

Por que Estender os Modelos Lineares

Desvendando a complexidade das séries temporais para uma análise mais profunda e precisa.

Eduardo Ogasawara

eduardo.ogasawara@cefet-rj.br

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

Limitações dos Modelos Lineares

Modelos ARIMA descrevem dependência temporal linear e supõem estacionariedade após diferenciação. Eles assumem variância constante ao longo do tempo, mas não capturam estruturas temporais complexas.

Séries reais exibem sazonalidade, não linearidade e variância variável. Modelos lineares são um ponto de partida poderoso, mas limitado. Quando a série apresenta padrões mais ricos — como ciclos, mudanças de variância ou relações não lineares — precisamos ampliar o modelo.

Sazonalidade como Estrutura Periódica

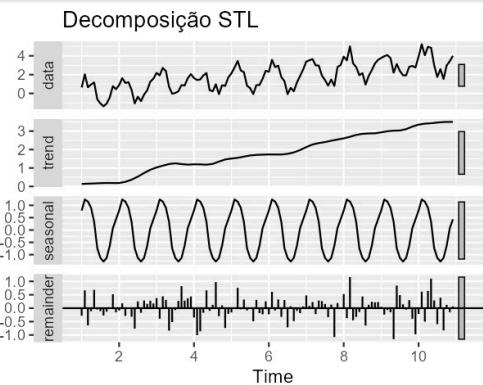
Sazonalidade representa a repetição regular de padrões no tempo. O comportamento estatístico se repete a cada s períodos.

Uma série é sazonal quando a média se repete ao longo do tempo:

$$E[X_t] = E[X_{t+s}]$$

A decomposição da série mostra três componentes distintos:

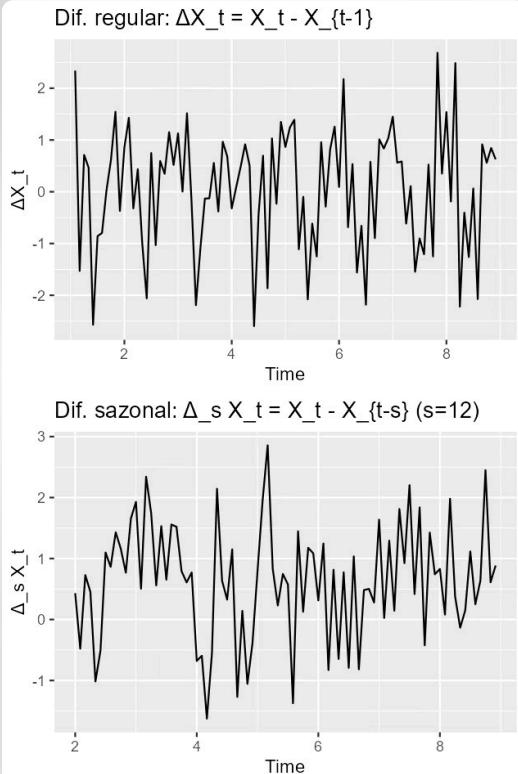
$$X_t = T_t + S_t + \varepsilon_t$$



Símbolos

- X_t = valor da série no tempo t
- T_t = componente de tendência
- S_t = componente sazonal
- ε_t = ruído aleatório
- s = período sazonal

Diferenciação Sazonal



Diferença Regular

Remove padrões persistentes e tendências lineares

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



Diferença Sazonal

Remove padrões periódicos que se repetem

$$\Delta_s X_t = X_t - X_{t-s}$$

Enquanto a diferença regular remove tendências, a diferença sazonal remove padrões periódicos. Se a série repete o comportamento a cada s períodos, subtrair X_{t-s} elimina esse padrão estrutural.

SARIMA: ARIMA com Estrutura Sazonal

SARIMA combina dinâmica regular e sazonal em um único modelo integrado. A forma conceitual do modelo integra componentes autorregressivos, diferenciação e médias móveis em duas escalas temporais:

$$(\text{parte AR regular} + \text{parte AR sazonal} + \text{diferenças}) = (\text{parte MA regular} + \text{parte MA sazonal} + \text{ruído})$$



Notação completa: **SARIMA(p, d, q)(P, D, Q)(s)**

O SARIMA estende o ARIMA ao incorporar dependência sazonal, descrevendo simultaneamente relações de curto prazo e padrões periódicos.

ACF e PACF em Séries Sazonais

A função de autocorrelação (ACF) mede correlação entre X_t e X_{t-h} . Em séries sazonais, há picos em múltiplos de s :



$$\rho(h) \text{ alto quando } h = k \cdot s$$

Onde $\rho(h)$ é a autocorrelação na defasagem h , k é um número inteiro e s é o período sazonal.

Se a série é sazonal, a autocorrelação reaparece nas defasagens múltiplas do período. Isso ajuda a identificar componentes sazonais AR e MA.

Identificação de Modelos SARIMA



Verificar Estacionariedade

Analisar se a série possui média e variância constantes ao longo do tempo

A série transformada é expressa como:

$$Y_t = \Delta^d \Delta_s^D X_t$$

Onde Y_t é a série transformada, d é a ordem de diferença regular e D é a ordem de diferença sazonal. A identificação do SARIMA segue a lógica de Box-Jenkins, mas considerando dois tipos de padrões: regulares e sazonais.



Aplicar Diferenças

Aplicar diferenças regulares e sazonais conforme necessário

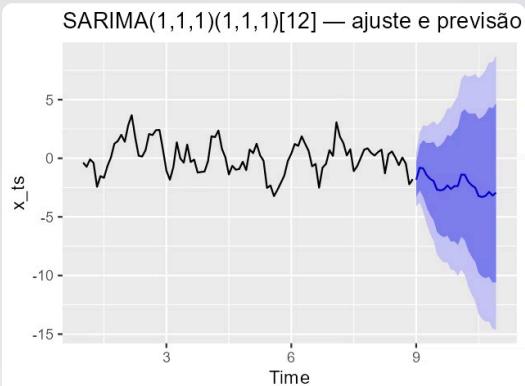


Analizar ACF e PACF

Examinar padrões de autocorrelação para identificar ordens do modelo

Exemplo de SARIMA

Considere o modelo SARIMA(1,1,1)(1,1,1)_s, que captura dependência local e sazonal simultaneamente.



AR Regular

Captura dependência de curto prazo entre observações consecutivas

AR Sazonal

Captura dependência periódica entre observações separadas por s períodos

Diferenças

Remove não estacionariedade regular e sazonal da série

Esse modelo combina quatro mecanismos fundamentais: dependência regular, dependência sazonal, diferença regular e diferença sazonal. Ele é uma generalização estrutural do ARIMA tradicional.

Regressão com Erros Temporais

A série pode depender de variáveis explicativas externas. O modelo combina regressão com dependência temporal:

$$X_t = \beta_0 + \beta \cdot Z_t + u_t$$

Símbolos

- X_t = série de interesse
- Z_t = vetor de variáveis explicativas
- β_0, β = coeficientes
- u_t = erro temporal

Aqui combinamos regressão com dependência temporal. O erro não é aleatório simples, mas segue um processo temporal (ARMA ou ARIMA).

Modelos Dinâmicos

O efeito das variáveis pode ser distribuído no tempo. O impacto de uma variável pode não ser imediato:

$$X_t = \beta_0 + \beta_0 Z_t + \beta_1 Z_{t-1} + \dots + \beta_k Z_{t-k} + u_t$$



Efeito Imediato

Impacto de Z_t no tempo presente

Efeitos Defasados

Impactos de Z_{t-1} , Z_{t-2} , etc.

Propagação Temporal

Efeitos acumulados ao longo do tempo

Onde Z_{t-k} é a variável explicativa defasada e β_k é o coeficiente da defasagem k . Modelos dinâmicos capturam efeitos retardados e propagação temporal.

SARIMA com Regressão

SARIMAX

Combinação de SARIMA e regressão, conhecida como SARIMAX
(SARIMA com variáveis exógenas)

Modelos modernos de séries temporais integram duas dimensões fundamentais: dinâmica própria da série e influência de fatores exógenos. Esta abordagem permite capturar tanto padrões internos quanto impactos de variáveis externas.

Duas Dimensões

ARIMA explica dependência interna, enquanto regressão explica efeitos externos

Heterocedasticidade Temporal

Em muitas séries, especialmente financeiras e econômicas, a variância muda ao longo do tempo:

$$\text{Var}(X_t) \neq \text{constante}$$



Variância Condisional

A variância depende da informação passada

$$\text{Var}(X_t | \text{passado}) = \sigma_t^2$$



Clusters de Volatilidade

Períodos de alta variância tendem a se agrupar no tempo

Onde σ_t^2 é a variância no tempo t . Em muitas séries, a média é estável, mas a variância muda significativamente.

Modelo ARCH

O modelo ARCH (Autoregressive Conditional Heteroskedasticity) modela a variância como função de choques passados:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \dots + \alpha_q \varepsilon_{t-q}^2$$

Símbolos

- ε_t = erro no tempo t
- σ_t^2 = variância condicional
- α_i = parâmetros do modelo

O ARCH modela clusters de volatilidade: grandes choques geram períodos de alta variância que persistem no tempo.

Modelo GARCH

O GARCH (Generalized ARCH) é uma extensão do ARCH. A variância depende de choques e variâncias passadas:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$



Termo ARCH

$\alpha_1 \varepsilon_{t-1}^2$ captura impacto de choques recentes



Termo GARCH

$\beta_1 \sigma_{t-1}^2$ captura persistência da volatilidade

Onde β_1 é o parâmetro da variância passada. O GARCH captura persistência da volatilidade: choques afetam a variância por longos períodos.

Propriedades do GARCH



Média Condicional Constante

A média esperada permanece estável ao longo do tempo



Variância Incondicional Finita

A variância de longo prazo converge para um valor finito



Caudas Pesadas

Maior probabilidade de eventos extremos comparado à distribuição normal



Clusters de Volatilidade

Períodos de alta volatilidade tendem a se agrupar temporalmente

O GARCH explica padrões empíricos típicos de séries financeiras, como alta volatilidade concentrada no tempo e distribuições com caudas mais pesadas que a normal.

Volatilidade e Eventos

Eventos podem alterar significativamente a variância da série:

$$\sigma_t^2 \neq \sigma_{t-h}^2$$

Onde h é a defasagem temporal.

Eventos não afetam apenas a média da série, mas também sua variância. Por isso, modelos de volatilidade são úteis para detecção de eventos e análise de risco.

- **Aplicação Prática:** Modelos GARCH são amplamente utilizados em finanças para gestão de risco, precificação de opções e detecção de eventos de mercado.

Domínio da Frequência

Domínio do Tempo

Observa a série ao longo do tempo: $X_t, X_{t+1}, X_{t+2} \dots$

Domínio da Frequência

Decompõe a série em componentes periódicas e ciclos

A análise no domínio da frequência revela ciclos e periodicidades ocultas. A ideia central é decompor a série em componentes periódicas, revelando estruturas que podem não ser evidentes no domínio do tempo.

Espectro e Autocovariância

Existe uma relação fundamental entre tempo e frequência. O espectro descreve como a variância se distribui entre diferentes frequências.

Autocovariância

Representa dependência temporal no domínio do tempo

Autocovariância e espectro são duas formas equivalentes de representar dependência temporal. A transformada de Fourier conecta essas duas representações.

Espectro

Representa dependência temporal no domínio da frequência

💡 NÃO LINEARIDADE

Modelos Não Lineares

Modelos não lineares capturam relações complexas no tempo. Um exemplo importante são modelos com regimes, onde a dinâmica muda conforme o estado do sistema.

A ideia geral é expressa como:

$$X_t = f(\text{passado}) + \text{erro}$$

Modelos de Limiar

Dinâmica muda quando a série cruza um limiar específico

Modelos de Markov

Transições probabilísticas entre diferentes regimes

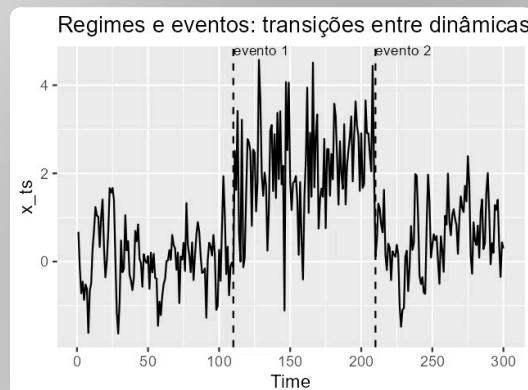
Transição Suave

Mudanças graduais entre diferentes dinâmicas

Onde $f(\cdot)$ é uma função não linear. Modelos não lineares permitem que a dinâmica da série mude conforme o estado do sistema.

Regimes e Eventos

Em modelos com múltiplos regimes, eventos correspondem a transições entre diferentes dinâmicas do processo:



$$k_t \neq k_{t-1}$$

Onde k_t representa o regime no tempo t .

Regime Normal

Baixa volatilidade e dinâmica estável

Recuperação

Retorno gradual ao regime normal



Transição

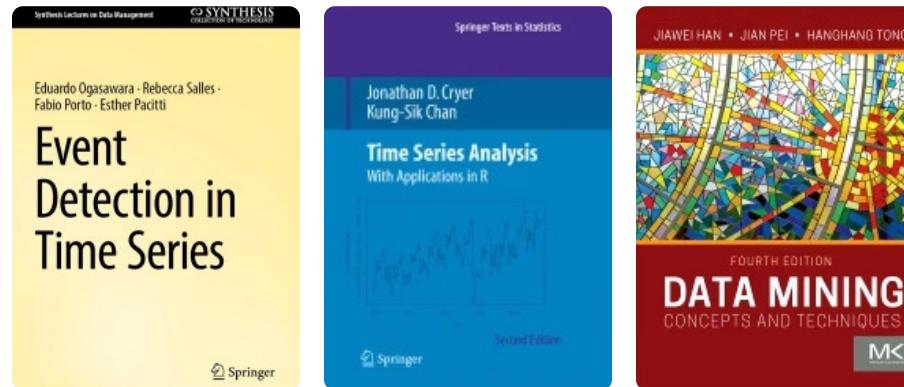
Mudança detectada no comportamento

Regime de Crise

Alta volatilidade e dinâmica alterada

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