Disciplina:

SÉRIES TEMPORAIS

Professora: Patrícia de Sousa Ilambwetsi

Bacharel em Estatística pela UFOP

Mestre em Estatística pela UFV

Doutora em Estatística pela UFV

Especialista em Ciências de Dados e Big Data pela PUC Minas

PUC Minas

PUC Minas

Roteiro

Unidade 4 – Modelagem de Séries temporais não Estacionárias

PUC Minas

4 Modelagem de Séries Temporais não Estacionárias

4.1 Modelo Autoregressivo Integrado de Média Móvel (ARIMA)

A modelagem de séries temporais pelo modelo autoregressivo integrado de médias móveis (ARIMA) é uma técnica estatística usada para analisar e prever observações coletadas ao longo do tempo que apresentam tendência.

O modelo ARIMA combina três componentes principais:

⇒ AR (Autoregressivo): refere-se à parte autoregressiva do modelo em que os valores passados da série são usados para prever o valor atual.

⇒ I (Integração): refere-se ao número de diferenciações aplicadas à série para torná-la estacionária.

⇒MA (Média Móvel): modela a série temporal como uma média ponderada dos erros passados (resíduos).

PUC Minas

4 Modelagem de Séries Temporais não Estacionárias

4.1 Modelo Autoregressivo Integrado de Média Móvel (ARIMA)

$\Rightarrow$ **Componente Autoregressivo (AR)**

O parâmetro p indica o número de lags (valores passados) que serão usados para prever o valor atual e se $\boldsymbol{p = 1}$ **indica que o valor atual da série depende linearmente do valor imediatamente anterior.**

- A equação básica do modelo $AR(p)$ é

$$Y_t = c + \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + \varepsilon_t$$

em que: Y_t o valor da série no tempo t; c é uma constante; $\phi_1, \cdots, \phi_p$ coeficientes dos termos autorregressivo que medem a influência dos valores passados; ε_t termo de erro sem autocorrelação com média zero e variância constante (ruído branco) .

4 Modelagem de Séries Temporais não Estacionárias

4.1 Modelo Autoregressivo Integrado de Média Móvel (ARIMA)

⇒ Componente Autoregressivo (AR)

Interpretação do coeficiente ϕ: cada coeficiente $\boldsymbol{\phi_i}$ mede a contribuição do valor defasado Y_{t-i} para o valor atual Y_t com $i = 1, \ldots, p$

- **Para $\boldsymbol{\phi_i}$ positivo:** existe uma relação linear positiva entre valores passados e presente. Um alto valor no passado tende a ser seguido por um alto valor no presente.
- **Para $\boldsymbol{\phi_i}$ negativo:** existe uma relação linear negativa entre os valores passados e presente. Um valor alto no passado tende a ser seguido por um valor baixo no presente.

4 Modelagem de Séries Temporais não Estacionárias

4.1 Modelo Autoregressivo Integrado de Média Móvel (ARIMA)

⇒ **Componente de Integração (I)**

Refere-se ao número de diferenciações aplicadas à série para torná-la estacionária, o que significa remover tendências ou sazonalidades.

A ordem de integração ($\boldsymbol{d}$) indica quantas vezes os dados foram diferenciados

A equação básica é

$$Y_t{}^* = \nabla^d Y_t$$

em que: $Y_t{}^*$ é a série diferenciada tempo t; Y_t é a série original, ∇^d é a diferenciação de ordem d e $\nabla^d Y_t = (1 - B)^d Y_t$, B é o operador de defasagem (lag), tal que $BY_t = Y_{t-1}$

4 Modelagem de Séries Temporais não Estacionárias

4.1 Modelo Autoregressivo Integrado de Média Móvel (ARIMA)

⇒ **Componente de Integração (I)**

- **Para $\boldsymbol{d = 0}$**, a série já é estacionária e não requer diferenciação.
- **Para $\boldsymbol{d = 1}$**, uma diferenciação foi aplicada, sugerindo que a série tinha uma tendência linear, e ao diferenciar uma vez, essa tendência foi removida.
- **Para $\boldsymbol{d = 2}$**, duas diferenciações foram necessárias, indicando uma tendência quadrática ou de segunda ordem.

Exemplo: Em um modelo ARIMA(0,1,0), o parâmetro $d = 1$ indica que a série temporal foi diferenciada uma vez para remover uma tendência linear e torná-la estacionária.

PUC Minas

4 Modelagem de Séries Temporais não Estacionárias

4.1 Modelo Autoregressivo Integrado de Média Móvel (ARIMA)

⇒ **Componente Média Móvel (MA)**

Em um modelo de média móvel de ordem q, o valor atual da série Y_t é representado como uma média ponderada dos q erros passados (ou choques) mais um termo de erro ε_t .

- A equação básica do modelo $MA(q)$ é

$$Y_t = c + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$$

em que: Y_t o valor da série no tempo t; c é uma constante; $\theta_1, \cdots, \theta_q$ coeficientes do modelo MA que que medem a influência dos erros passados sobre o valor atual; ε_t termo de erro no tempo t sem autocorrelação com média e variância zero (ruído branco); ε_{t-q} são os erros passados (ruído branco).

4 Modelagem de Séries Temporais não Estacionárias

4.1 Modelo Autoregressivo Integrado de Média Móvel (ARIMA)

⇒ **Componente Média Móvel (MA)**

Interpretação do coeficiente $\boldsymbol{\theta}$: cada coeficiente θ_i reflete o impacto que o erro ε_{t-i} no tempo $t-i$ tem sobre o valor atual Y_t em que $i = 1, \ldots, q$

- **Para $\boldsymbol{\theta_i}$ positivo:** indica que um erro positivo no período anterior tende a aumentar o valor atual de Y_t
- **Para $\boldsymbol{\theta_i}$ negativo:** indica que um erro positivo no período anterior tende a diminuir o valor atual de Y_t

4 Modelagem de Séries Temporais não Estacionárias

4.1 Modelo Autoregressivo Integrado de Média Móvel (ARIMA)

⇒ **Estrutura do Modelo ARIMA**

O modelo ARIMA é a combinação do componente AR(p), I(d) e MA (q) denotado por:

$$(1 - B)^d Y_t = c + \sum_{i=1}^{p} \phi_i Y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t$$

em que p é a ordem do componente autoregressivo (AR) e q é a ordem do componente de média móvel (MA), c é a constante, B o operador de defasagem.

PUC Minas

4 Modelagem de Séries Temporais não Estacionárias

4.1 Modelo Autoregressivo Integrado de Média Móvel (ARIMA)

⇒ **Pressuposições do Modelo ARIMA**

- **Estacionariedade:** a média, a variância e a autocorrelação devem ser constantes ao longo do tempo.

  Implicações: para série não estacionária se faz necessário aplicar a transformação de diferenciação na série.

- **Linearidade:** assume uma relação linear entre as observações passados com as observações atuais da série

  Implicações: relações não lineares o modelo ARIMA torna-se inadequado resultando em previsões menos precisas.

- **Tendência:** é capaz de modelar por meio do componente de diferenciação.

PUC Minas

4 Modelagem de Séries Temporais não Estacionárias

4.1 Modelo Autoregressivo Integrado de Média Móvel (ARIMA)

⇒ **Pressuposições do Modelo ARIMA**

- **Normalidade dos Resíduos:** os resíduos devem seguir uma distribuição normal, com média zero e variância constante.
- **Independência dos resíduos**: os resíduos devem ser independentes e não autocorrelacionados.

  Implicações: Se houver autocorrelação nos resíduos, o modelo pode não estar capturando toda a informação da série temporal.
- **Homocedasticidade dos Erros:** variância constante ao longo do tempo.

  Implicações: se a variância dos erros muda ao longo do tempo se faz necessário utilizar outros modelos para capturar essa característica, como o GARCH

PUC Minas

4 Modelagem de Séries Temporais não Estacionárias

4.1 Modelo Autoregressivo Integrado de Média Móvel (ARIMA)

⇒ **Ajuste do Modelo ARIMA**

O ajuste de um modelo ARIMA (AutoRegressive Integrated Moving Average) é fundamental na análise de séries temporais para capturar o comportamento dos dados e fazer previsões.

Etapas a seguir:

1) **Verificação da Estacionariedade**
2) **Identificação do modelo**
3) **Estimação dos parâmetros**
4) **Validação do Modelo**

4 Modelagem de Séries Temporais não Estacionárias

4.1 Modelo Autoregressivo Integrado de Média Móvel (ARIMA)

⇒ **Ajuste do Modelo ARIMA**

1) Verificação Estacionariedade

Visualização gráfica da Série Temporal: Examine a série temporal original para identificar tendências, sazonalidades, e padrões que possam indicar a necessidade de diferenciação ou ajustes nos parâmetros.

Teste de Estacionariedade: Use testes estatísticos, como o **teste de Dickey-Fuller aumentado (ADF)** para verificar se a série é estacionária. Se não for, determine o grau de diferenciação necessário (d)

Remoção de Sazonalidade: Se houver sazonalidade evidente na série, considere usar um modelo SARIMA (Seasonal ARIMA)

4 Modelagem de Séries Temporais não Estacionárias

4.1 Modelo Autoregressivo Integrado de Média Móvel (ARIMA)

⇒ **Ajuste do Modelo ARIMA**

2) Identificação do modelo

Consiste em identificar a estrutura do modelo, ou seja, conhecer os valores dos parâmetros p e q

- **Análise da Função de Autocorrelação (ACF)**

Picos significativos na ACF determinam a ordem do componente q do modelo MA

- **Análise da Função de Autocorrelação Parcial (PACF)**

Picos significativos na PACF identificam a ordem do componente p do modelo AR

4 Modelagem de Séries Temporais não Estacionárias

4.1 Modelo Autoregressivo Integrado de Média Móvel (ARIMA)

⇒ **Ajuste do Modelo ARIMA**

2) Identificação do modelo

- **Modelo ARIMA Inicial**: começar com uma combinação inicial de p, d, e q baseada nas análises ACF (Função de autocorrelação) e PACF (Função de autocorrelação parcial).

- **Critérios de Informação**: Use critérios como **AIC (Akaike Information Criterion)** ou **BIC (Bayesian Information Criterion)** para comparar diferentes modelos ARIMA. O modelo com o menor AIC ou BIC é geralmente preferido, pois esses critérios penalizam a complexidade do modelo.

4 Modelagem de Séries Temporais não Estacionárias

4.1 Modelo Autoregressivo Integrado de Média Móvel (ARIMA)

⇒ **Ajuste do Modelo ARIMA**

3) Estimação dos parâmetros

Método de Máxima Verossimilhança: verifique se os coeficientes estimados para os termos AR e MA são estatisticamente significativos.

Implicações: Coeficientes não significativos podem indicar que o termo correspondente pode ser removido ou que o modelo precisa ser reespecificado

Hipóteses:

H_0: o parâmetro não é significativo

H_1: o parâmetro é significativo

p-valor < nível de significância de 5% rejeita-se H_0

4 Modelagem de Séries Temporais não Estacionárias

4.1 Modelo Autoregressivo Integrado de Média Móvel (ARIMA)

⇒ **Ajuste do Modelo ARIMA**

4) Validação do Modelo

- **Análise dos Resíduos**

Resíduos: Verificar se são como <u>ruído branco</u> (independentes, normalmente distribuídos com média zero e variância contante)

- **ACF/PACF dos Resíduos**: verificar a autocorrelação significativa. Se houver, o modelo pode precisar ser reespecificado
- **Teste de Normalidade (não obrigatório):** teste de Shapiro-Wilk, Gráfico QQplot, Histograma
- **Homocedasticidade**: Teste Breusch-Pagan, White, ARCH, Gráfico resíduos vs tempo ou valores ajustados

PUC Minas

4 Modelagem de Séries Temporais não Estacionárias

4.1 Modelo Autoregressivo Integrado de Média Móvel (ARIMA)

⇒ **Refinamento do Modelo ARIMA**

- **Ajuste dos Parâmetros $\boldsymbol{p}$, $\boldsymbol{d}$, $\boldsymbol{q}$**: com base na análise dos resíduos e nos critérios de informação (AIC, BIC), refine os valores de p, d e q para melhorar o ajuste do modelo
- **Considerando-se Modelos Alternativos:** se o modelo ARMA não é suficiente, considere modelos mais complexos, como ARIMA (que inclui diferenciação) ou SARIMA (que inclui sazonalidade).

PUC Minas

4 Modelagem de Séries Temporais não Estacionárias

4.1 Modelo Autoregressivo Integrado de Média Móvel (ARIMA)

⇒ **Previsões do Modelo ARIMA**

Avaliar a previsão

- **Erro de Previsão**: Calcule métricas como o RMSE (Root Mean Square Error) ou MAE (Mean Absolute Error) para avaliar a precisão das previsões.
- **Cross-validation:** Outra técnica é a validação cruzada, onde a série temporal é dividida em duas partes e o modelo é treinado em uma parte e testado em outra.
- **Ajuste Iterativo**: Caso o desempenho preditivo não seja satisfatório, continue refinando o modelo ajustando os parâmetros ou considerando a inclusão de componentes sazonais, se aplicável.

PUC Minas

4 Modelagem de Séries Temporais não Estacionárias

4.1 Modelo Autoregressivo Integrado de Média Móvel (ARIMA)

⇒ **Vantagem de um Modelo ARIMA**

- **Flexibilidade**: pode ser adaptado a uma grande variedade de séries temporais estacionária ou transformada em estacionária através da diferenciação
- **Capacidade de Capturar Dinâmicas Complexas**: pode capturar as dinâmicas autoregressivas (AR) e de média móvel (MA) da série permitindo que ele modele comportamentos complexos
- **Modelagem Explícita de Estruturas Temporais**: ao decompor a série em componentes autoregressivos, de média móvel e de integração oferece estrutura clara para entender e modelar as dependências temporais.
- **Boa Performance para Séries Temporais Curtas e Médias:** é eficaz na previsão de séries temporais de curta a média duração.

4 Modelagem de Séries Temporais não Estacionárias

4.1 Modelo Autoregressivo Integrado de Média Móvel (ARIMA)

⇒ **Desvantagem do Modelo ARIMA**

- **Pressuposição de Estacionariedade**: exigência que a série temporal seja estacionária ou que possa ser transformada em estacionária.
- **Complexidade na Identificação do Modelo**: determinação dos parâmetros p, d e q pode ser desafiador e requer conhecimento especializado na análise de autocorrelação e autocorrelação parcial.
- **Não Captura Relações Não Lineares**: é um modelo linear, o que significa que pode falhar em capturar padrões não lineares nas séries temporais.
- **Incapacidade de Modelar Múltiplas Sazonalidades:** não captura séries com múltiplas sazonalidades complexas.
- **Sensibilidade a Ruído e Outliers**: é sensível a ruídos ou outliers na série temporal, o que pode afetar a precisão das previsões

PUC Minas

4 Modelagem de Séries Temporais não Estacionárias

4.2 Modelo Sazonal Autoregressivo Integrado e de Média Móvel (SARIMA)

O modelo SARIMA é a combinação de componentes de tendência e sazonalidade

1) Componentes Não Sazonais (ARIMA)

⇒ **AR** (Autoregressivo): captura a dependência entre a observação atual e anteriores

⇒ **I** (Integração): número de diferenciações para tornar a série em estacionária

⇒**MA** (Média Móvel): modela o efeito dos erros passados no valor atual da série

2) Componentes Sazonais

⇒ **AR_s** (AutoRegressivo Sazonal): captura a dependência sazonal.

⇒ **I_s** (Integração Sazonal): captura a diferenciação sazonal para tornar a série estacionária sazonalmente.

⇒ **MA_s** (Média Móvel Sazonal): captura o efeito dos erros sazonais passados.

PUC Minas

4 Modelagem de Séries Temporais não Estacionárias

4.2 Modelo Sazonal Autoregressivo Integrado e de Média Móvel (SARIMA)

⇒ **Estrutura do Modelo SARIMA**

O modelo SARIMA com a ordem não sazonal (p, d, q) e sazonal (P, D, Q, s) é denotado por:

$$\emptyset_p(B)\Phi_p(B^s)(1-B)^d(1-B^s)^D Y_t = \theta_q(B)\Theta_Q(B^s)\varepsilon_t + \varepsilon_t$$

em que Y_t é o valor da série no tempo t; $\emptyset(B)$ polinômio autoregressivo AR no operador de defasagem B; $\Phi(B^s)$ polinômio autorregressivo sazonal (AR_s) no operador de defasagem sazonal B^s; $\theta(B)$ polinômio de média móvel (MA) no operador de defasagem B; $\Theta(B^s)$ polinômio de média móvel sazonal (MA_s) no operador de defasagem sazonal B^s; ε_t termo de erro ou resíduo no tempo t.

4 Modelagem de Séries Temporais não Estacionárias

4.2 Modelo Sazonal Autoregressivo Integrado e de Média Móvel (SARIMA)

⇒ **Estrutura do Modelo SARIMA**

Definições dos polinômios:

- $\emptyset(B) = 1 - \emptyset_1 B - \cdots - \emptyset_p B^p$ sendo p a ordem do polinômio AR (não sazonal)
- $\theta(B) = 1 - \theta_1 B - \cdots - \theta_q B^q$ sendo q a ordem do polinômio MA (não sazonal)
- $\Phi(B^s) = 1 - \Phi_1 B^s - \cdots - \Phi_P B^{Ps}$ sendo que P a ordem do polinômio AR sazonal
- $\Theta(B^s) = 1 - \Theta_1 B^s - \cdots - \Theta_Q B^{Qs}$ sendo Q a ordem do polinômio MA sazonal

s o período sazonal (por exemplo, s=12 para dados mensais com sazonalidade anual)

D é o número de diferenciações sazonais necessárias

d é o número de diferenciações (não sazonal) necessárias para tornar a série estacionária

4 Modelagem de Séries Temporais não Estacionárias

4.2 Modelo Sazonal Autoregressivo Integrado e de Média Móvel (SARIMA)

⇒ **Pressuposições do Modelo SARIMA**

- **Estacionariedade:** média, variância e autocorrelação constantes ao longo do tempo.

  Implicações: para série não estacionária se faz necessário aplicar a transformação de diferenciação para remover tendência e sazonalidade.

- **Linearidade:** assume uma relação linear entre as observações

  Implicações: relações não lineares resulta em previsões menos precisas.

- **Ausência de Outliers Significativos:** não consegue suavizar os outliers significativos.

  Implicações: identifique e trate outliers antes de ajustar o modelo para não resultar na distorção do ajuste e das previsões

- **Sazonalidade e Tendência**: é capaz de modelar por meio dos componentes sazonais e de diferenciação.

4 Modelagem de Séries Temporais não Estacionárias

4.2 Modelo Sazonal Autoregressivo Integrado e de Média Móvel (SARIMA)

⇒ **Pressuposições do Modelo SARIMA**

- **Normalidade dos Resíduos:** distribuição normal com média zero e variância constante.
- **Independência dos Resíduos**: independentes e não autocorrelacionados.

  Implicações: Se houver autocorrelação nos resíduos, o modelo pode não estar capturando toda a informação da série temporal.
- **Homocedasticidade do Resíduos:** resíduos com variância constante ao longo do tempo

  Implicações: se a variância dos erros muda ao longo do tempo se faz necessário utilizar outros modelos para capturar essa característica, como o GARCH

4 Modelagem de Séries Temporais não Estacionárias

4.2 Modelo Sazonal Autoregressivo Integrado e de Média Móvel (SARIMA)

⇒ **Ajuste do Modelo SARIMA**

O ajuste de um modelo SARIMA (AutoRegressive Integrated Moving Average) requer as etapas a seguir:

1) **Verificação da Estacionariedade**
2) **Identificação dos Parâmetros**
3) **Estimação dos parâmetros**
4) **Validação do Modelo**

4 Modelagem de Séries Temporais não Estacionárias

4.2 Modelo Sazonal Autoregressivo Integrado e de Média Móvel (SARIMA)

⇒ **Ajuste do Modelo SARIMA**

1) Verificação Estacionariedade

Visualização gráfica da Série Temporal: Examine a série temporal original para identificar tendências, sazonalidades, e padrões que possam indicar a necessidade de diferenciação ou ajustes nos parâmetros.

Teste de Estacionariedade: Verifique a estacionaridade da série usando testes como o **Dickey-Fuller Aumentado (ADF)**. Se a série não for estacionária, aplique diferenciação (d ou D) até alcançar estacionaridade.

4 Modelagem de Séries Temporais não Estacionárias

4.2 Modelo Sazonal Autoregressivo Integrado e de Média Móvel (SARIMA)

⇒ **Ajuste do Modelo SARIMA**

2) Identificação do modelo

- **Função de Autocorrelação (ACF) e Função de Autocorrelação Parcial (PACF)**

  Picos significativos na ACF e na PACF identificam a ordem do componente q e p e para P e Q observe os lags sazonais nos gráficos

- **Ciclo sazonal observado:** determina o valor de s (sazonalidade anual: $s = 12$)
- **Critérios de Informação**: teste diferentes combinações de (p, d, q) e (P, D, Q, s) e avalie a performance do modelo usando critérios AIC ou BIC para comparar diferentes modelos. O modelo com o menor AIC ou BIC será o melhor.

4 Modelagem de Séries Temporais não Estacionárias

4.2 Modelo Sazonal Autoregressivo Integrado e de Média Móvel (SARIMA)

⇒ **Ajuste do Modelo SARIMA**

3) Estimação dos parâmetros

Método de Máxima Verossimilhança: verifique se os coeficientes estimados para os termos AR e MA são estatisticamente significativos.

Implicações: Coeficientes não significativos podem indicar que o termo correspondente pode ser removido ou que o modelo precisa ser reespecificado

Hipóteses:

H_0: o parâmetro não é significativo

H_1: o parâmetro é significativo

p-valor < nível de significância de 5% rejeita-se H_0

4 Modelagem de Séries Temporais não Estacionárias

4.2 Modelo Sazonal Autoregressivo Integrado e de Média Móvel (SARIMA)

⇒ **Ajuste do Modelo SARIMA**

4) Validação do Modelo

- **Análise dos Resíduos:** Verificar se são como ruído branco (independentes, normalmente distribuídos com média zero e variância contante)
  - **ACF/PACF dos Resíduos**: verificar a autocorrelação significativa. Se houver, o modelo pode precisar ser reespecificado
  - **Teste de Normalidade (não obrigatório):** teste de Shapiro-Wilk, Gráfico QQplot, Histograma
  - **Homocedasticidade**: Teste Breusch-Pagan, White, ARCH, Gráfico resíduos vs tempo ou valores ajustados

PUC Minas

4 Modelagem de Séries Temporais não Estacionárias

4.2 Modelo Sazonal Autoregressivo Integrado e de Média Móvel (SARIMA)

⇒ **Refinamento do Modelo SARIMA**

- **Ajuste dos Parâmetros** p, d, q, P, D, Q: com base na análise dos resíduos e nos critérios de informação (AIC, BIC), refine os valores dos parâmetros para melhorar o ajuste do modelo
- **Considerando-se Modelos Alternativos:** se o modelo SARIMA não é suficiente, considere modelos mais complexos

PUC Minas

4 Modelagem de Séries Temporais não Estacionárias

4.2 Modelo Sazonal Autoregressivo Integrado e de Média Móvel (SARIMA)

⇒ **Previsões do Modelo SARIMA**

Avaliar a previsão

- **Erro de Previsão**: calcule métricas como o RMSE (Root Mean Square Error) ou MAE (Mean Absolute Error) para avaliar a precisão das previsões.
- **Cross-validation:** teste o modelo em um conjunto de dados de teste ou usando validação cruzada para verificar a capacidade preditiva.

PUC Minas

4 Modelagem de Séries Temporais não Estacionárias

4.2 Modelo Sazonal Autoregressivo Integrado e de Média Móvel (SARIMA)

⇒ **Vantagem de um Modelo SARIMA**

- **Capacidade de Modelar Sazonalidade e Tendência**
- **Flexibilidade:** permite capturar uma ampla gama de padrões temporais por ser ajustado para diferentes ordens autoregressivas, diferenciação, média móvel e sazonalidade.
- **Interpretação Relativamente Simples:** os componentes do modelo apresentam significados claros o que facilita a análise dos resultados.

4 Modelagem de Séries Temporais não Estacionárias

4.2 Modelo Sazonal Autoregressivo Integrado e de Média Móvel (SARIMA)

⇒ **Desvantagem do Modelo SARIMA**

- **Pressuposição de Estacionariedade**: exigência que a série temporal seja estacionária ou que possa ser transformada em estacionária.
- **Complexidade na Escolha dos Parâmetros**: determinação dos parâmetros p, d, q e P, D, Q exige um conhecimento profundo da série e do comportamento dos resíduos.
- **Não Captura Relações Não Lineares**: é um modelo linear, o que significa que pode falhar em capturar padrões não lineares nas séries temporais.
- **Sensibilidade a Outliers:** a presença de outliers pode exigir tratamentos especiais nos dados antes do ajuste do modelo
- Dificuldade em Séries com Múltiplos Ciclos Sazonais: dois ou mais tipos, por exemplo, sazonalidade semanal e outra anual.

PUC Minas

4 Modelagem de Séries Temporais não Estacionárias

4.3 Análise de Intervenção

⇒ **Análise de intervenção em séries temporais**

É uma técnica utilizada para avaliar o quanto um impacto ou intervenção de uma mudança externa (como uma política governamental, uma crise econômica, uma nova regulamentação, ou outro tipo de evento significativo) pode afetar uma série temporal.

A ideia central da análise de intervenção é comparar o comportamento da série temporal antes e depois da intervenção.

A intervenção pode ser de natureza planejada (como uma nova política) ou não planejada (como um desastre natural).

O primeiro passo é verificar a existência de tendência, sazonalidade e identificar o ponto da intervenção na série temporal em estudo.

PUC Minas

4 Modelagem de Séries Temporais não Estacionárias

4.3 Análise de Intervenção

⇒ **Modelagem com Intervenção**

Modelo ARIMA ou SARIMA pode ser utilizado para modelar a série antes e depois da intervenção desde que a série seja estacionária ou possa torná-la estacionária através de diferenciação.

O modelo ARIMA é indicado para séries que apresentam tendência e o modelo SARIMA para séries com sazonalidade

- **Identificação da Intervenção**

  Define o ponto no tempo onde a intervenção ocorreu.

- **Modelagem Pré-Intervenção**

  Ajuste um modelo à série temporal utilizando apenas os dados até o ponto de intervenção que será a linha de base para descrever o comportamento da série sem a intervenção

PUC Minas

4 Modelagem de Séries Temporais não Estacionárias

4.3 Análise de Intervenção

⇒ **Modelagem com Intervenção**

O modelo ARIMA com intervenção de ordem (p, d, q) é dado por:

$$(1-B)^d Y_t = c + \sum_{i=1}^{p} \phi_i Y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \delta I_t$$

em que Y_t valor da série no tempo presente; $\emptyset(B)$ polinômio autorregressivo; $\theta(B)$ polinômio de média móvel; I_t variável indicadora de intervenção (dummy) que assume valor 0 antes da intervenção e 1 após; δ parâmetro que mede o impacto da intervenção.

- **Sem a Intervenção**: $I_t = 0$ o modelo se comporta como um ARIMA comum.

PUC Minas

4 Modelagem de Séries Temporais não Estacionárias

4.3 Análise de Intervenção

⇒ **Modelagem com Intervenção**

O modelo SARIMA com intervenção de ordem não sazonal (p, d, q) e sazonal (P, D, Q, s) é denotado por:

$$\emptyset_p(B)\Phi_P(B^s)(1-B)^d(1-B^s)^D Y_t = \theta_q(B)\Theta_Q(B^s)\varepsilon_t + \delta I_t$$

em que Y_t série observada; B operador de defasagem (lag operator) em que $BY_t = Y_{t-1}$; $\emptyset_p(B)$ polinômio autorregressivo não sazonal de ordem p; $\Phi_p(B^s)$ polinômio autorregressivo sazonal de ordem P e período s; $\theta_q(B)$ polinômio de média móvel de ordem q; $\Theta_Q(B^s)$ polinômio de média móvel sazonal de ordem Q e período s; d ordem de diferenciação não sazonal; D ordem de diferenciação sazonal; ε_t termo de erro (ruído branco) em t; I_t variável indicadora de intervenção (dummy) que é 0 antes da intervenção e 1 após; δ parâmetro que mede o impacto da intervenção

Sem a Intervenção: $I_t = 0$ o modelo se comporta como um SARIMA comum.

PUC Minas

4 Modelagem de Séries Temporais não Estacionárias

4.3 Análise de Intervenção

⇒ **Estrutura da Função de intervenção**

As funções de intervenção são usadas para incorporar a intervenção ao modelo e a sua escolha depende da natureza do impacto esperado

O objetivo da função de intervenção é descrever como o impacto se desenvolve na série ao longo do tempo.

As funções de intervenção são:

- **Intervenção Instantânea e Permanente**: O impacto ocorre imediatamente após a intervenção e persiste indefinidamente

$$Y_t = (Modelo\ ARIMA\ ou\ SARIMA) + \delta I_t$$

PUC Minas

4 Modelagem de Séries Temporais não Estacionárias

4.3 Análise de Intervenção

⇒ **Estrutura da Função de intervenção**

- **Intervenção Instantânea e Temporária**: o impacto ocorre imediatamente após a intervenção, mas desaparece depois de algum tempo

$$Y_t = (Modelo\ ARIMA\ ou\ SARIMA) + \delta I_t - \delta I_{t-k}$$

- **Intervenção Gradual e Permanente:** o impacto ocorre de forma gradual e persiste indefinidamente

$$Y_t = (Modelo\ ARIMA\ ou\ SARIMA) + \delta \sum_{i=1}^{t} I_i$$

- **Intervenção Gradual e Temporária:** o impacto ocorre de forma gradual e desaparece após algum tempo

$$Y_t = (Modelo\ ARIMA\ ou\ SARIMA) + \delta \sum_{i=1}^{t} I_i - \delta \sum_{i=1}^{t-k} I_i$$

PUC Minas

4 Modelagem de Séries Temporais não Estacionárias

4.3 Análise de Intervenção

⇒ **Ajuste do Modelo de Intervenção**

Após a escolha do modelo ARIMA/SARIMA e da função de intervenção apropriada, os parâmetros serão estimados utilizando a Função de Máxima Verossimilhança.

- **Teste de Significância:** Após a estimação, o parâmetro de intervenção δ é testado quanto à sua significância estatística.

$\boldsymbol{\delta}$ **significativo:** a intervenção teve um impacto estatisticamente detectável na série temporal (p-valor < 5%).

$\boldsymbol{\delta}$ **positivo**: a intervenção aumenta o valor esperado da série.

$\boldsymbol{\delta}$ **negativo**: a intervenção diminui o valor esperado da série.

PUC Minas

4 Modelagem de Séries Temporais não Estacionárias

4.3 Análise de Intervenção

⇒ **Ajuste do Modelo de Intervenção**

Magnitude do Coeficiente: indica a força do impacto da intervenção. Coeficientes maiores em valor absoluto indicam um impacto mais forte.

- **Validação do Modelo**

Análise dos Resíduos: os resíduos como ruído branco (não correlacionados, média zero e variância constante), o que sugere que o modelo é adequado. Utilizar teste estatísticos e análise gráficos.

Menor AIC/BIC: Indica que o modelo com intervenção se ajusta melhor aos dados do que o modelo sem intervenção.

Menor RMSE ou MAE: Indica a melhor precisão do modelo.

PUC Minas

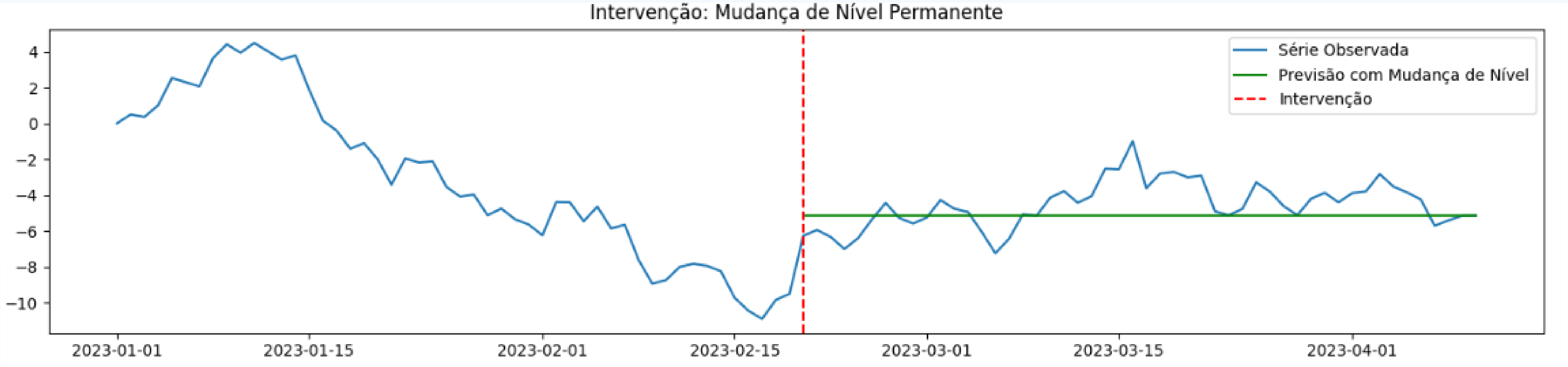
4 Modelagem de Séries Temporais não Estacionárias

4.3 Análise de Intervenção

⇒ **Interpretação do gráfico após a Intervenção**

Gráfico de Previsão: Mostra a série observada, a previsão a partir do ponto de intervenção, e a linha de intervenção. A diferença entre a previsão com intervenção e a linha de base pode ser interpretada como o impacto da intervenção.

- **Mudança de Nível Permanente:** uma mudança súbita no nível da série indica que a intervenção causou uma mudança no valor médio da série.

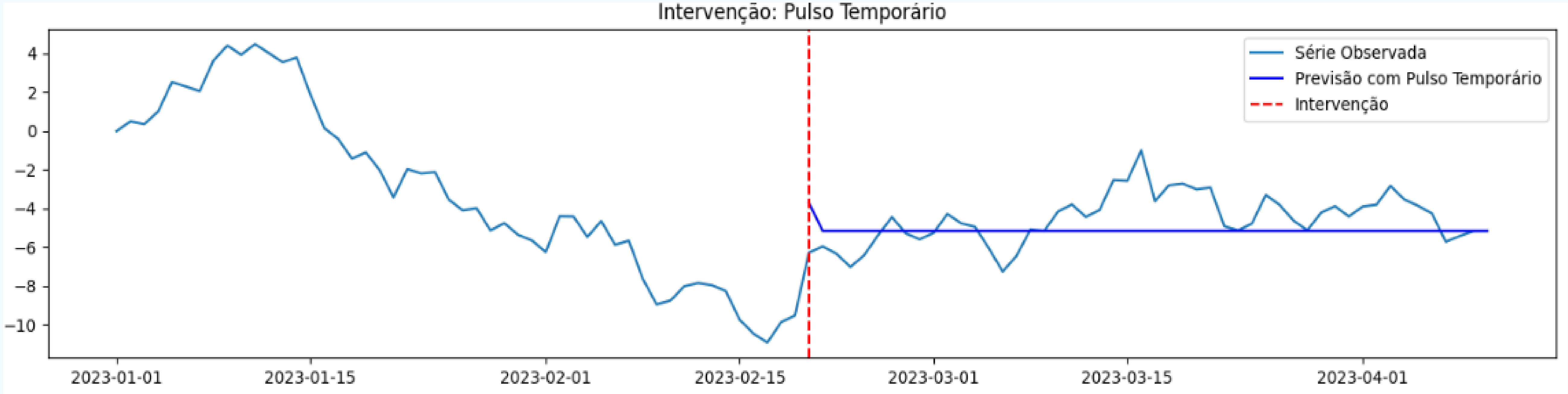


4 Modelagem de Séries Temporais não Estacionárias

4.3 Análise de Intervenção

⇒ **Interpretação do gráfico após a Intervenção**

- **Pulso Temporário:** um pico ou queda temporária após a intervenção, com um retorno ao nível original, isso indica que a intervenção teve um impacto breve e transitório na série.

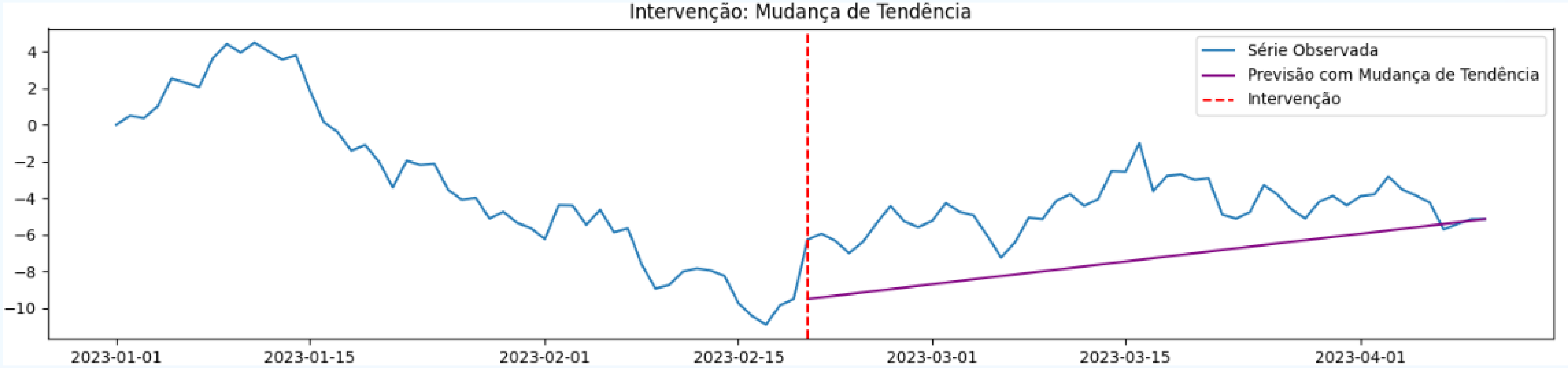


4 Modelagem de Séries Temporais não Estacionárias

4.3 Análise de Intervenção

⇒ **Interpretação do gráfico após a Intervenção**

- **Mudança de Tendência:** uma alteração gradual na inclinação da linha após a intervenção, isso sugere que a intervenção alterou a tendência de longo prazo da série, acelerando ou desacelerando seu crescimento ou declínio



PUC Minas

4 Modelagem de Séries Temporais não Estacionárias

4.4 Regressão em Séries Temporais

⇒ **Regressão em Séries Temporais**

Modelo de Regressão é usado para descrever a relação entre duas variáveis, ou seja, o quanto uma variável consegue explicar a variação (mudanças) que ocorre na outra.

- **Variável independente (explicativa/ exógena):** é a variável que vai explicar a variação no comportamento da outra, denotado por X.
- **Variável dependentes (resposta):** é a variável que temos o interesse de estudar a variação em seu comportamento, denotado por Y.

Objetivo do modelo de regressão: conseguir prever valores não observados de Y baseado nos valores observados de X com maior precisão.

PUC Minas

4 Modelagem de Séries Temporais não Estacionárias

4.4 Regressão em Séries Temporais

⇒ **Regressão em Séries Temporais**

Modelos de Regressão Linear: descrevem a relação linear entre as variáveis, ou seja, diretamente proporcional.

OBS: Relações não lineares exigem técnicas de modelagem não lineares tais como polinômios, logaritmos, ou outros tipos de funções para capturar a complexidade da relação.

Pressupostos:

- **Linearidade:** mudança na variável dependente é proporcional às mudanças nas variáveis independentes
- **Exogeneidade**: variáveis independentes devem ser não correlacionadas com o termo de erro.

PUC Minas

4 Modelagem de Séries Temporais não Estacionárias

4.4 Regressão em Séries Temporais

⇒ **Regressão em Séries Temporais**

Pressupostos:

- **Homoscedasticidade**: variância dos erros deve ser constante ao longo de todas as observações
- **Independência dos Erros**: o erro em uma observação não deve ser correlacionado com o erro em outra observação.
- **Normalidade dos Erros**
- **Ausência de Multicolinearidade**: não deve existir relação linear entre duas ou mais variáveis independentes.

PUC Minas

4 Modelagem de Séries Temporais não Estacionárias

4.4 Regressão em Séries Temporais

⇒ **Regressão em Séries Temporais**

Regressão em séries temporais: técnica usada para descrever a relação entre o valor da série temporal com fator exógeno ou o próprio valor passado da série temporal

- **Variável dependente:** valor da série temporal
- **Variável independente:** fator exógeno ou o próprio valor passado da série temporal

*__DESAFIOS:__ em modelagem de séries temporais alguns dos pressupostos do modelo de regressão linear clássico podem não ser válidos ou precisam ser ajustados.

Os erros são correlacionados no tempo

As variância não são constantes

Os erros não são normalmente distribuídos

E os valores da série são correlacionados com os erros

PUC Minas

4 Modelagem de Séries Temporais não Estacionárias

4.4 Regressão em Séries Temporais

⇒ **Regressão em Séries Temporais**

Técnicas para ajustar modelos de regressão linear em séries temporais:

- **Adicionar termo autoregressivo:** considerar como fator exógeno o próprio valor passado da série temporal.
- **Adicionar termo autoregressivo no erros:** os erros (resíduos) são modelados como uma série autorregressiva para tratar a autocorrelação nos resíduos.
- **Adicionar fator sazonal:** capturar padrões que se repetem regularmente ao longo do tempo.
- **Modelos dinâmicos:** permite que os coeficientes da regressão variem ao longo do tempo capturando as mudanças na relação entre as variáveis dependentes e independentes

4 Modelagem de Séries Temporais não Estacionárias

4.4 Regressão em Séries Temporais

⇒ **Regressão Linear Simples com termo autoregressivo**

Um modelo de regressão simples com um termo autorregressivo incorpora valores passados da própria série temporal como uma variável explicativa ou exógena.

- Modelo AR(1) (autorregressivo de ordem 1):

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \varepsilon_t$$

em que Y_t valor da série no tempo t; $X = Y_{t-1}$ é o valor da série no tempo anterior $t-1$; β_0 é o intercepto do modelo; β_1 o coeficiente de regressão que representa o impacto do valor passado Y_{t-1} sobre Y_t; ε_t o erro residual.

Interpretação:

- β_1 próximo de 1 sugere que os valores passados têm uma forte influência nos valores futuros
- Ausência de padrões significativos nos resíduos indica que o modelo capturou bem a estrutura da série temporal.

PUC Minas

4 Modelagem de Séries Temporais não Estacionárias

4.4 Regressão em Séries Temporais

⇒ **Regressão Linear Simples com termo autoregressivo nos resíduos**

Um modelo de regressão linear simples com termo autorregressivo nos erros é uma extensão da regressão linear onde os erros (resíduos) são modelados como uma série autorregressiva. É útil quando há autocorrelação nos resíduos.

- Modelo AR(1) nos resíduos:

$$Y_t = \beta_0 + \beta_1 X + \emptyset \varepsilon_{t-1} + u_t$$

em que Y_t valor da série no tempo t; X variável exógena ou preditora no tempo t; β_0 e β_1 o coeficiente de regressão; ε_{t-1} o erro residual no tempo t-1; $\emptyset$ coeficiente do termo autorregressivo; u_t ruído branco (distribuição normal com média zero e variância constante).

Interpretação:

- Incluir um termo AR(1) nos resíduos pode melhorar a acurácia das previsões
- Ausência de padrões significativos nos resíduos após incluir o termo AR(1) sugere que o modelo está capturando bem a estrutura dos dados.

4 Modelagem de Séries Temporais não Estacionárias

4.4 Regressão em Séries Temporais

⇒ **Regressão Linear Simples com termo autorregressivo e sazonalidade**

Um modelo de regressão simples com um fator sazonal é uma extensão da regressão linear capaz de modelar a variação periódica nos dados tis como ciclos anuais, mensais ou semanais.

- Modelo de regressão linear com fator sazonal

$$Y_t = \beta_0 + \beta_1 X + \gamma S_t + \varepsilon_t$$

em que Y_t valor da série no tempo t; X variável exógena ou preditora no tempo t; β_0 e β_1 o coeficiente de regressão; S_t fator sazonal; γ coeficiente do fator sazonal; ε_t o erro residual.

Coeficientes Sazonais: Cada coeficiente associado às variáveis dummies sazonais (γ) indica o impacto de um mês específico na série temporal, em relação ao mês de referência.

Validação do Modelo: a qualidade do ajuste pode ser validada observando o R^2 e outros indicadores de desempenho, além da análise visual das previsões.

PUC Minas

4 Modelagem de Séries Temporais não Estacionárias

4.4 Regressão em Séries Temporais

⇒ **Regressão Dinamica**

A regressão dinâmica desempenha um papel fundamental na modelagem de séries temporais permitindo a inclusão da estrutura de dependência temporal de maneira mais flexível e robusta do que as abordagens de regressão tradicional.

Esses modelos combinam o efeito de variáveis explicativas com a dinâmica de séries temporais em que o valor da série em estudo é explicada por seus valores defasados e pelos valores das variáveis exógenas

Deve ser usado quando existe dependência entre a série em estudo com outras variáveis exógenas e quando a estrutura de correlação da série em estudo não pode supor independência.

4 Modelagem de Séries Temporais não Estacionárias

4.5 Regressão Dinâmica

⇒ **Modelo ARMAX**

O modelo ARMAX é a combinação do componente AR(p), MA (q) com a variável exógena denotado por:

$$Y_t = c + \sum_{i=1}^{p} \phi_i Y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \sum_{j=1}^{k} B_k X_{t-k} + \varepsilon_t$$

em que ϕ_i coeficiente autorregressivo (AR) de ordem p aplicados as defasagens da variável dependente Y_{t-i}; θ_j coeficiente do componente de média móvel (MA) de aplicados a defasagens dos erros ε_{t-j}; c é a constante; X_{t-k} variáveis exógenas no tempo t; B_k coeficientes das variáveis exógenas; ε_t termo de erro ruído branco no tempo t.

Modelo ARIMAX: Ajusta a série temporal considerando a variável exógena e exige que a série seja estacionária

PUC Minas

4 Modelagem de Séries Temporais não Estacionárias

4.5 Regressão Dinâmica

⇒ **Modelo ARMAX**

Pressuposto do modelo

- **Estacionariedade**: a série deve ser estacionária ou transformada.
- **Independência dos Resíduos**: os resíduos do modelo devem ser independentes e seguir uma distribuição normal.
- **Linearidade**: relação linear entre a variável dependente e suas lags e os erros passados é linear.
- **Exogeneidade das Variáveis $\boldsymbol{X}$**: As variáveis exógenas X_t devem ser independentes do termo de erro ε_t .

PUC Minas

4 Modelagem de Séries Temporais não Estacionárias

4.5 Regressão Dinâmica

⇒ **Modelo ARMAX**

Validação do modelo

- **Diagnóstico de Resíduos**: verificar a autocorrelação dos resíduos para assegurar que eles são ruídos brancos (usando gráficos ACF/PACF)
- **Teste de Autocorrelação**: Use testes como o teste de Ljung-Box para confirmar a ausência de autocorrelação nos resíduos
- **Análise da Significância dos Parâmetros**: verificar se os coeficientes dos termos AR e MA são estatisticamente significativos.
- **Comparação de Modelos**: utilizar critérios como AIC, BIC para selecionar o melhor modelo.

Previsões

As previsões para curto prazo. As previsões dependem tanto das defasagens da série temporal quanto dos valores futuros das variáveis exógenas.

PUC Minas

4 Modelagem de Séries Temporais não Estacionárias

4.5 Regressão Dinâmica

⇒ **Modelo ARIMAX**

O modelo ARIMA é a combinação do componente AR(p), I(d) , MA (q) com a variável exógena denotado por:

$$(1-B)^d Y_t = c + \sum_{i=1}^{p} \phi_i Y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \sum_{j=1}^{k} B_k X_{t-k} + \varepsilon_t$$

em que ϕ_i coeficiente autorregressivo (AR) de ordem p aplicados as defasagens da variável dependente Y_{t-i}; θ_j coeficiente do componente de média móvel (MA) de aplicados a defasagens dos erros ε_{t-j}; c é a constante, B o operador de defasagem; X_{t-k} variáveis exógenas no tempo t; B_k coeficientes das variáveis exógenas; d ordem de diferenciação; ε_t termo de erro ruído branco no tempo t.

Modelo ARIMAX: Ajusta a série temporal considerando a variável exógena e inclui componentes de tendência e diferenciação

PUC Minas

4 Modelagem de Séries Temporais não Estacionárias

4.5 Regressão Dinâmica

⇒ **Modelo ARIMAX**

Pressuposto do modelo

- **Estacionariedade**: a série deve ser estacionária ou transformada.
- **Independência dos Resíduos**: os resíduos do modelo devem ser independentes e seguir uma distribuição normal.
- **Linearidade**: relação linear entre a variável dependente e suas lags e os erros passados é linear.
- **Multicolinearidade**: as variáveis exógenas não devem ser altamente correlacionadas entre si para evitar problemas de multicolinearidade

PUC Minas

4 Modelagem de Séries Temporais não Estacionárias

4.5 Regressão Dinâmica

⇒ **Modelo ARIMAX**

Validação do modelo

- **Diagnóstico de Resíduos**: verificar a autocorrelação dos resíduos para assegurar que eles são ruídos brancos (usando gráficos ACF/PACF e teste de Ljung-Box)
- **Teste de Estacionariedade**: testar se a série temporal ou os resíduos são estacionários usando o teste de Dickey-Fuller Aumentado
- **Análise da Significância dos Parâmetros**: verificar se os coeficientes dos termos AR, MA e sazonais são estatisticamente significativos.
- **Comparação de Modelos**: utilizar critérios como AIC, BIC para selecionar o melhor modelo.

Previsões

As previsões para curto e longo prazo incorporando a tendência e os efeitos das variáveis exógenas.

PUC Minas

4 Modelagem de Séries Temporais não Estacionárias

4.5 Regressão Dinâmica

⇒ **Modelo SARIMAX**

O modelo SARIMAX com intervenção é dado por:

$$\emptyset_p(B)\Phi_P(B^s)(1-B)^d(1-B^s)^D Y_t = \theta_q(B)\Theta_Q(B^s)\varepsilon_t + B_k X_{t-k}$$

em que Y_t é o valor da série no tempo t; $\emptyset(B)$ polinômio autoregressivo AR no operador de defasagem B; $\Phi(B^s)$ polinômio autorregressivo sazonal (AR_s) no operador de defasagem sazonal B^s; $\theta(B)$ polinômio de média móvel (MA) no operador de defasagem B; $\Theta(B^s)$ polinômio de média móvel sazonal (MA_s) no operador de defasagem sazonal B^s; X_{t-k} variáveis exógenas no tempo t; B_k coeficientes das variáveis exógenas; ε_t termo de erro no tempo t.

Modelo SARIMAX: Ajusta a série temporal considerando a variável exógena e inclui componentes sazonais e de tendência

PUC Minas

4 Modelagem de Séries Temporais não Estacionárias

4.5 Regressão Dinâmica

⇒ **Modelo SARIMAX**

Pressuposto do modelo

- **Estacionariedade**: a série deve ser estacionária ou transformada.
- **Sazonalidade**: a série pode apresentar padrões sazonais que são modelados.
- **Independência dos Resíduos**: os resíduos do modelo devem ser independentes e seguir uma distribuição normal.
- **Linearidade**: relação linear entre a variável dependente e suas lags e os erros passados é linear.

PUC Minas

4 Modelagem de Séries Temporais não Estacionárias

4.5 Regressão Dinâmica

⇒ **Modelo SARIMAX**

Validação do modelo

- **Diagnóstico de Resíduos**: verificar a autocorrelação dos resíduos para assegurar que eles são ruídos brancos (usando gráficos ACF/PACF e e teste de Ljung-Box)
- **Teste de Estacionariedade**: testar se a série temporal ou os resíduos são estacionários usando o teste de Dickey-Fuller Aumentado
- **Análise da Significância dos Parâmetros**: verificar se os coeficientes dos termos AR, MA e sazonais são estatisticamente significativos.
- **Comparação de Modelos**: utilizar critérios como AIC, BIC para selecionar o melhor modelo.

Previsões

As previsões para curto e longo prazo incorporando a sazonalidade e os efeitos das variáveis exógenas.