Modelo linear de Holt e modelo Holt+Damped

Análise de Séries Temporais

José Augusto Fiorucci

Aula anterior

- Modelo SES(ℓ_0, α), com $\ell_0 \in \mathbb{R}$ e $\alpha \in (0,1)$
  - Equação original

    $$\widehat{y}_{t+1} = \alpha y_t + \alpha(1-\alpha)y_{t-1} + \alpha(1-\alpha)^2 y_{t-2} + \alpha(1-\alpha)^3 y_{t-3} + \dots$$

    ou equivalentemente,

    $$\widehat{y}_{t+1} = \alpha y_t + (1-\alpha)\widehat{y}_t$$

    ou ainda,

    $$\begin{aligned} \text{Previsão:} \quad & \widehat{y}_t = \ell_{t-1} \\ \text{Level:} \quad & \ell_t = \alpha\, y_t + (1-\alpha)\ell_{t-1} \end{aligned}$$

  - Modelo Estocástico com erros aditivos ($y_t = \mu_t + \varepsilon_t$)

    $$\begin{aligned} y_t &= \ell_{t-1} + \varepsilon_t \\ \ell_t &= \alpha y_t + (1-\alpha)\ell_{t-1} \end{aligned}$$

  - Representação em MEE

    $$\begin{aligned} y_t &= \ell_{t-1} + \varepsilon_t \\ \ell_t &= \ell_{t-1} + \alpha\, \varepsilon_t \end{aligned}$$

Aula anterior

- Previsão SES(ℓ_0, α)

$$\widehat{y}_{n+h|n} = E_n[y_{n+h}] = \ell_n, \qquad h = 1, 2, 3, \ldots$$

com

$$\sigma^2_{n+h|n} = Var_n[y_{n+h}] = \sigma^2_\varepsilon + \alpha^2\sigma^2_\varepsilon(h-1), \qquad h = 1, 2, 3, \ldots$$

Section 1

Método de Holt

Método linear de Holt

Holt (1953) incluiu um termo de tendência ao método SES, para que a previsão fosse corrigida pela tendência, isto é

$$\widehat{y}_{n+h|n} = E_n[y_{n+h}] = \ell_n + b_n h, \qquad h = 1, 2, 3, \ldots$$

em que b_t é o termo de crescimento (ou decrescimento), obtido a partir de um alisamento exponencial da variação dos leveis passados.

- Método de Holt:

$$\begin{aligned} \text{Previsão:} \quad & \widehat{y}_t = \ell_{t-1} + b_{t-1} \\ \text{Level:} \quad & \ell_t = \alpha\, y_t + (1-\alpha)(\ell_{t-1} + b_{t-1}) \\ \text{Crescimento:} \quad & b_t = \beta^* \left(\ell_t - \ell_{t-1}\right) + \left(1 - \beta^*\right) b_{t-1} \end{aligned}$$

Note que β^* é um parâmetro de alisamento, e portanto, $\beta^* \in (0, 1)$.

Admitindo erros aditivos ($y_t = \mu_t + \varepsilon_t$), o método de Holt corresponde ao seguinte modelo estocástico

$$\begin{aligned}
y_t &= \ell_{t-1} + b_{t-1} + \varepsilon_t \\
\ell_t &= \alpha\, y_t + (1-\alpha)(\ell_{t-1} + b_{t-1}) \\
b_t &= \beta^* \,(\ell_t - \ell_{t-1}) + (1-\beta^*)\, b_{t-1}
\end{aligned}$$

- Note que a segunda equação pode ser reescrita em função da componente estocástica

$$\ell_t = \ell_{t-1} + b_{t-1} + \alpha\,(\,y_t - \ell_{t-1} - b_{t-1}) \quad \Rightarrow \quad \ell_t = \ell_{t-1} + b_{t-1} + \alpha\,\varepsilon_t$$

da mesma forma, também podemos reescrever a terceira equação

$$b_t = b_{t-1} + \beta^*\,(\ell_t - \ell_{t-1} - b_{t-1}) \quad \Rightarrow \quad b_t = b_{t-1} + \alpha\,\beta^* \varepsilon_t$$

Assim obtemos a representação do modelo de Holt como MEE linear

$$
\begin{aligned}
y_t &= \ell_{t-1} + b_{t-1} + \varepsilon_t \\
\ell_t &= \ell_{t-1} + b_{t-1} + \alpha\,\varepsilon_t \\
b_t &= b_{t-1} + \beta\,\varepsilon_t
\end{aligned}
$$

em que $\beta = \alpha\,\beta^*$. Observe que os parâmetros do modelo são $\ell_0 \in \mathbb{R}, b_0 \in \mathbb{R}, \alpha \in (0,1)$ e $\beta \in (0, \alpha)$.

- Note que a correspondencia com a equação de MEE linear que vimos na aula anterior é direta, basta tomarmos $\mathbf{x}_t = (\ell_t\ ,\ b_t)'$ como o vetor de espaço de estado, e reescrevermos a ultimas equações em uma forma matricial

$$
\begin{aligned}
y_t &= \begin{pmatrix} 1 & 1 \end{pmatrix} \mathbf{x}_{t-1} + \varepsilon_t \\
\mathbf{x}_t &= \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \mathbf{x}_{t-1} + \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \varepsilon_t
\end{aligned}
$$

- Lembrete: MEE linear: $\begin{cases} y_t = & \mathbf{w}'\,\mathbf{x_{t-1}} + \varepsilon_t \\ \mathbf{x_t} = & \mathbf{F}\,\mathbf{x_{t-1}} + \mathbf{g}\,\varepsilon_t \end{cases}$

Previsão do modelo de Holt

Note que

$$
\begin{aligned}
y_{n+h} &= \ell_{n+h-1} + b_{n+h-1} + \varepsilon_{n+h} \\
&= (\ell_{n+h-2} + b_{n+h-2} + \alpha\,\varepsilon_{n+h-1}) + (b_{n+h-2} + \beta\,\varepsilon_{n+h-1}) + \varepsilon_{n+h} \\
&= \ell_{n+h-2} + 2\,b_{n+h-2} + (\alpha+\beta)\,\varepsilon_{n+h-1} + \varepsilon_{n+h} \\
&= (\ell_{n+h-3} + b_{n+h-3} + \alpha\,\varepsilon_{n+h-2}) + 2\,(b_{n+h-3} + \beta\,\varepsilon_{n+h-2}) + (\alpha+\beta)\,\varepsilon_{n+h-1} + \varepsilon_{n+h} \\
&= \ell_{n+h-3} + 3\,b_{n+h-2} + (\alpha+2\beta)\,\varepsilon_{n+h-2} + (\alpha+\beta)\,\varepsilon_{n+h-1} + \varepsilon_{n+h} \\
&\;\vdots \\
&= \ell_n + h\,b_n + \sum_{i=1}^{h-1}(\alpha + i\beta)\,\varepsilon_{n+h-i} + \varepsilon_{n+h}
\end{aligned}
$$

Como

$$y_{n+h} = \ell_n + h\, b_n + \sum_{i=1}^{h-1} (\alpha + i\beta)\, \varepsilon_{n+h-i} + \varepsilon_{n+h}$$

segue que

- Previsão pontual:

$$\widehat{y}_{n+h|n} = E_n[y_{n+h}] = \ell_n + h\, b_n, \qquad h = 1, 2, 3, \ldots$$

  - Note que essas previões são sempre lineares.

- Variância:

$$\sigma^2_{n+h|n} = Var_n[y_{n+h}] = \left[1 + \sum_{i=1}^{h-1} (\alpha + i\beta)^2\right] \sigma^2_{\varepsilon}$$

Exercício

- Exercício 1: Suponha que as medianas das notas finais dos alunos que cursaram uma determinada disciplina nos ultimos 5 anos foram $4.7; 5.3; 4.6; 5.0; 4.5$. Utilizando $\alpha = 0.6, \beta = 0.2, \ell_0 = y_1, b_0 = 0$, calcule a previsão pontual e o erro quadrático médio de previsão do modelo Holt para os próximos 3 anos. Utilize o estimador usual para a variância do termo aleatório, $\widehat{\sigma}^2_\varepsilon = \frac{1}{n}\sum_{t=1}^n \varepsilon_t^2$.

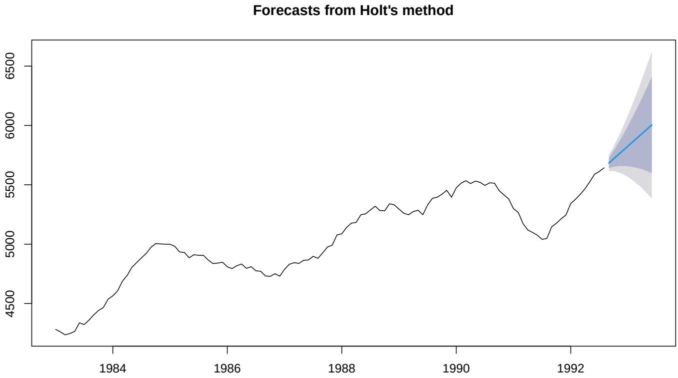
-Resposta:

  - Previsão pontual: 4.631, 4.574, 4.516
  - Variância da previsão: 0.189, 0.310, 0.499

R

- No R, utilize o comando *holt()* da biblioteca *forecast*.
- Exemplo:

```
holt(x, h=10) %>% plot()
```



Modelo SES com drift

Modelo SES com drift

- Um caso especial do modelo de Holt ocorre quanto fixamos $\beta = 0$. Neste caso, a componente de crescimento se torna constante, pois $b_t = b_0$, para todo $t = 1, 2, 3, \ldots$. Desta forma b_0 correspondere a uma taxa de crescimento fixa, geralmente chamada de *drift*.
- O modelo correspondente pode ser escrito como

$$\begin{aligned} y_t &= \ell_{t-1} + b_0 + \varepsilon_t \\ \ell_t &= \ell_{t-1} + b_0 + \alpha\, \varepsilon_t \end{aligned}$$

o qual é conhecido como *SES com drift*.

- Curiosidade:
  - Assimakopoulos & Nikolopoulos publicaram um artigo no *International Journal of Forecasting* descrevendo um algoritmo de previsão que eles chamaram de método Theta, o qual havia vendido a competição de previsão de séries temporais M3 (Makridakis 3), ocorrida no mesmo ano.
  - Em 2003, Hyndman & Billah publicaram um artigo no mesmo periódico, o qual demonstra que o método Theta esta diretamente relacionado ao modelo *SES com drift*.
  - O modelo DOTM (*Dynamic Optimized Theta Model*), proposto em Fiorucci et al (2016), possui a melhor performance já publicado para o banco de dados da M3, este consiste em uma evolução do método Theta. O modelo DOTM esta disponível no pacote *forecTheta* do R.

Section 2

Holt+Damped

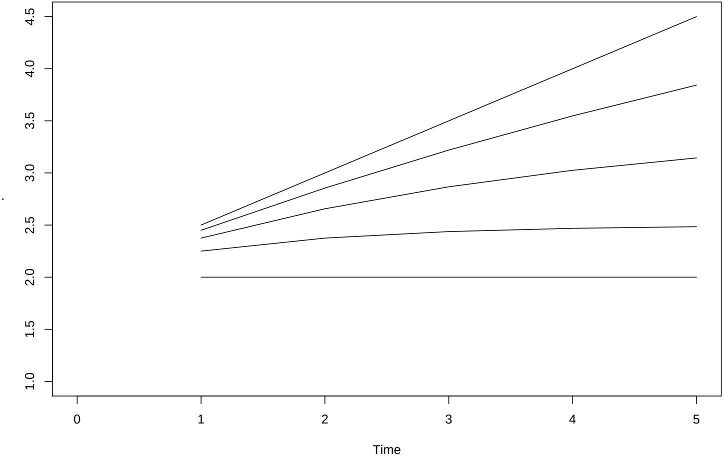
Método com tendência damped

- Proposto por Gardner e Mckenzie (1985)
- Modifica o método linear de Holt para possibilitar previsões com tendência amortizada, chama de "damping"
- Previsões com tendência damping

$$\widehat{y}_{n+h|n} = \ell_n \, + \, (\phi + \phi^2 + \phi^3 + \cdots + \phi^h) \, b_n$$

  em que $\phi \in (0, 1)$ é o parâmetro de damping.
- Note que
  - Se $\phi = 0$, as previsões coincidem com SES (previsões constantes)
  - Se $\phi = 1$, as previsões coincidem com Holt (previsões lineares)
  - Se $0 < \phi < 1$, então $0 < \phi + \phi^2 + \phi^3 + \cdots + \phi^h < h$ para qualquer $h = 1, 2, 3, \ldots$, ou seja, as previsões tem tendência amortizadas.

- Exemplo: considerando $\ell_n = 2$ e $b_n = 0.5$, cada linha da figura abaixo correspende a previsão pontual construida variando ϕ nos seguintes valores 0, 0.5, 0.75, 0.9 e 1.



Método de Holt+Damped

- No método de Holt, essa modificação foi feita multiplicando por ϕ a componente de crescimento

$$\begin{aligned} \text{Previsão:} \quad & \widehat{y}_t = \ell_{t-1} + \phi\, b_{t-1} \\ \text{Level:} \quad & \ell_t = \alpha\, y_t + (1-\alpha)(\ell_{t-1} + \phi\, b_{t-1}) \\ \text{Crescimento:} \quad & b_t = \beta^*\,(\ell_t - \ell_{t-1}) + (1-\beta^*)\,\phi\, b_{t-1} \end{aligned}$$

Note que β^* é um parâmetro de alisamento, e portanto, $\beta^* \in (0,1)$.

Admitindo erros aditivos ($y_t = \mu_t + \varepsilon_t$), e facil ver que o método de Holt com Damped corresponde ao seguinte modelo MEE linear

$$
\begin{aligned}
y_t &= \ell_{t-1} + \phi\, b_{t-1} + \varepsilon_t \\
\ell_t &= \ell_{t-1} + \phi\, b_{t-1} + \alpha\, \varepsilon_t \\
b_t &= \phi\, b_{t-1} + \beta\, \varepsilon_t
\end{aligned}
$$

em que $\beta = \alpha\, \beta^*$. Observe que os parâmetros do modelo são $\ell_0 \in \mathbb{R}, b_0 \in \mathbb{R}, \alpha \in (0,1), \beta \in (0,\alpha)$ e $\phi \in (0,1)$.

- ou equivalentemente, se $\mathbf{x}_t = (\ell_t \ ,\ b_t)'$, podemos reescrever na forma matricial

$$
\begin{aligned}
y_t &= \begin{pmatrix} 1 & \phi \end{pmatrix} \mathbf{x}_t + \varepsilon_t \\
\mathbf{x}_t &= \begin{pmatrix} 1 & \phi \\ 0 & \phi \end{pmatrix} \mathbf{x}_{t-1} + \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \varepsilon_t
\end{aligned}
$$

- Lembrete: MEE linear: $\begin{cases} y_t = & \mathbf{w}'\, \mathbf{x_{t-1}} + \varepsilon_t \\ \mathbf{x_t} = & \mathbf{F}\, \mathbf{x_{t-1}} + \mathbf{g}\, \varepsilon_t \end{cases}$

- Pode-se mostrar que a variância da previsão é dada por

$$\sigma^2_{n+h|n} = \sigma^2_\varepsilon \left\{ 1 + \alpha^2(h-1) + \frac{\beta\phi h}{(1-\phi)^2}[2\phi(1-\phi) + \beta\phi] \right.$$
$$\left. - \frac{\beta\phi(1-\phi^h)}{(1-\phi)^2(1-\phi^2)}[2\alpha(1-\phi^2) + \beta\phi(1+2\phi-\phi^h)] \right\}$$

Referência: Hyndman et al (2008)

Exercícios

- Exercício 2: Mostre que

$$\widehat{y}_{n+h|n} = \ell_n \; + \; (\phi + \phi^2 + \phi^3 + \cdots + \phi^h)\, b_n$$

- Exercício 3: Mostre que

$$\lim_{h \to \infty} \widehat{y}_{n+h|n} = \ell_n + \frac{\phi}{1-\phi} b_n$$

R

- No R, utilize o comando *holt(damped=TRUE)* da biblioteca *forecast*.
- Exemplo:

```
holt(x, h=10, damped=TRUE) %>% plot()
```