Análise de Séries Temporais:

Modelos ETS

Prof. Dr. José Augusto Fiorucci

Universidade de Brasília

Departamento de Estatística

Modelos ETS

- Error, Trend and Seasonal (ETS);
- Estrutura que permite descrever os modelos de alisamento exponencial em função dos tipos de suas componentes de erro, tendência e sazonalidade.
- Referências:

  - Hyndman, R. J., Koehler, A. B., Snyder, R. D., Grose, S. (2002). A state space framework for automatic forecasting using exponential smoothing methods. International Journal of forecasting, 18(3), 439-454.
  - Hyndman, R., Koehler, A. B., Ord, J. K., Snyder, R. D. (2008). Forecasting with exponential smoothing: the state space approach. Springer Science Business Media.

Tipos de decomposições

- Aditiva:

$$Y = T + S + E$$

- Multiplicativa:

$$Y = T \times S \times E$$

- Mista:

$$Y = (T + S) \times E$$

ou

$$Y = T \times S + E$$

Sendo E, T e S as componentes de erro, tendência e sazonalidade, respectivamente;

Classificação dos modelos de alisamento exponencial

Obs: Inicialmente focaremos apenas nas componentes de tendência e sazonalidade;

- Tendências: (5 tipos)

$$
\begin{aligned}
\text{None:} \quad & T_h = \ell \\
\text{Additive:} \quad & T_h = \ell + bh \\
\text{Additive damped:} \quad & T_h = \ell + (\phi + \phi^2 + \cdots + \phi^h)b \\
\text{Multiplicative:} \quad & T_h = \ell b^h \\
\text{Multiplicative damped:} \quad & T_h = \ell b^{(\phi + \phi^2 + \cdots + \phi^h)}
\end{aligned}
$$

- Sazonalidade: (3 tipos) Nenhuma, Aditiva e Multiplicativa

Classificação dos modelos de alisamento exponencial

- Combinações de tendência e sazonalidade

Trend component	Seasonal component		
	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

Métodos resultantes

Trend	Seasonal: N	Seasonal: A	Seasonal: M
N	$\ell_t = \alpha y_t + (1-\alpha)\ell_{t-1}$ $\hat{y}_{t+h\|t} = \ell_t$	$\ell_t = \alpha(y_t - s_{t-m}) + (1-\alpha)\ell_{t-1}$ $s_t = \gamma(y_t - \ell_{t-1}) + (1-\gamma)s_{t-m}$ $\hat{y}_{t+h\|t} = \ell_t + s_{t-m+h_m^+}$	$\ell_t = \alpha(y_t / s_{t-m}) + (1-\alpha)\ell_{t-1}$ $s_t = \gamma(y_t / \ell_{t-1}) + (1-\gamma)s_{t-m}$ $\hat{y}_{t+h\|t} = \ell_t s_{t-m+h_m^+}$
A	$\ell_t = \alpha y_t + (1-\alpha)(\ell_{t-1} + b_{t-1})$ $b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)b_{t-1}$ $\hat{y}_{t+h\|t} = \ell_t + hb_t$	$\ell_t = \alpha(y_t - s_{t-m}) + (1-\alpha)(\ell_{t-1} + b_{t-1})$ $b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)b_{t-1}$ $s_t = \gamma(y_t - \ell_{t-1} - b_{t-1}) + (1-\gamma)s_{t-m}$ $\hat{y}_{t+h\|t} = \ell_t + hb_t + s_{t-m+h_m^+}$	$\ell_t = \alpha(y_t / s_{t-m}) + (1-\alpha)(\ell_{t-1} + b_{t-1})$ $b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)b_{t-1}$ $s_t = \gamma(y_t / (\ell_{t-1} + b_{t-1})) + (1-\gamma)s_{t-m}$ $\hat{y}_{t+h\|t} = (\ell_t + hb_t)s_{t-m+h_m^+}$
$\mathrm{A_d}$	$\ell_t = \alpha y_t + (1-\alpha)(\ell_{t-1} + \phi b_{t-1})$ $b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)\phi b_{t-1}$ $\hat{y}_{t+h\|t} = \ell_t + \phi_h b_t$	$\ell_t = \alpha(y_t - s_{t-m}) + (1-\alpha)(\ell_{t-1} + \phi b_{t-1})$ $b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)\phi b_{t-1}$ $s_t = \gamma(y_t - \ell_{t-1} - \phi b_{t-1}) + (1-\gamma)s_{t-m}$ $\hat{y}_{t+h\|t} = \ell_t + \phi_h b_t + s_{t-m+h_m^+}$	$\ell_t = \alpha(y_t / s_{t-m}) + (1-\alpha)(\ell_{t-1} + \phi b_{t-1})$ $b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)\phi b_{t-1}$ $s_t = \gamma(y_t / (\ell_{t-1} + \phi b_{t-1})) + (1-\gamma)s_{t-m}$ $\hat{y}_{t+h\|t} = (\ell_t + \phi_h b_t)s_{t-m+h_m^+}$
M	$\ell_t = \alpha y_t + (1-\alpha)\ell_{t-1}b_{t-1}$ $b_t = \beta^*(\ell_t / \ell_{t-1}) + (1-\beta^*)b_{t-1}$ $\hat{y}_{t+h\|t} = \ell_t b_t^h$	$\ell_t = \alpha(y_t - s_{t-m}) + (1-\alpha)\ell_{t-1}b_{t-1}$ $b_t = \beta^*(\ell_t / \ell_{t-1}) + (1-\beta^*)b_{t-1}$ $s_t = \gamma(y_t - \ell_{t-1}b_{t-1}) + (1-\gamma)s_{t-m}$ $\hat{y}_{t+h\|t} = \ell_t b_t^h + s_{t-m+h_m^+}$	$\ell_t = \alpha(y_t / s_{t-m}) + (1-\alpha)\ell_{t-1}b_{t-1}$ $b_t = \beta^*(\ell_t / \ell_{t-1}) + (1-\beta^*)b_{t-1}$ $s_t = \gamma(y_t / (\ell_{t-1}b_{t-1})) + (1-\gamma)s_{t-m}$ $\hat{y}_{t+h\|t} = \ell_t b_t^h s_{t-m+h_m^+}$
$\mathrm{M_d}$	$\ell_t = \alpha y_t + (1-\alpha)\ell_{t-1}b_{t-1}^{\phi}$ $b_t = \beta^*(\ell_t / \ell_{t-1}) + (1-\beta^*)b_{t-1}^{\phi}$ $\hat{y}_{t+h\|t} = \ell_t b_t^{\phi_h}$	$\ell_t = \alpha(y_t - s_{t-m}) + (1-\alpha)\ell_{t-1}b_{t-1}^{\phi}$ $b_t = \beta^*(\ell_t / \ell_{t-1}) + (1-\beta^*)b_{t-1}^{\phi}$ $s_t = \gamma(y_t - \ell_{t-1}b_{t-1}^{\phi}) + (1-\gamma)s_{t-m}$ $\hat{y}_{t+h\|t} = \ell_t b_t^{\phi_h} + s_{t-m+h_m^+}$	$\ell_t = \alpha(y_t / s_{t-m}) + (1-\alpha)\ell_{t-1}b_{t-1}^{\phi}$ $b_t = \beta^*(\ell_t / \ell_{t-1}) + (1-\beta^*)b_{t-1}^{\phi}$ $s_t = \gamma(y_t / (\ell_{t-1}b_{t-1}^{\phi})) + (1-\gamma)s_{t-m}$ $\hat{y}_{t+h\|t} = \ell_t b_t^{\phi_h} s_{t-m+h_m^+}$

In each case, ℓ_t denotes the series level at time t, b_t denotes the slope at time t, s_t denotes the seasonal component of the series at time t, and m denotes the number of seasons in a year; α, β^*, γ and ϕ are constants, $\phi_h = \phi + \phi^2 + \cdots + \phi^h$ and $h_m^+ = \lfloor (h-1) \bmod m \rfloor + 1$.

Transformando os métodos em modelos estocásticos

- Em cada combinação de tendência e sazonalidade da tabela anterior, o termo de erro pode ser incluído de duas maneiras:

$$\text{Aditivo: } y_t = \mu_t + \varepsilon_t$$
$$\text{Multiplicativo: } y_t = \mu_t \times (1 + \varepsilon_t)$$

sendo $\mu_t = \hat{y}_{t|t-1} = E[y_t | y_{t-1}, y_{t-2}, \ldots]$ e $\{\varepsilon_t\}$ um ruído branco;

- A previsão pontual não é diretamente afetada pela forma como os erros são inseridos, pois,

$$\hat{y}_{n+h|n} = E_n[\mu_{n+h} + \varepsilon_{n+h}] = E_n[\mu_{n+h} \times (1 + \varepsilon_{n+h})] = \mu_{n+h|n}$$

- No entanto, a forma como os erros são incluídos pode levar a diferente estimativas de parâmetros;

Transformando os métodos em modelos estocásticos

- Nas aulas anteriores vimos que nos métodos de SES, Holt, Damped e Holt-Winters, a inclusão de erros aditivos possibilita a representação do modelo resultante em modelo de espaço de estado (MEE). Na prática, o mesmo pode ser feito quando erros multiplicativos são considerados;
- Nos dois slides a seguir são apresentados os modelos estocásticos resultantes. Estes já estão escritos na forma de modelos de espaço de estado;
- Em todos os casos, por simplificação, considere $\beta = \alpha\,\beta^*$;

Erro aditivo: modelos resultantes

Trend	Seasonal		
	N	A	M
N	$\mu_t = \ell_{t-1}$ $\ell_t = \ell_{t-1} + \alpha\varepsilon_t$	$\mu_t = \ell_{t-1} + s_{t-m}$ $\ell_t = \ell_{t-1} + \alpha\varepsilon_t$ $s_t = s_{t-m} + \gamma\varepsilon_t$	$\mu_t = \ell_{t-1}s_{t-m}$ $\ell_t = \ell_{t-1} + \alpha\varepsilon_t/s_{t-m}$ $s_t = s_{t-m} + \gamma\varepsilon_t/\ell_{t-1}$
A	$\mu_t = \ell_{t-1} + b_{t-1}$ $\ell_t = \ell_{t-1} + b_{t-1} + \alpha\varepsilon_t$ $b_t = b_{t-1} + \beta\varepsilon_t$	$\mu_t = \ell_{t-1} + b_{t-1} + s_{t-m}$ $\ell_t = \ell_{t-1} + b_{t-1} + \alpha\varepsilon_t$ $b_t = b_{t-1} + \beta\varepsilon_t$ $s_t = s_{t-m} + \gamma\varepsilon_t$	$\mu_t = (\ell_{t-1} + b_{t-1})s_{t-m}$ $\ell_t = \ell_{t-1} + b_{t-1} + \alpha\varepsilon_t/s_{t-m}$ $b_t = b_{t-1} + \beta\varepsilon_t/s_{t-m}$ $s_t = s_{t-m} + \gamma\varepsilon_t/(\ell_{t-1} + b_{t-1})$
$\mathrm{A_d}$	$\mu_t = \ell_{t-1} + \phi b_{t-1}$ $\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha\varepsilon_t$ $b_t = \phi b_{t-1} + \beta\varepsilon_t$	$\mu_t = \ell_{t-1} + \phi b_{t-1} + s_{t-m}$ $\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha\varepsilon_t$ $b_t = \phi b_{t-1} + \beta\varepsilon_t$ $s_t = s_{t-m} + \gamma\varepsilon_t$	$\mu_t = (\ell_{t-1} + \phi b_{t-1})s_{t-m}$ $\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha\varepsilon_t/s_{t-m}$ $b_t = \phi b_{t-1} + \beta\varepsilon_t/s_{t-m}$ $s_t = s_{t-m} + \gamma\varepsilon_t/(\ell_{t-1} + \phi b_{t-1})$
M	$\mu_t = \ell_{t-1}b_{t-1}$ $\ell_t = \ell_{t-1}b_{t-1} + \alpha\varepsilon_t$ $b_t = b_{t-1} + \beta\varepsilon_t/\ell_{t-1}$	$\mu_t = \ell_{t-1}b_{t-1} + s_{t-m}$ $\ell_t = \ell_{t-1}b_{t-1} + \alpha\varepsilon_t$ $b_t = b_{t-1} + \beta\varepsilon_t/\ell_{t-1}$ $s_t = s_{t-m} + \gamma\varepsilon_t$	$\mu_t = \ell_{t-1}b_{t-1}s_{t-m}$ $\ell_t = \ell_{t-1}b_{t-1} + \alpha\varepsilon_t/s_{t-m}$ $b_t = b_{t-1} + \beta\varepsilon_t/(s_{t-m}\ell_{t-1})$ $s_t = s_{t-m} + \gamma\varepsilon_t/(\ell_{t-1}b_{t-1})$
$\mathrm{M_d}$	$\mu_t = \ell_{t-1}b_{t-1}^{\phi}$ $\ell_t = \ell_{t-1}b_{t-1}^{\phi} + \alpha\varepsilon_t$ $b_t = b_{t-1}^{\phi} + \beta\varepsilon_t/\ell_{t-1}$	$\mu_t = \ell_{t-1}b_{t-1}^{\phi} + s_{t-m}$ $\ell_t = \ell_{t-1}b_{t-1}^{\phi} + \alpha\varepsilon_t$ $b_t = b_{t-1}^{\phi} + \beta\varepsilon_t/\ell_{t-1}$ $s_t = s_{t-m} + \gamma\varepsilon_t$	$\mu_t = \ell_{t-1}b_{t-1}^{\phi}s_{t-m}$ $\ell_t = \ell_{t-1}b_{t-1}^{\phi} + \alpha\varepsilon_t/s_{t-m}$ $b_t = b_{t-1}^{\phi} + \beta\varepsilon_t/(s_{t-m}\ell_{t-1})$ $s_t = s_{t-m} + \gamma\varepsilon_t/(\ell_{t-1}b_{t-1}^{\phi})$

Erro multiplicativo: modelos resultantes

Trend	Seasonal		
	N	A	M
N	$\mu_t = \ell_{t-1}$ $\ell_t = \ell_{t-1}(1+\alpha\varepsilon_t)$	$\mu_t = \ell_{t-1} + s_{t-m}$ $\ell_t = \ell_{t-1} + \alpha(\ell_{t-1} + s_{t-m})\varepsilon_t$ $s_t = s_{t-m} + \gamma(\ell_{t-1} + s_{t-m})\varepsilon_t$	$\mu_t = \ell_{t-1}s_{t-m}$ $\ell_t = \ell_{t-1}(1+\alpha\varepsilon_t)$ $s_t = s_{t-m}(1+\gamma\varepsilon_t)$
A	$\mu_t = \ell_{t-1} + b_{t-1}$ $\ell_t = (\ell_{t-1} + b_{t-1})(1+\alpha\varepsilon_t)$ $b_t = b_{t-1} + \beta(\ell_{t-1} + b_{t-1})\varepsilon_t$	$\mu_t = \ell_{t-1} + b_{t-1} + s_{t-m}$ $\ell_t = \ell_{t-1} + b_{t-1} + \alpha(\ell_{t-1} + b_{t-1} + s_{t-m})\varepsilon_t$ $b_t = b_{t-1} + \beta(\ell_{t-1} + b_{t-1} + s_{t-m})\varepsilon_t$ $s_t = s_{t-m} + \gamma(\ell_{t-1} + b_{t-1} + s_{t-m})\varepsilon_t$	$\mu_t = (\ell_{t-1} + b_{t-1})s_{t-m}$ $\ell_t = (\ell_{t-1} + b_{t-1})(1+\alpha\varepsilon_t)$ $b_t = b_{t-1} + \beta(\ell_{t-1} + b_{t-1})\varepsilon_t$ $s_t = s_{t-m}(1+\gamma\varepsilon_t)$
A_d	$\mu_t = \ell_{t-1} + \phi b_{t-1}$ $\ell_t = (\ell_{t-1} + \phi b_{t-1})(1+\alpha\varepsilon_t)$ $b_t = \phi b_{t-1} + \beta(\ell_{t-1} + \phi b_{t-1})\varepsilon_t$	$\mu_t = \ell_{t-1} + \phi b_{t-1} + s_{t-m}$ $\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha(\ell_{t-1} + \phi b_{t-1} + s_{t-m})\varepsilon_t$ $b_t = \phi b_{t-1} + \beta(\ell_{t-1} + \phi b_{t-1} + s_{t-m})\varepsilon_t$ $s_t = s_{t-m} + \gamma(\ell_{t-1} + \phi b_{t-1} + s_{t-m})\varepsilon_t$	$\mu_t = (\ell_{t-1} + \phi b_{t-1})s_{t-m}$ $\ell_t = (\ell_{t-1} + \phi b_{t-1})(1+\alpha\varepsilon_t)$ $b_t = \phi b_{t-1} + \beta(\ell_{t-1} + \phi b_{t-1})\varepsilon_t$ $s_t = s_{t-m}(1+\gamma\varepsilon_t)$
M	$\mu_t = \ell_{t-1}b_{t-1}$ $\ell_t = \ell_{t-1}b_{t-1}(1+\alpha\varepsilon_t)$ $b_t = b_{t-1}(1+\beta\varepsilon_t)$	$\mu_t = \ell_{t-1}b_{t-1} + s_{t-m}$ $\ell_t = \ell_{t-1}b_{t-1} + \alpha(\ell_{t-1}b_{t-1} + s_{t-m})\varepsilon_t$ $b_t = b_{t-1} + \beta(\ell_{t-1}b_{t-1} + s_{t-m})\varepsilon_t/\ell_{t-1}$ $s_t = s_{t-m} + \gamma(\ell_{t-1}b_{t-1} + s_{t-m})\varepsilon_t$	$\mu_t = \ell_{t-1}b_{t-1}s_{t-m}$ $\ell_t = \ell_{t-1}b_{t-1}(1+\alpha\varepsilon_t)$ $b_t = b_{t-1}(1+\beta\varepsilon_t)$ $s_t = s_{t-m}(1+\gamma\varepsilon_t)$
M_d	$\mu_t = \ell_{t-1}b_{t-1}^{\phi}$ $\ell_t = \ell_{t-1}b_{t-1}^{\phi}(1+\alpha\varepsilon_t)$ $b_t = b_{t-1}^{\phi}(1+\beta\varepsilon_t)$	$\mu_t = \ell_{t-1}b_{t-1}^{\phi} + s_{t-m}$ $\ell_t = \ell_{t-1}b_{t-1}^{\phi} + \alpha(\ell_{t-1}b_{t-1}^{\phi} + s_{t-m})\varepsilon_t$ $b_t = b_{t-1}^{\phi} + \beta(\ell_{t-1}b_{t-1}^{\phi} + s_{t-m})\varepsilon_t/\ell_{t-1}$ $s_t = s_{t-m} + \gamma(\ell_{t-1}b_{t-1}^{\phi} + s_{t-m})\varepsilon_t$	$\mu_t = \ell_{t-1}b_{t-1}^{\phi}s_{t-m}$ $\ell_t = \ell_{t-1}b_{t-1}^{\phi}(1+\alpha\varepsilon_t)$ $b_t = b_{t-1}^{\phi}(1+\beta\varepsilon_t)$ $s_t = s_{t-m}(1+\gamma\varepsilon_t)$

Identificação dos métodos clássicos como modelos ETS

- Alisamento Exponencial Simples (SES)

  Classificação: ETS(A,N,N)

- Holt

  Classificação: ETS(A,A,N)

- Holt+Damped

  Classificação: ETS(A,Ad,N)

- Holt-Winters Aditivo

  Classificação: ETS(A,A,A)

- Holt-Winters Multiplicativo

  Classificação: ETS(A,A,M)

- Exercício 1:

O modelo ETS(A,N,A) foi ajustado em uma série mensal, a tabela abaixo apresenta as datas, observações e os valores das componentes com $\alpha = 0.7$ e $\gamma = 0.1$

	y	ell	s
Sep 1977	8314	8438.04	-110.08
Oct 1977	8850	8563.66	250.46
Nov 1977	8265	8529.12	-254.26
Dec 1977	8796	8746.09	-12.08
Jan 1978	7836	8662.85	-803.07
Feb 1978	6892	8503.66	-1566.18
Mar 1978	7791	8553.65	-776.93

	y	ell	s
Apr 1978	8192	8694.51	-542.75
May 1978	9115	8790.32	297.31
Jun 1978	9434	8686.38	777.32
Jul 1978	10484	8782.77	1673.69
Aug 1978	9827	8866.88	936.09
Sep 1978	9110	9114.12	-74.76
Oct 1978	9070	8907.92	221.00

A) Sabendo que as próximas 2 observações da série foram 8633 e 9240, calcule os valores das componetes de level e sazonalidade.

- Resposta:

	y	ell	s
Nov 1978	8633	8893.46	-256.32
Dec 1978	9240	9144.49	23.79

B) Calcule as previsões pontuais até o horizonte $h = 5$.

- Resposta:

	Jan	Feb	Mar	Apr	May
1979	8341.416	7578.313	8367.563	8601.737	9441.796

- Exercício 2:

O modelo ETS(M,Ad,N) foi ajustado em uma série trimestral, a tabela abaixo apresenta as datas, observações e os valores das componentes com $\alpha = 0.8, \beta = 0.2$ e $\phi = 0.9$

		y	ell	b
1991	Q4	4959	4882.98	5.72
1992	Q1	4987	4967.22	24.92
1992	Q2	4998	4996.33	24.10
1992	Q3	5510	5411.60	120.08
1992	Q4	5212	5273.54	46.54
1993	Q1	4939	5014.28	-33.40

A) Sabendo que as próximas 2 observações da série foram 5002 e 4820, calcule os valores das componetes de level e crescimento.

- Resposta:

		y	ell	b
1993	Q2	5002	4998.45	-26.50
1993	Q3	4820	4850.92	-54.77

B) Calcule as previsões pontuais até o horizonte $h = 5$.

- Resposta:

	Qtr1	Qtr2	Qtr3	Qtr4
1993				4801.624
1994	4757.259	4717.330	4681.395	4649.053

Modelos de Espaço de Estado

- Seja $\mathbf{x}_t = (\ell_t, b_t, s_t, s_{t-1}, \ldots, s_{t-m+1})'$ e $\varepsilon_t \sim^{iid} N(0, \sigma^2)$. Assim todos os modelos podem ser escritos como modelos espaço de estado não linear

$$\begin{aligned} y_t &= \underbrace{h(\mathbf{x}_{t-1})}_{\mu_t} + k(\mathbf{x}_{t-1})\varepsilon_t \\ \mathbf{x}_t &= f(\mathbf{x}_{t-1}) + g(\mathbf{x}_{t-1})\varepsilon_t \end{aligned}$$

- Erros aditivos

$$y_t = \mu_t + \varepsilon_t \quad \Longrightarrow \quad k(\mathbf{x}_{t-1}) = 1$$

Resíduos: $\varepsilon_t = y_t - \mu_t = y_t - \hat{y}_t$

- Erros multiplicativos

$$y_t = \mu_t(1 + \varepsilon_t) \quad \Longrightarrow \quad k(\mathbf{x}_{t-1}) = \mu_t$$

Resíduos: $\varepsilon_t = (y_t - \mu_t)/\mu_t = (y_t - \hat{y}_t)/\hat{y}_t$ (erros relativos)

Estimação de modelos ETS

- Parâmetros de alisamento: $\boldsymbol{\theta} = (\alpha, \beta, \gamma, \phi)'$;
- Inicialização dos estados: $\mathbf{x}_0 = (\ell_0, b_0, s_0, s_{-1}, \ldots, s_{-m+1})'$;
- A função de verossimilhança é escrita como

$$L(\boldsymbol{\theta}, \mathbf{x}_0) = p(y_1) \prod_{t=2}^{n} p(y_t | y_{t-1}, \ldots, y_1)$$

  em que $p(y_t | y_{t-1}, \ldots, y_1)$ é a densidade da distribuição $N(\mu_t\,,\ k(\mathbf{x}_{t-1})^2 \sigma^2)$.

- Pode-se mostrar que

$$\begin{aligned} L^*(\boldsymbol{\theta}, \mathbf{x}_0) &= n \log \left(\sum_{t=1}^{n} \varepsilon_t^2 / k^2(\mathbf{x}_{t-1}) \right) + 2 \sum_{t=1}^{n} \log |k^2(\mathbf{x}_{t-1})| \\ &= -2 \log(\, L(\boldsymbol{\theta}, \mathbf{x}_0)\,) + constant \end{aligned}$$

- Assim $\boldsymbol{\theta}$ e $\mathbf{x}_0$ são estimados de forma a minimizar L^*.

Estimação de modelos ETS

- Em modelos com erros aditivos, o estimador de máxima verossimilhança (EMV) coincide com o de mínimos quadrados (EMQ);
- Em modelos com erros multiplicativos, o EMV e o EMQ não coincidem;

Previsão intervalar: método paramétrico analítico

- Sob a suposição de normalidade, a previsão intervalar com $100(1-\alpha)\%$ são dadas por

$$\hat{y}_{n+h|n} \pm z_{(1-\alpha/2)}\,\sigma_{n+h|n}$$

- A dificuldade se resume em calcular $\sigma^2_{n+h|n} = Var_n[y_{n+h}]$ analiticamente;
- Nas aulas anteriores derivamos as equações para os modelos SES, Holt e HW Aditivo;
- Para a maioria dos modelos aditivos é possível obter a expressão analiticamente, veja Hyndman et. al. (2008);
- Desvantagens:
  - Não aplicável quando $\sigma^2_{n+h|n}$ não pode ser obtido analiticamente;
  - Depende da verificação de normalidade dos resíduos;

Previsão intervalar: método bootstrap

- Todos os modelos podem ser escritos como MEE (não linear)

$$
\begin{aligned}
y_t &= h(\mathbf{x}_{t-1}) + k(\mathbf{x}_{t-1})\varepsilon_t \\
\mathbf{x}_t &= f(\mathbf{x}_{t-1}) + g(\mathbf{x}_{t-1})\varepsilon_t
\end{aligned}
$$

- Diversos possíveis caminhos para o processo são simulados, digamos N caminhos, dados por:

$$
\begin{array}{cccc}
y_{n+1}^{(1)}, & y_{n+2}^{(1)}, & \ldots, & y_{n+h}^{(1)} \\
y_{n+1}^{(2)}, & y_{n+2}^{(2)}, & \ldots, & y_{n+h}^{(2)} \\
\vdots \quad , & \vdots \quad , & \ldots, & \vdots \\
y_{n+1}^{(N)}, & y_{n+2}^{(N)}, & \ldots, & y_{n+h}^{(N)}
\end{array}
$$

- Previsões intervalares para y_{n+h} são obtidas a partir do quantis empíricos de $y_{n+h}^{(1)}, \ldots, y_{n+h}^{(N)}$;

- **Bootstrap paramétrico:** A simulação é feita admitindo que o termo de erro tem distribuição normal (ou outra distribuição conhecida);

- **Bootstrap não paramétrico:** A simulação é feita a partir da distribuição empírica dos resíduos;

- OBS: Vale notar que a variância dos estimadores não é considerada por essas abordagens, de modo que o intervalo de previsão tende a ser mais fechado que o ideal. Alternativa: "bagging"
  - Bergmeir, C., Hyndman, R. J., Benítez, J. M. (2016). Bagging exponential smoothing methods using STL decomposition and Box-Cox transformation. International Journal of Forecasting, 32(2), 303–312. https://robjhyndman.com/publications/bagging-ets/

Critérios Seleção de modelos

- Critério de informação de Akaike

$$AIC = -2\log(L) + 2\,k$$

onde L é a função de verossimilhança e k o número de parâmetros e estados iniciais do modelo;

- AIC corrigido

$$AIC_c = AIC + \frac{2(k+1)(k+2)}{n-k}$$

corrigido para séries pequenas;

- Critério de informação Bayesiano

$$BIC = -2\log(L) + k\,\log(n)$$

Algoritmo automático

- Proposto em Hyndman et al. (2002)
- Algoritmo

  Passo 1: Aplique cada modelo que é apropriado a séries;

  Passo 2: Selecione o melhor usando AICc;

  Passo 3: Produza previsões pontuais utilizando o melhor método;

  Passo 4: Obtenha previsões intervalares utilizando o modelo de espaço de estado equivalente;

- Não considera nenhum tipo de análise sobre os resíduos (o modelo menos ruim é escolhido);
- Apresentou boa performance no conjunto de dados da competição M3;

Algoritmo automático

- Esta abordagem está implementada no pacote *forecast* do R na função ets();

ets(y, model = "ZZZ", damped = NULL, alpha = NULL, beta = NULL, gamma = NULL, phi = NULL, allow.multiplicative.trend = FALSE)

- Alguns exemplos:
  - SES via ETS: ets(y, model = "ANN", damped=FALSE)
  - Holt via ETS: ets(y, model = "AAN", damped=FALSE)
  - Holt+Damped via ETS: ets(y, model = "AAN", damped=TRUE)
  - Holt Winter Aditivo: ets(y, model = "AAA", damped=FALSE)
  - Holt Winter Multip.: ets(y, model = "MAM", damped=FALSE)
- OBS: Se model não é especificado então é feita a seleção automática;

Alguns modelos instáveis

- Algumas combinações de erro, tendência e sazonalidade podem apresentar problemas numéricos, em especial aquelas que misturam erro aditivo com sazonalidade multiplicativa: ETS(A, N, M), ETS(A, A, M), ETS(A, A_d, M);
- Estas são proibidas na função ets() do pacote *forecast*;
- Modelos com erros multiplicativos são aplicáveis apenas para séries estritamente positivas;

Exercício

- Verifique o comportamento do algoritmo automático na seguintes séries do pacote *fpp2*: bicoal, chicken, dole, usdeaths, bricksq, ausbeer