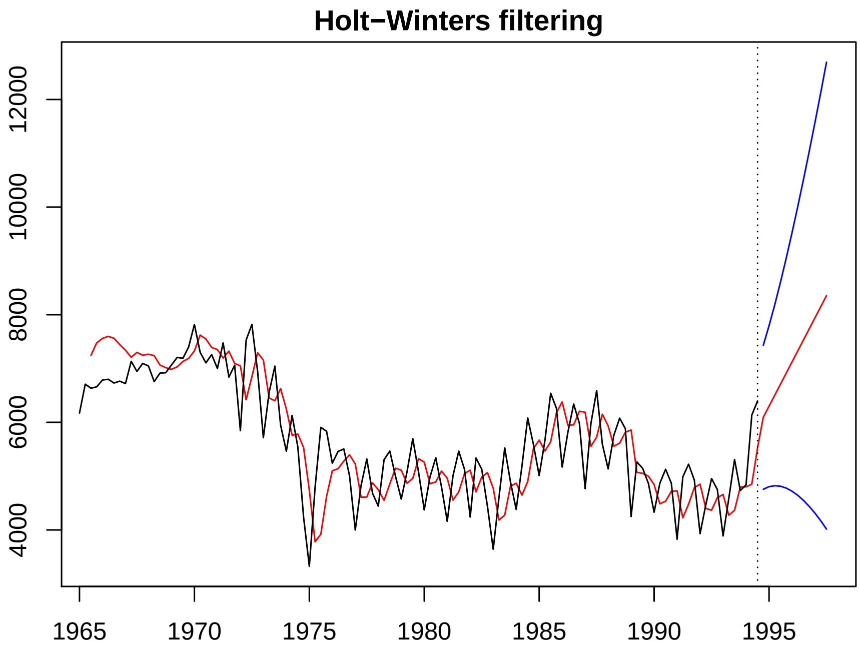
$$\hat{Y}_{t+h|t} = \mathsf{a} + \mathsf{b}t = 5888.1247 + 205.5479t.$$

Výsledky jednoduchého exponenciálního vyrovnávání i predikci vykreslíme do grafu.

```
> pred2 <- predict(model2, n.ahead = 12, prediction.interval = TRUE)
> summary(pred2)
```

```
      fit            upr              lwr
 Min.   :6094   Min.   : 7432   Min.   :4016
 1st Qu.:6659   1st Qu.: 8503   1st Qu.:4406
 Median :7224   Median : 9767   Median :4682
 Mean   :7224   Mean   : 9880   Mean   :4568
 3rd Qu.:7789   3rd Qu.:11173   3rd Qu.:4784
 Max.   :8355   Max.   :12694   Max.   :4822
```

```
> par(mar = c(2, 2, 1, 0) + 0.5)
> plot(model2, predicted.values = pred2)
```



Obrázek 3: Holtova metoda pomocí funkce `HoltWinters()` pro časovou řadu: *Produkce vlny v Austrálii v letech 1965-1994*

Porovnáme-li tento graf s grafem pro jednoduché vyrovnávání, můžeme si všimnout rozdílu v predikci. U jednoduchého exponenciálního vyrovnávání je predikce konstantní, zatímco u Holtovy metody je to rostoucí lineární funkce. To souvisí s tím, že u jednoduchého exponenciálního vyrovnávání pracujeme s lokálním polynomem nultého stupně, u Holtovy metody s lokálním polynomem prvního stupně.

Protože data, která máme k dispozici, jsou čtvrtletní, je možné (a při pohledu „od oka“ to vypadá hodně pravděpodobně), že bychom v nich mohli objevit sezónní chování. Proto vyzkoušíme ještě třetí metodu – Holtovo–Wintersovo vyrovnávání, které bere v úvahu i sezónnost.

Model bude ve tvaru

$$\hat{Y}_{t+h} = L_t + T_t h + S_{t-p+h_p^+},$$

kde L_t, T_t a S_t jsou dány

$$\begin{aligned} L_t &= \alpha(Y_t - s_{t-p}) + (1-\alpha)(L_{t-1} + T_{t-1}) \\ T_t &= \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1} \\ S_t &= \gamma(Y_t - L_t) + (1-\gamma)S_{t-p} \end{aligned}$$

Holtovo-Wintersovo vyrovnávání vyžaduje tři vyrovnávací konstanty α, β a γ. K jejich určení opět použijeme funkci `HoltWinters`, tentokrát již bez argumentů udávajících vynechané konstanty.

```
> model3 <- HoltWinters(x)
> summary(model3)
```

```
             Length Class  Mode
fitted       460    mts    numeric
x            119    ts     numeric
alpha          1    -none- numeric
beta           1    -none- numeric
gamma          1    -none- numeric
coefficients   6    -none- numeric
seasonal       1    -none- character
SSE            1    -none- numeric
call           2    -none- call
```

```
> model3$alpha
```

```
    alpha
0.6521521
```

```
> model3$beta
```

```
       beta
0.007690488
```

```
> model3$gamma
```

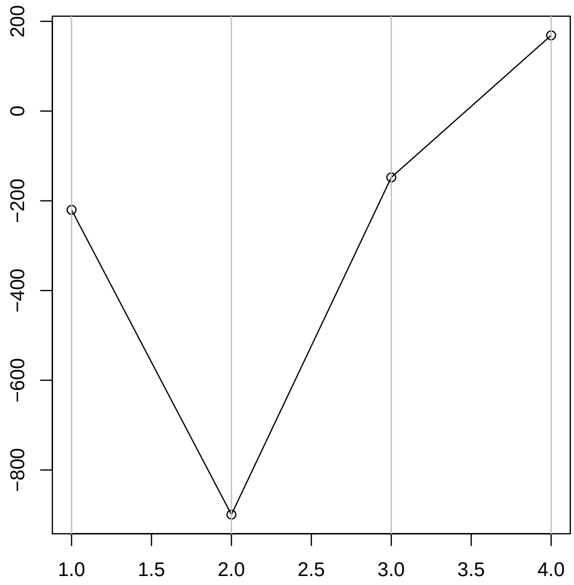
```
    gamma
0.6478487
```

```
> coefficients(model3)
```

```
         a          b         s1         s2         s3         s4
6222.72882   18.95975 -220.20379 -899.45453 -147.85037  168.78680
```

Hodnoty sezónních složek vykreslíme do grafu.

```
> par(mar = c(2, 2, 0, 0) + 0.5)
> plot(1:4, coefficients(model3)[3:6], type = "o")
> for (k in 1:4) abline(v = k, col = "gray", lty = 1)
```



Obrázek 4: Hodnoty sezónních složek z Holtova–Wintersova exponenciálního vyrovnávání pomocí funkce `HoltWinters()` pro časovou řadu: *Produkce vlny v Austrálii v letech 1965-1994*

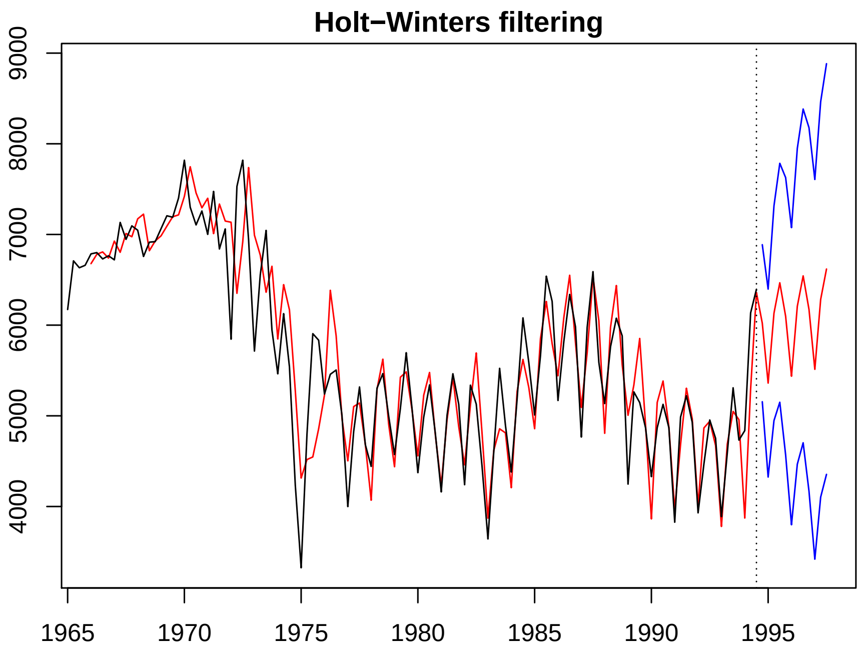
Na základě tohoto modelu znovu provedeme predikci na dva roky dopředu. Výsledky Holtova-Wintersova exponenciálního vyrovnání i predikci vykreslíme do jednoho grafu. Predikci dostaneme ze vztahu

$$\hat{Y}_{t+h|t} = \mathtt{a} + \mathtt{b}h + \mathtt{s}_{t+p+1+(h-1) \mod p},$$

```
> pred3 <- predict(model3, n.ahead = 12, prediction.interval = T)
> summary(pred3)
```

```
      fit             upr             lwr
 Min.   :5361   Min.   :6397   Min.   :3419
 1st Qu.:5894   1st Qu.:7256   1st Qu.:4152
 Median :6152   Median :7706   Median :4410
 Mean   :6071   Mean   :7713   Mean   :4430
 3rd Qu.:6329   3rd Qu.:8229   3rd Qu.:4764
 Max.   :6619   Max.   :8883   Max.   :5156
```

```
> par(mar = c(2, 2, 1, 0) + 0.5)
> plot(model3, predicted.values = pred3)
```



Obrázek 5: Holtovo–Wintersovo exponenciální vyrovnávání pomocí funkce `HoltWinters()` pro časovou řadu: *Produkce vlny v Austrálii v letech 1965-1994*

Dále se podíváme, jak dopadnou výsledky v případě použití funkce `ets()` z balíčku `forecast`.

```
> library(forecast)
```

Funkce `ets()` umožňuje vyrovnávání jak v aditivních, tak v multiplikativních modelech (to umí i `HoltWinters()`), a navíc umí odhadnout i exponenciální vyrovnávání s tlumícím faktorem. My zatím nebudeme uvažovat tak komplikované modely, provedeme obyčejné Holtovo-Wintersovo exponenciální vyrovnávání bez tlumícího faktoru a budeme chtít pouze aditivní model, to znamená, že uděláme přibližně totéž co v předchozím modelu (`model3`).

Odhad vyrovnávacích konstant bez tlumícího faktoru provedeme v případě funkce `ets()` přidáním argumentu `damped=FALSE`, argumentem `additive.only=TRUE` vyloučíme možnost multiplikativních modelů.

```
> x <- woolTS
> modelETS1 <- ets(x, damped = FALSE, additive.only = TRUE)
> summary(modelETS1)

ETS(A,N,A)

Call:
 ets(y = x, damped = FALSE, additive.only = TRUE)

  Smoothing parameters:
    alpha = 0.7488
    gamma = 1e-04

  Initial states:
    l = 6644.9513
    s=31.5458 408.2945 129.0941 -568.9344

  sigma:  436.0035

     AIC     AICc      BIC
2027.196 2027.946 2043.871

In-sample error measures:
         ME        RMSE         MAE         MPE        MAPE
 -7.8706447 436.0035123 340.1038656  -0.5567476   6.0128422
       MASE
  0.6055871
```

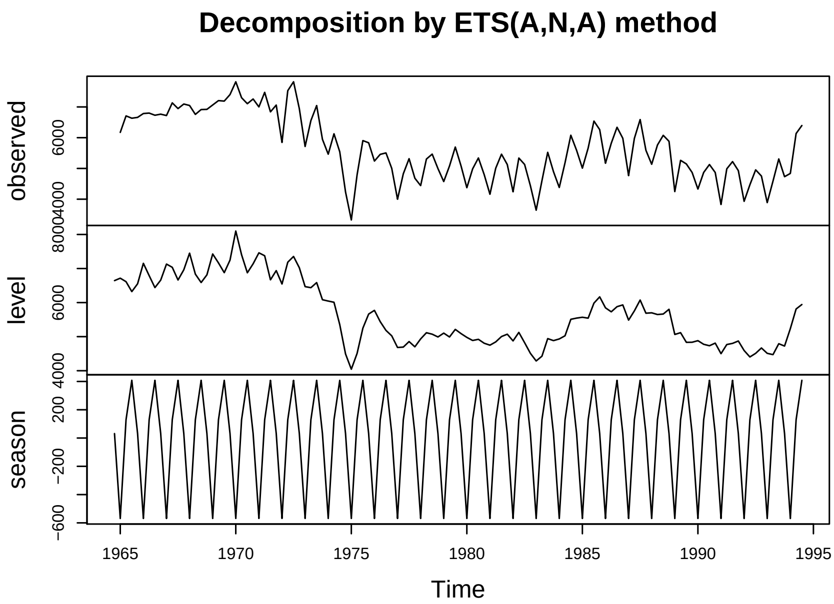
Prohlédneme si odhadnuté parametry.

```
> modelETS1$par

       alpha        gamma            l           s0           s1
7.487666e-01 1.000492e-04 6.644951e+03 3.154583e+01 4.082945e+02
          s2
1.290941e+02
```

Na základě AIC kritéria funkce `ets()` vybrala pro naše data jednoduché exponenciální vyrovnávání se sezónní složkou – tj. lokálně konstantní trend s odchylkami odpovídajícím čtvrtletím (značení $ETS(A, N, A)$). Všechny komponenty jsou aditivní. Odhadnutý model se liší od modelu 3 v tom, že se v něm nevyskytuje lineární trend.

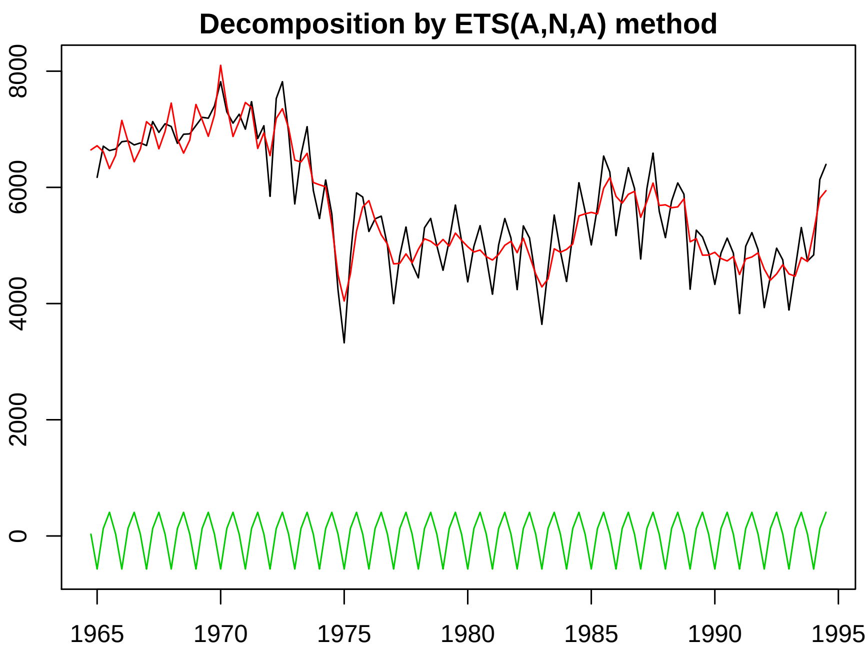
Výsledky exponenciálního vyhlazení opět vykreslíme. Pokud použijeme funkci `plot` na model, který je výstupem funkce `ets`, dostaneme graf, v němž jsou vykreslena zvlášť původní data, odhadnuté úrovně L_t, trendová komponenta T_t a cyklická komponenta S_t. Pokud zadáme argument `plot.type=single`, vše se zobrazí do jediného grafu.

```
> par(mar = c(2, 2, 1, 0) + 0.5)
> plot(modelETS1)
```



Obrázek 6: Dekompozice pomocí funkce `ets()` pro časovou řadu: *Produkce vlny v Austrálii v letech 1965-1994*

```
> par(mar = c(2, 2, 1, 0) + 0.5)
> plot(modelETS1, plot.type = "single", col = 1:3, ylab = "")
```



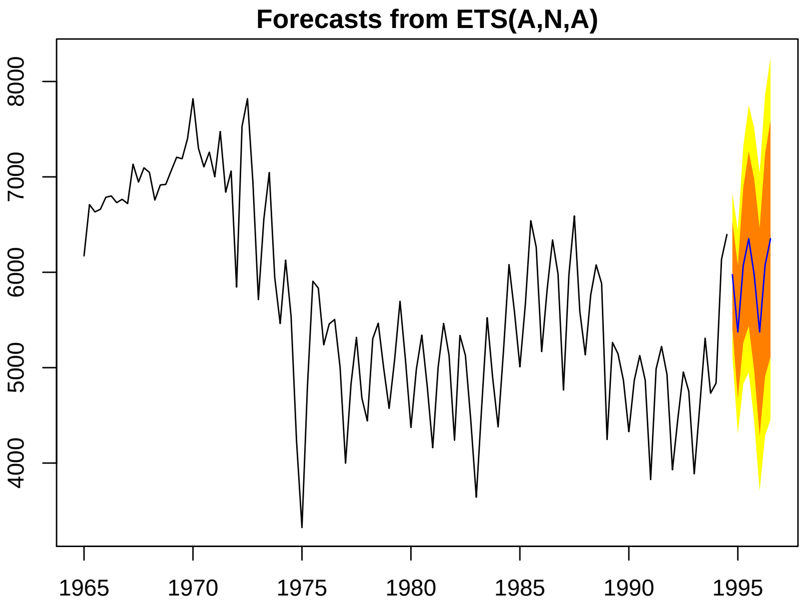
Obrázek 7: Holtovo–Wintersovo exponenciální vyrovnávání pomocí funkce `ets()` pro časovou řadu: *Produkce vlny v Austrálii v letech 1965-1994*

Opět se podíváme na predikované hodnoty pro dva následující roky.

```
> predETS1 <- forecast(modelETS1)
> summary(predETS1)
```

```
Forecast method: ETS(A,N,A)

Model Information:
ETS(A,N,A)

Call:
 ets(y = x, damped = FALSE, additive.only = TRUE)

  Smoothing parameters:
    alpha = 0.7488
    gamma = 1e-04

  Initial states:
    l = 6644.9513
    s=31.5458 408.2945 129.0941 -568.9344

  sigma:  436.0035

     AIC     AICc      BIC
2027.196 2027.946 2043.871

In-sample error measures:
         ME        RMSE         MAE         MPE        MAPE
 -7.8706447 436.0035123 340.1038656  -0.5567476   6.0128422
       MASE
  0.6055871

Forecasts:
        Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
1994 Q4       5975.187 5416.426 6533.948 5120.636 6829.738
1995 Q1       5374.703 4676.665 6072.741 4307.146 6442.260
1995 Q2       6072.705 5258.858 6886.552 4828.034 7317.377
1995 Q3       6351.917 5436.827 7267.007 4952.407 7751.427
1995 Q4       5975.187 4968.989 6981.384 4436.341 7514.033
1996 Q1       5374.703 4284.989 6464.417 3708.129 7041.277
1996 Q2       6072.705 4905.415 7239.995 4287.489 7857.921
1996 Q3       6351.917 5111.913 7591.921 4455.495 8248.339
```

```
> par(mar = c(2, 2, 1, 0) + 0.5)
> plot(predETS1)
```

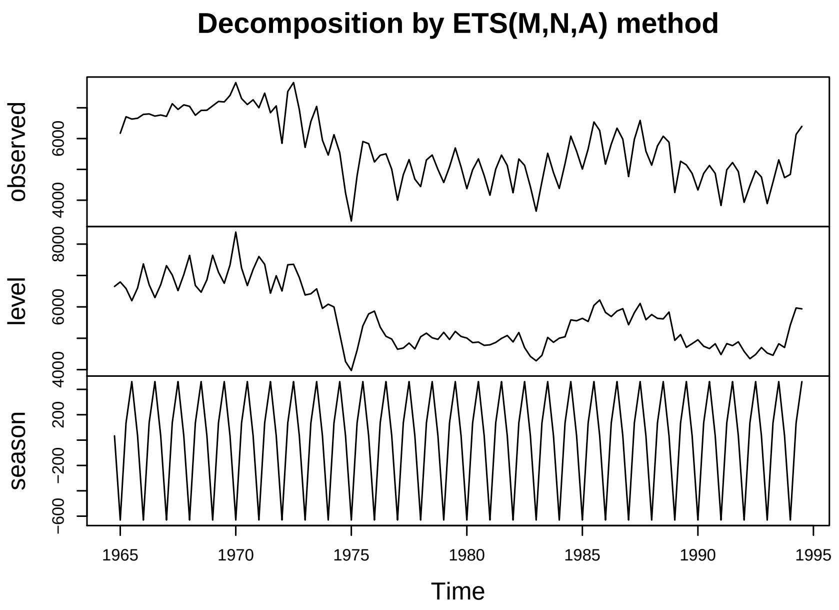


Obrázek 8: Predikce pomocí funkce `forecast()` pro časovou řadu: *Produkce vlny v Austrálii v letech 1965-1994*

Pro zajímavost si můžeme ještě na závěr vyzkoušet, jak by dopadlo exponenciální vyrovnání naší časové řady v případě, že bychom funkci `ets()` povolili použít všechny možnosti, které umí, včetně tlumícího faktoru a multiplikativních modelů.

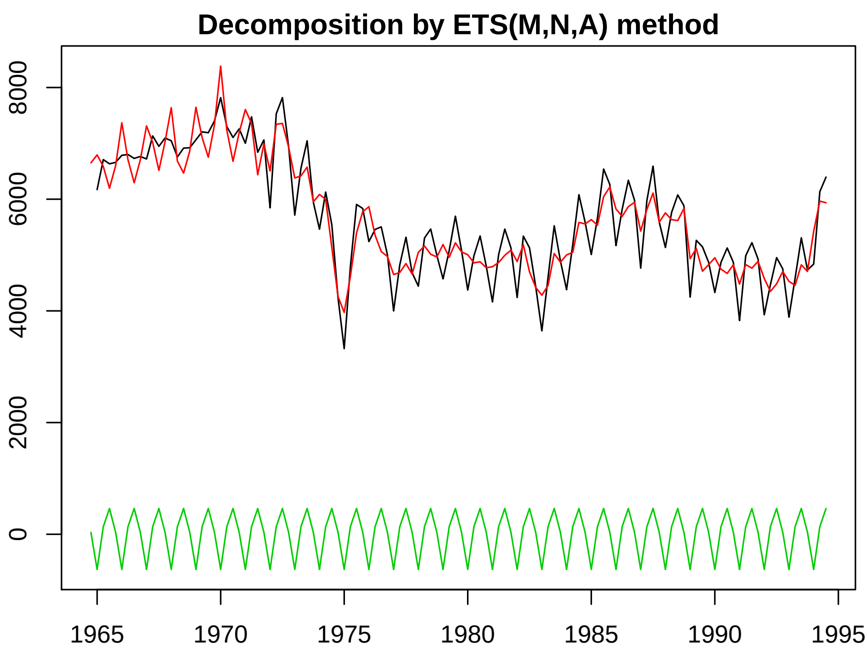
```
> x <- woolTS
> modelETS2 <- ets(x)
> summary(modelETS2)


ETS(M,N,A)

Call:
 ets(y = x)

  Smoothing parameters:
    alpha = 0.941
    gamma = 1e-04

  Initial states:
    l = 6650.654
    s=33.2949 461.4852 135.7959 -630.576

  sigma:  0.0748

     AIC     AICc      BIC
2016.480 2017.230 2033.155

In-sample error measures:
         ME        RMSE          MAE          MPE         MAPE
 -6.3777596 443.0083548 349.5425906  -0.4032783   6.0751771
       MASE
  0.6223936
```

```
> par(mar = c(2, 2, 1, 0) + 0.5)
> plot(modelETS2)
```



Obrázek 9: Dekompozice pomocí funkce `ets()`: *Produkce vlny v Austrálii v letech 1965-1994*

```
> par(mar = c(2, 2, 1, 0) + 0.5)
> plot(modelETS2, plot.type = "single", col = 1:3, ylab = "")
```

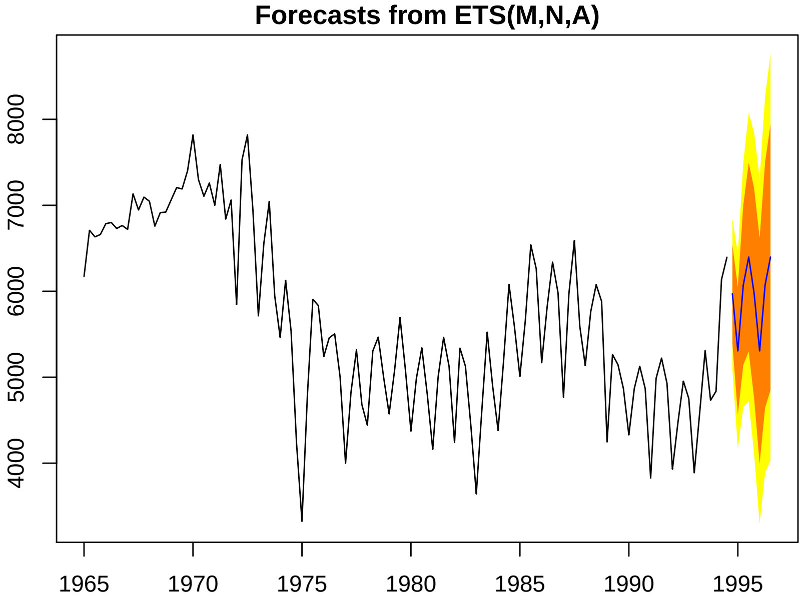


Obrázek 10: Holtovo–Wintersovo exponenciální vyrovnávání pomocí funkce `ets()` pro časovou řadu: *Produkce vlny v Austrálii v letech 1965-1994*

```
> predETS2 <- forecast(modelETS2)
> summary(predETS2)
```

```
Forecast method: ETS(M,N,A)

Model Information:
ETS(M,N,A)

Call:
 ets(y = x)

  Smoothing parameters:
    alpha = 0.941
    gamma = 1e-04

  Initial states:
    l = 6650.654
    s=33.2949 461.4852 135.7959 -630.576

  sigma:  0.0748

     AIC     AICc      BIC
2016.480 2017.230 2033.155

In-sample error measures:
         ME        RMSE          MAE         MPE        MAPE
 -6.3777596 443.0083548 349.5425906  -0.4032783   6.0751771
       MASE
  0.6223936

Forecasts:
        Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
1994 Q4       5969.934 5397.324 6542.543 5094.203 6845.664
1995 Q1       5306.109 4563.842 6048.377 4170.909 6441.310
1995 Q2       6072.065 5142.938 7001.192 4651.088 7493.042
1995 Q3       6397.812 5299.973 7495.651 4718.813 8076.812
1995 Q4       5969.934 4746.830 7193.037 4099.358 7840.509
1996 Q1       5306.109 3992.769 6619.450 3297.529 7314.690
1996 Q2       6072.065 4642.732 7501.397 3886.090 8258.040
1996 Q3       6397.812 4851.590 7944.034 4033.069 8762.555
```

```
> par(mar = c(2, 2, 1, 0) + 0.5)
> plot(predETS2)
```



Obrázek 11: Predikce pomocí funkce `forecast()` pro časovou řadu: *Produkce vlny v Austrálii v letech 1965-1994*

7 Úkol:

Aplikujte exponenciální vyhlazování na časovou řadu s měsíčními údaji o počtu smrtelných úrazů v USA v letech 1973-1978 v souboru `deaths.dat`.