

Aula 003: Séries Temporais – Parte 3

Testes Estatísticos e ARIMA

[Notebook do Projeto.](#)

Séries Estacionárias

- Séries Temporais precisam ser estacionárias para que possamos fazer previsões.
- Uma série estacionária significa ter a média, variância e covariância permanecendo constante ao longo da dimensão temporal.

Como verificar se uma TS é estacionária?

Olhar os Plots

- Identificar visualmente durante a análise exploratória se a TS é estacionária ou não.

Resumo Estatístico

- Gerar um resumo com as principais informações estatísticas da TS, e verificar se há algum indício de estacionaridade.

Teste Estacionários

- São testes que possuem a capacidade de nos dizer se uma TS é estacionária ou não, sempre se baseando em métodos científicos. Um desses testes, utilizado nessa aula, é o *Augmented Dickey-Fuller (ADF)*.

Teste ADF

- ADF é um dos testes estatísticos mais utilizados para TS. Suas hipóteses nula e alternativa são:
 - A Hipótese Nula dos testes é que a TS não é estacionária. Ou seja, possui alguma dependência em relação ao tempo.
 - A Hipótese Alternativa rejeita a hipótese nula, ou seja, que a TS é estacionária.
- Normalmente, rejeitamos a hipótese nula quando o *p-value* é menor ou igual a um determinado nível de significância. Normalmente, esse nível de significância é igual a 0.01 (1%), 0.05 (5%) ou até mesmo 0.1 (10%).
- Quanto menor o *p-value*, mais evidências contra a hipótese nula você tem.
 - Por exemplo, caso seja usado $p\text{-value} = 0.05$, rejeitamos a Hipótese Nula com um nível de confiança de 95%.
- Caso o valor do teste fique acima do *threshold*, falharemos em rejeitar a hipótese nula.
- **Atenção! Isso não significa que a Hipótese Nula é verdadeira.** Significa apenas que o experimento não encontrou evidências para rejeitá-la.

Transformando uma TS em Estacionária

- Para transformar uma TS em estacionária, precisamos remover a tendência e sazonalidade que há nos dados originais. Para isso, existem diversas técnicas que podem ser usadas.
- Aplicar o *log* a fim de reduzirmos a magnitude dos valores da TS é um bom começo.
- Em seguida, subtrair a média móvel de um determinado período em relação ao *log* da TS.

Diferenciação

- Uma outra técnica simples que ajuda a tornar TS mais estacionárias é a diferenciação.
- Apenas calculando a diferença entre duas observações possíveis, é possível incrementar ainda mais o desempenho no ADF.

$$value(t) = observation(t) - observation(t - 1)$$

Forecasting usando ARIMA

Arima

- **ARIMA** significa *AutoRegressive Integrated Moving Average* e é um dos modelos mais usados para fazer previsões em cima de Séries Temporais.
- É uma classe de modelo que consegue capturar um conjunto de diferentes estruturas temporais dados de TS. O acrônimo ARIMA significa:
 - **AR**: *AutoRegression*, pois usa a relação entre dependência entre uma observação e um número de *lagged observations*.
 - **I**: *Integrated*, pois usa a diferenciação das observações *raw* para transformar a TS em estacionária.
 - **MA**: *Moving Average*, pois usa a dependência entre uma observação e o erro residual proveniente de uma média móvel aplicada a *lagged observations*.
- Cada um desses componentes são representados por um parâmetro do modelo. Por esse fato, é muito comum ver a notação **ARIMA** (p, d, q).
- Para encontrar os parâmetros otimizados de $p, d, e q$, usaremos funções ACF (Autocorrelação) e PACF (Autocorrelação Parcial).

Sazonalidade

- Sazonalidade é qualquer mudança ou padrão previsível em uma série temporal que seja recorrente ou se repita ao longo de um determinado período de tempo.

Padrões Cíclicos

- Esses padrões existem quando os dados demonstram altas e quedas que não acontecem em um período fixo.
 - Bolsa de Valores
 - Tempos de expansão
 - Tempos de recessão
 - Sem padrão de repetição
- O período está em mudança e não está associado com aspectos do calendário.
- Bem mais difíceis de prever.

Análise de Séries Temporais

- Como importar dados de Séries Temporais
- Como visualizar os dados
- Decompondo uma série temporal
 - Para analisar Séries Temporais, pode ser necessário separar sazonalidade e tendência da nossa série.