

*Modelovanie trendu: Exponenciálne zhladzovanie,
Holt-Wintersova metóda, Hodrick-Prescottov filter*

Beáta Stehlíková

Cvičenia z časových radov, FMFI UK

Exponenciálne zhladzovanie

- Máme dáta x_1, \dots, x_n a chceme predikovať hodnotu x_{n+k}
- V tejto časti predpokladáme, že dáta nemajú trend, ani sezónne efekty

- Model:

$$x_t = \mu_t + w_t,$$

kde

- μ_t je stredná hodnota procesu, môže závisieť od času t
 - w_t sú nezávislé náhodné odchýlky s nulovou strednou hodnotou
- Označme a_t náš odhad strednej hodnoty μ_t

Exponenciálne zhladzovanie

- Základná myšlienka exponenciálneho zhladzovania: ďalší odhad strednej hodnoty bude váženým priemerom predchádzajúceho odhadu (teda a_{t-1}) a realizovanej hodnoty x_t :

$$a_t = \alpha x_t + (1 - \alpha)a_{t-1},$$

pričom $\alpha \in (0, 1)$

- Parameter α - parameter zhladzovania:
 - $\alpha \approx 1$: slabé zhladzovanie, $a_t \approx x_t$
 - $\alpha \approx 0$: silné zhladzovanie, malý vplyv pozorovanej hodnoty procesu x_t
- Predikcie: keďže nemáme trend ani sezónnosť, vieme spraviť iba $\hat{x}_{n+k|n} = a_n$

Exponenciálne zhladzovanie

- Iný zápis a_t :

$$a_t = \alpha x_t + \alpha(1 - \alpha)x_{t-1} + \alpha(1 - \alpha)^2 x_{t-2} + \dots$$

- z toho názov exponenciálne zhladzovanie (váhy exponenciálne klesajú, v súčte dávajú 1)

- Pre daný parameter α :
 - $a_1 = x_1$; ostatné rekurentne
 - predikčné chyby: $e_t = x_t - \hat{x}_{t|t-1} = x_t - a_{t-1}$
- Optimálny parameter zhladzovania: $\sum_{t=2}^n e_t^2 \rightarrow \min$

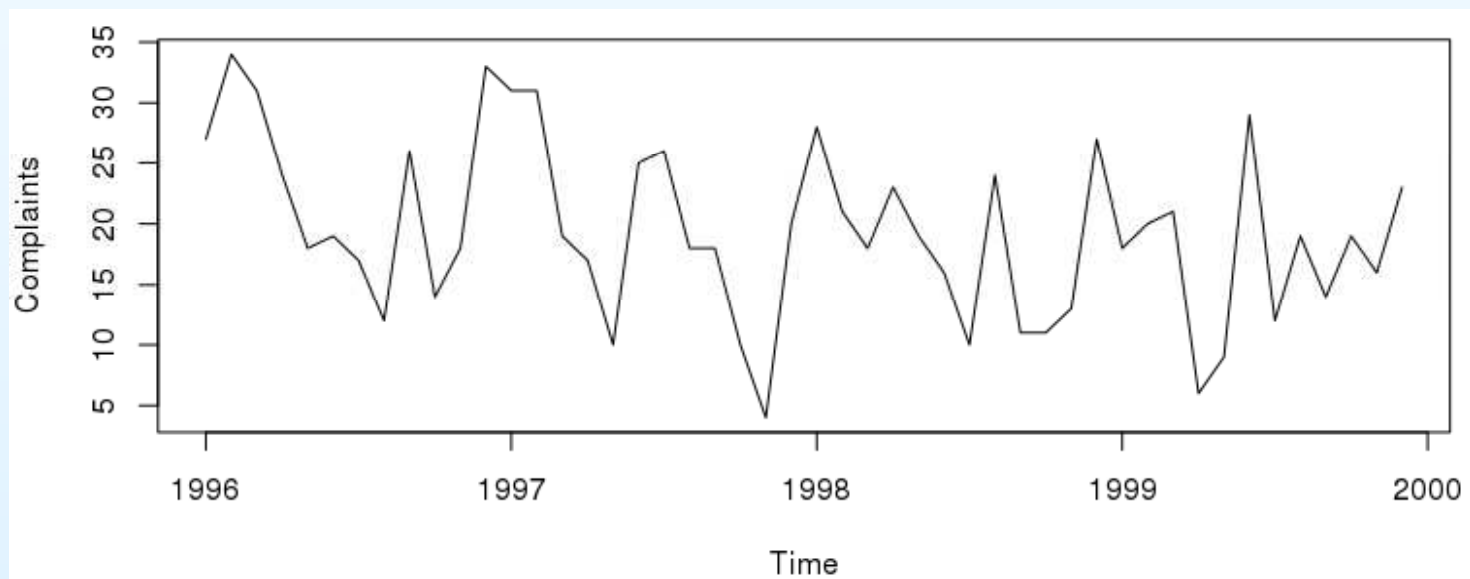
Exponenciálne zhladzovanie v R-ku: príklad

PRÍKLAD:

P. S. P.Cowpertwait, A. V. Metcalfe: Introductory Time Series with R. Springer, 2009.

Complaints to a motoring organization, pp. 56-58.

- **Dáta:**
 - `motor.txt` na stránke cvičení
 - počet sťažností, mesačné dáta, 1996/01 - 1999/12
- **Priebeh:**



Exponenciálne zhladzovanie v R-ku: príklad

- Samotné exponenciálne zhladzovanie:

```
model1 = HoltWinters(x, beta=FALSE, gamma=FALSE)
```

(lebo je špeciálny to prípad všeobecnejšieho modelu, pre ktorý máme funkciu `HoltWinters` - uvidíme neskôr)

```
> model1=HoltWinters(x, beta=FALSE,gamma=FALSE)
> model1
Holt-Winters exponential smoothing without trend and without seasonal component

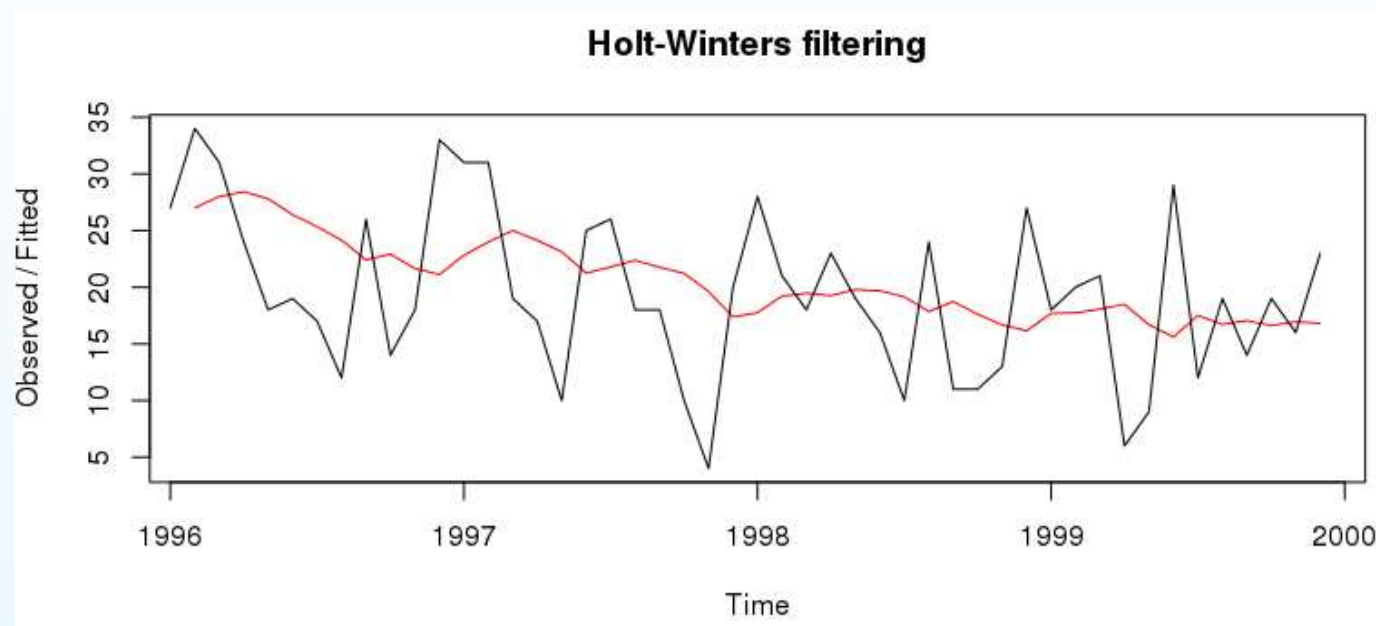
Call:
HoltWinters(x = x, beta = FALSE, gamma = FALSE)

Smoothing parameters:
alpha: 0.1429622
beta : FALSE
gamma: FALSE

Coefficients:
      [,1]
a 17.70343
```

Exponenciálne zhladzovanie v R-ku: príklad

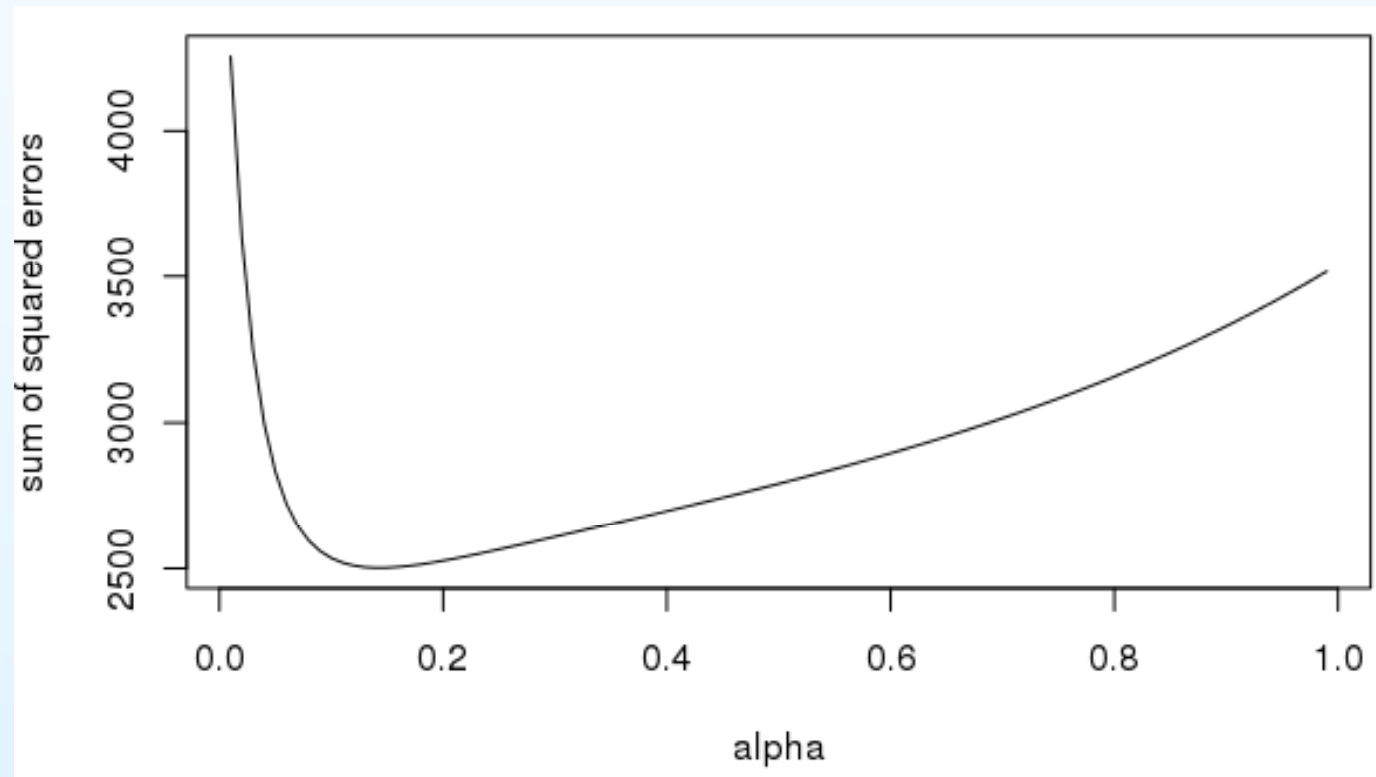
- Vykreslenie: `plot(model1)`



- Prístup k hodnote *sum of squared errors*, podľa ktorej sa vyberala optimálna α : príkazom `model1$SSE`
- Ak chceme použiť našu hodnotu α : napr. `model1 = HoltWinters(x, alpha=0.2, beta=FALSE, gamma=FALSE)`

Exponenciálne zhladzovanie v R-ku: príklad

- CVIČENIE:
Vykreslite závislosť SSE od parametra α pre tieto dáta.
Váš výpočet má potvrdiť optimálnu hodnotu α , ktorú našlo R-ko.



Holt-Wintersova metóda

- Charakteristiky časového radu:
 - $a_t = \text{level}$ - sezónne očistená stredná hodnota
 - $b_t = \text{slope}$ - zmena hodnoty *level* z jednej periódy na druhú (zachytáva rôzne, aj krátkodobé trendy)
 - $s_t = \text{seasonal component}$ - sezónna zložka (závisí napr. od mesiaca)
- Predikcia pri aditívnej sezónnosti:

$$\hat{x}_{n+k|n} = a_n + kb_n + s_{n+k-p}$$

pre $k \leq p$ (napr. pre mesačné dáta je $p = 12$)

- Pri mutiplikatívnej sezónnosti:

$$\hat{x}_{n+k|n} = (a_n + kb_n)s_{n+k-p}$$

Holt-Wintersova metóda

- Algoritmus pre aditívnu sezónnosť:

$$a_t = \alpha(x_t - s_{t-p}) + (1 - \alpha)(a_{t-1} + b_{t-1}),$$

$$b_t = \beta(a_t - a_{t-1}) + (1 - \beta)b_{t-1},$$

$$s_t = \gamma(x_t - a_t) + (1 - \gamma)s_{t-p},$$

pričom $\alpha, \beta, \gamma \in (0, 1)$.

- Pre multiplikatívnu sezónnosť analogicky, napr. V R-ku `help(HoltWinters)`
- Parametre α, β, γ sa určia minimalizáciou SSE

Holt-Wintersova metóda v R-ku: príklad

PRÍKLAD:

P. S. P.Cowpertwait, A. V. Metcalfe: Introductory Time Series with R. Springer, 2009.

Sales of Australian wine, pp. 60-62

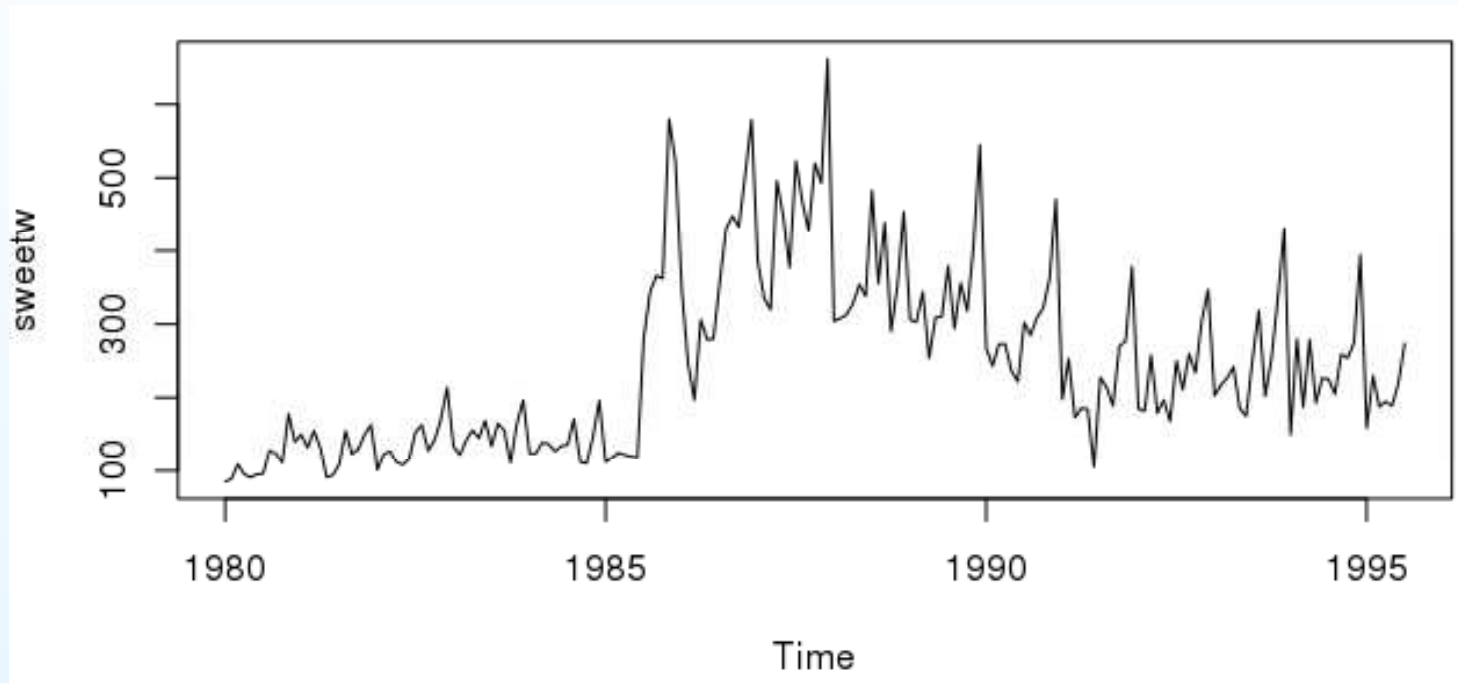
- Dáta:
 - `wine.txt` na stránke cvičení
 - predaj austrálskeho vína v tisícoch litrov, mesačné dáta, 1980/01 - 1995/07

```
> wine=read.table("wine.txt", header=T)
> attach(wine)
>
> wine
  winet fortw dryw sweetw  red  rose spark
1      1  2585 1954     85  464  112 1686
2      2  3368 2302     89  675  118 1591
3      3  3210 3054    109  703  129 2304
4      4  3111 2414     95  887   99 1712
```

a spravíme z toho časový rad: `sweetw=ts(sweetw, ...)`

Holt-Wintersova metóda v R-ku: príklad

- Priebeh:



- Očakávame multiplikatívnu sezónnosť

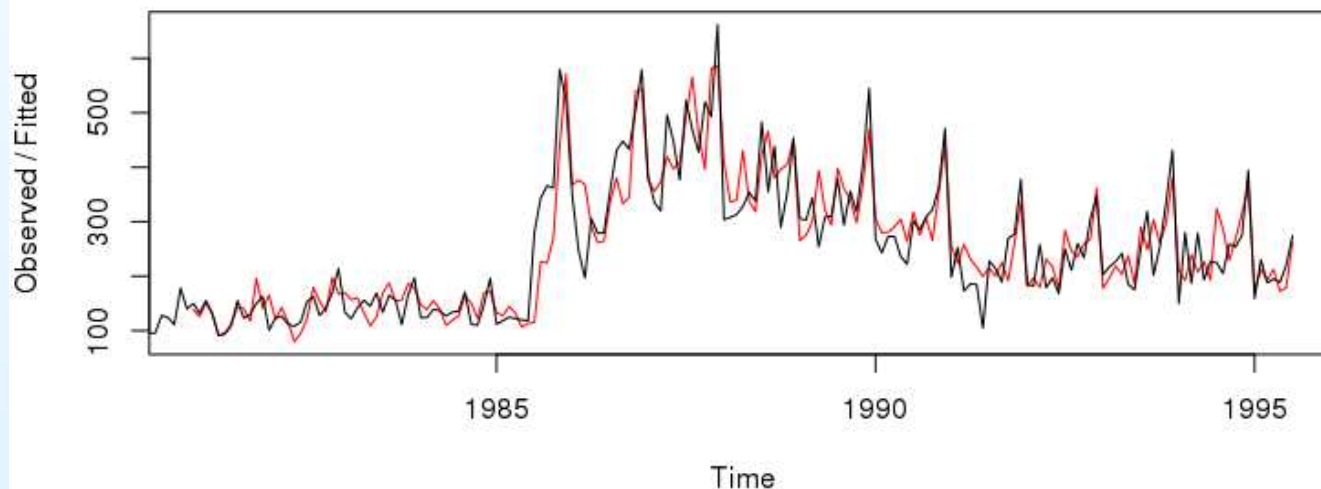
Holt-Wintersova metóda v R-ku: príklad

```
> HWSweet=HoltWinters(sweetw, seasonal="mult")
> HWSweet
Holt-Winters exponential smoothing with trend and multiplicative seasonal component.

Call:
HoltWinters(x = sweetw, seasonal = "mult")

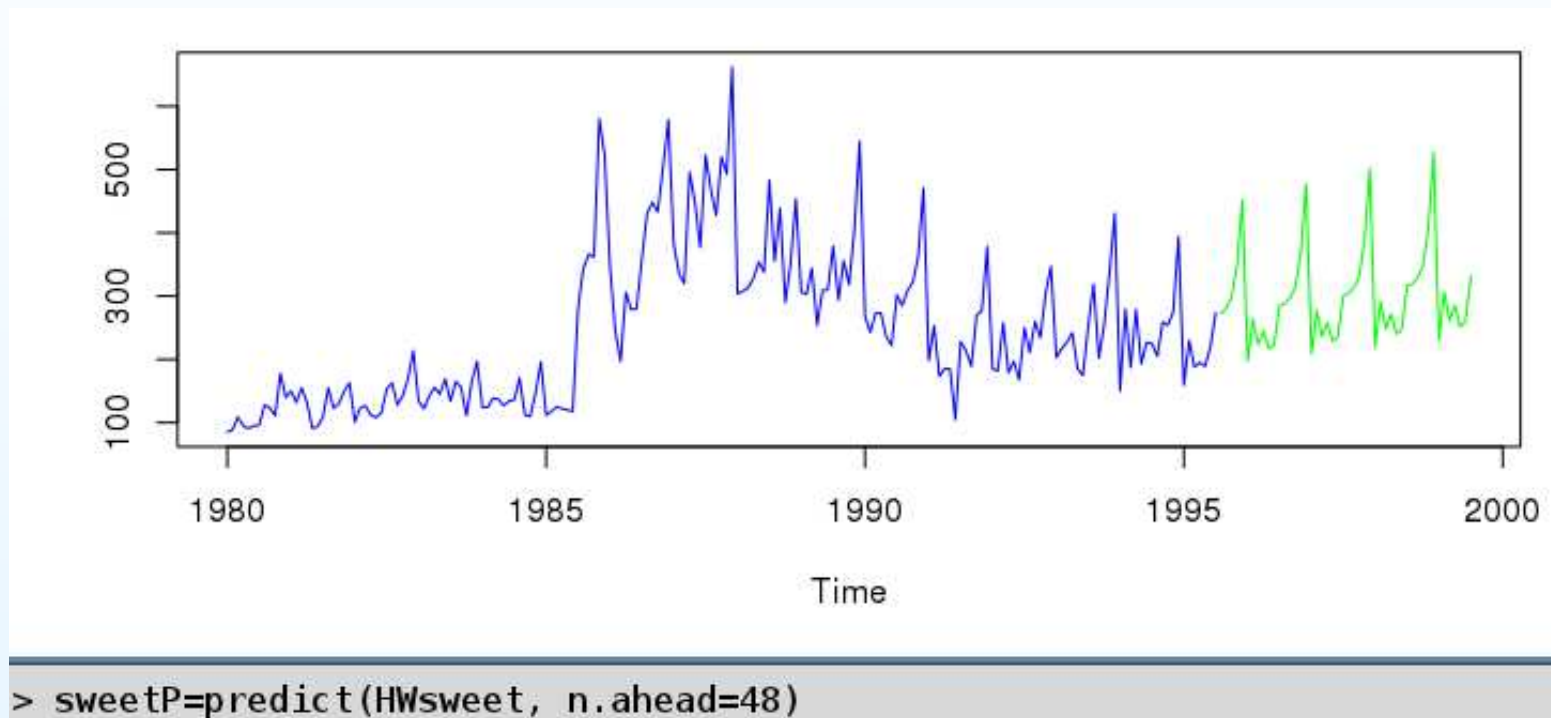
Smoothing parameters:
alpha: 0.4086698
beta : 0
gamma: 0.4929402
```

Holt-Winters filtering



Holt-Wintersova metóda v R-ku: príklad

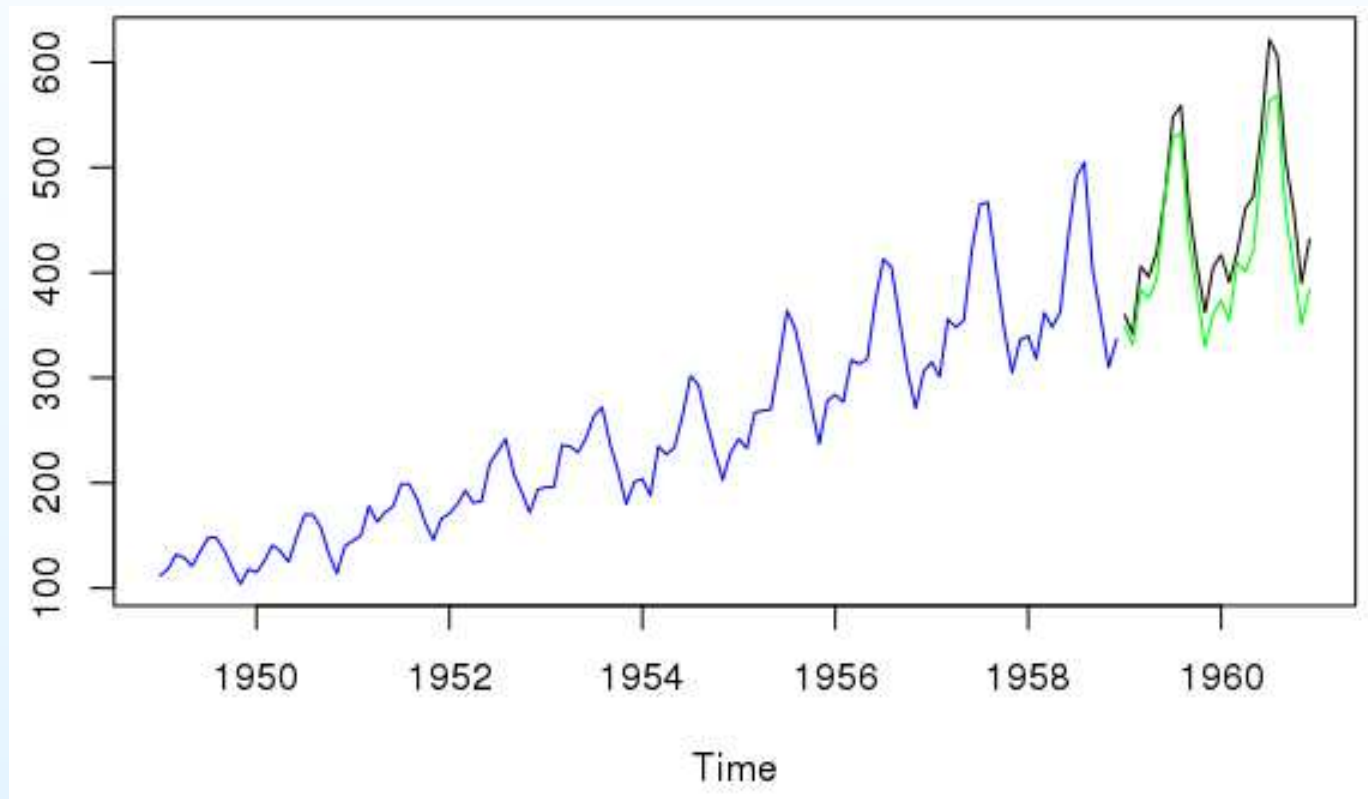
Konštrukcia predikcií - funkcia `predict`:



Kreslenie viacerých časových radov do jedného grafu:
`ts.plot(sweetw, sweetP, gpars=list(col=c("blue","green")))`

Holt-Wintersova metóda v R-ku: cvičenie

Zoberte dáta o počtoch cestujúcich aerolinkami, vynechajte posledné hodnoty a porovnajte ich s predikciou získanou Holt-Wintersovou metódou:



Výber časti časového radu: napr. `window(air, end=c(1958))`

Hodrick-Prescottov filter

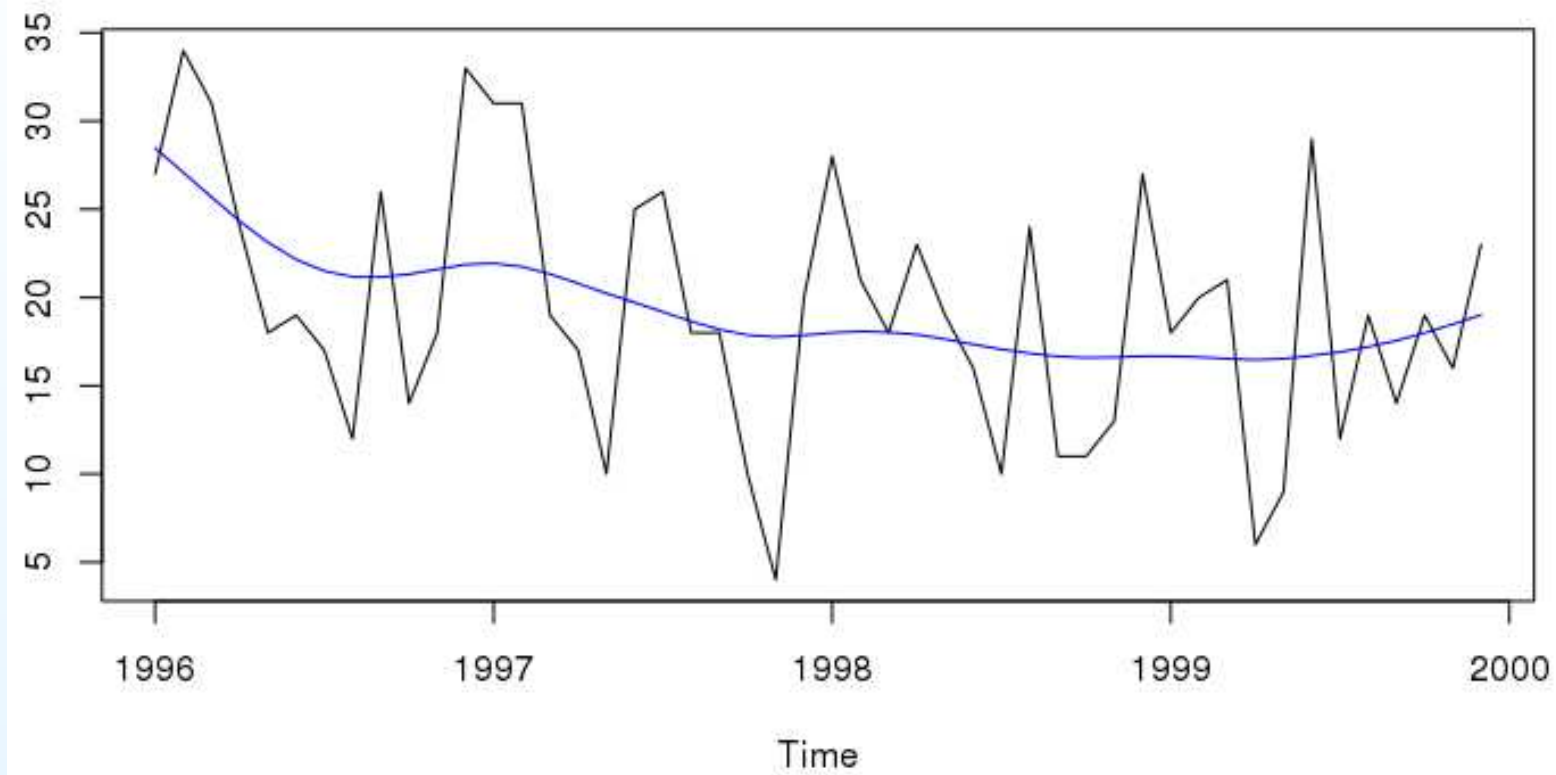
- PREDPOKLAD: V dátach nie je sezónnosť
- CIEL': znovu chceme vyhladiť dáta a získať trend
- Myšlienka: potrebujeme dosiahnuť dve kritériá, ktoré sú v protiklade:
 - vyhladené hodnoty by mali byť bízko skutočných
 - hladkosť, malá krivosť grafu vyhladených hodnôt (nie veľké fluktuácie), tú vieme merať druhými diferenciami (analógia s druhou deriváciou)
- priradíme im váhy
- Optimalizačná úloha (y_1, \dots, y_n sú dáta, λ parameter):

$$\sum_{t=1}^n (y_t - \tilde{y}_t)^2 + \lambda \sum_{t=2}^{n-1} (\tilde{y}_{t+1} - 2\tilde{y}_t + \tilde{y}_{t-1})^2 \rightarrow \min_{\tilde{y}_1, \dots, \tilde{y}_n}$$

Hodrick-Prescottov filter v R-ku: príklad

- Hodrick-Prescottov filter v R-ku:
 - knižnica `library(mFilter)`
 - potom napr. `hpf1=hpfilter(x,freq=100)`, kde `freq` je parameter λ
 - odhadnutý trend: `hpf1$trend`
- Vplyv parametra λ :
 - zoberme dáta o počte sťažností
 - použime HP-filter napr. takto: `hpf1=hpfilter(x,freq=100)`
 - a vykreslime: `lines(hpf1$trend)`

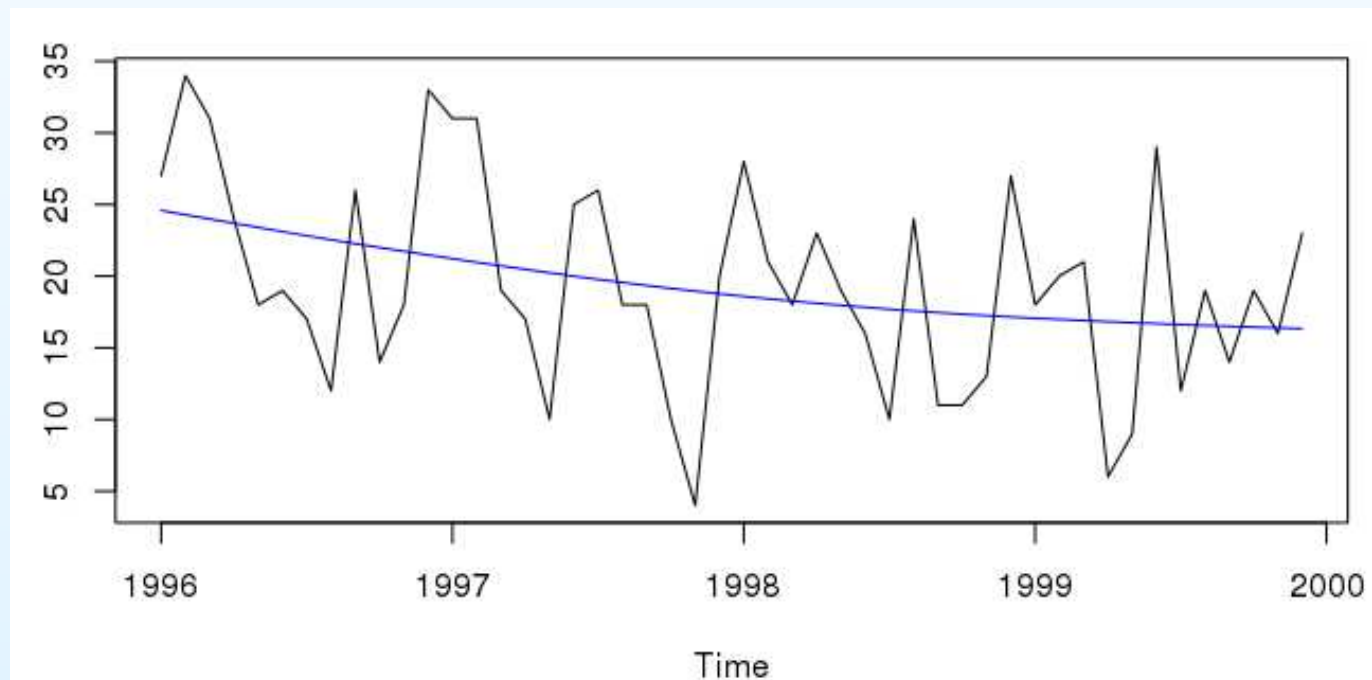
Hodrick-Prescottov filter v R-ku: príklad



Dokreslite vyhladené dáta pre iné hodnoty λ . Čo sa deje pre $\lambda \rightarrow 0$ a pre $\lambda \rightarrow \infty$? Prečo?

Hodrick-Prescottov filter v R-ku: príklad

- Obvyklé hodnoty parametra λ :
 - $\lambda = 100$ pre ročné dáta
 - $\lambda = 1600$ pre kvartálne dáta
 - $\lambda = 14400$ pre mesačné dáta
- V našom prípade:



Hodrick-Prescottov filter: aplikácia

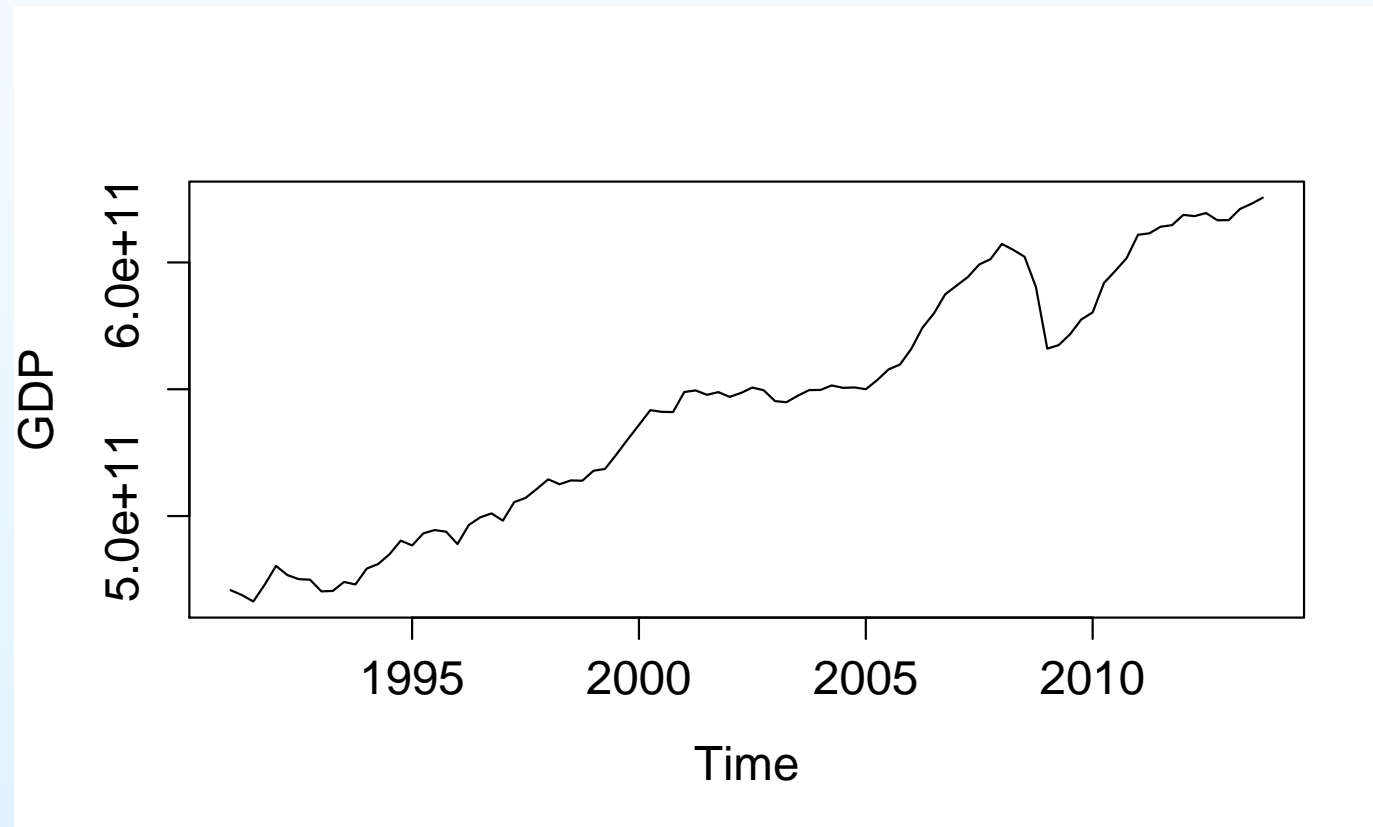
- Potenciálny HDP a produkčná medzera:
 - potenciálny HDP: maximálny výstup, ktorý vie ekonomika pri daných faktoroch vyprodukovať bez inflačných tlakov
 - skutočný HDP osciluje okolo potenciálneho (hospodárske cykly)
 - produkčná medzera: rozdiel medzi potenciálnym a reálnym výstupom
- Aplikácia HP filtra na skúmanie produkčnej medzery:
 - trendom HP filtra odhadneme potenciálny HDP
 - je to jeden z prístupov z nedávnej diplomovej práce
<http://www.iam.fmph.uniba.sk/studium/efm/diplomovky/2012/silanic/diplomovka.pdf>

Hodrick-Prescottov filter: aplikácia

- Nemecko, HDP v konštantných cenách, sezónne očistené kvartálne dáta (posledná hodnota zo 4. kvartálu 2013)

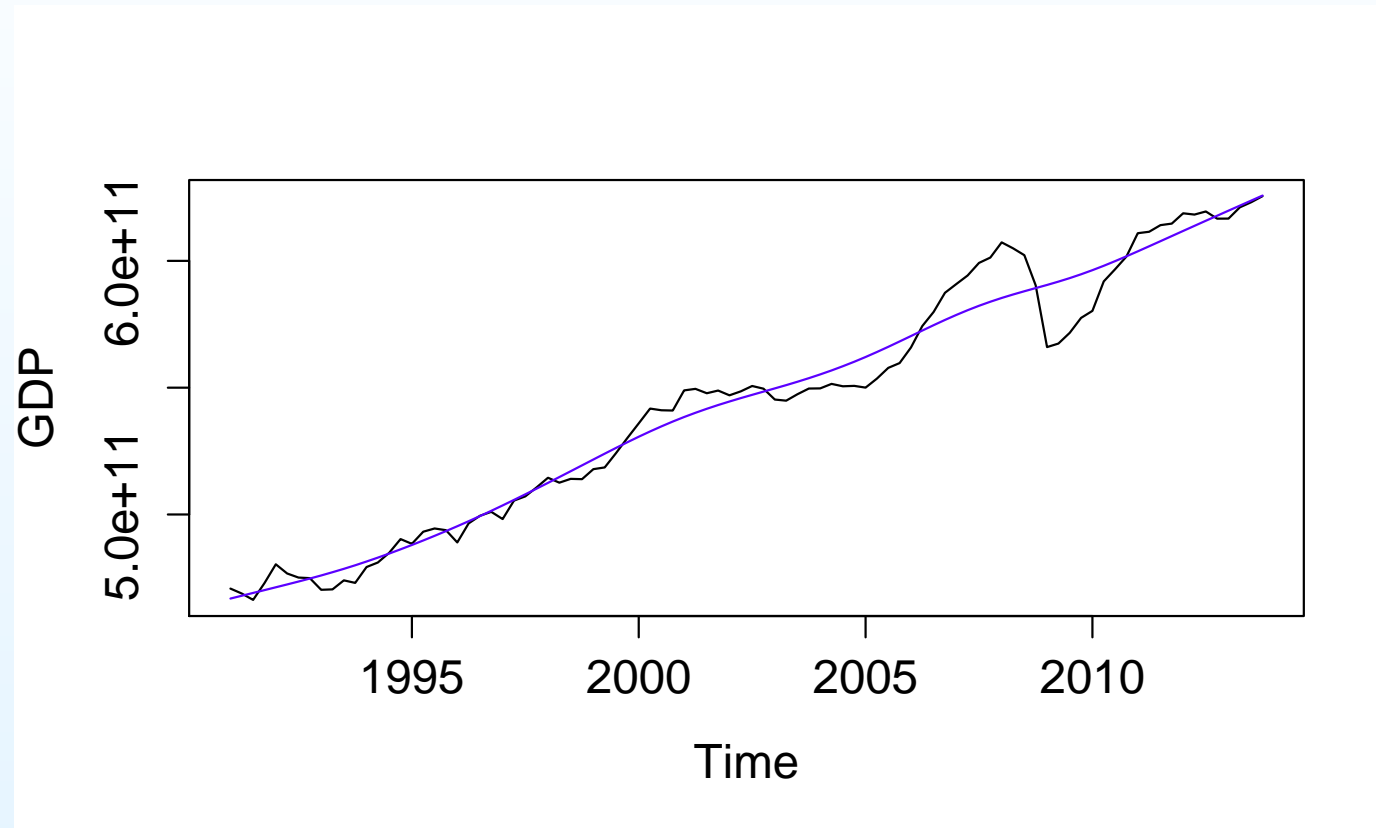
<http://research.stlouisfed.org/fred2/series/>

NAEXKP01DEQ189S/downloaddata?cid=32273



Hodrick-Prescottov filter: aplikácia

- Odhadnite HP filtrom potenciálny HDP:



Hodrick-Prescottov filter: aplikácia

- Zobrazte priebeh produkčnej medzery. Kedy je kladná a kedy záporná?
- Aby sme mohli porovnávať viaceré štáty: zobrazte produkčnú medzeru ako percento z potenciálneho HDP (DP, obrázok 3.4 na str. 33)

Domáca úloha

Len jednu z možností, každá je za **20 bodov**.

MOŽNOSŤ 1

Nájdite dáta pre iný štát (potrebujete sezónne očistené kvartálne hodnoty HDP v stálych cenách), zjenuťte časový interval a porovnajte priebeh produkčnej medzery vyjdrenej ako percento z potenciálneho HDP - pre tento štát a Nemecko

MOŽNOSŤ 2

Nájdite dáta, ktoré obsahujú sezónnosť (dobrý zdroj dát je napr. <https://datamarket.com/data/list/?q=provider:tsdl>). Vynechajte niekoľko posledných dát (aspoň jednu periódu) a zostávajúce modelujte Holt-Wintersovou metódou. Spravte predikcie a porovnajte ich so skutočnými hodnotami.

Domáca úloha

ODOVZDÁVANIE:

- Mailom na adresu **beata.ulohy@gmail.com** s predmetom **CR cvicenia 2014 - DU2 - meno/mena**, úlohu môžete robiť samostatne alebo v dvojici
- Úloha v **pdf** formáte + dáta v **txt** formáte (nie spolu ako zip/rar/... archív)
- Termín: **15. 10. 2014**
- Každý má iné dáta, link pre rezerváciu dát na stránke cvičení