

ESPECIALIZAÇÃO
INTELIGÊNCIA ARTIFICIAL
& CIÊNCIA DE DADOS

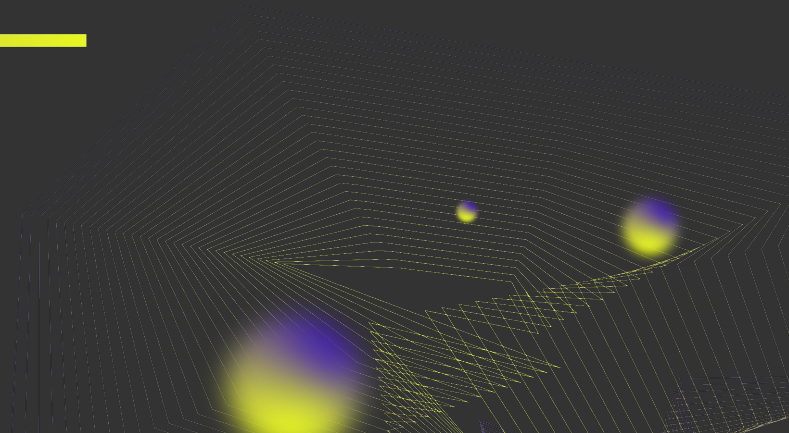
SEAD
UFES | Superintendência de
Educação a Distância

MÓDULO III

Séries temporais

Alexandre Loureiros Rodrigues

Professor do departamento de Estatística - UFES



ÍNDICE



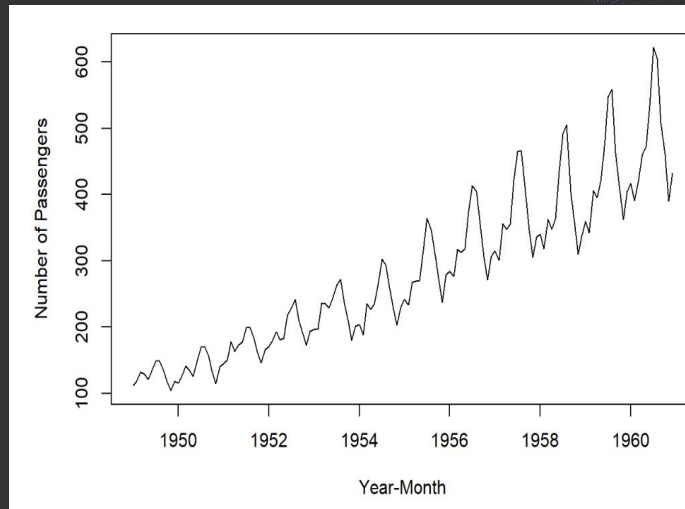
1. Introdução
2. Conceitos Básicos
3. Modelos AR, MA e ARMA
4. Métricas de avaliação



1. Introdução

Séries temporais

- Uma série temporal é um conjunto de observações coletadas sequencialmente ao longo do tempo.
- Cada ponto de dados é registrado em intervalos regulares (horários, diários, mensais, anuais etc.).
- Exemplos: temperatura diária, preços de ações, volume de vendas mensais.



Séries temporais

Característica	Séries Temporais	Regressão Clássica
Dependência Temporal	Observações são correlacionadas no tempo	Assume independência entre observações
Variáveis Explicativas	Pode depender apenas da própria variável ao longo do tempo	Usa variáveis preditoras (X) para prever Y
Objetivo	Modelar padrões temporais e fazer previsões futuras	Encontrar relações entre variáveis independentes e dependentes
Exemplos	Previsão de demanda, temperatura futura	Estimar impacto de preço na demanda de um produto

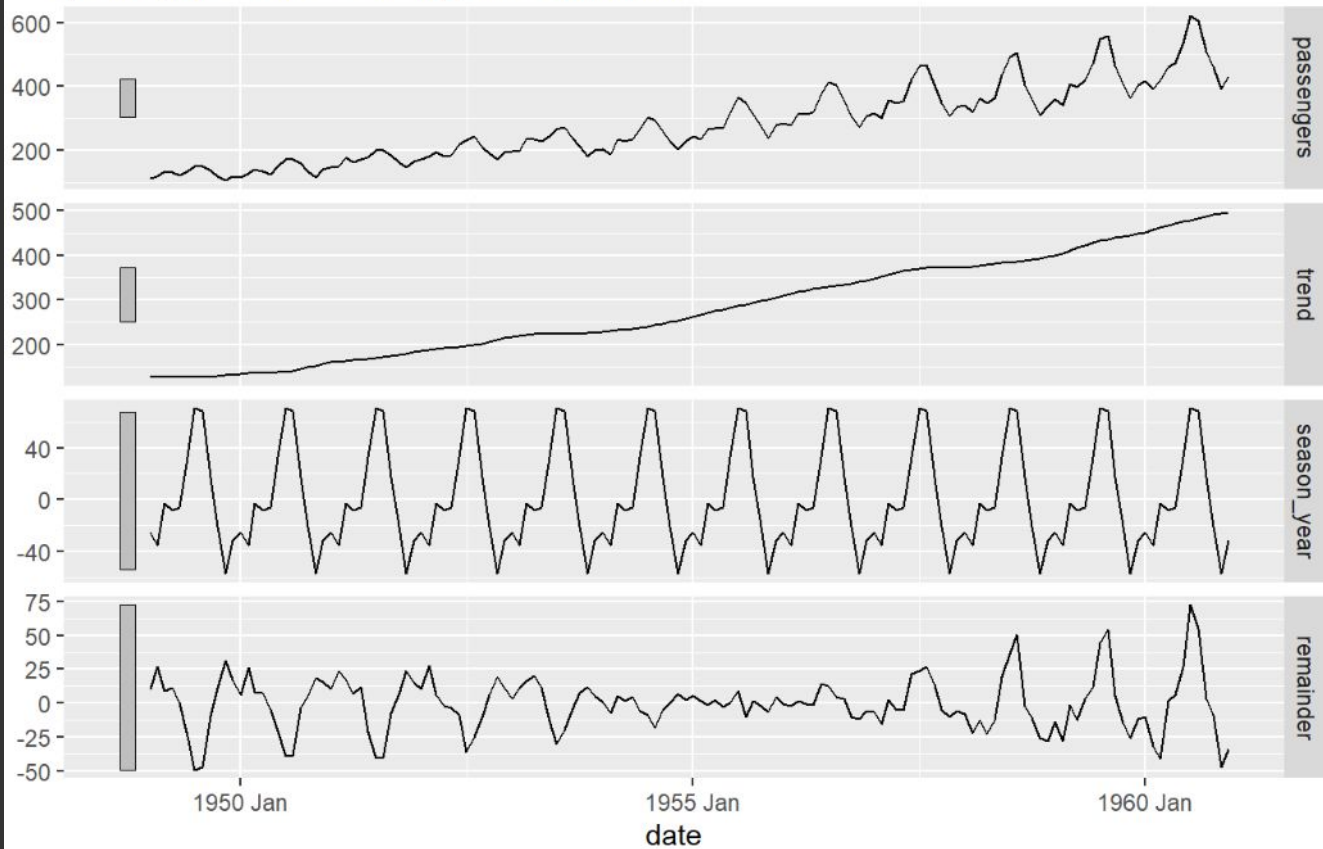


2. Conceitos básicos

Componentes série temporal

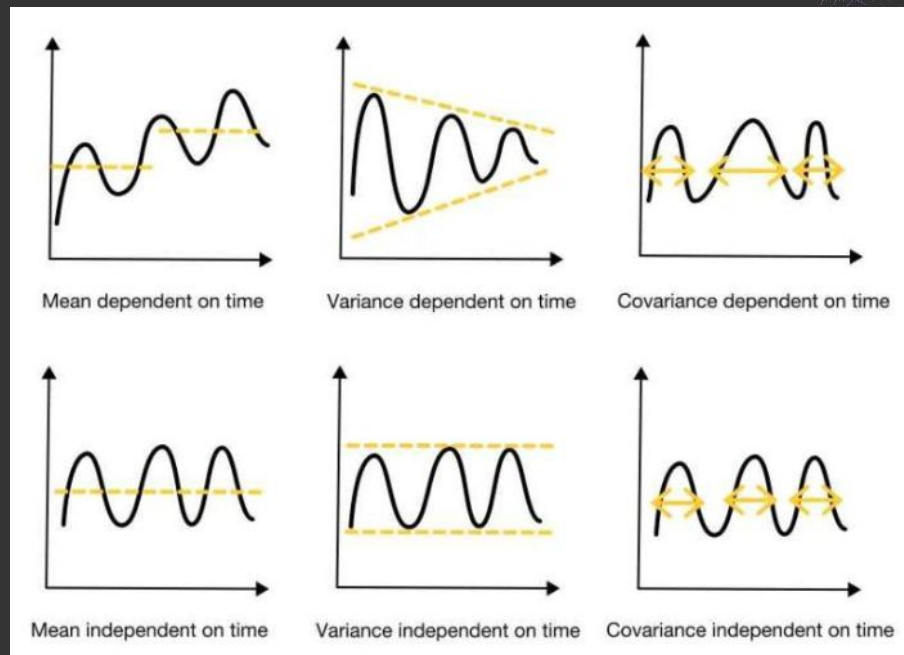
- **Tendência** : Movimento de longo prazo da série (crescente ou decrescente).
- **Sazonalidade**: Padrões que se repetem em intervalos fixos (diários, mensais, anuais).
- **Ciclos** : Flutuações de longo prazo sem periodicidade fixa, influenciadas por fatores externos.
- **Ruído** : Variações aleatórias sem padrão identificável.
- **Modelo**
 - Aditivo: $Y_t = T_t + S_t + C_t + R_t$
 - Multiplicativo $Y_t = T_t \cdot S_t \cdot C_t \cdot R_t$

passengers = trend + season_year + remainder



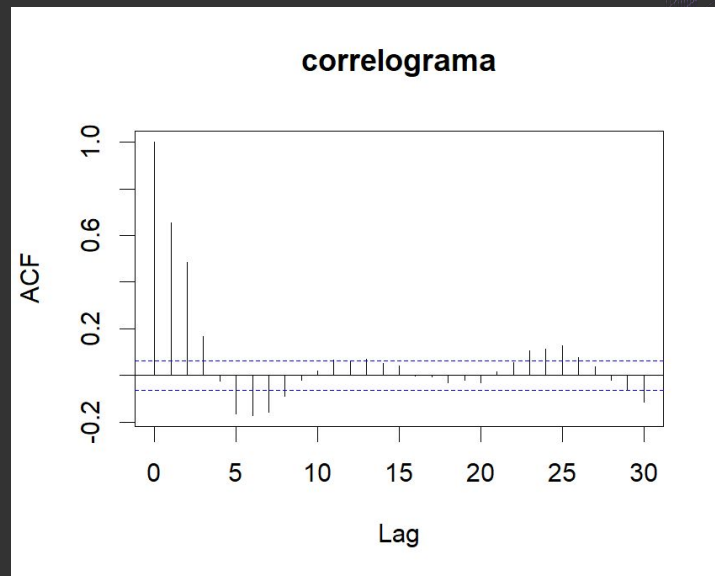
Conceitos Básicos

- Para cada tempo t , a série temporal é a realização de uma variável X_t
 - μ_t : média ($T_t + S_t + C_t$)
 - σ_t^2 : variância do processo
 - $\rho_{t,v}$: correlação entre X_t e X_v
- Série estacionária
 - $\mu_t = \mu$
 - $\sigma_t^2 = \sigma^2$
 - $\rho_{t,v}$ só depende de $|t - v|$



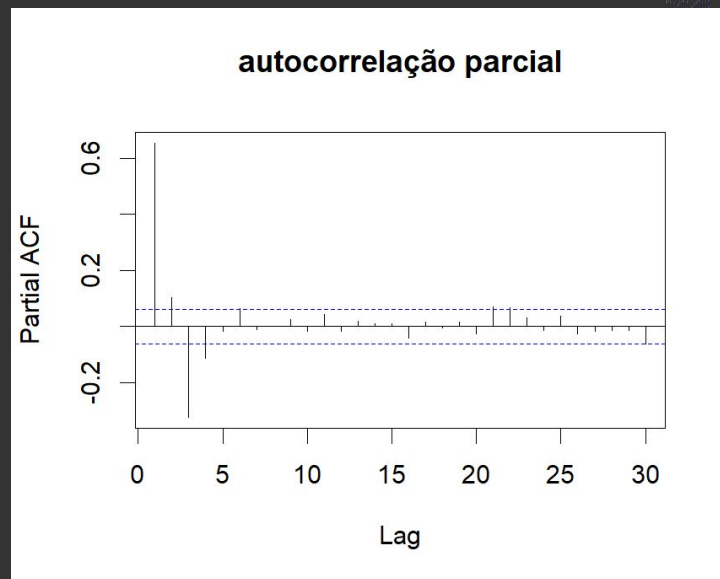
Autocorrelação

- Correlação de uma série temporal com ela própria (defasada no tempo)
- Se os valores passados influenciam os valores futuros, há autocorrelação
- Podemos calcular a autocorrelação para diferentes lags (tempos de defasagem) e plotar em um gráfico (correlograma)
- Usar somente em séries estacionárias



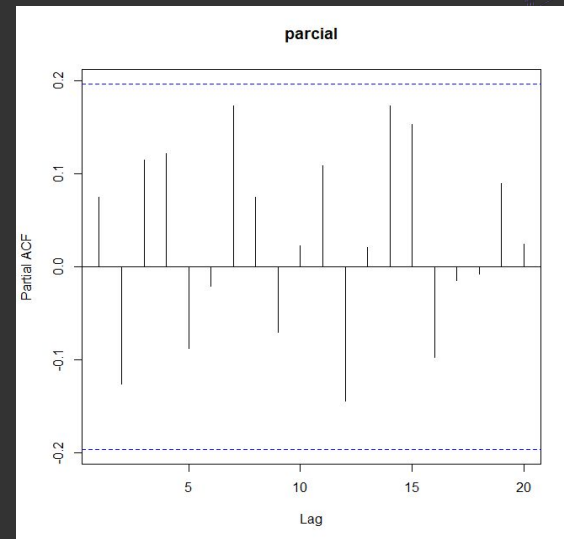
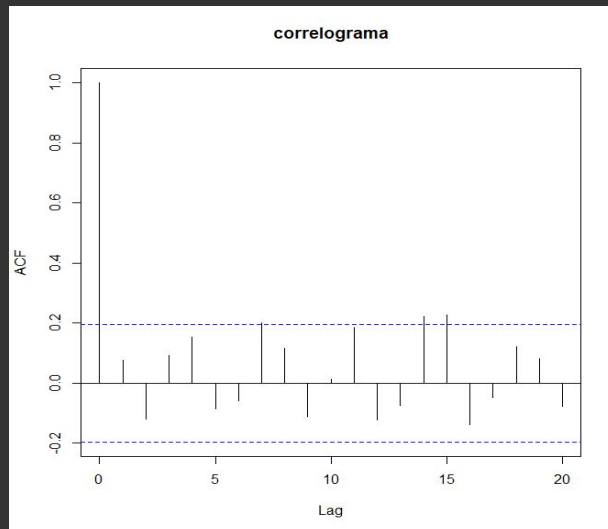
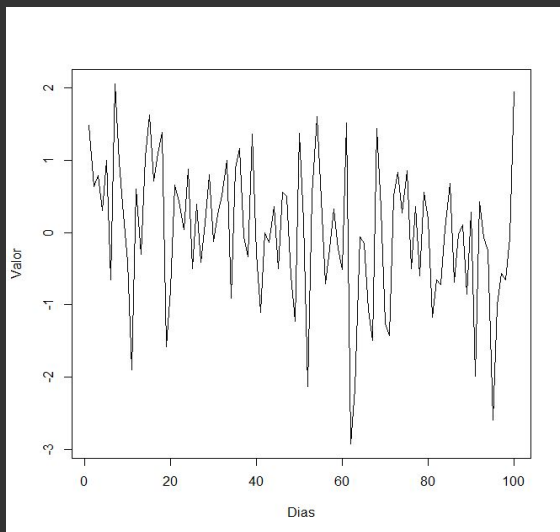
Autocorrelação parcial

- Mede a correlação entre um valor da série e seus valores defasados (**lags**), eliminando o efeito intermediário das defasagens anteriores.
- Ajuda a identificar a influência direta de cada lag sem considerar as influências indiretas.
- PAC e PACF serão utilizadas para utilizar os modelos de uma série temporal que iremos estudar
- Usar somente em séries estacionárias



Ruído Branco

- Série sem dependência temporal com média 0 e variância constante.
- Usado como base para: (i) construção de modelos e (ii) verificação do ajuste





3. Modelos



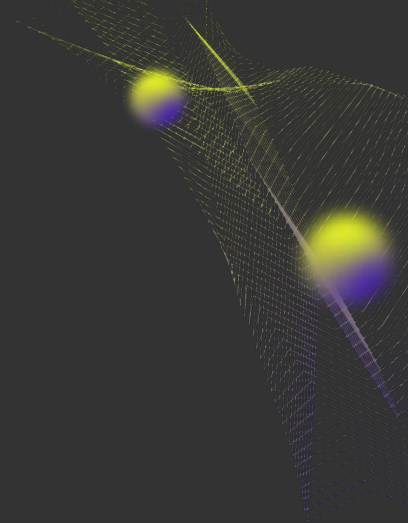
Modelo AutoRegressivo - AR(p)

- Um modelo **Autorregressivo (AR)** expressa um valor presente da série como uma combinação linear dos seus valores passados.

- Formalmente, um modelo **AR(p)** é escrito como

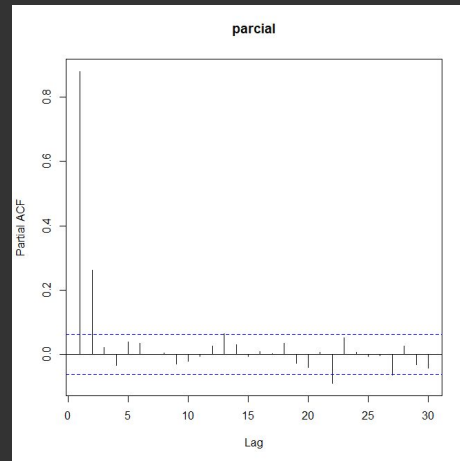
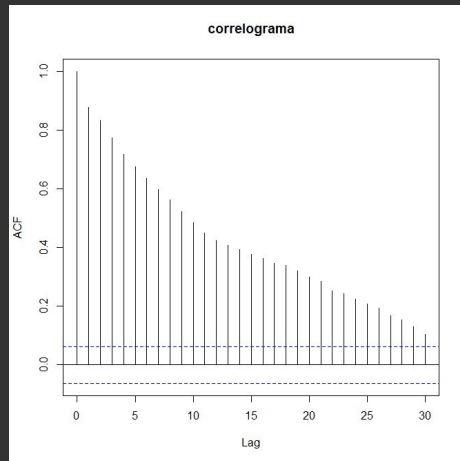
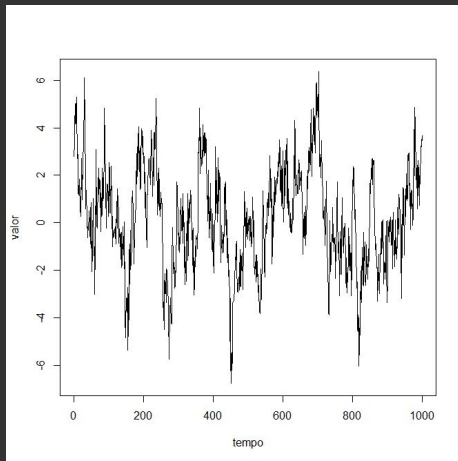
$$Y_t = \mu + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + e_t$$

- **Ordem p:** Número de defasagens (lags) consideradas no modelo.
- Modelo de regressão onde as covariáveis são os valores passados



Modelo AutoRegressivo - AR(p)

- Em um modelo **AR(p)**
 - ACF decai lentamente
 - PACF corta abruptamente após p lags





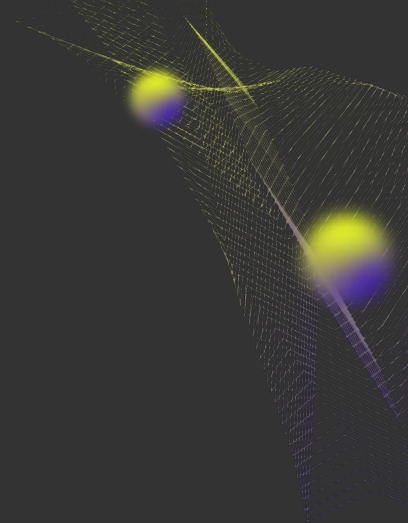
Modelo Médias móveis - MA(q)

- Um modelo **Médias móveis (MA)** expressa um valor presente da série como uma combinação linear dos seus erros passados.

- Formalmente, um modelo **MA(q)** é escrito como

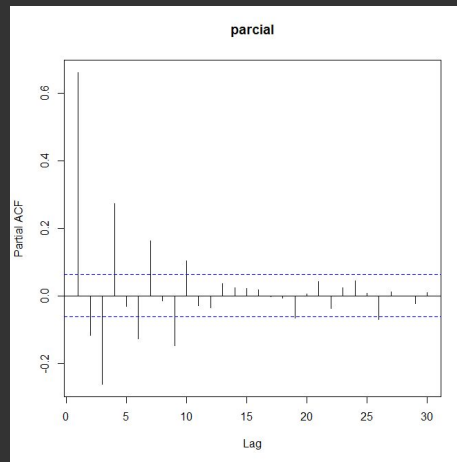
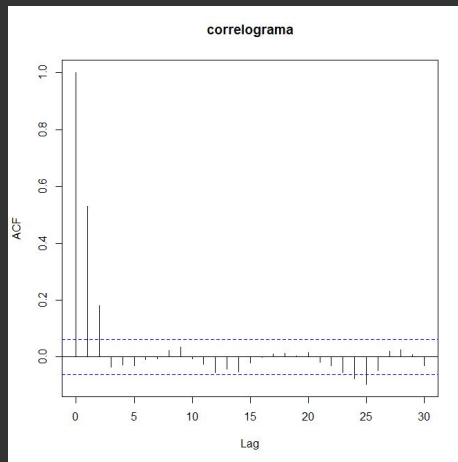
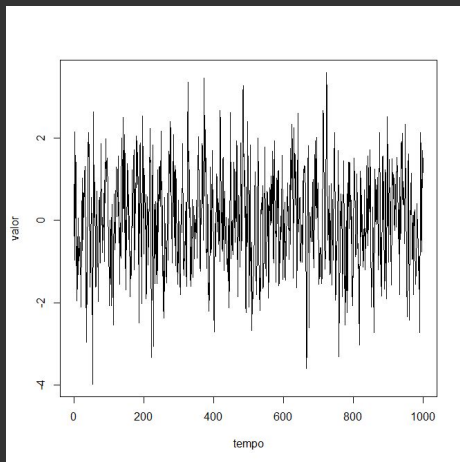
$$Y_t = \mu + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \dots + \theta_q e_{t-q} + e_t$$

- **Ordem q:** Número de defasagens (lags) consideradas no modelo.
- Modelo de regressão onde as covariáveis são os erros passados



Modelo Médias Móveis - MA(q)

- Em um modelo **MA(q)**
 - PACF decai lentamente
 - ACF corta abruptamente após q lags

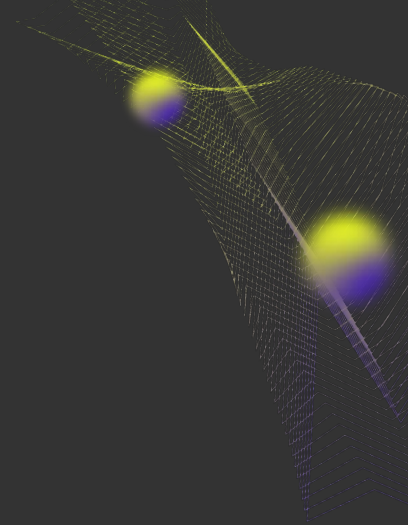




Modelo ARMA(p,q)

- Combinação dos modelos **AR(p)** e **MA(q)**
 - Captura relações mais complexas
 - O componente AR captura dependências de valores passados.
 - O componente MA captura dependências de erros passados.
- Ordens **p** e **q** podem ser baseadas na ACF e PACF, mas visualização é um pouco mais mascarada

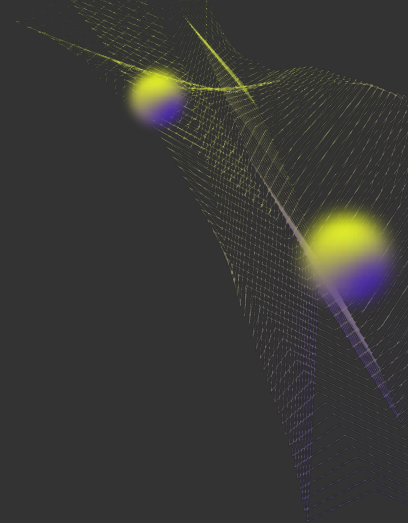
$$Y_t = \mu + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \theta_1 e_{t-1} + \dots + \theta_q e_{t-q} + e_t$$





Estimação e previsão

- Estimação dos parâmetros pode ser feita por mínimos quadrados
 - Verificar os resíduos para ver se toda estrutural temporal foi capturada
 - Os resíduos devem se comportar como um ruído branco
- Evitar previsões de muitos passos a frente
- Atualizar as previsões conforme novas observações ficam disponíveis





3. Métrica de avaliação

Métricas de avaliação

- Mesma de problemas de regressão

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i|$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right| \times 100$$

\hat{Y}

\hat{Y}

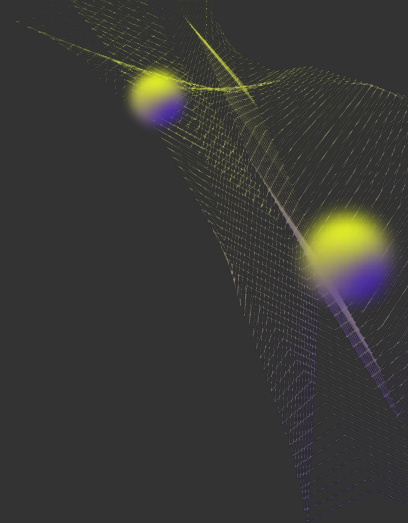


Comparação de modelos

- Existem vários outros modelos para fazer previsões de séries temporais
 - Modelos estilo ARMA mais complexos
 - Alisamento exponencial
 - Redes neurais
 - Entre outros
- Necessário procedimento de validação cruzada para comparar modelos
 - Note que não podemos dividir os dados de forma aleatória
 - Isto quebra estrutura temporal

\hat{Y}

\hat{Y}



Comparação de modelos

Validação cruzada



\hat{Y}



Séries não estacionárias

- Necessário medidas para retirar tendência / sazonalidade
 - Aplicar diferenças
 - $\Delta Y_t = Y_t - Y_{t-1}$
 - $\Delta^s Y_t = Y_t - Y_{t-s}$
 - Usar modelos de regressão tradicionais para μ_t
 - Usar ARMA nos resíduos $r_t = Y_t - \widehat{\mu}_t$
- Usar transformação de Y_t para estabilizar variância
 - Transformações log e raiz quadrada são as mais tradicionais
- Modelos mais complexos

