Previsão em Séries Temporais

Uma série temporal é um processo aleatório que evolui ao longo do tempo. A previsão consiste em estimar valores futuros condicionados à informação disponível até o momento atual. O horizonte de previsão determina se estamos falando de curto, médio ou longo prazo.

Eduardo Ogasawara

eduardo.ogasawara@cefet-rj.br

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

CONCEITO FUNDAMENTAL

A Matemática da Previsão

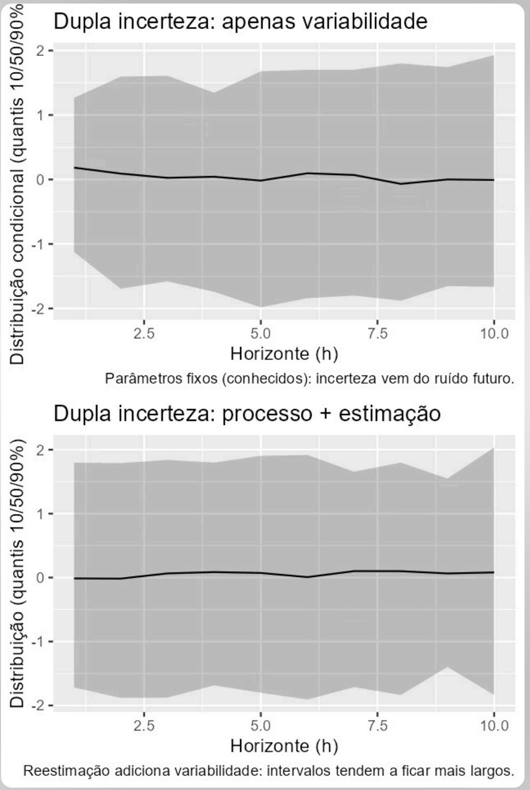
A previsão de um valor futuro é definida como a esperança condicional do processo no futuro, dada a informação disponível no presente:

$$\hat{X}_{t+h|t} = \mathbb{E}[X_{t+h} \mid \mathcal{F}_t]$$

Onde X_t representa o valor da série no tempo t, h é o horizonte de previsão, $\mathcal{F}_t$ é o conjunto de informações disponíveis até o tempo t, e $\mathbb{E}[\cdot]$ denota a esperança matemática.

Ideia Central

Prever significa estimar o futuro usando tudo o que sabemos no presente. A definição é probabilística e independe do modelo específico.



INFERÊNCIA

Previsão como Problema de Inferência

A previsão depende de um modelo ajustado aos dados observados. Este modelo possui parâmetros que precisam ser estimados, criando duas fontes de incerteza: a variabilidade natural do processo e a incerteza na estimação dos parâmetros.

Função de Previsão

A previsão é uma função dos dados observados e dos parâmetros estimados

Parâmetros Estimados

Usamos $\hat{\theta}$ para representar as estimativas dos parâmetros do modelo

Dupla Incerteza

Incerteza vem do comportamento aleatório da série e da estimação dos parâmetros

$$\hat{X}_{t+h|t} = g(X_t, X_{t-1}, \ldots; \hat{\theta})$$

Previsão Ótima e Erro Quadrático

O erro de previsão é definido como a diferença entre o valor observado e o valor previsto. A qualidade da previsão é medida pelo erro quadrático médio.

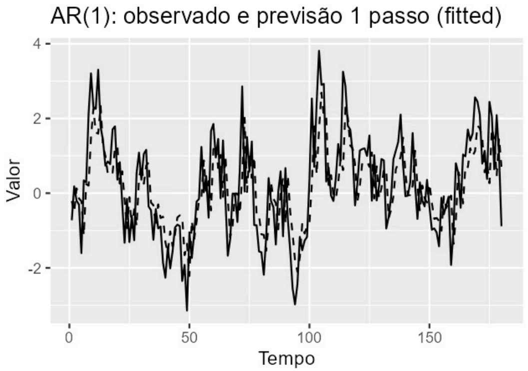
Erro de previsão:

$$e_{t+h} = X_{t+h} - \hat{X}_{t+h|t}$$

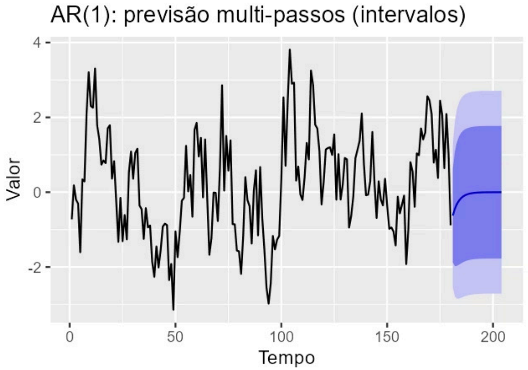
A previsão ótima minimiza:

$$\mathbb{E}[e_{t+h}^2]$$

Se queremos minimizar o erro quadrático médio, a melhor previsão possível é a esperança condicional. Este é um resultado fundamental da teoria de previsão.



Tracejado: \hat{X}_{t|t-1} (1 passo).



Intervalos aumentam com h: mais incerteza no longo prazo.

MODELO AR(1)

Previsão em Modelos Autoregressivos

Em modelos lineares com dependência do passado, a previsão é obtida substituindo valores futuros por suas esperanças. A estrutura é recursiva e elegante.

1

Modelo AR(1)

$$X_t = \phi X_{t-1} + \varepsilon_t$$

2

Previsão 1 passo

$$\hat{X}_{t+1|t} = \phi X_t$$

3

Previsão h passos

$$\hat{X}_{t+h|t} = \phi^h X_t$$

O valor futuro depende do último valor observado multiplicado pelo parâmetro ϕ do modelo. À medida que o horizonte aumenta, o efeito do passado é elevado a potências sucessivas de ϕ.

Estrutura Temporal da Série

Uma série temporal é composta por componentes estruturais distintas: tendência, sazonalidade e ruído. A previsão consiste em extrapolar essas componentes para o futuro.

Tendência

Movimento de longo prazo da série, representado por T_t

Sazonalidade

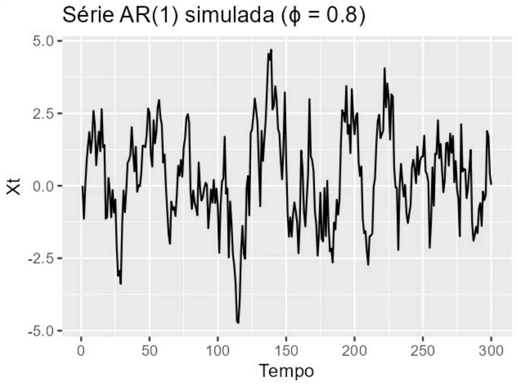
Padrões que se repetem em intervalos regulares, representado por S_t

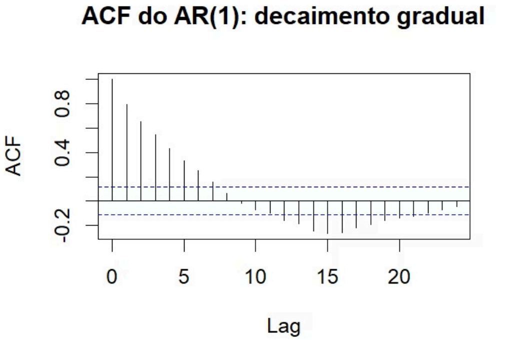
Componente Aleatória

Parte imprevisível da série, representada por ε_t

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

$$\hat{X}_{t+h|t} = \hat{T}_{t+h} + \hat{S}_{t+h} + \hat{\varepsilon}_{t+h}$$





MULTI-PASSOS

Previsão para Múltiplos Horizontes

Na previsão multi-passos, usamos previsões anteriores para calcular previsões mais distantes no futuro. A dependência temporal é propagada através do uso recursivo das previsões.

01

Um passo à frente

$$\hat{X}_{t+1|t} = \phi X_t$$

02

Dois passos à frente

$$\hat{X}_{t+2|t} = \phi \hat{X}_{t+1|t} = \phi^2 X_t$$

03

h passos à frente

$$\hat{X}_{t+h|t} = \phi^h X_t$$

Quanto maior o horizonte, mais a previsão depende do modelo e menos dos dados observados diretamente.

Propriedades do Erro de Previsão

O erro de previsão é definido como a diferença entre o valor observado e o valor previsto. Uma propriedade fundamental é que a esperança condicional do erro é zero.

$$\mathbb{E}[e_{t+h} \mid \mathcal{F}_t] = 0$$

Porém, a variância do erro aumenta com o horizonte de previsão, refletindo maior incerteza no futuro distante.

Interpretação: Se a previsão é a esperança condicional, o erro médio condicionado à informação presente é zero. A incerteza, no entanto, cresce com o tempo.

VARIÂNCIA

Variância do Erro no Modelo AR(1)

O erro de previsão resulta da acumulação de choques futuros. Em processos estacionários, a variância cresce até um limite, mostrando que a incerteza não explode indefinidamente.

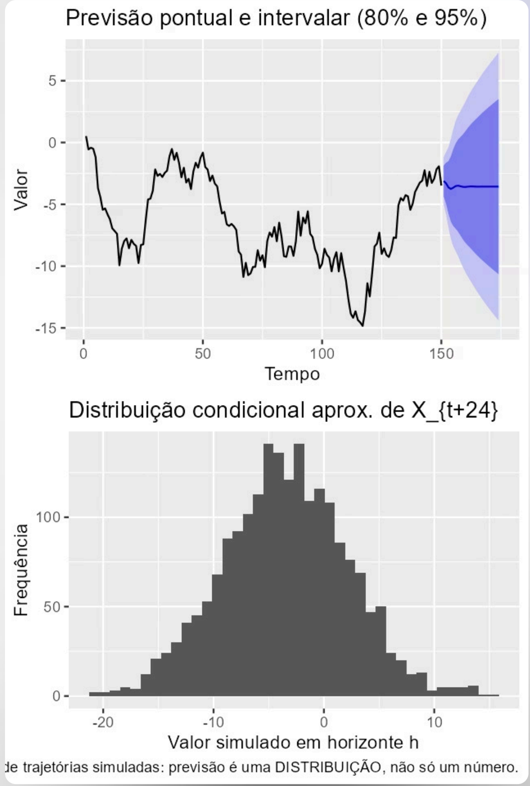
Erro Acumulado

$$e_{t+h} = \sum_{j=0}^{h-1} \phi^j \varepsilon_{t+h-j}$$

Variância do Erro

$$\mathrm{Var}(e_{t+h}) = \sigma^2 \sum_{j=0}^{h-1} \phi^{2j}$$

Onde σ^2 representa a variância do termo aleatório. Os choques futuros são ponderados pelo parâmetro do modelo, e sua contribuição para a variância total depende do horizonte de previsão.



Previsão Pontual e Intervalar

A previsão pontual fornece um único valor esperado, mas não captura toda a incerteza. O intervalo de previsão incorpora essa incerteza, mostrando a faixa provável dos valores futuros.

Previsão Pontual

Um único valor: a média esperada do processo futuro

Intervalo de Previsão

Faixa de valores prováveis baseada na incerteza do modelo

Base Probabilística

Traduz a incerteza em termos de probabilidade

$$\hat{X}_{t+h|t} \pm z_{\alpha/2}\sqrt{\text{Var}(e_{t+h})}$$

Onde $z_{\alpha/2}$ é o quantil da distribuição normal padrão e α é o nível de significância escolhido.

DISTRIBUIÇÃO

Previsão como Distribuição Condicional

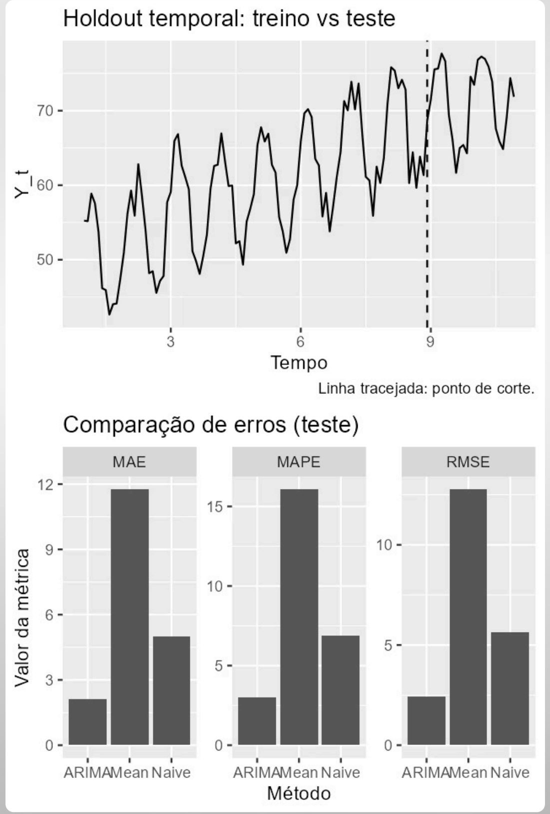
Em vez de pensar apenas em um número, a previsão deve ser vista como uma distribuição completa de possíveis valores futuros. A média, mediana ou quantis são apenas diferentes formas de resumir essa distribuição.

Distribuição condicional:

$$f(X_{t+h} \mid \mathcal{F}_t)$$

Onde $f(\cdot)$ representa a densidade de probabilidade condicional.

> Diferentes funções de perda geram diferentes previsões ótimas. A escolha depende do objetivo e do contexto da aplicação.



Métricas de Erro de Previsão

As métricas de erro permitem comparar modelos e avaliar qualidade preditiva. Cada métrica enfatiza um tipo de erro diferente, como magnitude média, variabilidade ou sensibilidade a valores extremos.

MSE - Erro Quadrático Médio

$$\mathrm{MSE} = \frac{1}{n}\sum_{t=1}^{n} e_t^2$$

Penaliza fortemente erros grandes

MAE - Erro Absoluto Médio

$$\mathrm{MAE} = \frac{1}{n}\sum_{t=1}^{n} |e_t|$$

Mais robusto a outliers, trata todos os erros de forma linear

MAPE - Erro Percentual

$$\mathrm{MAPE} = \frac{100}{n}\sum_{t=1}^{n} \left|\frac{e_t}{y_t}\right|$$

Expressa o erro como porcentagem do valor observado

FUNÇÕES DE PERDA

A Escolha da Função de Perda

Escolher uma métrica de erro é escolher implicitamente uma função de perda. Se usamos MSE, buscamos a média condicional; se usamos MAE, buscamos a mediana condicional. Não existe uma única previsão "melhor" em todos os contextos.

1. **Perda Quadrática**

$$L(e) = e^2$$

Leva à média como previsão ótima

2. **Perda Absoluta**

$$L(e) = |e|$$

Leva à mediana como previsão ótima

3. **Perda Assimétrica**

Penaliza erros positivos e negativos de forma diferente

Validação Temporal

Em séries temporais, não podemos embaralhar os dados. O tempo impõe restrições: o passado treina o modelo e o futuro testa a previsão. Métodos como janelas deslizantes e janelas expansivas respeitam essa estrutura temporal.

Separação Temporal

Dividir dados em treino e teste respeitando a ordem cronológica

Janela Deslizante

Mover a janela de treino ao longo do tempo

Janela Expansiva

Aumentar progressivamente o conjunto de treino

Previsão Recursiva

Gerar previsões sequenciais: $\hat{X}_{t+1|t}$ para $t = T_0, \ldots, T-1$

> COMPARAÇÃO

Comparação Estatística de Modelos

Comparar modelos exige mais do que olhar métricas médias. O teste de Diebold-Mariano avalia se a diferença de desempenho é estatisticamente significativa, considerando a dependência temporal dos erros.

Diferença de perdas:

Onde L(·) é a função de perda escolhida (MSE, MAE, etc.).

$$d_t = L(e_t^{(1)}) - L(e_t^{(2)})$$

Onde $e_t^{(1)}$ e $e_t^{(2)}$ representam os erros de dois modelos concorrentes.

A hipótese nula é de igualdade de desempenho preditivo. Rejeitá-la indica superioridade estatisticamente significativa de um modelo sobre o outro.

Overfitting em Séries Temporais

Overfitting ocorre quando o modelo ajusta ruído em vez de estrutura verdadeira. É um problema comum em modelos complexos e se manifesta quando o desempenho no treino é muito melhor que no teste.

Sintoma Principal

$$\text{Erro}_{\text{treino}} \ll \text{Erro}_{\text{teste}}$$

Causa

Modelo captura flutuações aleatórias do passado que não se repetem

Solução

Validação temporal rigorosa e regularização adequada

LIMITES TEÓRICOS

Os Limites da Previsão

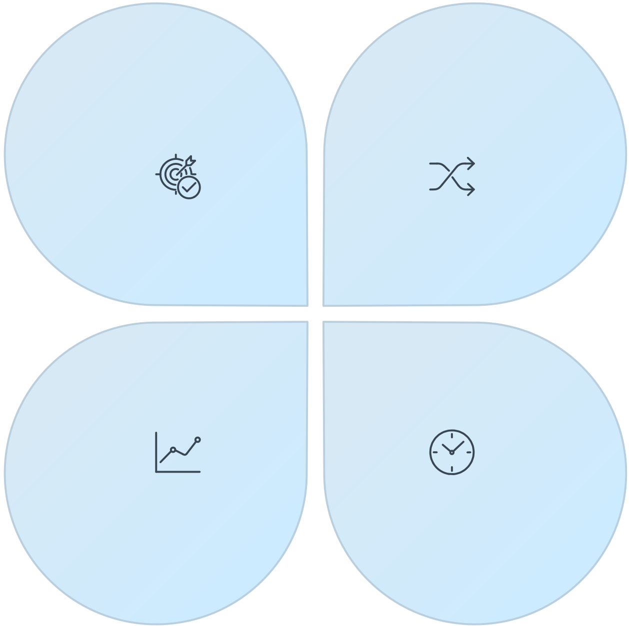
Sempre existe uma componente imprevisível em qualquer série temporal. Mesmo o melhor modelo possível tem um erro residual que não pode ser eliminado. Parte da variabilidade é inerente ao processo e define um limite teórico para a previsibilidade.

Componente Previsível

Estrutura sistemática que pode ser modelada

Ruído Irredutível

Variabilidade aleatória inerente ao processo

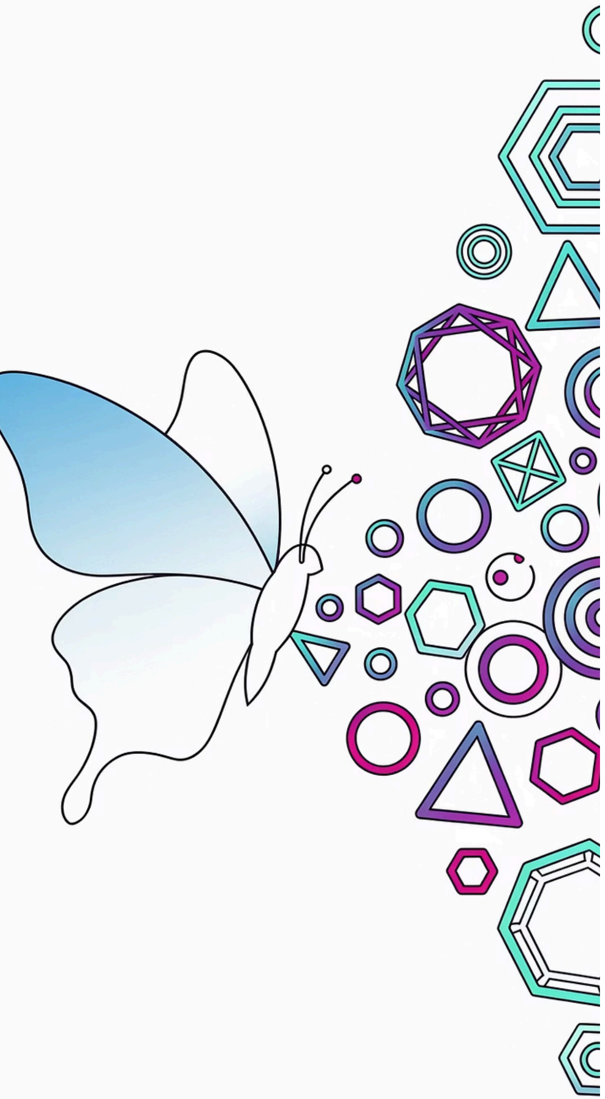


Decomposição

$$X_t = \hat{X}_{t|t-1} + \varepsilon_t$$

Horizonte Temporal

Previsibilidade diminui com o tempo



Previsibilidade e Complexidade do Sistema

A previsibilidade depende fundamentalmente da dinâmica do processo gerador da série. Processos simples são mais previsíveis, enquanto sistemas complexos têm previsibilidade limitada.

Sistemas Lineares

- Comportamento previsível
- Erros crescem linearmente
- Horizonte de previsão longo

Sistemas Caóticos

- Sensibilidade a condições iniciais
- Divergência exponencial
- Horizonte de previsão curto

$$|\delta X_t| \approx e^{\lambda t}|\delta X_0|$$

Onde λ representa a taxa de divergência do sistema. Pequenas diferenças iniciais podem gerar grandes divergências futuras.

ESTATÍSTICA VS ML

Abordagens Complementares

Apesar das diferenças metodológicas, estatística e machine learning resolvem o mesmo problema fundamental: aproximar a relação entre o passado e o futuro da série. A escolha depende do contexto e dos dados disponíveis.

Modelos Estatísticos

- Base probabilística explícita
- Alta interpretabilidade
- Inferência formal
- Intervalos de confiança rigorosos

Machine Learning

- Flexibilidade para não linearidades
- Capacidade de lidar com alta dimensionalidade
- Menos suposições sobre a estrutura
- Foco em desempenho preditivo

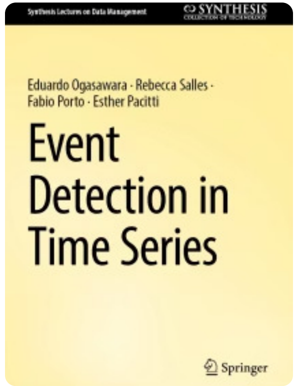
Função alvo comum:

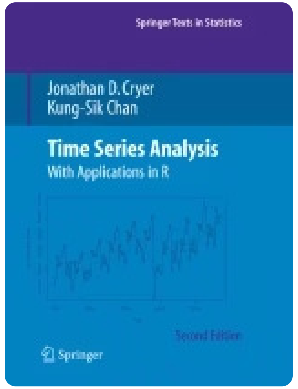
$$\mathbb{E}[X_{t+1} \mid \mathcal{F}_t]$$

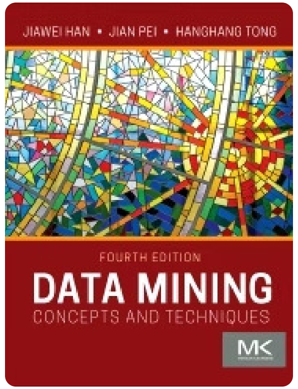
RECURSOS ACADÊMICOS

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