Introdução aos modelos de alisamento exponencial

Análise de Séries Temporais

José Augusto Fiorucci

Section 1

Introdução

Introdução

- Objetivo: Previsão de séries temporais
  - Poucas suposições
- Origem: Robert Brown, 1944
  - Funcionário das forças armadas dos EUA
  - O método (SES) foi criado com a intenção de predizer a posição futura de submarinos inimigos
- Principal referência:
  - Hyndman, R., Koehler, A. B., Ord, J. K., & Snyder, R. D. (2008). Forecasting with exponential smoothing: the state space approach. Springer Science & Business Media.

Modelos de espaço de estado

No que segue, será util escrever os modelos de alisamento exponencial como modelos de espaço de estado (MEE).

- MEE linear

$$\begin{aligned} y_t &= \mathbf{w}' \, \mathbf{x_{t-1}} + \varepsilon_t & (1) \\ \mathbf{x_t} &= \mathbf{F} \, \mathbf{x_{t-1}} + \mathbf{g} \, \varepsilon_t & (2) \end{aligned}$$

  - $\mathbf{w}'$, $\mathbf{F}$ e g são coeficientes
  - (1) é chamada de equação de observação (ou medida)
  - (2) é chamada de equação de estado (ou transição)
  - $\mathbf{x_t}$ é o vetor de estados
  - $\{\varepsilon_t\}$ é um ruído branco

- MEE não linear

$$\begin{aligned} y_t &= w(\mathbf{x_{t-1}}) + r(\mathbf{x_{t-1}})\varepsilon_t \\ \mathbf{x_t} &= f(\mathbf{x_{t-1}}) + g(\mathbf{x_{t-1}})\varepsilon_t \end{aligned}$$

  - neste caso, $w(.)$, $r(.)$, $f(.)$ e $g(.)$ são funções.

Notação

A menos que especificado o contrário, vamos assumir:

- $y_1, \ldots, y_n$ como sendo a série temporal observada
- $\mu_t = E[y_t | y_{t-1}, y_{t-2}, \ldots, y_1]$, média condicional
- $\sigma_t^2 = Var[y_t | y_{t-1}, y_{t-2}, \ldots, y_1]$, variância condicional
- $\widehat{y}_t = \mu_t$, previsão 1 passo a frente
- $\widehat{y}_{t+h|t} = E_t[y_{t+h}] = E[y_{t+h} | y_t, y_{t-1}, y_{t-2}, \ldots, y_1]$, previsão h passos a frente de t
- $\sigma_{t+h|t}^2 = Var_t[y_{t+h}] = Var[y_{t+h} | y_t, y_{t-1}, y_{t-2}, \ldots, y_1]$, erro quadrático da previsão de h passos a frente de t

Section 2

Modelo SES(ℓ_0, α)

Alisamento exponencial simples - SES(ℓ_0, α)

Para $0 < \alpha < 1$, o método de alisamento exponencial simples (*simple exponential smoothing*, SES) consiste em escrever a previsão de y_{t+1}, dado a informação prévia, da seguinte forma

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

Note que

- a previsão de y_{t+1} é uma ponderação dos pontos anteriores em que os pesos decaem exponencialmente
- quanto maior o valor de α maior é o peso atribuido as observações mais recentes
- essa previsão considera nem tendência e nem sazonalidade

A equação (3) pode ser reescrita como

$$
\begin{aligned}
\widehat{y}_{t+1} &= \sum_{i=0}^{\infty} \alpha(1-\alpha)^i y_{t-i} \\
&= \alpha\, y_t + \sum_{i=1}^{\infty} \alpha(1-\alpha)^i y_{t-i} \\
&= \alpha\, y_t + (1-\alpha) \sum_{i=1}^{\infty} \alpha(1-\alpha)^{i-1} y_{t-i} \\
&= \alpha\, y_t + (1-\alpha) \sum_{j=0}^{\infty} \alpha(1-\alpha)^{j} y_{t-1-j} \qquad\qquad < j = i-1 >
\end{aligned}
$$

De onde, obtemos a relação

$$
\widehat{y}_{t+1} = \alpha\, y_t + (1-\alpha)\, \widehat{y}_t \tag{4}
$$

Para escrevermos o SES como um modelo de espaço de estados, vamos denotar $\widehat{y}_{t+1}$ por ℓ_t na equação (4), assim obtemos

$$\ell_t = \alpha y_t + (1-\alpha)\ell_{t-1}$$

o qual pode ser computado recursivamente para todo $t = 1, 2, \ldots, n$, desde que ℓ_0 seja conhecido.

- O termo ℓ_t é chamado de level, este é utilizado no contexto de modelos de alisamento exponencial para descrever a previsão pontual em séries sem tendência e sem sazonalide (ou com tendência e sazonalidade ajustadas).
- Das equações anteriores, note que o método SES pode ser estruturado pelo seguinte par de equações

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

- Originalmente a aproximação $\ell_0 = \widehat{y}_1 \approx y_1$ foi proposta para inicializar o processo.
- Abordagens modernas tem tratado $\ell_0 \in \mathbb{R}$ como um parâmetro a ser estimado em conjunto com $\alpha \in (0, 1)$.

Escrevendo o método SES como um MEE

- Admitindo erros aditivos, isto é

$$y_t = \mu_t + \varepsilon_t$$

em que $\mu_t = \widehat{y}_t = E[y_t | y_{t-1}, y_{t-2}, \ldots, y_1]$, o método SES pode ser escrito como um modelo estocástico pelo seguinte par de equações

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

- Reescrevendo a segunda equações, obtemos a representação do modelo SES como um MEE linear

$$y_t = \ell_{t-1} + \varepsilon_t \quad (5)$$
$$\ell_t = \ell_{t-1} + \alpha\,\varepsilon_t \quad (6)$$

  - 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 usando SES

Das equações (5) e (6) note que

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

- Previsão pontual

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

  - Note que essas previsões são sempre constantes.

- Variância - erro quadrático médio de previsão

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

Exercício

- Exercício 1: Suponha que as medianas das notas finais dos alunos que cursaram uma determinada disciplina nos ultimos 7 anos foram $6.1; 5.1; 4.7; 5.3; 4.6; 5.0; 4.5$. Utilizando $\alpha = 0.6$ e $\ell_0 = y_1$, calcule a previsão pontual e o erro quadrático médio de previsão do modelo SES para os próximos 3 anos. Utilize o estimador usual para a variância do termo aleatório, $\widehat{\sigma}_{\varepsilon}^2 = \frac{1}{n} \sum_{t=1}^{n} \varepsilon_t^2$.

-Resposta:

  - Previsão pontual: 4.674, 4.674, 4.674
  - Variância da previsão: 0.326, 0.443, 0.560

R

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

```
ses(x, h=5) %>% plot()
```