Especialização em *Data Science* e Estatística Aplicada

Módulo IV - Análise de Séries Temporais

Prof. Dr. Eder Angelo Milani

Goiânia, 2025

IME
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA

FEN
FACULDADE DE ENFERMAGEM

UFG
UNIVERSIDADE FEDERAL DE GOIÁS

Conteúdo Programático

- Principais conceitos de séries temporais. (Aula 1)
- Funções de autocovariância e de autocorrelação. (Aula 1)
- Definição e estimação da tendência e da sazonalidade. (Aula 2)
- Métodos de suavização.
- Modelagem Box-Jenkins: modelos AR, MA, ARMA, ARIMA e SARIMA.
- Análise de séries temporais interrompidas.

Conteúdo - Aula 3

1. Definição e estimação de tendência e sazonalidade

- Tendência - Suavização
- Sazonalidade - Suavização

2. Modelos de Suavização

- Introdução
- Modelo para séries localmente constantes
  Suavização Exponencial Simples (SES)
- Modelos para séries que apresentam tendência
  Suavização exponencial de Holt (SEH)
- Modelos para séries sazonais
  Suavização exponencial de Holt-Winters

Tendência

Quando supomos que a tendência possa ser representada por um polinômio de baixo grau, isto implica que usamos todas as observações $Z_t, \ t = 1, \ldots, N$, para estimar o polinômio, que representará T_t sobre todo o intervalo de tempo considerado.

A ideia de se usar algum tipo de suavização é que a tendência num instante t será estimada usando-se observações Z_s, com s ao redor de t, por exemplo usamos as observações $Z_{t-n}, Z_{t-n+1}, \ldots, Z_{t+n}$ para estimar T_t. Existem vários métodos de suavização, iremos apresentar alguns deles.

Suavização - Médias Móveis

O que fazemos é usar um *filtro linear*, ou seja, uma operação que transforma a série Z_t na série Z_t^*, ou seja,

$$Z_t^* = \mathfrak{F}[Z_t], \; t = 1, \ldots, N.$$

Dado o modelo $Z_t = T_t + a_t$, transformando-o através de $\mathfrak{F}$, obtemos

$$Z_t^* = T_t^* + a_t^*,$$

sendo que $T_t^* = \mathfrak{F}[T_t]$, $a_t^* = \mathfrak{F}[a_t]$. Queremos que $\mathfrak{F}$ seja tal que $T_t^* \approx T_t$ e $E(a_t^*) = 0$, de modo que, suavizando-se as observações Z_t, obtenhamos $\mathfrak{F}[Z_t] = Z_t^* \approx T_t$.

Suavização - Médias Móveis

Dadas as observações $Z_1, \dots, Z_N$, o filtro $\mathfrak{F}$ comumente utilizado é da forma

$$Z_t^* = \sum_{j=-n}^{n} c_j Z_{t+j},\ t = n+1, \dots, N-n,$$

sendo que $\sum_{j=-n}^{n} c_j = 1.$

Observe que:

- perdemos n observações no início e n no final da série original;
- Z_t^* será uma estimativa da tendência no instante t;
- dizemos que a equação acima é um filtro de médias móveis.

Suavização - Médias Móveis

O caso mais simples é aquele em que $c_j = \frac{1}{2n+1}$, para todo j, de modo que

$$Z_t^* = \frac{1}{2n+1} \sum_{j=-n}^{n} Z_{t+j}, \; t = n+1, \ldots, N-n.$$

Segundo Anderson (1971), há três desvantagens principais neste processo de suavização:

(i) inferências estatísticas derivadas do método são limitadas, dado que ele não é baseado em nenhum modelo probabilístico;

(ii) não podemos obter as estimativas da tendência nos instantes $t = 1, \ldots, n$ e $t = N-n+1, \ldots, N$;

(iii) não fornece um meio de fazer previsões.

Suavização - Médias Móveis - Exemplo

Para ilustrar o procedimento, consideremos os dados da Tabela 1 (parte da série temporal com os dados de energia) e $n = 1$, ou seja, teremos médias móveis centradas de três termos, sendo dada por

$$Z_t^* = \frac{1}{3}\sum_{j=-1}^{1} Z_{t+j}, t = 2, \ldots, 5.$$

t	1	2	3	4	5	6
Z_t	84,6	89,9	81,9	95,4	91,2	89,8

Tabela 1: Dados para exemplificar o método de médias móveis

Assim,

$$Z_2^* =$$

Suavização - Médias Móveis - Exemplo

Para ilustrar o procedimento, consideremos os dados da Tabela 1 (parte da série temporal com os dados de energia) e $n = 1$, ou seja, teremos médias móveis centradas de três termos, sendo dada por

$$Z_t^* = \frac{1}{3}\sum_{j=-1}^{1} Z_{t+j}, t = 2, \ldots, 5.$$

t	1	2	3	4	5	6
Z_t	84,6	89,9	81,9	95,4	91,2	89,8

Tabela 1: Dados para exemplificar o método de médias móveis

Assim,

$$Z_2^* = \frac{1}{3}(Z_1 + Z_2 + Z_3) = \frac{84,6 + 89,9 + 81,9}{3} = 85,5$$

Suavização - Médias Móveis - Exemplo

$$Z_3^* =$$

Suavização - Médias Móveis - Exemplo

$$Z_3^* = \frac{1}{3}(Z_2 + Z_3 + Z_4) = \frac{89,9 + 81,9 + 95,4}{3} = 89,1$$

t	1	2	3	4	5	6
Z_t	84,6	89,9	81,9	95,4	91,2	89,8
Z_t^*		85,5	89,1	89,5	92,1	

Suavização - Medianas Móveis

Em vez de tomar médias móveis, podemos calcular medianas móveis. Assim, o valor suavizado $Z_t^{(md)}$ será

$$Z_t^{(md)} = mediana(Z_{t-n}, Z_{t-n+1}, ..., Z_{t+n}).$$

Se Z_t seguir o modelo $Z_t = T_t + a_t$, então $Z_t^{(md)}$ será uma estimativa da tendência no instante t e é denominada mediana móvel centrada de ordem $2n+1$.

Exemplo: Para ilustrar o procedimento, vamos utilizar os dados anteriores e calcular medianas móveis utilizando $n = 1$.

A primeira mediana móvel é $Z_2^{(md)} =$

Suavização - Medianas Móveis

Em vez de tomar médias móveis, podemos calcular medianas móveis. Assim, o valor suavizado $Z_t^{(md)}$ será

$$Z_t^{(md)} = mediana(Z_{t-n}, Z_{t-n+1}, ..., Z_{t+n}).$$

Se Z_t seguir o modelo $Z_t = T_t + a_t$, então $Z_t^{(md)}$ será uma estimativa da tendência no instante t e é denominada mediana móvel centrada de ordem $2n + 1$.

Exemplo: Para ilustrar o procedimento, vamos utilizar os dados anteriores e calcular medianas móveis utilizando $n = 1$.

A primeira mediana móvel é $Z_2^{(md)} =$ mediana$(Z_1, Z_2, Z_3) = 84,6$.

Aplicação

Aplicação no *software* R.

Sazonalidade - Método de Médias Móveis

O método de médias móveis é apropriado quando temos uma série temporal cuja componente sazonal varia com o tempo, ou seja, para séries cuja sazonalidade é estocástica.

Todavia, este procedimento é aplicado usualmente mesmo para padrão sazonal constante. Vamos considerá-lo aqui para o caso em que temos um padrão sazonal constante.

Sem perda de generalidade consideremos o caso que temos dados mensais e o número total de observações, N, é múltiplo de 12, isto é, $N = 12p$, $p =$ número de anos, de modo que os dados podem ser representados como na Tabela 2.

Sazonalidade

Anos	Meses jan 1	fev 2	mar 3	...	dez 12	Médias
1	Z_{11}	Z_{12}	Z_{13}	$\dots$	$Z_{1,12}$	$\overline{Z}_{1.}$
2	Z_{21}	Z_{22}	Z_{23}	$\dots$	$Z_{2,12}$	$\overline{Z}_{2.}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\dots$	$\vdots$	$\vdots$
p	Z_{p1}	Z_{p2}	Z_{p3}	$\dots$	Z_{p12}	$\overline{Z}_{p.}$
Médias	$\overline{Z}_{.1}$	$\overline{Z}_{.2}$	$\overline{Z}_{.3}$	$\dots$	$\overline{Z}_{.12}$	$\overline{Z}$

Tabela 2: Observações mensais de uma série temporal com p anos

Sendo que

$$\overline{Z}_{i.} = \frac{1}{12}\sum_{j=1}^{12} Z_{ij}, \; i = 1, \dots, p, \;\; \overline{Z}_{.j} = \frac{1}{p}\sum_{j=1}^{p} Z_{ij}, \; j = 1, \dots, 12, \;\; \overline{Z} = \frac{1}{12p}\sum_{i=1}^{p}\sum_{j=1}^{12} Z_{ij}$$

Sazonalidade - Método de Médias Móveis

Considere o modelo dado por

$$Z_t = T_t + S_t + a_t,\ t = 1, 2, \ldots, N$$

estimamos T_t através de

$$\widehat{T}_t = \sum_{j=-n}^{n} c_j Z_{t+j},\ t = n+1, 2, \ldots, N-n$$

e consideramos $Y_t = Z_t - \widehat{T}_t$, agora calculamos as médias dos meses: janeiro, fevereiro, ..., utilizando

$$\overline{Y}_{.j} = \frac{1}{n_j} \sum_{i=1}^{n_j} Y_{ij}, j = 1, 2, ..., 12$$

Sazonalidade - Método de Médias Móveis

Como a soma dos $\overline{Y}_{.j}$ em geral não é zero, tomamos como estimativas das constantes sazonais

$$\widehat{S}_j = \overline{Y}_{.j} - \overline{Y},$$

sendo que

$$\overline{Y} = \frac{1}{12} \sum_{j=1}^{12} \overline{Y}_{.j}.$$

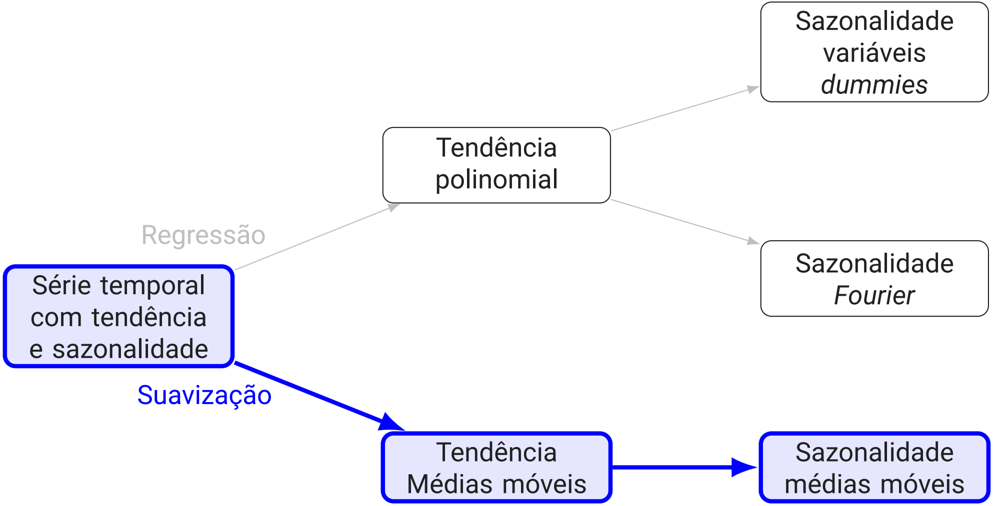
A série livre de sazonalidade é

$$Z_t^{SA} = Z_t - \widehat{S}_t.$$

Aplicação

Aplicação no *software* R.

Aula de hoje


Sazonalidade variáveis *dummies*
Tendência polinomial
Regressão
Sazonalidade *Fourier*
Série temporal com tendência e sazonalidade
Suavização
Tendência Médias móveis
Sazonalidade médias móveis

Introdução

A maioria dos métodos de previsão baseia-se na ideia de que observações passadas contêm informações sobre o padrão de comportamento da série temporal. O propósito dos métodos é distinguir o padrão de qualquer ruído que possa estar contido nas observações e então usar esse padrão para prever valores futuros da série.

Uma grande classe de métodos de previsão, que tenta tratar ambas as causas de flutuações em séries de tempo, é a das suavizações. Técnicas específicas desse tipo assumem que os valores extremos da série representam a aleatoriedade e, assim, por meio da suavização desses extremos, pode-se identificar o padrão básico.

Veremos a seguir, a grande popularidade atribuída aos métodos de suavização é devida à simplicidade, à eficiência computacional e à sua razoável precisão.

Modelo para séries localmente constantes

Vamos considerar agora um série temporal $Z_1, \ldots, Z_N$, localmente composta de seu nível mais um ruído aleatório, isto é,

$$Z_t = \mu_t + a_t, \; t = 1, \ldots, N,$$

sendo que $E(a_t) = 0$, $Var(a_t) = \sigma_a^2$ e μ_t é um parâmetro desconhecido, que pode variar lentamente com o tempo.

Suavização Exponencial Simples (SES)

Procedimento

A SES pode ser descrita matematicamente por

$$\overline{Z}_t = \alpha Z_t + (1-\alpha)\overline{Z}_{t-1}, \; \overline{Z}_0 = Z_1, \; t = 1, \dots, N,$$

sendo que $\overline{Z}_t$ é denominado valor exponencial suavizado e α é a constante de suavização, $0 \leq \alpha \leq 1$.

Efetuando a expansão temos que

$$\overline{Z}_t = \alpha Z_t + \alpha(1-\alpha)Z_{t-1} + \alpha(1-\alpha)^2 Z_{t-2} + \dots,$$

o que significa que a SES é uma média ponderada que dá pesos maiores às observações mais recentes.

Suavização Exponencial Simples (SES)

A previsão de todos os valores futuros é dada pelo último valor exponencialmente suavizado, isto é,

$$
\begin{aligned}
\widehat{Z}_t(h) &= \overline{Z}_t, \forall h > 0, \\
\widehat{Z}_t(h) &= \alpha Z_t + (1-\alpha)\overline{Z}_{t-1} \\
&= \alpha Z_t + (1-\alpha)\widehat{Z}_{t-1}(h+1),
\end{aligned}
$$

que pode ser interpretada como uma equação de atualização de previsão, quando tivermos uma nova observação.

Além disso, a previsão feita de acordo com a última equação reduz o problema de armazenagem de observações, pois pode ser calculada utilizando apenas a observação mais recente, a previsão imediatamente anterior e o valor de α.

Suavização Exponencial Simples (SES)

Determinação de α

Quanto menor for o valor de α mais estáveis serão as previsões finais, uma vez que a utilização de baixo valor de α implica que pesos maiores serão dados às observações passadas e, consequentemente, qualquer flutuação aleatória, no presente, exercerá um peso menor no cálculo da previsão. Em geral, quanto mais aleatória for a série estudada, menores serão os valores da constante de suavização.

Um procedimento objetivo é selecionar o valor de α que fornece a "melhor previsão" a um passo das observações já obtidas, ou seja, encontrar o valor de α que minimize

$$S = \sum_{t=l+1}^{N} (Z_t - \widehat{Z}_{t-1}(1))^2,$$

sendo l escolhido de tal modo que o valor inicial não influencie a previsão.

Suavização Exponencial Simples (SES)

Vantagens e Desvantagens

A SES é um método muito utilizado devido às seguintes vantagens:

i) fácil entendimento;
ii) aplicação não dispendiosa;
iii) grande flexibilidade permitida pela variação da constante de suavização α;
iv) necessidade de armazenar somente Z_t, $\overline{Z}_t$ e α.

A principal desvantagem é a dificuldade em determinar o valor mais apropriado da constante de suavização. Mas alguns *softwares* já fazem isso, um exemplo é o R.

Aplicação

Aplicação no *software* R.

Modelos para séries que apresentam tendência

Considerando, agora, o caso de uma série temporal não sazonal, que é composta localmente da soma de nível, tendência e resíduo aleatório com média zero e variância constante (σ_a^2), isto é,

$$Z_t = \mu_t + T_t + a_t, \; t = 1, \ldots, N.$$

A técnica vista anteriormente não é adequada para analisar séries temporais que apresentam tendência!

O método de SES quando aplicado a uma série que apresenta tendência linear positiva (ou negativa), fornece previsões que subestimam (ou superestimam) continuamente os valores reais. Para evitar esse erro sistemático, um dos métodos aplicáveis é o Suavização Exponencial de Holt (SEH).

Suavização exponencial de Holt (SEH)

Procedimento
Esse método é similar, em princípio, à SES, a diferença é que em vez de suavizar só o nível, ele utiliza uma nova constante de suavização para "modelar" a tendência da série.

Os valores do nível e da tendência da série, no instante t, serão estimados por

$$\overline{Z}_t = \alpha Z_t + (1-\alpha)(\overline{Z}_{t-1} + \widehat{T}_{t-1}),\ 0 \leq \alpha \leq 1 \text{ e } t = 2, \ldots, N$$

$$\widehat{T}_t = \beta(\overline{Z}_t - \overline{Z}_{t-1}) + (1-\beta)\widehat{T}_{t-1},\ 0 \leq \beta \leq 1 \text{ e } t = 2, \ldots, N,$$

respectivamente, sendo que α e β são denominadas constantes de suavização.

As fórmulas como definidas acima, como todas os métodos de suavização, modificam estimativas prévias quando uma nova observação é obtida.

Suavização exponencial de Holt (SEH)

Previsão

A previsão para o valor Z_{t+h}, com origem em t é dada por

$$\widehat{Z}_t(h) = \overline{Z}_t + h\widehat{T}_t, \forall h > 0,$$

ou seja, a previsão é feita adicionando-se ao valor básico $(\overline{Z}_t)$ a tendência multiplicada pelo número de passos à frente que se deseja prever (h).

Suavização exponencial de Holt (SEH)

Determinação das constantes de suavização

O procedimento é análogo ao de determinação da constante de suavização de uma SES, só que em vez de escolhermos o valor de α que torna a soma dos erros quadráticos de previsão mínimo, escolhemos o valor do vetor (α, β) tal que isto ocorra.

Suavização exponencial de Holt (SEH)

Vantagens e Desvantagens

As vantagens são semelhantes às do método anterior.

A desvantagem principal é a dificuldade em determinar os valores mais apropriados das duas constantes de suavização, α e β.

Aplicação

Aplicação no *software* R.

Suavização exponencial de Holt-Winters

Para séries temporais que apresentam um padrão de comportamento mais complexo, existe o método de Holt-Winters (HW).

Existem dois tipos de procedimentos cuja utilização depende das características da série considerada. Tais procedimentos são baseados em três equações com constantes de suavização diferentes, que são associadas a cada uma das componentes do padrão da série: nível, tendência e sazonalidade. Os dois procedimentos são:

a) série sazonal multiplicativa,
b) série sazonal aditiva.

Suavização exponencial sazonal de Holt-Winters

Procedimento - série sazonal multiplicativa

Considere uma série sazonal com período s. A variante do método que considera o fator sazonal (F_t) como sendo multiplicativo, enquanto a tendência (T_t) permanece aditiva, isto é

$$Z_t = \mu_t F_t + T_t + a_t, \; t = 1, \ldots, N.$$

Suavização exponencial sazonal de Holt-Winters

Procedimento - série sazonal multiplicativa

As três equações de suavização são dadas por

$$
\begin{aligned}
\overline{Z}_t &= \alpha\left(\frac{Z_t}{\widehat{F}_{t-s}}\right) + (1-\alpha)(\overline{Z}_{t-1} + \widehat{T}_{t-1}),\ 0 \leq \alpha \leq 1,\\
\widehat{T}_t &= \beta(\overline{Z}_t - \overline{Z}_{t-1}) + (1-\beta)\widehat{T}_{t-1},\ 0 \leq \beta \leq 1,\\
\widehat{F}_t &= \gamma\left(\frac{Z_t}{\overline{Z}_t + \widehat{T}_{t-1}}\right) + (1-\gamma)\widehat{F}_{t-s},\ 0 \leq \gamma \leq 1,
\end{aligned}
$$

com $t = s+1, \ldots, N$, representando estimativas do fator de nível, da tendência e do sazonal, respectivamente, e α, β e γ são constantes de suavização e estão restritos no intervalo 0 a 1.

Suavização exponencial sazonal de Holt-Winters

Previsão - série sazonal multiplicativa

As previsões dos valores futuros da série são dadas por

$$\widehat{Z}_t(h) \quad = \quad (\overline{Z}_t + h\widehat{T}_t)\widehat{F}_{t+h-s}, \; h = 1, 2, \ldots, s.$$

onde $\overline{Z}_t$, $\widehat{T}_t$ e $\widehat{F}_t$ são dados como anteriormente.

Suavização exponencial sazonal de Holt-Winters

Procedimento - série sazonal aditiva

O procedimento anterior pode ser modificado para tratar com situações onde o fator sazonal é aditivo,

$$Z_t = \mu_t + T_t + F_t + a_t,\ t = 1, \ldots, N.$$

As estimativas do fator sazonal, nível e tendência da série são dadas por

$$\begin{aligned}
\overline{Z}_t &= \alpha(Z_t - \widehat{F}_{t-s}) + (1-\alpha)(\overline{Z}_{t-1} + \widehat{T}_{t-1}),\ 0 \leq \alpha \leq 1, \\
\widehat{T}_t &= \beta(\overline{Z}_t - \overline{Z}_{t-1}) + (1-\beta)\widehat{T}_{t-1},\ 0 \leq \beta \leq 1, \\
\widehat{F}_t &= \gamma(Z_t - \overline{Z}_t - \widehat{T}_{t-1}) + (1-\gamma)\widehat{F}_{t-s},\ 0 \leq \gamma \leq 1,
\end{aligned}$$

com $t = s+1, \ldots, N$, e α, β e γ são constantes de suavização e estão restritos no intervalo 0 a 1.

Suavização exponencial sazonal de Holt-Winters

Previsão - série sazonal aditiva

As previsões dos valores futuros da série são dadas por

$$\widehat{Z}_t(h) \quad = \quad \overline{Z}_t + h\widehat{T}_t + \widehat{F}_{t+h-s}, \; h = 1, 2, \ldots, s$$

onde $\overline{Z}_t$, $\widehat{T}_t$ e $\widehat{F}_t$ são dados como anteriormente.

Suavização exponencial sazonal de Holt-Winters

Vantagens e desvantagens

As vantagens são semelhantes às da utilização do método de Holt, sendo que os métodos de HW são adequados à análise de séries com padrão de comportamento mais geral.

As desvantagens são as dificuldades em determinar os valores mais apropriados das constantes de suavização.

A determinação das constantes de suavização (α, β, γ) é realizada de modo a tornar mínima a soma dos quadrados dos erros de ajustamento.

Aplicação

Aplicação no *software* R.

Especialização em *Data Science* e Estatística Aplicada

Módulo IV - Análise de Séries Temporais

Prof. Dr. Eder Angelo Milani

edermilani@ufg.br

IME INSTITUTO DE MATEMÁTICA E ESTATÍSTICA

FEN FACULDADE DE ENFERMAGEM

UFG UNIVERSIDADE FEDERAL DE GOIÁS