

Séries Temporais como Objeto de Processamento

Da observação bruta à preparação da séries para análise.

Eduardo Ogasawara

eduardo.ogasawara@cefet-rj.br

<https://eic.cefet-rj.br/~eogasawara>

Processamento de Séries Temporais

Na análise de séries temporais, a série é transformada antes de modelar, gerando uma nova representação que facilita a predição e identificação de desvios.

Uma transformação leva a série original para outra série:

$$Y_t = \mathcal{T}(X_t)$$

Onde X_t é o valor observado no tempo t , Y_t é o valor transformado, e \mathcal{T} é a transformação aplicada.

- ❑ A ideia central é mudar o ponto de vista: em vez de analisar diretamente os dados brutos, tratamos a série como matéria-prima para construir representações mais úteis.

Pipeline de Processamento Temporal

Um pipeline é uma sequência de transformações onde cada etapa gera uma nova série. O resultado final é uma representação mais informativa dos dados originais.

Uma sequência de transformações pode ser escrita como:

$$X_t^{(k)} = \mathcal{T}_k \circ \mathcal{T}_{k-1} \circ \dots \circ \mathcal{T}_1(X_t)$$

Onde X_t é a série original, $X_t^{(k)}$ é a série após k transformações, \mathcal{T}_i é a i -ésima transformação, e \circ representa a composição de funções.

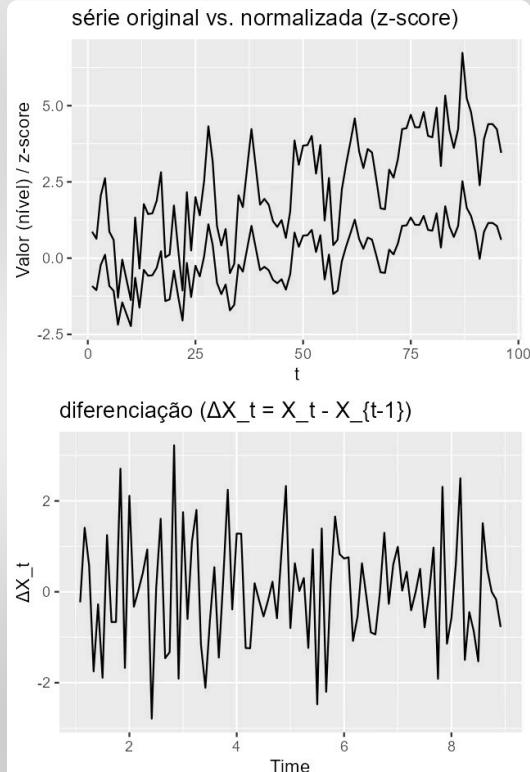
Estrutura do Pipeline de Séries Temporais

Dados brutos passam por etapas sucessivas, cada uma produzindo uma nova representação. Desvios e insights são identificados ao final do processo.



Esse esquema mostra que as análises podem não ocorrer diretamente nos dados brutos, mas em representações intermediárias. Cada etapa altera a forma como a série é interpretada.

TRANSFORMAÇÕES BÁSICAS



Pré-processamento Temporal

Normalização

Ajusta a escala dos valores para tornar a série mais adequada à análise:

$$X_t^{(norm)} = \frac{X_t - \mu}{\sigma}$$

Onde μ é a média da série e σ é o desvio padrão.

Diferenciação

Destaca variações no tempo:

$$\Delta X_t = X_t - X_{t-1}$$

Onde ΔX_t é a diferença entre instantes consecutivos.

- ❑ O pré-processamento altera a escala e a estrutura da série. Essas transformações mudam a forma como desvios e eventos aparecem nos dados.

Engenharia Temporal e Modelagem de Desvios

Modelos geram uma representação da série, e desvios são identificados em relação ao modelo. A escolha da representação é decisiva.

Cálculo do Resíduo

$$R_t = X_t - \hat{X}_t$$

Onde X_t é o valor observado e \hat{X}_t é o valor estimado pelo modelo.

Identificação de Desvios

$$|R_t| > \tau$$

Um desvio ocorre quando o resíduo R_t ultrapassa o limiar τ .

Primeiro definimos o comportamento esperado da série por meio de um modelo. Em seguida, medimos o quanto os dados reais se afastam desse comportamento.

Transformações de Escala

Algumas séries têm variância dependente do nível. Transformações de escala alteram a amplitude dos valores para tornar a variância mais estável.

Uma transformação de escala é dada por:

$$Y_t = g(X_t)$$

Onde $g(\cdot)$ é uma função de transformação, como logaritmo ou raiz quadrada.

Transformações de escala mudam a forma como a série cresce ou varia. O logaritmo, por exemplo, reduz a influência de valores muito grandes, facilitando a comparação de variações ao longo do tempo.

TENDÊNCIA

Remoção de Tendência

Séries podem ter crescimento ou declínio de longo prazo. A tendência pode ser removida antes da análise para destacar variações locais.

Se a série tem uma tendência estimada \hat{T}_t , define-se:

$$Y_t = X_t - \hat{T}_t$$

Série Original

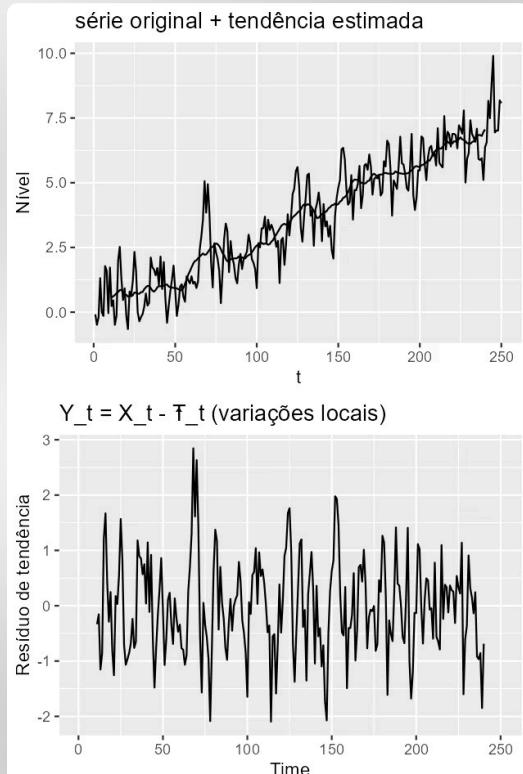
Contém tendência e variações locais misturadas

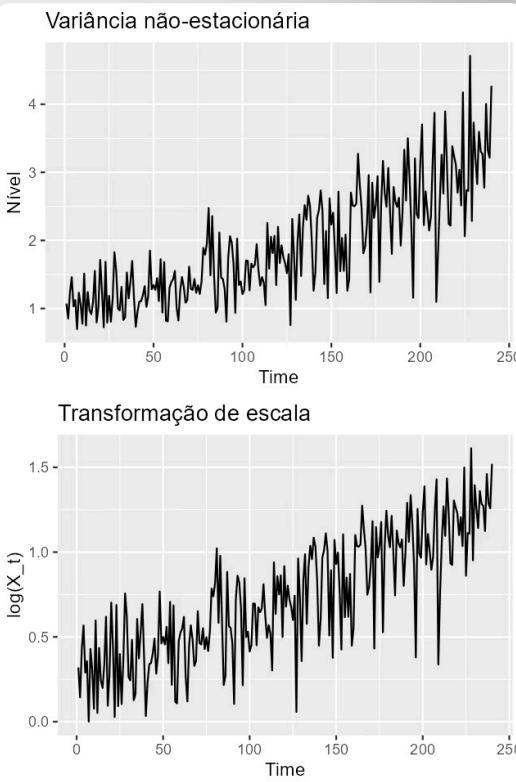
Tendência Estimada

Representa o comportamento global de longo prazo

Série Sem Tendência

Destaca apenas variações locais





Estabilização da Variância

Algumas séries têm variância que muda no tempo. Transformações buscam tornar a variância aproximadamente constante.

Objetivo da transformação:

$$\text{Var}(Y_t) \approx \text{constante}$$

Onde Y_t é a série transformada e $\text{Var}(Y_t)$ é sua variância.

- Quando a variância muda ao longo do tempo, eventos podem ser confundidos com períodos de alta volatilidade. Ao estabilizar a variância, separamos variações estruturais de mudanças realmente excepcionais.

Transformações Temporais

Muitas análises usam variações em vez de níveis. Diferenças e retornos destacam mudanças abruptas na série temporal.

%

Retorno Simples

$$R_t = \frac{X_t - X_{t-1}}{X_{t-1}}$$

Mede a variação percentual entre períodos consecutivos.



Retorno Logarítmico

$$r_t = \log(X_t) - \log(X_{t-1})$$

Fornece propriedades matemáticas mais convenientes para análise.

Essas transformações substituem os níveis da série por variações relativas, tornando mais visíveis mudanças bruscas que são frequentemente interpretadas como eventos.

Transformações e Análise

Padrões são definidos na série transformada. A transformação altera o tipo de padrão detectável.

Um critério de padrão pode ser:

$$|Y_t - \mathbb{E}[Y_t]| > \tau$$

Onde Y_t é a série transformada, $\mathbb{E}[Y_t]$ é o valor esperado, e τ é o limiar de detecção.

Comportamento Normal

Definido pela transformação aplicada e pelo modelo escolhido

Desvio Detectado

Desvio significativo em relação ao comportamento esperado

Decomposição como Transformação

A série pode ser separada em componentes, onde cada componente representa um aspecto do comportamento temporal.



Tendência

 T_t

Comportamento de longo prazo



Sazonalidade

 S_t

Padrões cíclicos regulares



Resíduo

 R_t

Variações não explicadas

Define-se o resíduo como:

$$R_t = X_t - T_t - S_t$$

O resíduo concentra variações que não são explicadas pela tendência ou sazonalidade, sendo um espaço natural para detectar eventos.

Transformada de Fourier

A série pode ser representada por frequências. A transformada de Fourier muda o domínio da análise de tempo para frequência.

A transformada discreta é dada por:

$$C_k = \sum_t X_t e^{-i2\pi kt/n}$$

Onde C_k é o coeficiente de frequência k , n é o número de observações, e i é a unidade imaginária.

A transformada de Fourier descreve a série em termos de componentes periódicos, permitindo separar padrões cíclicos de ruído e facilitando a identificação de comportamentos anômalos.

Decomposição por Wavelets

Wavelets combinam tempo e frequência, permitindo analisar desvios localizados com precisão temporal e de escala.

A transformada wavelet pode ser representada por:

$$W(a, b) = \int X(t) \psi_{a,b}(t) dt$$

Onde $\psi_{a,b}(t)$ é a wavelet escalada e deslocada, a é a escala, e b é o deslocamento temporal.

Localização Temporal

Preserva informação sobre quando os desvios ocorrem

Análise Multi-escala

Detecta desvios em diferentes escalas de tempo simultaneamente

Decomposição EMD

A Decomposição por Modos Empíricos (EMD) decompõe a série em componentes adaptativas sem exigir hipóteses fortes sobre a estrutura dos dados.

A decomposição é dada por:

$$X_t = \sum_{k=1}^K IMF_k(t) + r(t)$$

Onde $IMF_k(t)$ é o k -ésimo modo intrínseco, $r(t)$ é o resíduo, e K é o número de modos.

01

Extração Adaptativa

Componentes extraídos diretamente dos dados

02

Sem Modelos Pré-definidos

Não impõe estrutura prévia aos dados

03

Multi-escala

Revela padrões em diferentes escalas temporais

Decomposição e Análise

Desvios podem ser definidos no resíduo da decomposição. Diferentes métodos geram diferentes resíduos, influenciando diretamente os desvios detectados.

Critério de análise:

$$|R_t| > \tau$$

Onde R_t é o resíduo e τ é o limiar de detecção.

- Ao definir desvios no resíduo, isolamos variações inesperadas. Como cada método de decomposição produz um resíduo diferente, a escolha do método influencia diretamente os desvios detectados.

ANÁLISE LOCAL

Representação Local da Série

A série global pode ser dividida em subsequências, onde cada subsequência representa um comportamento local específico.

Uma subsequência é definida por:

$$X_{t:t+w-1} = (X_t, X_{t+1}, \dots, X_{t+w-1})$$

Onde w é o tamanho da janela e $X_{t:t+w-1}$ é a subsequência iniciada em t .

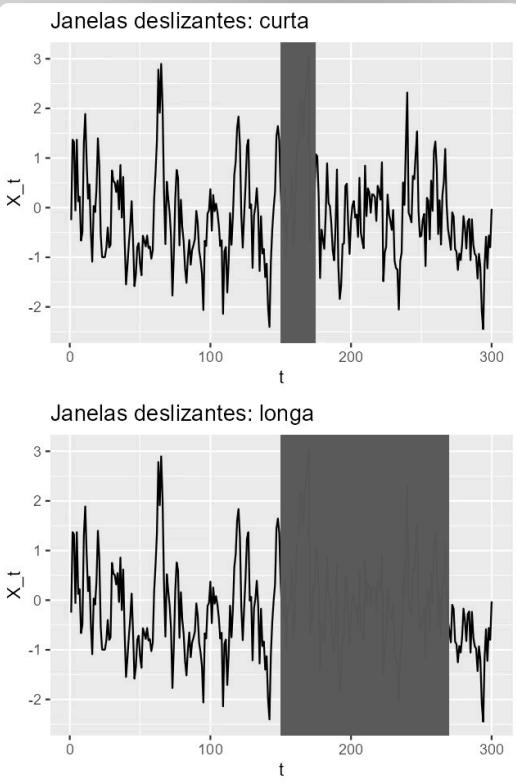
Análise Local

Eventos entendidos como desvios em janelas específicas

Contexto Temporal

Cada janela captura comportamento em um período definido

Janelas Deslizantes



Janelas deslizantes geram múltiplas subsequências sobrepostas. O tamanho da janela define a escala da análise e o tipo de desvio analisável.

Definição da transformação:

$$\mathcal{W}_w(X_t) = \{X_{t:t+w-1}\}$$

Onde \mathcal{W}_w é o operador de janela de tamanho w .

Janelas Curtas

Destacam eventos rápidos e mudanças abruptas

Janelas Longas

Capturam padrões amplos e tendências de longo prazo

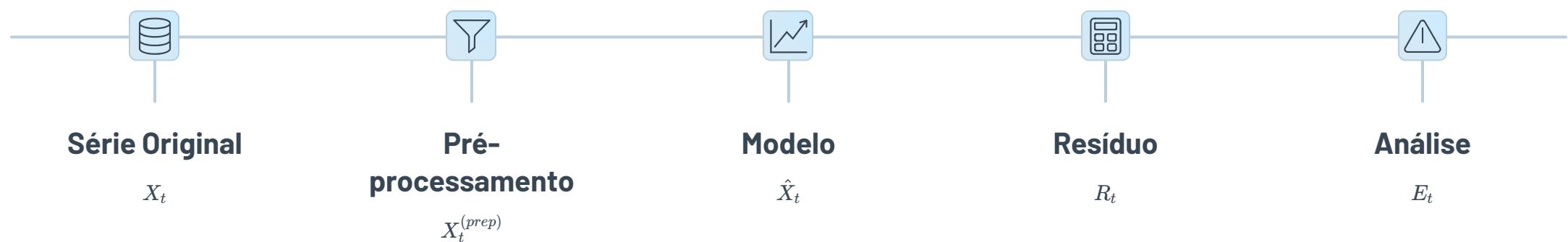


Janelas Médias

Equilibram detalhes locais e contexto temporal

Workflow Temporal

O processamento segue etapas bem definidas. A análise é realizada ao final do fluxo como resultado de todas as transformações anteriores.



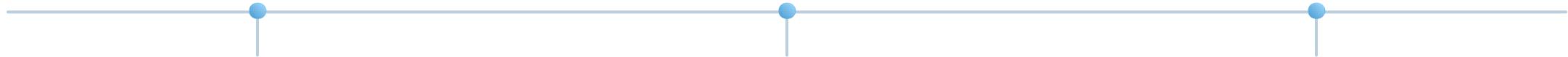
Esse fluxo mostra que a análise de séries temporais depende de todas as etapas anteriores. A análise é o resultado final de uma cadeia de transformações e modelagens.

Modelagem Preditiva

A previsão define o comportamento esperado da série. Desvios são identificados em relação ao esperado.

Definição do indicador de desvio:

$$E_t = \begin{cases} 1, & |X_t - \hat{X}_t| > \tau \\ 0, & \text{caso contrário} \end{cases}$$



Valor Observado

X_t — dado real coletado no tempo t

Valor Previsto

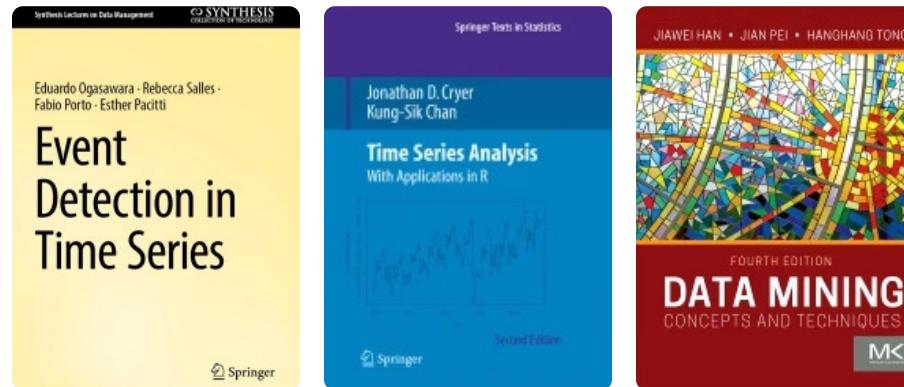
\hat{X}_t — referência de normalidade gerada pelo modelo

Limiar

τ — define o que é considerado desvio significativo

Referências Bibliográficas

Uma coleção cuidadosamente selecionada de obras fundamentais que abordam análise de séries temporais e mineração de dados.



Event Detection in Time Series

Ogasawara, E.; Salles, R.; Porto, F.; Pacitti,

E. (2025). Publicação recente da Springer Nature Switzerland que explora técnicas avançadas de detecção de eventos em séries temporais.

Time Series Analysis: With Applications in R

Cryer, J. D.; Chan, K.-S. (2008). Obra clássica da Springer que combina fundamentação teórica sólida com implementações práticas.

Data Mining: Concepts and Techniques

Han, J.; Pei, J.; Tong, H. (2022). Quarta edição publicada pela Morgan Kaufmann que consolida conceitos fundamentais e técnicas avançadas de mineração de dados