

Representação de Eventos em Séries Temporais

Explore como diferentes representações de dados de séries temporais moldam fundamentalmente os eventos que podemos detectar e como os definimos.

Eduardo Ogasawara

eduardo.ogasawara@cefet-rj.br

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

Representação de séries temporais

Uma série temporal observada ao longo do tempo pode ser transformada em diferentes formas de representação. Tradicionalmente tratada como sequência de valores numéricos, em mineração temporal a série pode ser convertida em outro espaço representacional.

Exemplos incluem representações vetoriais, simbólicas, espectrais e probabilísticas. Essa transformação é fundamental porque determina o tipo de informação que conseguimos extrair e quais eventos podem ser detectados.

$$\mathcal{R}(X_t) \rightarrow Z$$

Onde X_t é o valor da série no instante t , \mathcal{R} é a função de representação, e Z é o espaço de representações.

Representações Globais e Locais

Representação Global

Utiliza toda a série temporal para análise, capturando padrões gerais e tendências de longo prazo.

$$\mathcal{R}_G(X_t)$$

Representação Local

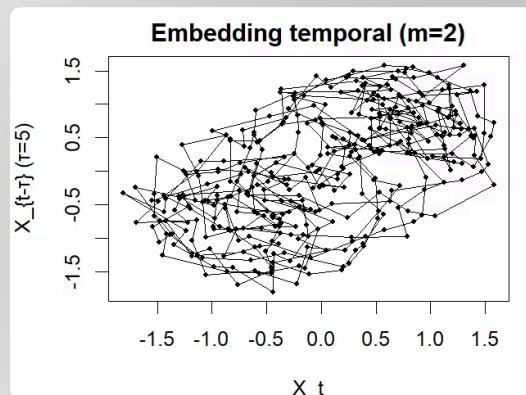
Usa apenas partes específicas da série, essencial para detectar eventos que ocorrem em trechos temporais específicos.

$$\mathcal{R}_L(X_{t:t+w})$$

Eventos são fenômenos que ocorrem em trechos específicos da série, exigindo representações locais. Enquanto modelos clássicos analisam a série como um todo, eventos geralmente aparecem em janelas específicas do tempo, onde w representa o tamanho da janela temporal.

Embeddings Temporais

Uma série temporal pode ser representada em um espaço de estados, onde cada ponto é construído a partir de valores defasados da série. O conjunto desses pontos forma uma trajetória no espaço de estados.



01

Coleta de Valores Defasados

Valores da série em diferentes momentos temporais são coletados.

02

Construção do Vetor

Os valores defasados formam um vetor de embedding multidimensional.

03

Trajetória no Espaço

A sequência de vetores cria uma trajetória geométrica que revela padrões.

$$\mathbf{X}_t = (X_t, X_{t-\tau}, X_{t-2\tau}, \dots, X_{t-(m-1)\tau})$$

Onde \mathbf{X}_t é o vetor de embedding no instante t , m é a dimensão do embedding, e τ é o atraso temporal. A série deixa de ser vista como sequência unidimensional e passa a ser interpretada como trajetória geométrica.

Representação e Conceito de Evento

Múltiplas Perspectivas

A mesma série pode ter representações completamente diferentes, cada uma revelando aspectos distintos dos dados.

Resultados Distintos

Representações diferentes geram conjuntos diferentes de eventos detectados.

Dependência Metodológica

O conjunto de eventos detectados depende fundamentalmente da representação escolhida.

$$\mathcal{R}_1(X_t) \neq \mathcal{R}_2(X_t)$$

Representar uma série é, na prática, escolher um ponto de vista. Cada representação enfatiza certos aspectos dos dados e ignora outros. Por isso, a própria definição de evento depende da forma como a série é representada.

TIPICIDADE

Conceito de Tipicidade Temporal

A tipicidade depende de um modelo ou referência estabelecida. Pode ser definida por modelos estatísticos ou empíricos. Sem uma definição clara de tipicidade, não existe conceito de evento.

Modelo de Referência

Estabelece o comportamento esperado da série temporal através de padrões históricos ou teóricos.

Valor Esperado

Define o comportamento típico como a expectativa condicional dado o modelo.

Base para Eventos

Eventos são definidos como desvios significativos em relação a essa referência de tipicidade.

$$N_t = \mathbb{E}[X_t | M]$$

Onde N_t é o comportamento típico no instante t , $\mathbb{E}[\cdot]$ é o valor esperado, e M é o modelo ou representação da série.

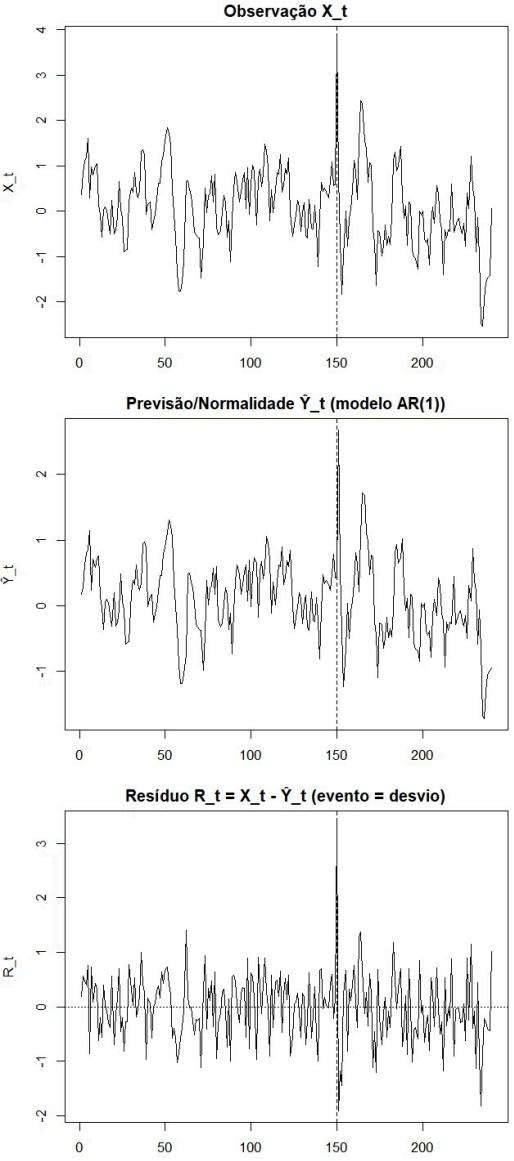
Tipicidade como Modelo Temporal

Um modelo define o comportamento esperado da série através de previsão condicional. A tipicidade é dada pela previsão baseada em informações passadas. Diferentes modelos geram diferentes noções de tipicidade.

$$\hat{X}_t = \mathbb{E}[X_t | \mathcal{F}_{t-1}, M]$$

Onde \hat{X}_t é o valor previsto no instante t , \mathcal{F}_{t-1} é a informação disponível até $t - 1$, e M é o modelo temporal.

A tipicidade é formalizada como previsão. Se o modelo descreve bem a série, então o comportamento típico é aquilo que ele prevê. Eventos aparecem quando a série se afasta sistematicamente dessa previsão.



Resíduos Temporais

1

Observação

Valor real medido na série temporal

2

Previsão

Valor esperado pelo modelo

3

Resíduo

Diferença não explicada pelo modelo

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

O resíduo R_t é a diferença entre o valor observado X_t e o valor previsto \hat{X}_t . Resíduos representam a parte não explicada pelo modelo.

Se o modelo captura bem o comportamento típico, os resíduos deveriam ser aleatórios. Quando aparecem padrões, rupturas ou valores extremos nos resíduos, isso sugere a presença de eventos.

Evento como Desvio Estatístico

Um evento é definido como um desvio significativo em relação à tipicidade. O critério depende de um limiar estatístico, que pode ser definido por desvio padrão ou quantis da distribuição.



Limiar Estatístico

Define o ponto de corte para considerar um desvio como significativo



Detecção Binária

Classifica cada instante como evento ou não-evento



Decisão Metodológica

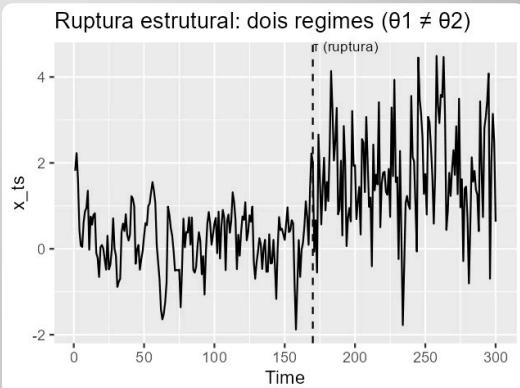
A escolha do limiar é crucial e depende do contexto

$$E_t = \begin{cases} 1, & |R_t| > \tau \\ 0, & \text{caso contrário} \end{cases}$$

Onde E_t é o indicador de evento, R_t é o resíduo, e τ é o limiar estatístico. Nesta visão, evento é um conceito estatístico: ocorre quando o desvio em relação à tipicidade é grande o suficiente.

Evento como Ruptura Estrutural

Um evento pode representar uma mudança no processo gerador da série. Antes e depois do evento, os parâmetros do modelo são diferentes, caracterizando uma ruptura estrutural.



Regime Anterior

Parâmetros θ_1 governam o processo

Novo Regime

Parâmetros θ_2 definem novo comportamento



Ponto de Ruptura

Momento de transição estrutural

$$\theta_1 \neq \theta_2$$

O evento não é apenas um ponto extremo, mas uma mudança no próprio mecanismo que gera a série. Essa visão inclui change points e mudanças de regime, representando alterações fundamentais no processo temporal.

Dependência da Representação

Representação Influencia Eventos

O conjunto de eventos detectados depende fundamentalmente da representação escolhida para analisar a série temporal.

Representações diferentes geram conjuntos completamente diferentes de eventos, mesmo analisando os mesmos dados brutos.

$$\mathcal{E}(\mathcal{R}_1) \neq \mathcal{E}(\mathcal{R}_2)$$

Onde $\mathcal{E}(\mathcal{R})$ representa o conjunto de eventos sob a representação \mathcal{R} .

- ❑ **Implicação Epistemológica:** Não existe evento absoluto. O que é considerado evento depende do espaço onde a série é analisada. Isso reforça o caráter epistemológico da detecção de eventos.

Ruído, Sinal e Evento

Sinal Estruturado
Componente S_t com padrões sistemáticos e previsíveis



Ruído Aleatório

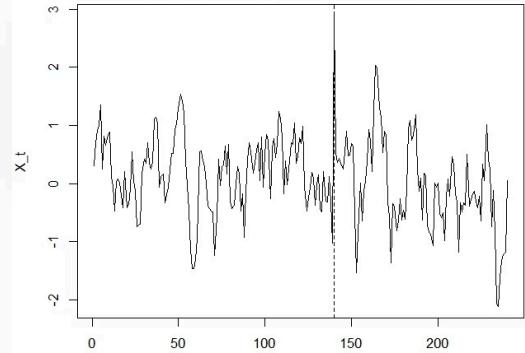
Componente ε_t com flutuações estocásticas naturais

Componente de Evento

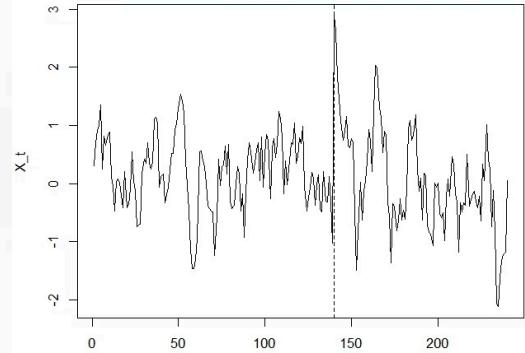
Componente δ_t com perturbações estruturadas significativas

A série pode ser decomposta em sinal, ruído e evento. Ruído é flutuação aleatória, enquanto evento é uma alteração estruturada. Essa distinção é fundamental para não confundir variação natural com fenômenos relevantes.

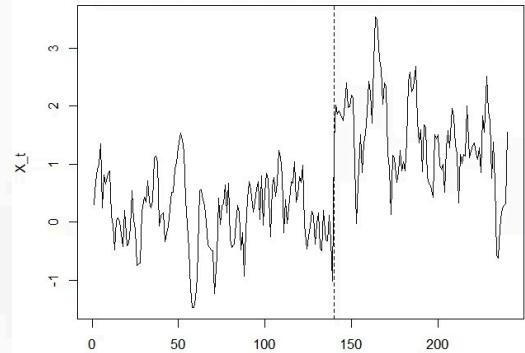
Outlier Aditivo (AO): afeta um único ponto



Outlier Inovacional (IO): choque que se propaga



Mudança de Nível (LS): alteração permanente após t0



Outliers em Séries Temporais

Outlier é uma observação atípica isolada que se destaca significativamente do padrão geral. Existem diferentes tipos de outliers temporais, e nem todo outlier representa um evento estrutural.

1

Outlier Aditivo

Afeta apenas uma observação específica sem impactar valores subsequentes

2

Outlier Inovacional

Afeta a observação e propaga efeitos para valores futuros

3

Mudança de Nível

Altera permanentemente o nível da série após o ponto de ocorrência

$$|R_t| \gg \sigma_R$$

Onde R_t é o resíduo e σ_R é o desvio padrão dos resíduos. Outliers são casos extremos, mas podem não representar mudanças estruturais. Eventos, em sentido amplo, podem envolver padrões mais complexos do que observações isoladas.

Change Points e Eventos

Change point é uma mudança nos parâmetros do processo gerador da série. Eventos podem ser locais ou globais, e change points representam uma classe particular de eventos estruturais.

Regime Inicial

Parâmetros θ_1 até o instante τ

Regime Posterior

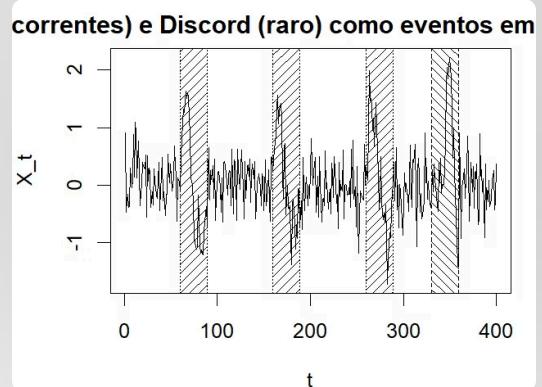
Parâmetros θ_2 após o instante τ

$$\theta(t) = \begin{cases} \theta_1, & t \leq \tau \\ \theta_2, & t > \tau \end{cases}$$

Change points representam mudanças globais na série, enquanto eventos podem ser mais localizados. Assim, change points são uma classe particular de eventos que afetam o comportamento de longo prazo da série.

Motifs, Discords e Eventos

Padrões Temporais



Motifs são padrões recorrentes que aparecem frequentemente na série temporal.

Discords são padrões raros que se destacam por sua singularidade.

Eventos podem ser interpretados como padrões raros ou discords no conjunto de subsequências.

$$\text{discord} = \arg \max_s \min_{u \neq s} d(s, u)$$

Onde s e u são subsequências, e $d(\cdot, \cdot)$ é uma medida de distância. Nesta abordagem, eventos são vistos como padrões incomuns no conjunto de subsequências. Isso conecta detecção de eventos com mineração de padrões temporais.

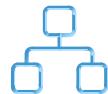
Tipologia de Eventos

Eventos podem ser classificados em diferentes categorias, cada uma capturando aspectos distintos de anomalias temporais.



Eventos Estatísticos

Desvios significativos em medidas estatísticas como média, variância ou distribuição



Eventos Estruturais

Mudanças nos parâmetros ou regime do processo gerador da série



Eventos Geométricos

Alterações na trajetória ou forma da série no espaço de estados



Eventos de Padrão

Aparição ou desaparecimento de padrões específicos na série temporal

$$\mathcal{E} = \mathcal{E}_{stat} \cup \mathcal{E}_{struct} \cup \mathcal{E}_{geom} \cup \mathcal{E}_{pattern}$$

Essa taxonomia mostra que event detection é um problema multifacetado. Diferentes métodos capturam diferentes tipos de eventos.

Arquitetura da Detecção de Eventos

A detecção de eventos é um processo em múltiplos níveis, onde cada nível influencia o tipo de evento detectável. Não é uma etapa única, mas uma cadeia de decisões metodológicas.

Série Temporal Bruta

Dados observados no domínio temporal original

Representação

Transformação dos dados em espaço adequado

Modelo de Tipicidade

Definição do comportamento esperado

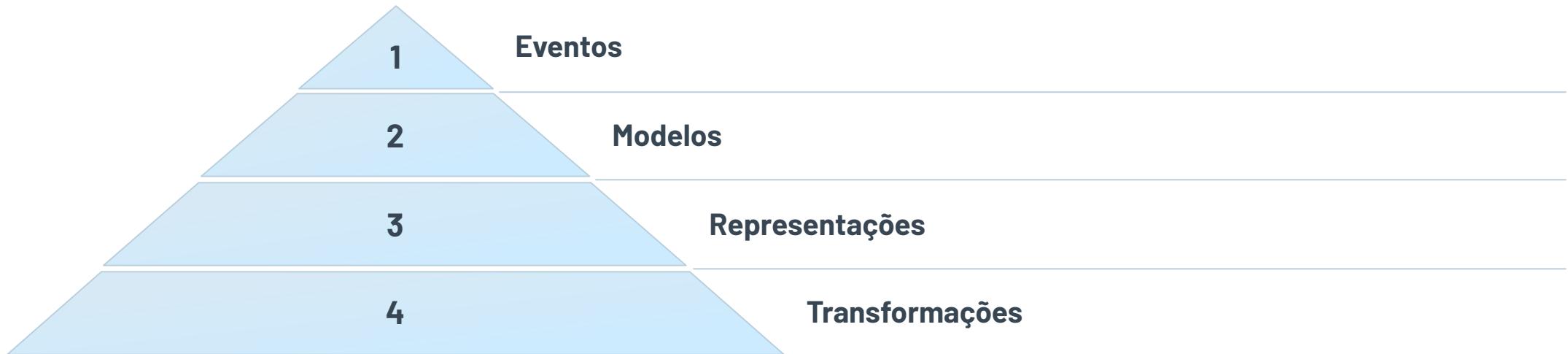
Detecção de Evento

Identificação de desvios significativos

série → representação → modelo → evento

Cada escolha metodológica molda o resultado final. A arquitetura revela que eventos emergem de um processo de construção teórica e não são propriedades intrínsecas dos dados.

Fundamentos Teóricos



Transformações, representações e modelos definem eventos. Eventos não surgem diretamente dos dados, mas de uma combinação estruturada desses três elementos fundamentais.

$$\mathcal{E} = f(\mathcal{T}, \mathcal{R}, \mathcal{M})$$

Onde \mathcal{T} representa transformações, \mathcal{R} representa representações, e \mathcal{M} representa modelos. Esta visão formaliza o papel do arcabouço teórico na detecção de eventos.

A fórmula expressa que eventos são funções das escolhas metodológicas, não descobertas neutras nos dados.

Tipicidade e Conceito de Evento

O conceito de evento depende fundamentalmente da definição de tipicidade. Primeiro definimos o que é comportamento típico; depois, definimos o que é desvio.

Portanto, o conceito de evento emerge diretamente da noção de tipicidade estabelecida pelo modelo ou referência escolhida.



Definir Tipicidade

Estabelecer modelo de referência

Identificar Desvios

Comparar observações com tipicidade

Classificar Eventos

Determinar significância dos desvios

$$\text{tipicidade} \Rightarrow \text{evento}$$

A relação de implicação mostra que eventos são conceitos derivados, não primitivos. Sem tipicidade, não há evento.

Métodos como Instâncias de Definições

Cada método de detecção assume uma representação específica e um modelo particular. Métodos diferentes detectam eventos diferentes porque incorporam escolhas teóricas distintas.



Métodos Estatísticos

Baseados em distribuições e testes de hipótese



Métodos de Aprendizado

Utilizam modelos adaptativos e redes neurais



Métodos Geométricos

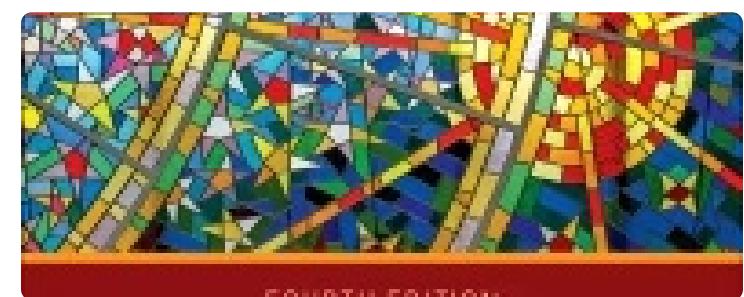
Analisam trajetórias no espaço de estados

$$\text{método}_i = g_i(\mathcal{T}, \mathcal{R}, \mathcal{M})$$

Onde g_i representa o método de detecção de eventos. Algoritmos não são neutros: cada um incorpora escolhas teóricas implícitas. Por isso, métodos distintos produzem resultados distintos ao analisar os mesmos dados.

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.