

Motifs em Séries Temporais

Um motif é um padrão que aparece várias vezes na série temporal. Diferente de anomalias ou pontos de mudança, motifs representam eventos de recorrência, não de ruptura.

Eduardo Ogasawara

eduardo.ogasawara@cefet-rj.br

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

Definição de Motif

São definidos sobre subsequências (janelas) de tamanho w , não sobre pontos isolados.

Considere a série $X = \{x_t\}$ com $t = 1$ até T . Um motif é um padrão temporal que reaparece em diferentes trechos da série, formando um conjunto de subsequências semelhantes $X_{t:t+w}$.

Motifs versus Outros Eventos

Anomalias

Eventos raros que representam desvios locais do padrão esperado. Definidas por $|r_t| > \tau$, onde o resíduo excede um limiar.

Change Points

Representam rupturas entre regimes, definidas por $\theta_1 \neq \theta_2$. Segmentam a série em diferentes fases.

Motifs

Eventos frequentes que representam similaridade entre janelas. Definidos por $d(X_{i:i+w}, X_{j:j+w}) < \varepsilon$, onde a distância entre subsequências é menor que um limiar.

Enquanto anomalias buscam desvios e change points identificam rupturas, motifs capturam regularidades recorrentes. São perspectivas complementares na análise temporal.

Ruído e Padrões

Decomposição da Série

Toda série temporal pode ser decomposta em: $x_t = S_t + \varepsilon_t$, onde S_t é o componente estruturado (sinal) e ε_t é o ruído aleatório.

Motifs pertencem ao sinal, representando estruturas recorrentes. O ruído pode dificultar a detecção, mas não define o motif. A identificação de padrões foca na regularidade estrutural, não nas variações aleatórias.

Motifs no Sistema de Eventos

O livro apresenta uma taxonomia completa de eventos temporais: $\mathcal{E} = \mathcal{A} \cup \mathcal{C} \cup \mathcal{M} \cup \mathcal{D}$, onde cada símbolo representa um tipo específico de evento.



Anomalias (\mathcal{A})

Desvios e exceções



Change Points (\mathcal{C})

Rupturas e transições



Motifs (\mathcal{M})

Padrões recorrentes

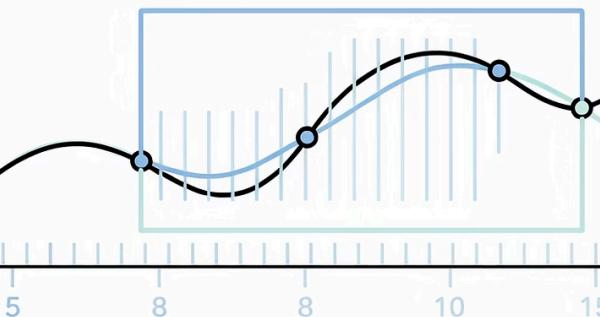


Discords (\mathcal{D})

Padrões únicos

Motifs representam a camada de regularidades recorrentes, focando em frequência e repetição, não em exceção.

Subsequências como Objetos Fundamentais



Motifs são definidos em janelas de tamanho w , não em pontos isolados. Cada janela se torna um "objeto" de análise independente.

Uma subsequência é definida como $X_{t:t+w-1} = (x_t, x_{t+1}, \dots, x_{t+w-1})$, onde t varia de 1 até $T - w + 1$. O conjunto $\mathcal{S} = \{X_{t:t+w-1}\}$ reúne todas as janelas possíveis da série.

Esta abordagem transforma a análise: em vez de olhar ponto a ponto, examinamos janela a janela, capturando padrões temporais completos.

Espaço de Padrões

Interpretação Geométrica

Cada subsequência $X_{t:t+w-1}$ é interpretada como um vetor em \mathbb{R}^w , onde w é a dimensão do espaço.

Motifs aparecem como "nuvens" ou grupos de pontos próximos nesse espaço multidimensional. A geometria depende de como medimos distância entre vetores.

Esta visão geométrica permite aplicar técnicas de agrupamento e análise espacial para identificar padrões recorrentes.

Similaridade entre Subsequências

Motifs dependem fundamentalmente de uma noção de distância d entre subsequências. A métrica escolhida determina o que conta como "mesmo padrão".



Distância Euclidiana

$d_E(s_i, s_j) = \sqrt{\sum (s_i^{(k)} - s_j^{(k)})^2}$. Compara valor a valor na mesma posição.



DTW

Dynamic Time Warping permite desalinhamento temporal entre padrões similares.



Correlação

$d_{\text{corr}}(s_i, s_j) = 1 - \rho(s_i, s_j)$. Foca na forma, não no nível absoluto.

Mudar a métrica muda fundamentalmente quais padrões serão considerados recorrentes na análise.

Definição Matemática de Motif

Um motif é um conjunto de subsequências próximas entre si, onde proximidade é definida por uma distância d e um limiar ε .

Condição Básica

Duas subsequências formam um motif se:

$$d(s_i, s_j) < \varepsilon$$

Generalização para Grupos

Um motif \mathcal{M}_k é um conjunto onde todas as subsequências são mutuamente similares: $\mathcal{M}_k = \{s \in \mathcal{S} | d(s, s') < \varepsilon, \forall s' \in \mathcal{M}_k\}$

A definição começa com pares, mas generaliza para grupos, conectando diretamente com a etapa de agrupamento do pipeline.

Papel do Parâmetro w



w Pequeno

Captura padrões curtos e detalhes locais

w Médio

Equilibra escala e repetições

w Grande

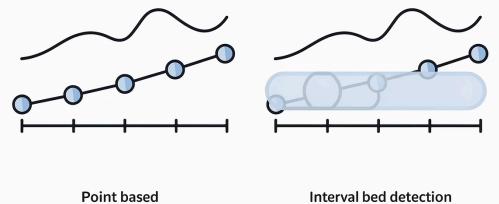
Captura estruturas longas, menos repetições

O tamanho da janela w controla a escala do padrão detectável. Padrões encontrados com w podem ser completamente diferentes dos encontrados com w' : $\mathcal{M}(w) \neq \mathcal{M}(w')$.

A escolha de w define o conceito operacional de motif, na prática, afetando quantas janelas existem e o custo computacional da análise.

Motifs Pontuais e Intervalares

Duas Abordagens

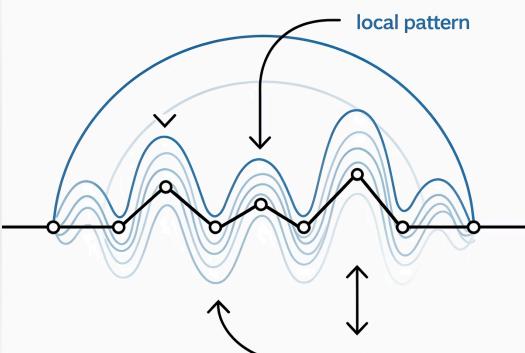


Com $w = 1$ (pontual), o "padrão" é um único valor que se repete - uma noção limitada de recorrência.

Com $w > 1$ (intervalar), o motif se torna um fenômeno temporal verdadeiro, envolvendo forma e dinâmica ao longo de vários passos.

Motifs em séries temporais são, na essência, intervalares, pois capturam padrões que se desenvolvem ao longo do tempo.

Escala Temporal dos Motifs



Motifs Locais

Aparecem em uma região específica da série $X_{t_0:t_1}$. Sugerem padrões típicos de um período, regime ou fase particular do processo.

Motifs Globais

Aparecem ao longo de toda a série $X_{1:T}$. Sugerem regularidade estrutural do processo como um todo, repetindo-se consistentemente ao longo do tempo.

A escala temporal muda fundamentalmente a interpretação: regime específico versus regularidade geral do processo.

Motifs Determinísticos e Estocásticos

Motifs podem ter diferentes origens: estrutura determinística explícita no sinal S_t , ou regularidade probabilística que emerge do processo estocástico.

Teste de Significância

Um padrão é considerado motif significativo quando: $\mathbb{P}(X_{t:t+w} \in \mathcal{M}) > \mathbb{P}_0$

Onde \mathbb{P} é a probabilidade sob o processo observado e \mathbb{P}_0 é a probabilidade sob um modelo de aleatoriedade de referência (baseline).

Esta comparação distingue padrões estruturais de coincidências aleatórias.

Motifs nos Componentes Temporais

Motifs podem ser buscados em diferentes componentes da série: valor observado, tendência ou volatilidade. Isso amplia o conceito para além do sinal bruto.



Motifs de Valor (M_{val})

Padrões recorrentes no nível observado x_t



Motifs de Tendência (M_{trend})

Padrões em $tr(x_t)$: subidas/descidas similares



Motifs de Volatilidade (M_{vol})

Padrões em $v(x_t)$: períodos de alta/baixa variância

O conjunto completo de motifs é a união: $M = M_{val} \cup M_{trend} \cup M_{vol}$. Esta abordagem conecta motifs com a engenharia temporal.

Estrutura Conceitual dos Motifs

Motifs emergem de três escolhas fundamentais que definem o procedimento operacional. Não existe "motif universal" - tudo depende destas definições.

01

Subsequências (S)

Definir as janelas de análise

02

Distância (d)

Escolher métrica de similaridade

03

Janela (w)

Determinar escala temporal

O resultado é sintetizado como: $\mathcal{M} = f(S, d, w)$, onde f representa o procedimento de detecção e agrupamento. Trocar qualquer elemento muda fundamentalmente o resultado.

Motifs e Periodicidade

Repetição em Defasagem Fixa

Em um processo periódico com período T_p , janelas separadas por T_p passos tendem a ser semelhantes:

$$X_{t:t+w-1} \approx X_{t+T_p:t+T_p+w-1}$$

Motifs capturam essa repetição localmente através das janelas, servindo como evidências locais da periodicidade global.

A definição por janelas torna concreta a noção abstrata de periodicidade, transformando-a em padrões detectáveis e mensuráveis.

Motifs e Sazonalidade

Sazonalidade representa repetição com período sazonal T_s , tipicamente ligada a calendário. Motifs podem ser definidos diretamente no componente sazonal.

Decomposição

$$x_t = \beta_t + \pi_t + \omega_t$$

Onde π_t é o componente sazonal

Recorrência Sazonal

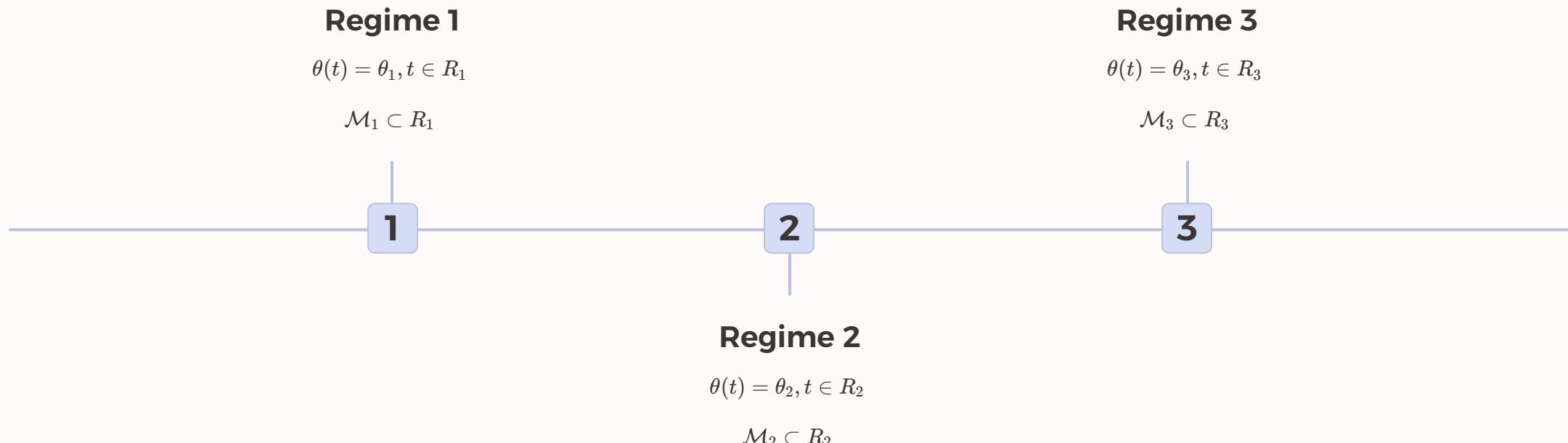
$$\pi_{t:t+w-1} \approx \pi_{t+T_s:t+T_s+w-1}$$

Janelas do componente sazonal separadas por T_s são muito parecidas

Isolar o componente sazonal permite interpretar motifs como padrões recorrentes associados diretamente à sazonalidade.

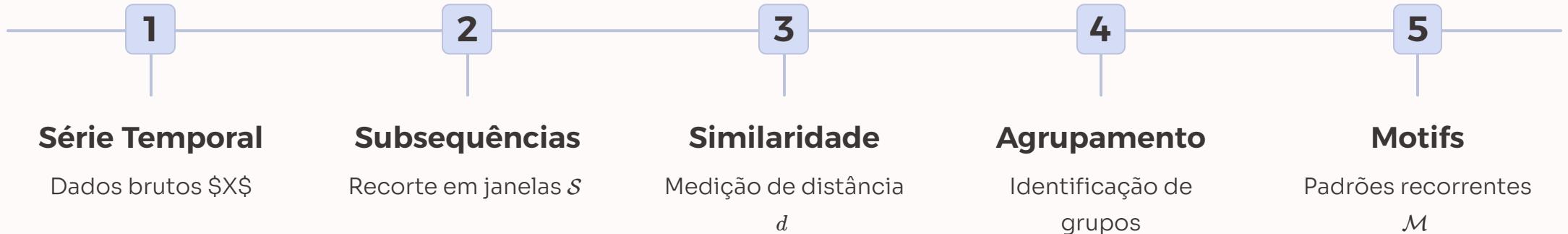
Motifs e Regimes

Regimes são definidos por parâmetros θ_i em intervalos R_i . Motifs podem ser específicos de um regime ou globais ao longo da série.



Motifs podem caracterizar "como é" cada regime, enquanto change points marcam "quando mudou". São conceitos complementares: regimes delimitam fases, motifs descrevem padrões típicos dentro delas.

Arquitetura Conceitual dos Motifs



Motifs surgem da geometria do espaço de subsequências. Esta cadeia conceitual conecta representação, regimes e padrões recorrentes em uma arquitetura unificada.

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.