

MAE 0325 - aula 04

Tendência e Sazonalidade

$\{Z_t, t = 1, \dots, N\}$: observações de uma série temporal.

Um modelo de decomposição consiste em escrever Z_t como uma soma de 3 componentes não-observáveis:

$$Z_t = T_t + S_t + a_t, \quad (1)$$

T_t : tendência;

S_t : sazonalidade;

a_t : componente aleatória, de média zero e variância constante σ_a^2 .

Se $\{a_t\}$ for um ruído branco, então

$$E(a_t a_s) = 0, \quad s \neq t;$$

ou

tomando $\{a_t\}$ como um processo estacionário.

$\{Z_t\}$ será uma série não estacionária.

Interesse:

estimar S_t e construir a série livre de sazonalidade, ou sazonalmente ajustada:

$$Z_t^{SA} = Z_t - \hat{S}_t$$

As componentes T_t e S_t são, em geral, bastante relacionadas e a influência da tendência sobre a componente sazonal pode ser muito forte, por duas razões:

- a) métodos de estimação de S_t podem ser bastante afetados se não levarmos em conta a tendência;
- b) a especificação de S_t depende da especificação de T_t .

Por isso, não poderemos isolar uma das componentes sem tentar isolar a outra. Estimando-se T_t e S_t e subtraindo de Z_t obteremos uma estimativa da componente aleatória a_t .

Tendência

Inicialmente vamos supor que a componente sazonal S_t não esteja presente. O modelo:

$$Z_t = T_t + a_t,$$

a_t : ruído branco, de **média zero** e **variância constante** σ_a^2

Há vários métodos para estimar T_t . Os mais utilizados:

- a) **ajustar uma função do tempo, como um polinômio, uma exponencial ou outra função suave de t ;**
- b) **suavizar (ou filtrar) os valores da série ao redor de um ponto, para estimar a tendência naquele ponto;**
- c) **suavizar os valores da série através de sucessivos ajustes de retas de mínimos quadrados ponderados ("lowess").**

Estimando-se a tendência por meio de $\hat{T}_t$, podemos obter a série ajustada para tendência ou livre de tendência,

$$Y_t = Z_t - \hat{T}_t.$$

Um procedimento que é também utilizado para eliminar a tendência de uma série é aquele de tomar diferenças. Normalmente para séries econômicas, por exemplo, a primeira diferença já é estacionária:

$$\Delta Z_t = Z_t - Z_{t-1}$$

Tendência polinomial

$$T_t = \beta_0 + \beta_1 t + \cdots + \beta_m t^m, \quad (3.4)$$

m: grau do polinômio. $m \ll N$.

estimar β_j usando o métodos de MQ:

$$f(\beta_0, \dots, \beta_m) = \sum_{t=1}^N (Z_t - \beta_0 - \beta_1 t - \cdots - \beta_m t^m)^2, \quad (3.5)$$

Exemplo 3.1. Na Tabela 3.1, apresentamos parte dos dados da série Energia. São 24 observações, referentes aos anos 1977 e 1978 e arredondadas. Notamos que, *para este período*, um polinômio de primeiro grau é adequado para representar T_t .

Tabela 3.1: Série Energia - Consumo de Energia Elétrica no Espírito Santo, jan./1977 a dez./1978

t	Z_t	t	Z_t
1	84,6	13	100,3
2	89,9	14	118,1
3	81,9	15	116,5
4	95,4	16	134,2
5	91,2	17	134,7
6	89,8	18	144,8
7	89,7	19	144,4
8	97,9	20	159,2
9	103,4	21	168,2
10	107,6	22	175,2
11	120,4	23	174,5
12	109,6	24	173,7

Fonte: Série Energia

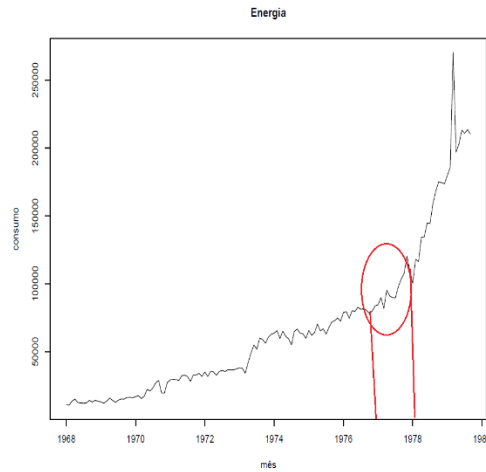


Figura 1.11: Série de consumo de energia elétrica no Espírito Santo.

O modelo (3.4) reduz-se a

$$Z_t = \beta_0 + \beta_1 t + a_t \quad (3.6)$$

e minimizando a soma dos quadrados dos resíduos obtemos

$$\hat{\beta}_0 = \bar{Z} - \hat{\beta}_1 \bar{t}, \quad (3.7)$$

$$\hat{\beta}_1 = \frac{\sum_{t=1}^N t Z_t - \left[\left(\sum_{t=1}^N t \right) \left(\sum_{t=1}^N Z_t \right) \right] / N}{\sum_{t=1}^N t^2 - \left(\sum_{t=1}^N t \right)^2 / N} \quad (3.8)$$

sendo

$$\bar{Z} = \frac{1}{N} \sum_{t=1}^N Z_t$$

a média amostral das $N = 24$ observações e

$$\bar{t} = \frac{1}{N} \sum_{t=1}^N t.$$

Levando em conta que

$$\sum_{t=1}^N t = \frac{N(N+1)}{2} = 300, \quad \sum_{t=1}^N Z_t = 2.905,2, \quad \bar{Z} = 121,05, \quad \bar{t} = 12,5,$$

$$\sum_{t=1}^N t^2 = \frac{N(N+1)(2N+1)}{6} = 4.900, \quad \sum_{t=1}^N t Z_t = 41.188,6,$$

obtemos

$$\hat{\beta}_0 = 68,076, \quad \hat{\beta}_1 = 4,238.$$

Logo, um estimador de T_t é

$$\hat{T}_t = 68,076 + 4,238t. \quad (3.9)$$

Utilizando o modelo (3.9) podemos prever valores futuros da série. Na Tabela 3.2, temos os valores reais e previstos para janeiro, fevereiro, março e abril de 1979. Observe que o valor para março de 1979 é bastante atípico e o erro de previsão, neste caso, é grande.

De modo geral, o valor previsto h passos à frente, dadas as observações até o instante $t = N$, é $\hat{Z}_N(h)$, e o erro de previsão correspondente é

$$e_N(h) = Z_{N+h} - \hat{Z}_N(h). \quad (3.10)$$

Por sua vez,

$$\hat{Z}_N(h) = \hat{T}_{N+h}, \quad (3.11)$$

para $h = 1, 2, 3, \dots$

Previsão:

Tabela 3.2: Valores reais e previstos para a Série Energia, janeiro a março de 1979.

h	Z_t	$\hat{Z}_N(h)$	erro de previsão ($e_N(h)$)
1	179,8	174,0	5,8
2	185,8	178,3	7,5
3	270,3	182,5	87,8
4	196,9	186,7	10,2

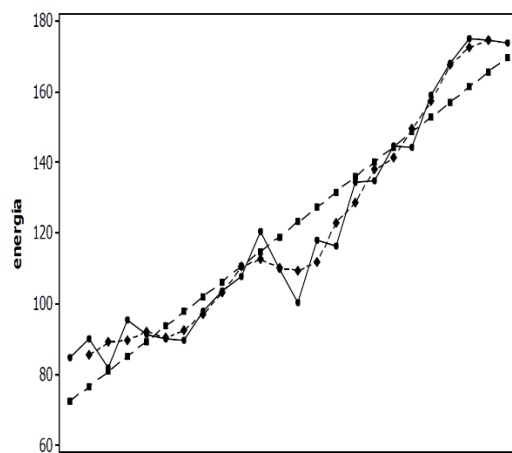


Figura 3.1: Gráfico de Z_t (linha cheia), Z_t^* (linha pontilhada) e $\hat{T}_t$ (linha tracejada).