



UNIVERSIDADE FEDERAL DE MINAS GERAIS
CENTER OF WHAT
DEPARTAMENTO DE ENGENHARIA ELETRÔNICA
UNDERGRADUATE COURSE IN ELECTRICAL ENGINEERING

GABRIEL TEIXEIRA LARA CHAVES

PROCESSAMENTO E REPRESENTAÇÃO DE SÉRIES TEMPORAIS
MULTIVARIADAS E NÃO ESTACIONÁRIAS

BELO HORIZONTE

2023

GABRIEL TEIXEIRA LARA CHAVES

PROCESSAMENTO E REPRESENTAÇÃO DE SÉRIES TEMPORAIS MULTIVARIADAS E
NÃO ESTACIONÁRIAS

Undergraduate Thesis submitted to the Electrical Engineering Course of the Center of What of the Universidade Federal de Minas Gerais, as a partial requirement for obtaining the Bachelors Degree in Electrical Engineering.

Advisor: Prof. Dr. Frederico Gualberto Coelho

Co-advisor: Prof. Dr. ?

BELO HORIZONTE

2023

GABRIEL TEIXEIRA LARA CHAVES

PROCESSAMENTO E REPRESENTAÇÃO DE SÉRIES TEMPORAIS MULTIVARIADAS E
NÃO ESTACIONÁRIAS

Undergraduate Thesis submitted to the Electrical Engineering Course of the Center of What of the Universidade Federal de Minas Gerais, as a partial requirement for obtaining the Bachelors Degree in Electrical Engineering.

Approved on: 31th August 2023

EXAMINATION BOARD

Prof. Dr. Frederico Gualberto Coelho (Advisor)
Universidade Federal de Minas Gerais

Prof. Dr. ? (Co-advisor)
Universidade Federal de Minas Gerais

Prof. Dr. Antônio de Pádua Braga

ACKNOWLEDGEMENTS

Obrigado !

“ Alice sighed wearily. ‘I think you might do something better with the time,’ she said, ‘than waste it in asking riddles that have no answers.’ ‘If you knew time as well as I do,’ said the Hatter, ‘you wouldn’t talk about wasting it. It’s him.’ ‘I don’t know what you mean,’ said Alice. ‘Of course you don’t!’ the Hatter said, tossing his head contemptuously. ‘I dare say you never even spoke to Time!’ ‘Perhaps not,’ Alice cautiously replied, ‘but I know I have to beat time when I learn music.’ ‘Ah! That accounts for it,’ said the Hatter. ‘He won’t stand beating.’ ”

(Lewis Carrol, *Alice’s Adventures in Wonderland*)

ABSTRACT

Abstract here.

Keywords: Time series analysis, Discrete-time signal processing, Machine learning, Frequency analysis.

RESUMO

Resumo

Palavras-chave: Análise de Séries Temporais, Processamento de Sinais de Tempo Discreto, Aprendizado de Máquina, Análise em Frequência

LIST OF FIGURES

Figure 1 – Visualização no tempo de ruído branco	17
Figure 2 – Série temporal com tendência linear	20
Figure 3 – Série temporal com tendência logarítmica	20
Figure 4 – Série temporal com sazonalidade senoidal de período arbitrário	21
Figure 5 – Série temporal com sazonalidade aditiva	21
Figure 6 – Série temporal com sazonalidade multiplicativa	22
Figure 7 – Visualização de correlatograma de série com tendência linear. Observe que as autocorrelações decaem lentamente ao decorrer dos atrasos, comportamento típico de tendências determinísticas.	24
Figure 8 – Visualização de correlatograma de série com sazonalidade. Observe que o padrão senoidal da série é reproduzido nas autocorrelações.	24
Figure 9 – Sinal periódico imerso em ruído e sua correspondente autocorrelação.	25
Figure 10 – Conjunto de séries demonstrando diferentes níveis de estacionariedade.	27
Figure 11 – Visualização de caminhada aleatória com $y_0 = 5$	29

LIST OF TABLES

LIST OF SYMBOLS

ω Frequência angular

ρ Autocorrelação

CONTENTS

	Contents	10
1	INTRODUÇÃO	13
2	DEFINIÇÕES E PROPRIEDADES	15
2.1	Processo Estocástico	15
2.2	Série Temporal	16
2.3	Operador de Atraso(<i>Lag</i>)	17
2.4	Operador de Diferença	19
2.5	Tendência(<i>Trend</i>)	19
2.6	Sazonalidade	20
2.7	Autocorrelação	22
2.7.1	<i>Autocorrelação Parcial</i>	23
2.7.2	<i>Correlalograma</i>	24
2.8	Estacionariedade	26
2.8.1	<i>Importância de Estacionariedade</i>	27
2.8.2	<i>Categorias Básicas de Não Estacionariedade</i>	28
2.9	Ruído Branco	31
3	TEORIA UNIVARIADA	32
3.1	Decomposição de Séries Temporais	32
3.2	Modelos Estacionários	32
3.2.1	<i>Modelo Linear Generalizado</i>	32
3.2.2	<i>Raízes Unitárias</i>	32
3.2.3	<i>Modelo Autoregressivo</i>	32
3.2.4	<i>Modelo Média Móvel</i>	32
3.2.5	<i>Modelo ARMA</i>	32
3.3	Modelo ARIMA	32
3.3.1	<i>Estabilidade e Invertibilidade</i>	32
3.4	Modelos de Sazonalidade	32
3.4.1	<i>Diferenciação Sazonal</i>	32
3.4.2	<i>SARIMA</i>	32
4	TEORIA DE VOLATILIDADE	33
4.1	Definições Adicionais	33

4.2	Modelo de Heterocedacidade Condicional Autoregressiva Generalizada	33
4.3	Modelo de Volatilidade Estocástica	33
4.4	Modelos Adicionais	33
5	TEORIA MULTIVARIADA	34
5.1	Definições Adicionais	34
5.1.1	<i>Correlação Cruzada</i>	34
5.1.2	<i>Cointegração</i>	34
5.2	Modelos em Espaço de Estados	34
5.3	Modelos Multivariados	34
5.4	Modelos de Variáveis Exógenas	34
6	TESTES ESTATÍSTICOS	35
6.1	Considerações sobre testes em séries temporais	35
6.2	Teste de <i>Dickey-Fuller</i>	35
6.2.1	<i>Teste Kwiatkowski-Phillips-Schmidt-Shin (KPSS)</i>	36
6.3	Causalidade de Granger	37
6.4	Box-Pierce	37
6.5	Ljung-Box-Pierce	37
6.6	<i>Convergence Cross Mapping</i>	37
6.7	Johansen	37
7	VALIDAÇÃO DE MODELOS	38
7.1	Análise de Resíduos	38
7.2	Validação Cruzada	38
8	ANÁLISE ESPECTRAL	39
9	MODELOS NÃO LINEARES	40
10	ANÁLISE DE SÉRIES NÃO ESTACIONÁRIAS	41
10.1	Definições Adicionais	41
10.1.1	<i>Autocorrelação Instantânea</i>	41
10.2	Indução de Estacionariedade	41
11	TÓPICOS ADICIONAIS	42
12	PROCESSAMENTO DE SINAIS ESTOCÁSTICOS	43
12.1	Introdução	43
13	APLICAÇÕES	44

14	CONCLUSÕES E TRABALHOS FUTUROS	45
14.1	Contribuições do Trabalho	45
14.2	Limitações	45
14.3	Trabalhos Futuros	45
	REFERENCES	46
	APPENDICES	47
	APPENDIX A – Exemplo de apêndice	47
	ANNEXES	47
	ANNEX A – Exemplo de um anexo	48

1 INTRODUÇÃO

INTRODUÇÃO

2 DEFINIÇÕES E PROPRIEDADES

2.1 Processo Estocástico

Dado um conjunto arbitrário \mathcal{T} um processo estocástico é uma família $X(t, \omega)$, $t \in \mathcal{T}$ e $\omega \in \Omega$ de forma que para cada $t \in T$, $\omega \in \Omega$ $X(t, \Omega)$ é uma variável aleatória. As variáveis aleatórias podem ser reais ou complexas. Esse trabalho aborda apenas processos estocásticos reais exceto quando explicitamente mencionado.

Supõe-se que a família de variáveis aleatórias seja definida em um mesmo espaço de probabilidades (Ω, \mathcal{A}, P) com Ω representando um espaço amostral, \mathcal{A} uma σ -álgebra e P uma medida de probabilidade. Para propósitos desse trabalho podemos tomar o conjunto \mathcal{T} como \mathbb{R} , resultando em processos de tempo contínuo, e \mathbb{Z} , resultando em processos de tempo discreto.

Para cada $t \in \mathcal{T}$ temos uma função de densidade de probabilidade associada à variável aleatória $X(t_k, \omega)$ (assumindo que essa função exista). Na prática um processo aleatório no mundo real é observado ao longo de t , tal que ω seja fixado ao universo em que a observação ocorre. Sob essa condição $X(t, \omega_k)$ é considerada uma realização do processo estocástico. Realizações também são chamadas de *sample record* em alguns livros de engenharia e séries temporais na literatura estatística.

Para ilustrar os conceitos acima podemos pensar em um exemplo proposto por (BENDAT; PIERSOL, 2000) em que um gerador de ruído térmico é construído e sua tensão ao longo de um intervalo do tempo é medida. Se um outro gerador fosse construído sob condições e com propriedades idênticas sua tensão medida no mesmo intervalo de tempo não seria idêntica, assim como a tensão medida sob qualquer outro gerador idêntico. De fato cada registro de tensão é um exemplo de infinitos registros que poderiam ter ocorrido. Nessa situação os registros ou séries temporais de tensão são as realizações de um processo estocástico representativo de todas as possíveis realizações.

O adequado estudo de séries temporais é consequência de um primeiro adequado estudo sobre processos estocásticos, geradores dessas séries temporais. Essa não é a intenção desse trabalho. Como em grande parte da teoria de séries temporais estamos preocupados com o que podemos compreender ou inferir sobre o processo estocástico gerador de uma realização a partir apenas de seu único registro. Essa abordagem é essencial e de fato mais aplicável que um estudo que se preocupa excessivamente com os processos geradores devido ao fato de dados do mundo real frequentemente representarem realizações únicas. Não é possível realizar novamente

o índice Ibovespa entre 1970 e 2020 e muito menos ter acesso à realização desses índices em universos paralelos.

Ao longo desse trabalho a distinção e referência ao processo gerador de uma série temporal será feita quando necessário.

2.2 Série Temporal

Uma série temporal é um conjunto de observações realizadas sequencialmente no tempo, indexadas de acordo com o momento em que foram observadas. As observações representam a realização de um processo estocástico. Em alguns contextos, como análise de processos industriais, a natureza do processo subjacente é relevante para análise e modelagem de qualquer série temporal. Em outros, como análise de séries financeiras, o sistema gerador das séries é tão complexo que dificilmente conhecimento sobre sua dinâmica seja útil.

Assume-me, na linguagem de (HAMILTON, 1994), que um conjunto de amostras $\mathbf{y}_t = (y_1, y_2, y_3 \dots y_T)$ pode ser interpretado como um segmento finito de uma sequência duplamente infinita:

$$\mathbf{y}_{t=-\infty}^{\infty} = (\dots, y_{-1}, y_0, \overbrace{y_1, y_2, y_3, \dots, y_T}^{\text{Série Observada}}, y_{T+1}, y_{T+2} \dots)$$

Apesar de parecer pouco tangível, de fato qualquer série observada é satisfatoriamente descrita dessa forma. Em um contexto industrial, por exemplo, o valor de uma variável de processo é zero até a planta ser construída e entrar em operação, assume valores representativos ou não da dinâmica de interesse (assumindo valores irrelevantes quando a planta não está em operação plena), e tendendo ao infinito retorna a zero quando a planta for desativada.

Essa interpretação de uma série temporal é importante ao implicitamente insinuar que o processo existe em um intervalo temporal mais abrangente do que o observado. É portanto necessário se questionar sobre quanto os dados representam o processo analisado e em quais intervalos de tempo.

Séries temporais são inerentemente diferentes de dados tabulares por representarem amostras de um mesmo processo estocástico ao invés de amostras aleatórias de uma população. Não faz sequer sentido discutir uma população no contexto de séries temporais uma vez que estamos restritos à realizações observadas de um processo estocástico. Essa relevante diferença

impede propriedades estatísticas agradáveis consequentes da independência amostral associada à dados tabulares apropriadamente amostrados como a lei do grandes números e o teorema do limite central.

A forma mais natural de analisar uma série temporal é visualizar seus valores no tempo, como ilustra a imagem 1.

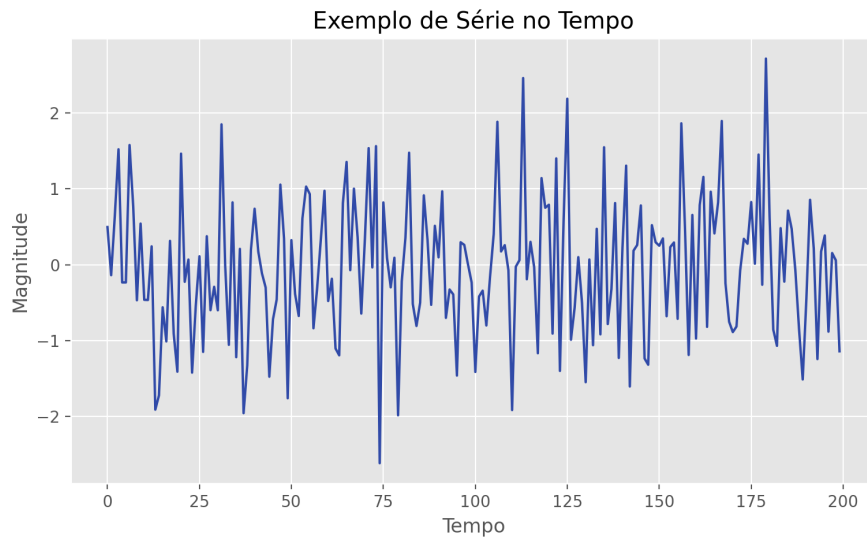


Figure 1 – Visualização no tempo de ruído branco

2.3 Operador de Atraso(Lag)

É importante introduzir o operador de atraso ou *lag*.

Dadas as séries $\mathbf{y}_t = (y_1, y_2, y_3 \dots y_T)$ e $\mathbf{x}_t = (x_0, x_1, x_2 \dots x_{T-1})$ tal que

$$\mathbf{y}_t = \mathbf{x}_{t-1}$$

isso é,

$$y_1 = x_0$$

$$y_2 = x_1$$

$$\vdots$$

$$y_T = x_{T-1}$$

Podemos definir \mathbf{x}_t em função de \mathbf{y}_t como:

$$\mathbf{x}_t = L\mathbf{y}_t$$

tal que

$$\mathbf{y}_{t-1} = L\mathbf{y}_t$$

Observamos que o operador de atraso atrasa uma série temporal em uma unidade de tempo. Uma breve divagação matemática (HAMILTON, 1994) permite definir o operador com propriedades muito semelhantes às de multiplicação dos números reais, como associatividade, comutatividade e distribuição. Para atrasar múltiplas unidades de tempo temos que:

$$L(L(\mathbf{y}_t)) = L(\mathbf{y}_{t-1}) = \mathbf{y}_{t-2} = L^2\mathbf{y}_t$$

de forma que

$$L^n\mathbf{y} = \mathbf{y}_{t-n}$$

Um uso importante do operador, decorrente de suas propriedades algébricas, é exemplificado na seguinte expansão

$$(aL^2 + bL^3)\mathbf{y}_t = a\mathbf{y}_{t-2} + b\mathbf{y}_{t-3}$$

Conhecimento do operador de atraso é importante para compreender a literatura de séries temporais e facilita comunicação objetiva de análises cotidianas. O presente trabalho usa do operador para descrever uma série de modelos.

Como nota final é importante mencionar que alguns livros (CHATFIELD; XING, 2019) (STOFFER, 2017) usam a letra B para denotar o operador de atraso e que na maior parte dos recursos *online* o operador é referido por seu nome em inglês, *lag*.

A analogia entre o operador de atraso e a variável complexa $e^{-j\omega} = z^{-1}$ é clara, com a relevante diferença que z^{-1} atrasa um sinal em uma unidade de tempo se a operação for realizada no domínio z enquanto o operador de atraso atua diretamente no domínio do tempo. Essa característica permite que filtros com equações de recorrência complexas sejam representados de forma compacta no domínio do tempo por meio de polinômios de atraso.

Uma propriedade interessante decorrente da equivalência entre z^{-1} e L é que pode se pensar em um plano L cuja análise é igualmente informativa à do plano z , notando que o espaço é de certa forma invertido. Uma análise da posição dos polos de um sistema representado por meio de um polinômio em L conclui que o sistema é instável se tais polos estiverem fora do círculo unitário, contrário do que conhecemos do plano z .

Por fim vale mencionar que alguns autores como (AGUIRRE, 2015) usam a notação q^{-1} para esse operador.

2.4 Operador de Diferença

O operador de diferenças ∇ ou Δ é o equivalente discreto da operação contínua de diferenciação e opera sob uma série temporal \mathbf{y}_t da seguinte forma:

$$\nabla \mathbf{y}_t = (1 - L)\mathbf{y}_t = \mathbf{y}_t - L\mathbf{y}_t = \mathbf{y}_t - \mathbf{y}_{t-1}$$

O operador possui propriedades de associatividade e distribuição, tal que

$$\nabla^2 \mathbf{y}_t = \nabla(\nabla(\mathbf{y}_t)) = \nabla(\mathbf{y}_t - \mathbf{y}_{t-1}) = \nabla \mathbf{y}_t - \nabla \mathbf{y}_{t-1} = \mathbf{y}_t - 2\mathbf{y}_{t-1} + \mathbf{y}_{t-2}$$

2.5 Tendência(*Trend*)

A variação do valor esperado de um processo estocástico é denominado tendência. A partir de uma série temporal definimos tendência como a variação de sua média amostral. A imagem 2 ilustra uma série com tendência linear. Observa-se que ao longo do tempo a média das observações cresce linearmente. Tendências de séries reais frequentemente seguem um perfil

logarítmico (CHATFIELD; XING, 2019), como na figura 3. Nesse caso uma transformação exponencial da série, isso é, a aplicação de uma função exponencial a cada observação, tornaria a tendência linear.

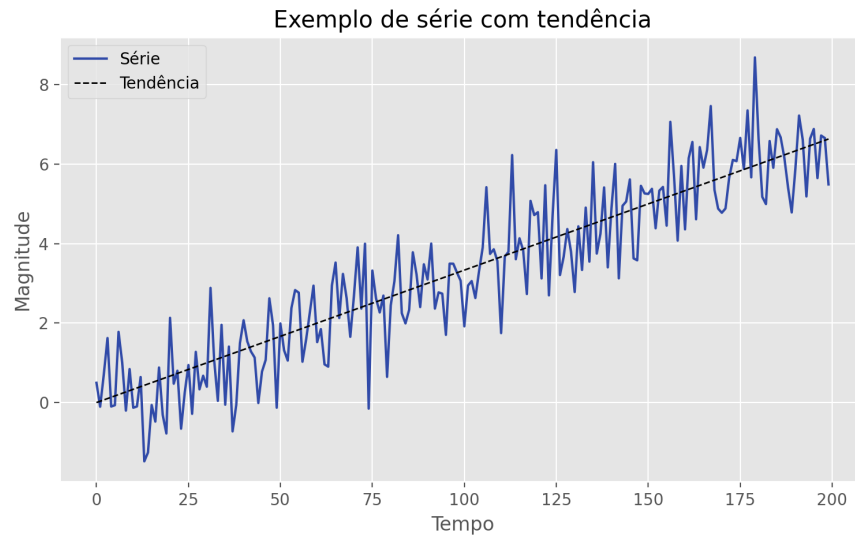


Figure 2 – Série temporal com tendência linear

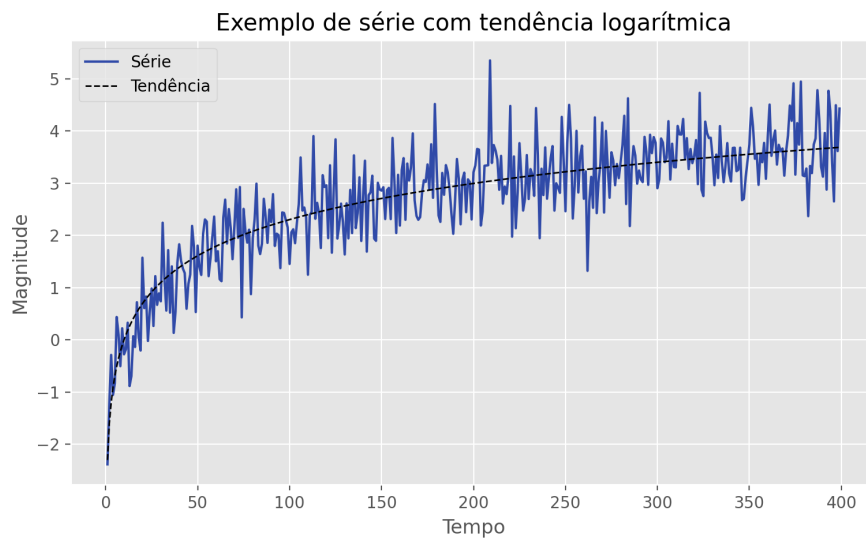


Figure 3 – Série temporal com tendência logarítmica

2.6 Sazonalidade

A variação periódica de média móvel das observações de uma série temporal é denominada sazonalidade. Em séries no contexto de finanças sazonalidade frequentemente segue ciclos de calendário como anual, mensal, semestral, etc. No contexto mais amplo de

séries temporais sazonalidade apresenta período arbitrário, apesar da linguagem em torno dessa propriedade estar muito associada aos períodos anteriormente mencionados.

Um exemplo de série temporal com sazonalidade é ilustrado na figura 4.

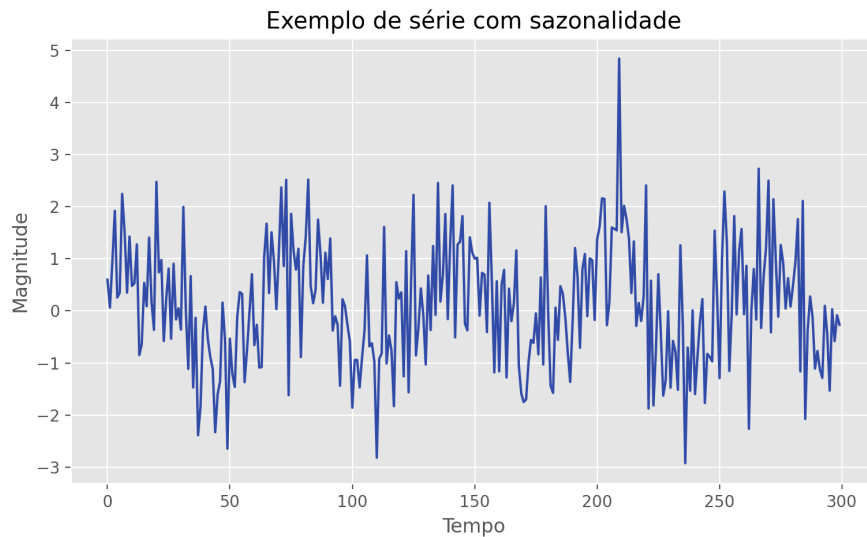


Figure 4 – Série temporal com sazonalidade senoidal de período arbitrário

Na presença de tendência sazonalidade pode ser considerada aditiva, se sua variação for constante em torno da tendência, e multiplicativa, se sua variação depender o valor da tendência. Exemplos de sazonalidade aditiva e multiplicativa são dados pelas figuras 5 e 6, respectivamente. Discernir entre os dois tipos de sazonalidade é importante para modelagem.

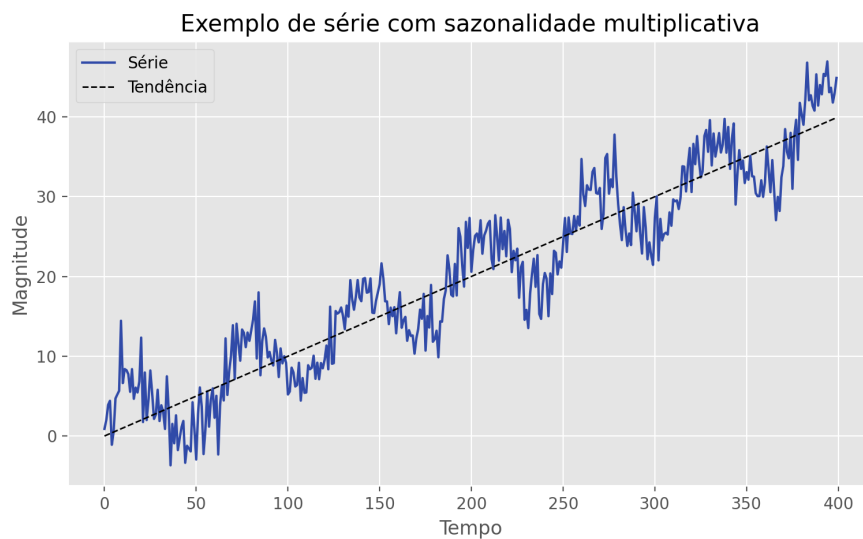


Figure 5 – Série temporal com sazonalidade aditiva

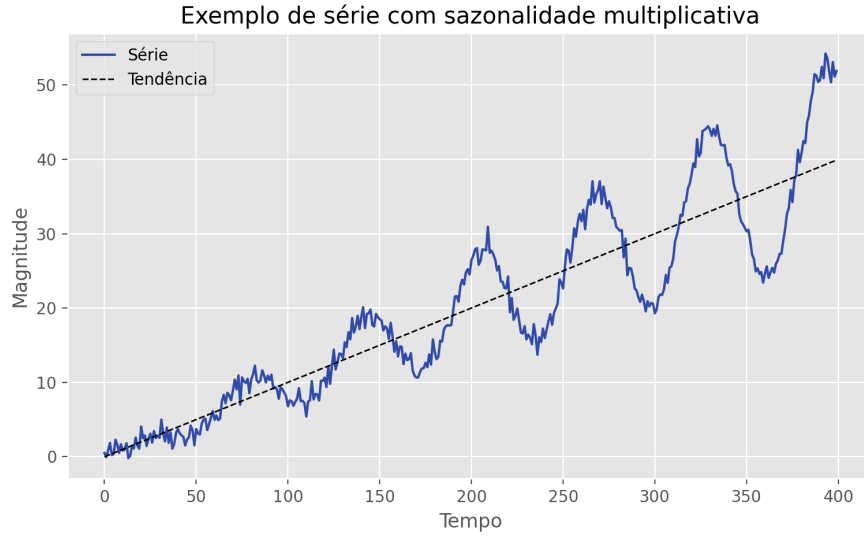


Figure 6 – Série temporal com sazonalidade multiplicativa

2.7 Autocorrelação

TODO: later analogy with convolution would be cool

A função de autocorrelação é definida para processos estocásticos como a correlação de Pearson entre valores do processo em instantes de tempo diferentes. A função de autocovariância entre os instantes de tempo t_1 e t_2 é dada pela seguinte equação

$$K_{xx}(t_1, t_2) = E[(X_{t_1} - \mu_{t_1})(X_{t_2} - \mu_{t_2})] \quad (2.1)$$

Normalizando a autocovariância obtemos a autocorrelação

$$\rho_{xx}(t_1, t_2) = \frac{K_{xx}(t_1, t_2)}{\sigma_{t_1} \sigma_{t_2}}$$

Para processos estacionários (seção 2.8) a autocovariância, e consequentemente a autocorrelação, é função apenas do atraso $\tau = |t_1 - t_2|$. Temos então que

$$\rho_{xx}(\tau) = \frac{K_{xx}(\tau)}{\sigma_{t_1} \sigma_{t_2}}$$

Para uma série temporal, isso é, uma única realização de um processo estocástico, a função de autocorrelação estacionária (tipicamente chamada apenas de função de autocorrelação)

pode ser definida diretamente a partir da definição de correlação amostral sob as seguintes premissas

1. O processo estocástico gerador da série temporal é estacionário
2. O número de observações N de \mathbf{y}_t é suficientemente grande ($N \approx 100$)

resultando na equação 2.2, onde o subscrito duplo é omitido. Note que o atraso é discreto, indicado por k .

$$\rho_y(k) = \frac{\sum_{t=1}^{N-k} (y_t - \bar{y})(y_{t+k} - \bar{y})}{\sum_{t=1}^N (y_t - \bar{y})^2}, \quad k = 0, 1, 2, \dots \quad (2.2)$$

Alguns comentários sobre a nomenclatura da função são apropriados. A literatura de engenharia tende a usar os termos autocovariância e autocorrelação de forma intercambiável para designar a definição de autocovariância apresentada. A literatura estatística assume as definições abordadas acima, que serão usada no restante desse trabalho.

Além disso, a função de autocorrelação como apresentada pela equação 2.1 é definida para um processo estocástico não necessariamente estacionário, apesar do termo ser usado para descrever a equação 2.2. A aplicação da equação que assume estacionariedade em uma série não estacionária resulta em correlações informativas praticamente apenas disso. Uma modificação estratégica da função dada por 2.1 resulta na chamada função de autocorrelação instantânea (seção 10.1.1), que é usada para representar séries temporais não estacionárias.

2.7.1 Autocorrelação Parcial

É interessante mencionar a existência de autocorrelação parcial nessa seção juntamente de uma descrição em alto nível do que esse valor representa. Sua definição formal será apresentada na seção 3.2.3.

Autocorrelação parcial foi introduzida por Box e Jenkins em (BOX; JENKINS, 2016) como uma ferramenta auxiliar na identificação de modelos. O valor ϕ_{kk} é definido como a correlação parcial entre \mathbf{y}_t e \mathbf{y}_{t-k} , isso é, a correlação restante entre \mathbf{y}_t e \mathbf{y}_{t-k} após levar em consideração a contribuição de $\mathbf{y}_t, \mathbf{y}_{t-1} \dots \mathbf{y}_{t-k+1}$.

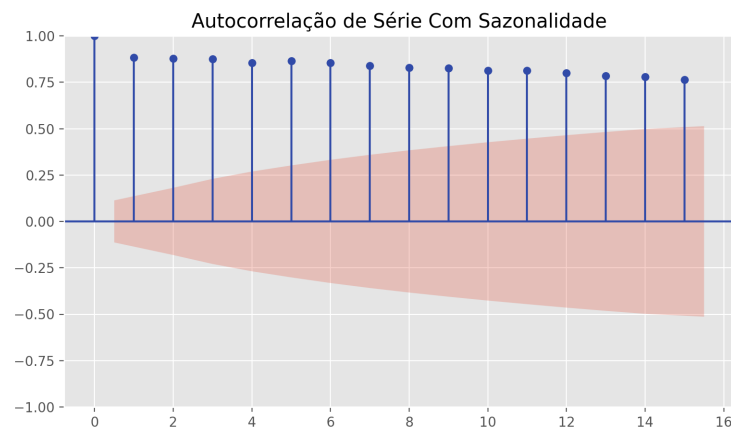


Figure 7 – Visualização de correlatograma de série com tendência linear. Observe que as autocorrelações decaem lentamente ao decorrer dos atrasos, comportamento típico de tendências determinísticas.

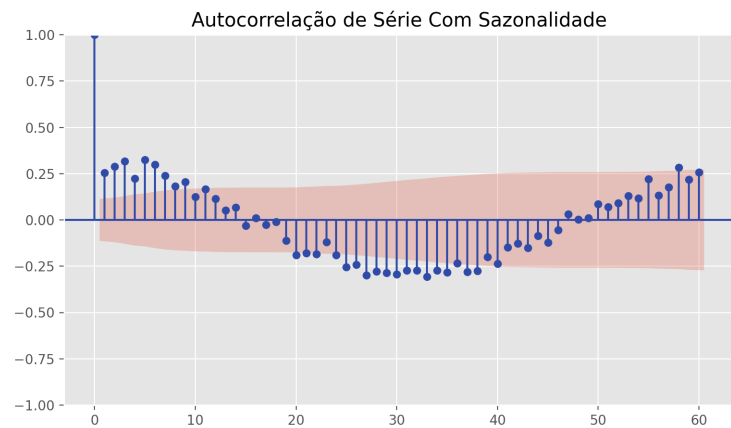


Figure 8 – Visualização de correlatograma de série com sazonalidade. Observe que o padrão senoidal da série é reproduzido nas autocorrelações.

2.7.2 Correlatograma

Um correlatograma é um gráfico de barras representativo da autocorrelação ou autocovariância em k amostras de uma série temporal ($k = 0, 1, 2, \dots$), de forma que a primeira barra represente a autocorrelação entre y_t e si mesmo (sempre igual à 1), a segunda entre y_t e y_{t-1} , a terceira entre y_t e y_{t-2} , e assim por diante. O correlatograma da série visualizada pela figura 2 é ilustrado na figura 7.

A figura 7 informa um intervalo de relevância dado por um sombreamento vermelho. Qualquer valor de autocorrelação dentro desse intervalo é estatisticamente insignificante e pode ser considerado igual a zero.

O correlatograma é uma ferramenta indispensável em análise de séries temporais

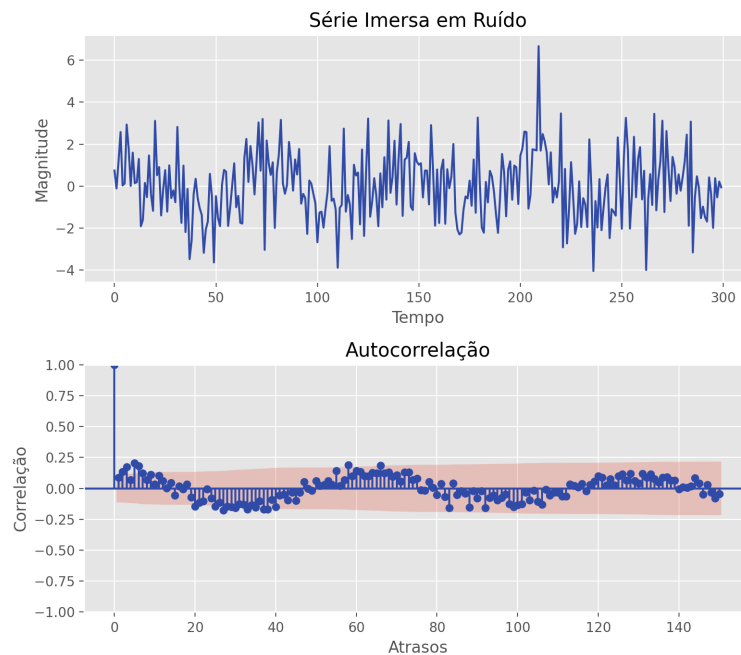


Figure 9 – Sinal periódico imerso em ruído e sua correspondente autocorrelação.

para tarefas como detecção de estacionariedade, identificação de sazonalidade, análise de resíduo, engenharia de características, escolha de modelo e identificação de ruído branco (seção 2.9), entre outros. No contexto de identificação de sistemas autocovariância e autocorrelação e portanto o correlalograma desempenha um importante papel na identificação de propriedades de sinais e sistemas imersos em ruído devido à robustez ao ruído da operação de correlação cruzada (AGUIRRE, 2015). A figura 9 apresenta um sinal imerso em ruído cuja natureza periódica subjacente se torna mais visível por meio de seu correlalograma.

É importante mencionar que o correlalograma de uma série com tendência determinística, como a da figura 7, apresenta o comportamento observado de autocorrelações altas com pouca atenuação ao longo dos atrasos. De forma análoga o correlalograma de uma série com sazonalidade apresenta periodicidade que reproduz seu padrão temporal, como ilustra a figura 8, correlalograma da série da figura 4. O primeiro correlalograma é informativo até certo ponto: informa simplesmente que a série apresenta tendência. Para analisar tais séries de forma mais produtiva, a fim de elaborar um possível modelo, por exemplo, é importante que a série seja estacionária (seção 2.8). É inclusive afirmado em alguns textos da literatura estatística (CHATFIELD; XING, 2019), que um correlalograma só faz sentido se a série associada for estacionária, observação mais geral e rigorosa das restrições de uso da equação 2.2.

Na engenharia a análise do correlalograma de sinais não estacionários é utilizada para investigação da adequação de tempo de amostragem, onde uma autocovariância com valores

lentamente decrescentes e um mínimo local indica que o sinal pode estar superamostrado, propriedade indesejável que pode resultar em problemas computacionais além de desperdício de memória.

TODO: add sampling analysis of silica series? later maybe

2.8 Estacionariedade

Uma série temporal \mathbf{x}_t gerada por um processo $\mathbf{X}(t)$ é considerada estacionária se atender às seguintes três condições:

1. $E(\mathbf{X}(t)) = \mu$
2. $Var(\mathbf{X}(t)) = \sigma^2$
3. $Cov[\mathbf{X}(t), \mathbf{X}(t + \tau)] = \gamma(\tau)$

Que podem ser interpretadas da seguinte forma

1. A média da série \mathbf{x}_t é constante ao longo do tempo
2. A variância da série \mathbf{x}_t é constante ao longo do tempo
3. A autocorrelação de \mathbf{x}_t depende apenas do atraso

Um bom exemplo de graus de estacionariedade em séries temporais dado por Athanasopoulos e Hyndman (ATHANASOPOULOS; HYNDMAN, 2018) é ilustrado pela figura 10. As séries das figuras 10.a, 10.e e 10.i demonstram clara tendência, sendo portanto não estacionárias. As séries das figuras 10.d, 10.h, 10.i são igualmente não estacionárias por apresentarem clara sazonalidade, enquanto a série da figura 10.g aparenta ser sazonal mas apresenta picos aperiódicos de intensidade muito distinta, sendo considerada estacionária pelos autores. O caso da série da figura 10.g é um complicado por aparentar também violar o requisito de variância constante. Para concluir assim como os autores que a série é estacionária é necessário mais do que uma investigação visual; o texto que discute a série leva em conta seu processo gerador (ATHANASOPOULOS; HYNDMAN, 2018).

As séries das figuras 10.c e 10.f são aparentemente não estacionárias no intervalo observado por demonstrarem variação em sua média móvel mas podem ser um caso de raízes unitárias.

Por eliminação temos que apenas as séries ilustradas pelas figuras 10.b e 10.g são estacionárias, o caso de 10.b contendo uma clara anomalia.

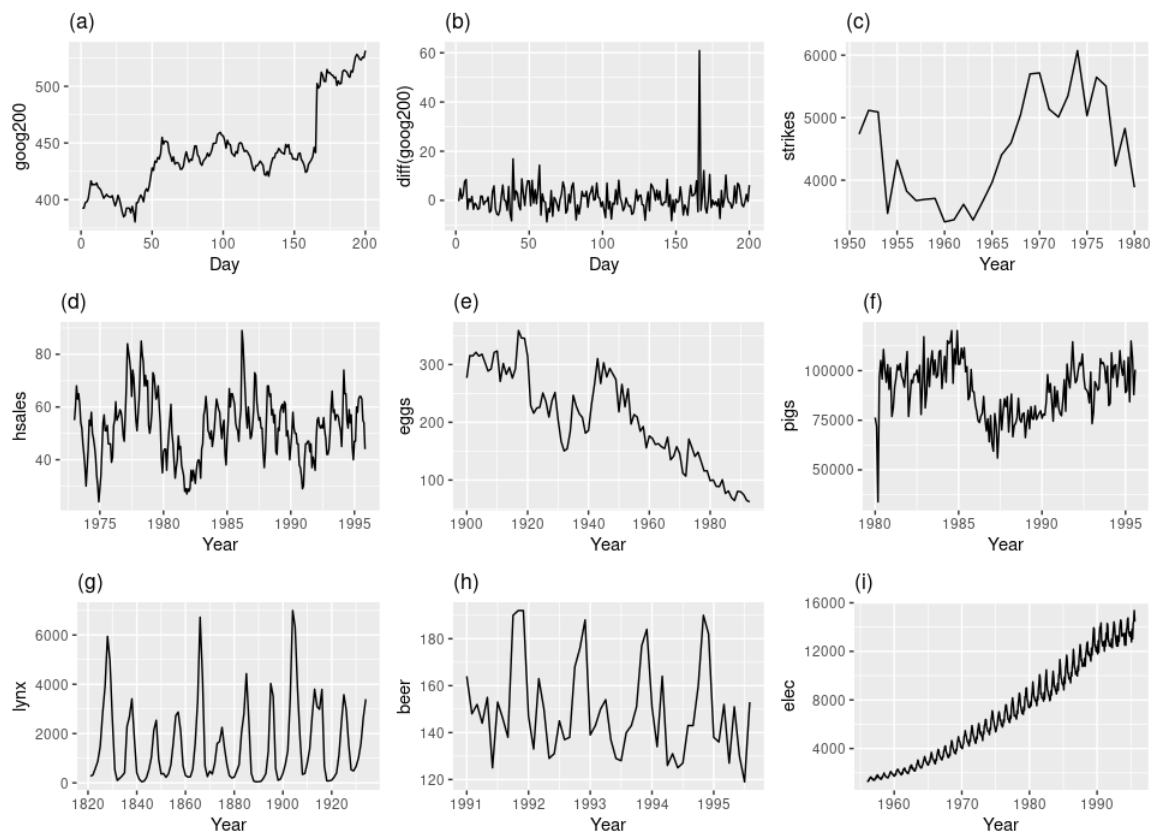


Figure 10 – Conjunto de séries demonstrando diferentes níveis de estacionariedade.

O exemplo de análise de estacionariedade da figura 10 demonstra a imprecisão da abordagem visual para essa tarefa. Torna-se necessário o estabelecimento de procedimentos mais objetivos para detecção de estacionariedade (seção ??).

2.8.1 Importância de Estacionariedade

Estacionariedade é uma propriedade desejável de se observar em uma série temporal para fins analíticos e de modelagem.

Há algumas formas de compreender como estacionariedade colabora para a modelagem bem sucedida de uma série temporal.

Podemos pensar na propriedade de estacionariedade como um tipo de estrutura de dependência. Se as amostras X_1, X_2, \dots, X_N de um conjunto \mathbf{X} forem independentes entre si temos formas interessantes de modelar a função geradora de \mathbf{X} como o teorema do limite central, lei dos grandes números, etc. Há apenas uma forma de um conjunto amostral ser independente mas muitas formas de ser dependente, tornando difícil o estabelecimento de recursos eficientes para modelagem geral de processos dependentes. Séries temporais sendo naturalmente observações de

processos dependentes (seção 2.2) é interessante definir estruturas de dependência que permitam o uso de tais recursos. Estacionariedade é uma estrutura de dependência que permite aplicar propriedades úteis de independência em séries temporais. Abordando a mesma ideia mais intuitivamente podemos pensar no seguinte exemplo: se um processo possui valor esperado e variância constante e autocorrelação invariante ao tempo podemos por meio da lei dos grandes números estimar seu valor esperado e variância com cada vez mais confiança a partir da média e variância amostral, respectivamente. O mesmo argumento intuitivo se estende analogamente para o caso de aprendizado de máquina, no qual por meio de teoria de aprendizado estatístico é possível argumentar que uma série estacionária é “mais fácil” de aprender.

De forma mais quantitativa o teorema de decomposição de Wold (CHATFIELD; XING, 2019) conclui que qualquer série temporal estacionária pode ser representada pela seguinte combinação linear

$$y_t = \sum_{j=0}^{\infty} b_j Z_{t-j} + \eta_t$$

No qual η representa uma série determinística e Z_t um processo puramente aleatório (seção 2.9). O leitor reconhecerá parte da expressão acima como um processo $MA(\infty)$ (seção 3.2.4). Esse resultado tem como consequência a importante conclusão que qualquer série estacionária é possivelmente aproximável por um modelo MA e portanto, via invertibilidade, modelos AR e ARMA (seção 3.3.1).

Por fim estacionariedade permite o uso de uma série de modelos que serão discutidos na seção sobre modelos estacionários. Esses métodos são bem compreendidos e implementados, facilitando sua interpretação, uso e sustentação.

2.8.2 Categorias Básicas de Não Estacionariedade

Como extensão do argumento sobre estruturas de dependência na seção anterior podemos afirmar que, sendo estacionariedade um padrão de dependência, temos infinitas formas de não estacionariedade, retornando ao caso de dependência generalizada. É interessante identificar nesse universo de dependência padrões de séries não estacionárias que são facilmente transformadas em séries estacionárias.

Uma série temporal com presença de tendência determinística, como ilustrada na figura 2, pode ser representada pela seguinte expressão:

$$y_t = e_t + f(t) + \varepsilon_t \quad \text{onde } \varepsilon_t \sim \text{i.i.d. } \mathcal{N}(0, \sigma^2)$$

Na qual e_t representa uma série estacionária, $f(t)$ uma função determinística do tempo e ε_t ruído branco (seção 2.9). Nota-se que $f(t)$ é uma função monotônica arbitrária tal que y_t seja uma série não estacionária. No caso da figura 2 temos $f(t)$ linear e na figura 3 logarítmica. Uma série temporal demonstrando esse tipo de não estacionariedade é considerada **tendência-estacionária**, uma vez que simplesmente removendo a tendência $f(t)$ temos estacionariedade. Isso pode ser feito de várias formas, talvez com maior simplicidade diferenciando a série. Métodos mais sofisticados incluem decomposição ETS (seção 3.1) e regressão com finalidade de modelar $f(t)$ de forma que o resíduo represente uma série estacionária $e_t + \varepsilon_t$.

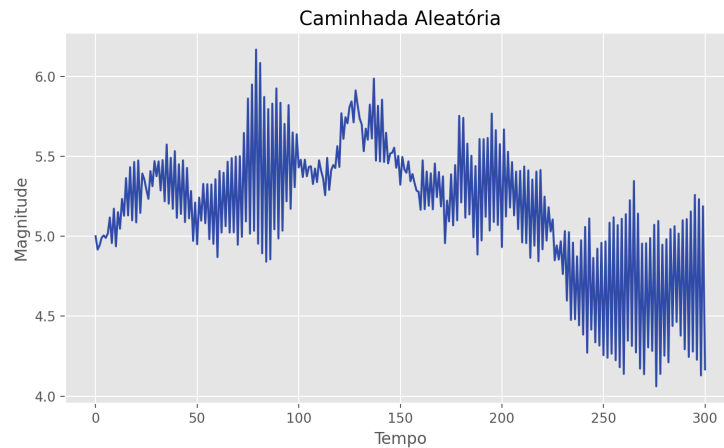


Figure 11 – Visualização de caminhada aleatória com $y_0 = 5$

Uma série com presença de tendência estocástica pode ser classificada de maneira semelhante. O exemplo mais simples de tal série é gerada por um passeio aleatório, definido pelo seguinte processo, visualizado pela figura 11:

$$y_t = y_{t-1} + \varepsilon_t \quad \text{onde } \varepsilon_t \sim \text{i.i.d. } \mathcal{N}(0, \sigma^2)$$

Por meio de um desenvolvimento recursivo do processo podemos escrever:

$$y_t = (y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t$$

$$y_t = ((y_{t-3} + \varepsilon_{t-2}) + \varepsilon_{t-1}) + \varepsilon_t$$

\vdots

$$y_t = \sum_{j=0}^{N-1} \varepsilon_{t-j} + y_0$$

Resultado a partir do qual a não estacionariedade de y_t se torna evidente, uma vez que

$$\text{var}(y_t) = \sigma^2 t$$

Além da covariância ser dependente do tempo.

Uma forma simples de tornar a série estacionária é diferenciá-la em primeira ordem, isso é, aplicar o operador de diferença primeira:

$$\nabla y_t = y_t - y_{t-1}$$

$$y_t - y_{t-1} = \varepsilon_t$$

$$\nabla y_t = \varepsilon_t$$

Sabemos pela seção 2.9 que ruído branco é um processo estacionário.

A caminhada aleatória é denominada uma série **diferença-estacionária** pelo fato da operação de diferença introduzir estacionariedade. Essa é uma forma tão comum de não estacionariedade que a ideia de “diferenciar uma série antes de fazer qualquer coisa” é proeminente entre profissionais de dados, apesar de que geralmente necessita-se apenas de estacionariedade. Essa prática é parcialmente justificada considerando que a maior parte das séries temporais “reais” são não estacionárias e frequentemente diferencialmente estacionárias.

É igualmente possível que uma série diferença-estacionária seja estacionária em sua n -ésima diferença, tal que estacionariedade seja observada por uma operação de diferenças de ordem n . A ideia de tirar sucessivas diferenças até atingir estacionariedade é fundamental no método de Box-Jenkins, por exemplo.

Séries diferença-estacionárias apresentam raízes unitárias e os dois termos são frequentemente usados nos mesmos contextos.

Podemos resumir as definições das categorias de não estacionariedade abordadas nessa seção assim como suas implicações como segue:

1. **Estacionariedade em Tendência:** Uma série é considerada tendência-estacionária se apresentar uma tendência determinística. No caso de anomalias ou eventos de perturbação séries com esse tipo de tendência retornam ao valor da tendência ao longo do tempo, “esquecendo” o evento perturbador. Esse tipo de série se torna estacionária pela remoção da tendência determinística, processo realizado por meio de regressão da tendência, por diferenciação, por decomposição, etc.
2. **Estacionariedade Diferenciável:** Uma série é considerada diferença-estacionária se apresentar uma tendência estocástica. No caso de anomalias ou eventos de perturbação séries com esse tipo de tendência são irreversivelmente afetadas, “lembrando” do evento perturbador. Esse tipo de série se torna estacionária por diferenciação em ordem n . Possui raízes unitárias e é frequentemente discutida nessa linguagem.

2.9 Ruído Branco

Uma série temporal \mathbf{x}_t gerada por um processo $\mathbf{X}(t)$ é considerada ruído branco ou um processo puramente aleatório se atender às seguintes três condições:

1. $E(\mathbf{X}(t)) = 0$
2. $Var(\mathbf{X}(t)) = \sigma^2$
3. $Cov[\mathbf{X}(t), \mathbf{X}(t + \tau)] = 0$

Que podem ser interpretadas da seguinte forma

1. A média da série \mathbf{x}_t é nula ao longo do tempo
2. A variância da série \mathbf{x}_t é constante ao longo do tempo
3. Não há correlação entre as amostras de \mathbf{x}_t

Observa-se que ruído branco é um caso específico de estacionariedade, se diferenciando pela especificação do valor esperado e autocorrelação entre quaisquer amostras em zero. A compreensão da definição e capacidade de identificação de ruído branco é importante para análise de resíduos, detalhado na seção 7.1.

3 TEORIA UNIVARIADA

3.1 Decomposição de Séries Temporais

3.2 Modelos Estacionários

3.2.1 Modelo Linear Generalizado

3.2.2 Raízes Unitárias

3.2.3 Modelo Autoregressivo

3.2.4 Modelo Média Móvel

3.2.5 Modelo ARMA

3.3 Modelo ARIMA

3.3.1 Estabilidade e Invertibilidade

3.4 Modelos de Sazonalidade

3.4.1 Diferenciação Sazonal

3.4.2 SARIMA

4 TEORIA DE VOLATILIDADE

4.1 Definições Adicionais

4.2 Modelo de Heterocedacidade Condicional Autoregressiva Generalizada

4.3 Modelo de Volatilidade Estocástica

4.4 Modelos Adicionais

5 TEORIA MULTIVARIADA

5.1 Definições Adicionais

Adicionar medidas de similaridade? Possivelmente em capítulo de análise moderna

5.1.1 *Correlação Cruzada*

5.1.2 *Cointegração*

5.2 Modelos em Espaço de Estados

5.3 Modelos Multivariados

5.4 Modelos de Variáveis Exógenas

6 TESTES ESTATÍSTICOS

6.1 Considerações sobre testes em séries temporais

6.2 Teste de *Dickey-Fuller*

Os testes de *Dickey-Fuller* testam a hipótese nula de presença de raízes unitárias no processo gerador de uma série temporal com a hipótese alternativa de estacionariedade.

O teste mais simples de *Dickey-Fuller* assume que o processo gerador da série temporal em questão é dado pela equação 6.1, na qual ε é ruído branco.

$$y_t = \phi y_{t-1} + \varepsilon_t \quad (6.1)$$

A hipótese nula do teste é a presença de raízes unitárias em y_t , isso é, $\phi = 1$, e a hipótese alternativa é $\phi < 1$, correspondente à estacionariedade.

$$\begin{cases} H_0 : \phi = 1, \text{ não estacionariedade (raízes unitárias, possível diferença-estacionariedade)} \\ H_1 : \phi < 1, \text{ estacionariedade} \end{cases}$$

Em seguida Δy_t é subtraído de ambos os lados da equação 6.1, resultando no desenvolvimento a seguir:

$$y_t - y_{t-1} = \phi y_{t-1} - y_{t-1} + \varepsilon$$

$$\Delta y_t = (\phi - 1)y_{t-1} + \varepsilon$$

$$\Delta y_t = \delta y_t + \varepsilon$$

As hipóteses do teste são agora reformuladas para as seguintes

$$\begin{cases} H_0 : \delta = 0, \text{ não estacionariedade (raízes unitárias, diferença-estacionariedade)} \\ H_1 : \delta < 0, \text{ estacionariedade} \end{cases}$$

TODO: definir estatística de teste

A estatística de teste é computada e comparada com um valor crítico proveniente da distribuição de *Dickey-Fuller* (geralmente sob $p = 0.5$) para rejeição ou não da hipótese nula.

O modelo da série temporal, dado explicitamente pela equação 6.1, pode ser alterado para testar raízes unitárias com constante e com constante e tendência determinística no tempo por meio das equações 6.2 e 6.3, respectivamente. Note que ambas as equações são apresentadas em função de δ . É mais comum testar por apenas raízes unitárias, uma vez que uma análise subjetiva visual, por correlatograma ou decomposição juntamente de remoção de tendência ou sazonalidade é tipicamente realizada antes de um teste estatístico de estacionariedade.

$$\nabla y_t = \delta y_{t-1} + u_t + a_0 \quad (6.2)$$

$$\nabla y_t = \delta y_{t-1} + u_t + a_0 + a_1 t \quad (6.3)$$

O teste aumentado de *Dickey-Fuller* (ADF) modela o processo de forma mais geral, incluindo na equação 6.1 termos representativos de processos estacionários arbitrários. O teste aumentado é projetado para remover autocorrelação do processo de validação de hipótese. De forma identicamente análoga ao teste de DF temos expansões do ADF para incluir constantes e tendências como nas equações 6.2 e 6.3 apesar de que, como no teste de DF, essas variações são pouco usadas. A estatística de teste do ADF é negativa, isso é, quanto menor seu valor maior a rejeição da hipótese nula de não estacionariedade (maior certeza de estacionariedade).

Como em qualquer teste de hipótese um valor p maior que 0.05 indica falha em rejeitar a hipótese nula, nesse caso correspondendo à impossibilidade de constatar estacionariedade. Um valor p menor ou igual a 0.05 indica rejeição da hipótese nula, correspondendo à conclusão que a série sob análise é estacionária.

Na prática o teste mais usado é o ADF que é uma simples extensão mais robusta do teste de *Dickey-Fuller*. Implementações eficiente e populares existem para R e Python.

6.2.1 Teste Kwiatkowski-Phillips-Schmidt-Shin (KPSS)

TODO: check this

O teste KPSS desempenha uma função semelhante ao ADF com a relevante diferença de inerentemente modelar uma tendência linear no tempo por meio da equação 6.4. Seu

desenvolvimento matemático é análogo porém mais trabalhoso que o caso do ADF e será portanto omitido.

$$y_t = \phi y_{t-1} + \varepsilon_t + \beta t \quad (6.4)$$

Com ε_t representando ruído branco. O teste em seguida define as seguintes hipóteses:

$$\begin{cases} H_0 : \text{a série apresenta tendência-estacionariedade} \\ H_1 : \text{a série apresenta raízes unitárias} \end{cases}$$

Como em qualquer teste de hipótese um valor p maior que 0.05 indica falha em rejeitar a hipótese nula, nesse caso correspondendo à impossibilidade de constatar que a série não apresenta tendência estacionariedade, boa evidência de que a série é tendência-estacionária. Um valor p menor ou igual a 0.05 indica rejeição da hipótese nula, correspondendo à conclusão que a série sob análise possui raízes unitárias e é portanto não estacionária.

Observe que há uma diferença crítica: a alternativa nula não postula não estacionariedade, como no caso do ADF, mas sim tendência-estacionariedade (seção 2.8.2) decorrente diretamente da inclusão de tendência linear no modelo da equação 6.4. A diferença principal da alteração da hipótese nula é que o KPSS é usado para investigar presença de estacionariedade sob uma tendência determinística (tendência-estacionariedade) e o ADF (tipicamente) de estacionariedade propriamente dita. O KPSS é bem implementado em R e Python.

6.3 Causalidade de Granger

6.4 Box-Pierce

6.5 Ljung-Box-Pierce

6.6 *Convergence Cross Mapping*

6.7 Johansen

7 VALIDAÇÃO DE MODELOS

7.1 Análise de Resíduos

7.2 Validação Cruzada

8 ANÁLISE ESPECTRAL

9 MODELOS NÃO LINEARES

10 ANÁLISE DE SÉRIES NÃO ESTACIONÁRIAS

10.1 Definições Adicionais

10.1.1 Autocorrelação Instantânea

10.2 Indução de Estacionariedade

11 TÓPICOS ADICIONAIS

12 PROCESSAMENTO DE SINAIS ESTOCÁSTICOS

12.1 Introdução

13 APLICAÇÕES

14 CONCLUSÕES E TRABALHOS FUTUROS

Conclusao

14.1 Contribuições do Trabalho

14.2 Limitações

14.3 Trabalhos Futuros

REFERENCES

AGUIRRE, L. A. **Introdução à Identificação de Sistemas. Técnicas Lineares e não Lineares Aplicadas a Sistemas. Teoria e Aplicação.** [S.l.]: UFMG, 2015.

ATHANASOPOULOS, G.; HYNDMAN, R. **Forecasting: Principles and Practice.** [S.l.]: O'Reily, 2018.

BENDAT, J.; PERSOL, G. **Random Data Analysis and Measurement Procedures.** [S.l.]: John Wiley & Sons, 2000.

BOX, G.; JENKINS, G. **Time Series Analysis: Forecasting and control.** [S.l.]: Wiley, 2016.

CHATFIELD, C.; XING, H. **The Analysis of Time Series: An introduction with r.** [S.l.]: CRC Press, 2019.

HAMILTON, J. **Time Series Analysis.** [S.l.]: Princeton University Press, 1994.

STOFFER, R. S. D. **Time Series Analysis and Its Applications.** [S.l.]: Springer, 2017.

APPENDIX A – EXEMPLO DE APÊNDICE

Um apêndice é um documento elaborado pelo autor, diferentemente do anexo. Geralmente, se coloca como apêndice, questionários, códigos de programação, tabelas que tomariam muito espaço no meio do trabalho. Artigos, resumos ou qualquer publicação relacionada ao trabalho podem ser utilizados como apêndice.

ANNEX A – EXEMPLO DE UM ANEXO

Um anexo é um documento que não foi elaborado pelo autor, ou seja, o autor apenas anexa. Anexos podem ser tabelas, mapas, diagramas, *datasheets*, manuais e etc.