

Previsão do Preço do Petróleo Brent via Séries Temporais Clássicas e Deep Learning

Luiz Tiago Wilcke
UNISOCIESC

Junho de 2025

Resumo

Este Trabalho de Conclusão de Curso (TCC) investiga a modelagem e previsão do preço do Brent (petróleo) empregando técnicas de séries temporais clássicas (ARIMA/SARIMA, GARCH e extensões) e métodos de aprendizado profundo (LSTM). Dados históricos reais mensais e diários de janeiro de 2015 a dezembro de 2024 são obtidos via U.S. Energy Information Administration (EIA) e outras fontes. Realiza-se análise exploratória (testes de estacionaridade ADF/KPSS, ACF, PACF), ajustam-se modelos ARIMA/SARIMA segundo Box-Jenkins com detalhamento de estimadores e verossimilhança, modela-se volatilidade condicional via GARCH, EGARCH, GJR-GARCH e GARCH-MIDAS, e implementa-se LSTM univariado com validação walk-forward. Apresentam-se equações detalhadas: Yule-Walker, representações em espaço de estado e Kalman Filter, funções de verossimilhança exata e condicional, equações de forecast, derivadas espectrais, bem como formulas internas de redes recorrentes. Resultados numéricos ilustrativos são descritos de forma genérica e tabelas de métricas em exemplo. Discutem-se limitações, implicações práticas para hedge e gestão de risco, e propostas de trabalhos futuros, incluindo inclusão de variáveis exógenas, ensembles e modelos de mudança de regime.

Sumário

1	Introdução	1
1.1	Objetivos	1
1.2	Estrutura do TCC	2
2	Revisão de Literatura e Fundamentação Teórica	3
2.1	Conceitos de Séries Temporais	3
2.1.1	Definições básicas	3
2.1.2	Operador defasagem e polinômios	3
2.1.3	Condições de estacionaridade e invertibilidade	3
2.1.4	Função Autocorrelação (ACF) e Autocorrelação Parcial (PACF)	4
2.1.5	Transformada de Fourier e densidade espectral	4
2.2	Modelos ARIMA/SARIMA	4
2.2.1	Modelo ARIMA(p,d,q)	4
2.2.2	Testes de raiz unitária	5
2.2.3	Identificação de ordens via ACF/PACF	5
2.2.4	Função de verossimilhança	5
2.2.5	Estimadores e propriedades	5
2.2.6	Diagnóstico de resíduos	6
2.2.7	Previsão ARIMA	6
2.2.8	Representação em Espaço de Estado e Kalman Filter	6
2.3	Modelos de Volatilidade	7
2.3.1	Heterocedasticidade condicional	7
2.3.2	Modelo ARCH(q)	7
2.3.3	Modelo GARCH(p,q)	7
2.3.4	Extensões de GARCH	7
2.3.5	Diagnóstico em modelos de volatilidade	8
2.4	Redes Neurais Recorrentes e LSTM	8
2.4.1	Motivação para séries temporais	8
2.4.2	Equações internas de LSTM	8
2.4.3	Preparação de dados para LSTM	9
2.4.4	Função de perda e otimização	9
2.4.5	Validação walk-forward	9
2.4.6	Explainability e interpretação	9
2.5	Modelos de Mudança de Regime	10
2.5.1	Modelos Markov-Switching	10
2.5.2	Modelos de intervenção e dummies	10
2.6	Avaliação de Modelos e Critérios de Seleção	10
2.6.1	Critérios de informação	10

2.6.2	Testes de comparação de previsões	10
2.6.3	Backtesting de VaR	11
2.7	Referências Teóricas Ampliadas	11
3	Dados e Pré-processamento	12
3.1	Fontes de Dados	12
3.2	Período e Frequência	12
3.3	Pré-processamento	12
4	Metodologia de Modelagem	13
4.1	Modelos ARIMA/SARIMA	13
4.2	Modelos de Volatilidade	13
4.3	Modelos de Mudança de Regime	14
4.4	Redes Neurais Recorrentes e LSTM	15
4.5	Crítérios de Comparação e Testes Estatísticos	15
5	Resultados e Estrutura de Apresentação	16
5.1	Teste de Estacionaridade	16
5.2	Ajuste ARIMA/SARIMA	16
5.2.1	Modelos Testados e Seleção	16
5.2.2	Parâmetros Estimados	16
5.2.3	Diagnóstico de Resíduos	17
5.2.4	Forecast ARIMA/SARIMA	17
5.3	Ajuste de Modelos de Volatilidade	17
5.3.1	Parâmetros GARCH	17
5.3.2	Diagnóstico	17
5.3.3	Forecast de Volatilidade	17
5.4	Resultados de LSTM	17
5.4.1	Treinamento e Validação	17
5.4.2	Desempenho Fora da Amostra	18
5.4.3	Validação Walk-Forward	18
5.5	Modelos de Regime e Intervenção	18
6	Discussão Geral	19
7	Conclusão e Trabalhos Futuros	20
	Referências	22
A	Código em Python	25
A.1	Coleta e Pré-processamento	25
A.2	Ajuste ARIMA/SARIMA	25
A.3	Ajuste GARCH e extensões	25
A.4	Implementação LSTM	25
A.5	Testes Estatísticos Detalhados	25

B	Derivações Matemáticas Complementares	26
B.1	Equações de Yule-Walker	26
B.2	Representação em Espaço de Estado	26
B.3	Kalman Filter	26
B.4	Função de Verossimilhança	26
B.5	Whittle Estimation	26

Capítulo 1

Introdução

A previsão do preço do petróleo Brent é vital para decisões de hedge em mercados financeiros, planejamento orçamentário de empresas de energia e formulação de políticas públicas, devido ao forte impacto macroeconômico de flutuações de preços [EIA, 2025, TE, 2025, MacroTrends, 2025]. Choques exógenos (pandemia de COVID-19, crises geopolíticas) podem introduzir mudanças abruptas de regime, exigindo análise cuidadosa de estacionaridade, sazonalidade e heterocedasticidade condicional. Este TCC detalha, em nível matemático e computacional, como aplicar modelos clássicos de séries temporais (ARIMA/SARIMA, GARCH e extensões), representações em espaço de estado, e redes LSTM, avaliando desempenho fora da amostra e discutindo robustez em diferentes condições de mercado.

1.1 Objetivos

- Desenvolver pipeline de coleta e pré-processamento de dados históricos de preço do Brent (mensal e diário) para o período jan/2015–dez/2024.
- Realizar análise exploratória: testes de estacionaridade (ADF, KPSS), análise de autocorrelação (ACF) e autocorrelação parcial (PACF), estatísticas descritivas (média, variância, skewness, kurtosis) e identificação de outliers e eventos extremos.
- Ajustar modelos ARIMA/SARIMA via metodologia Box-Jenkins: seleção de ordens $(p,d,q)(P,D,Q)_{12}$ por inspeção de ACF/PACF e critérios de informação (AIC, BIC), detalhando funções de verossimilhança condicional e exata, estimadores e propriedades de consistência.
- Ajustar modelos de volatilidade ARCH/GARCH sobre resíduos do modelo de média: estimar parâmetros via máxima verossimilhança condicional, diagnosticar persistência, e prever volatilidade condicional. Explorar extensões: EGARCH, GJR-GARCH, GARCH-MIDAS, volatilidade estocástica.
- Implementar modelo LSTM univariado para previsão de curto prazo: preparação de janelas, normalização, construção de arquitetura, treinamento com early stopping, validação walk-forward e explicabilidade.
- Avaliar desempenho fora da amostra (RMSE, MAE, MAPE) em diferentes janelas, inclusive janelas de crise (março–abril/2020, etc.). Incluir comparações teóricas de viés e variância dos estimadores.

- Discutir implicações práticas para hedge e gestão de risco, limitações dos métodos, e propor direções futuras: inclusão de variáveis exógenas (ARIMAX, GARCH-X, LSTM multivariado), ensembles, modelos de regime (Markov-Switching), GARCH-MIDAS, volatilidade estocástica, representações espectrais, Whittle estimation.

1.2 Estrutura do TCC

- Capítulo 2: Revisão de Literatura e Fundamentação Teórica (muitas equações adicionais).
- Capítulo 3: Dados e Pré-processamento.
- Capítulo 4: Metodologia de Modelagem (detalhamento ARIMA, espaço de estados, GARCH e extensões, LSTM).
- Capítulo 5: Estrutura de Resultados e Exemplos de Tabelas (ilustrativos), Testes Estatísticos.
- Capítulo 6: Discussão Geral.
- Capítulo 7: Conclusão e Trabalhos Futuros.
- Apêndices: Códigos em Python, Testes Estatísticos Detalhados, Tabelas Extensas, Derivações Matemáticas Complementares.
- Referências Bibliográficas.

Capítulo 2

Revisão de Literatura e Fundamentação Teórica

Esta seção aprofunda conceitos e equações de séries temporais e volatilidade, além de redes LSTM.

2.1 Conceitos de Séries Temporais

2.1.1 Definições básicas

Seja $\{X_t\}_{t \in \mathbb{Z}}$ uma série temporal. - *Estacionaridade estrita*: a distribuição conjunta de $(X_{t_1}, \dots, X_{t_k})$ independe de (t_1, \dots, t_k) por translação de tempo. - *Estacionaridade fraca* (ou em segundo momento): $E[X_t] = \mu$ constante e $\text{Cov}(X_t, X_{t+h}) = \gamma(h)$ depende apenas do lag h . - *Ergodicidade*: médias de amostra convergem a médias teóricas sob certas condições.

2.1.2 Operador defasagem e polinômios

Defina o operador de defasagem B tal que $BX_t = X_{t-1}$, $B^k X_t = X_{t-k}$. - Polinômio autoregressivo:

$$\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p.$$

- Polinômio de média móvel:

$$\theta(B) = 1 + \theta_1 B + \dots + \theta_q B^q.$$

- Operador de diferenciação:

$$(1 - B)^d = \sum_{j=0}^d \binom{d}{j} (-1)^j B^j,$$

que em séries inteiras define ARIMA.

2.1.3 Condições de estacionaridade e invertibilidade

- Modelo AR(p): $\phi(B)X_t = \varepsilon_t$. A condição de estacionaridade é que todas as raízes da equação característica $\phi(z) = 0$ satisfaçam $|z| > 1$. - Modelo MA(q): $X_t = \theta(B)\varepsilon_t$. Condição de invertibilidade: todas as raízes de $\theta(z) = 0$ satisfazem $|z| > 1$. - Em ARMA, ambas condições devem valer para representações estacionárias/invertíveis.

2.1.4 Função Autocorrelação (ACF) e Autocorrelação Parcial (PACF)

- Defina autocovariância $\gamma(h) = \text{Cov}(X_t, X_{t+h})$. ACF:

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)}.$$

- PACF: denotada $\alpha(h)$, é o coeficiente de X_{t-h} em regressão de X_t contra $\{X_{t-1}, \dots, X_{t-h}\}$.

- Propriedades de corte/decrescimento: - AR(p) estacionária: PACF corta após lag p ($\alpha(h) = 0$ para $h > p$); ACF decai exponencialmente ou oscila amortecido. - MA(q): ACF corta após lag q ; PACF decai exponencial/oscila amortecido. - ARMA: ambos decaem de forma mista. - Fórmulas de Yule-Walker para AR(p):

$$\gamma(h) = \sum_{i=1}^p \phi_i \gamma(h-i) + \sigma^2 \delta_{h0}, \quad h \geq 0,$$

onde δ_{h0} é delta de Kronecker. - Para $h = 1, \dots, p$, sistema linear:

$$\begin{pmatrix} \gamma(0) & \gamma(1) & \cdots & \gamma(p-1) \\ \gamma(1) & \gamma(0) & \cdots & \gamma(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma(p-1) & \gamma(p-2) & \cdots & \gamma(0) \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_p \end{pmatrix} = \begin{pmatrix} \gamma(1) \\ \gamma(2) \\ \vdots \\ \gamma(p) \end{pmatrix}.$$

- Estimação via Yule-Walker: $\hat{\phi}$ = solução do sistema amostral de autocovariâncias.

2.1.5 Transformada de Fourier e densidade espectral

- A densidade espectral de uma série estacionária é definida por

$$f(\omega) = \frac{1}{2\pi} \sum_{h=-\infty}^{+\infty} \gamma(h) e^{-i\omega h}, \quad \omega \in [-\pi, \pi].$$

- Para ARMA(p,q):

$$f(\omega) = \frac{\sigma^2 |\theta(e^{-i\omega})|^2}{2\pi |\phi(e^{-i\omega})|^2},$$

onde $\phi(e^{-i\omega}) = 1 - \sum_{j=1}^p \phi_j e^{-i\omega j}$ e $\theta(e^{-i\omega}) = 1 + \sum_{j=1}^q \theta_j e^{-i\omega j}$. - Estimação espectral (periodograma, suavizado), estimador Whittle: aproximar log-verossimilhança no domínio de frequência:

$$\ell_W(\theta) \approx - \sum_{k=1}^N \left[\log f_\theta(\omega_k) + \frac{I(\omega_k)}{f_\theta(\omega_k)} \right],$$

onde $I(\omega_k)$ é o periodograma observado em frequências ω_k .

2.2 Modelos ARIMA/SARIMA

2.2.1 Modelo ARIMA(p,d,q)

- Definição: a série X_t satisfaz

$$\phi(B)(1-B)^d X_t = c + \theta(B)\varepsilon_t,$$

com ε_t ruído branco de variância σ^2 . - Se $c \neq 0$, inclui tendência de ordem d . - Em notação: ARIMA(p,d,q) com constantes. Se incluir termo determinístico μ , pode escrever:

$$\phi(B)(1-B)^d(X_t - \mu) = \theta(B)\varepsilon_t.$$

- Caso sazonal mensal com período $s = 12$, modelo SARIMA(p,d,q) \times (P,D,Q) $_s$:

$$\Phi(B^s)\phi(B)(1-B)^d(1-B^s)^D(X_t - \mu) = \Theta(B^s)\theta(B)\varepsilon_t,$$

com $\Phi(B^s) = 1 - \Phi_1 B^s - \dots - \Phi_P B^{Ps}$ e $\Theta(B^s) = 1 + \Theta_1 B^s + \dots + \Theta_Q B^{Qs}$.

2.2.2 Testes de raiz unitária

- **ADF**: Modelo:

$$\Delta X_t = \alpha + \beta t + \gamma X_{t-1} + \sum_{i=1}^k \delta_i \Delta X_{t-i} + u_t.$$

Hipótese nula $H_0 : \gamma = 0$ (raiz unitária); $H_1 : \gamma < 0$. Estatística t de $\hat{\gamma}$. - **KPSS**: Testa H_0 : série estacionária em nível (ou estacionária em tendência, dependendo de versão), contra H_1 : raiz unitária. Baseia-se em estatística de soma cumulativa dos resíduos de regressão contra constante/ tendência.

2.2.3 Identificação de ordens via ACF/PACF

- Observe ACF e PACF de $\Delta^d X_t$. - Para componente sazonal (lag múltiplos de 12), observar picos em lags 12, 24 etc., sugerindo ordens sazonais P, Q, D. - Exemplo: ACF de ΔX_t decai lentamente \rightarrow possível AR de alta ordem; PACF cortando em lag p . - Critérios automáticos (AIC/BIC) comparando vários (p,d,q) e sazonais.

2.2.4 Função de verossimilhança

- Para ARMA, assume $\varepsilon_t \sim \mathcal{N}(0, \sigma^2)$. A verossimilhança condicional (inicializando $\varepsilon_1, \dots, \varepsilon_{\max(p,q)} = 0$ ou usando períodos de “burn-in”):

$$L(\theta) = \prod_{t=t_0+1}^T \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\varepsilon_t(\theta)^2}{2\sigma^2}\right),$$

onde $\varepsilon_t(\theta)$ são resíduos calculados dado $\theta = (\phi_i, \theta_j, \sigma^2)$. - Verossimilhança exata via representação de estado (Kalman), que permite considerar distribuição exata dos primeiros p observações.

2.2.5 Estimadores e propriedades

- Estimação via máxima verossimilhança (ou profiled likelihood). Em ARIMA, parâmetros não lineares, requer otimização numérica. - Consistência e normalidade assintótica sob condições de estacionaridade/invertibilidade e erros brancos ou leves dependências.

2.2.6 Diagnóstico de resíduos

- Resíduos ajustados: $\hat{\varepsilon}_t = X_t - \hat{X}_t$. Verificar: - ACF dos resíduos: devem estar dentro de bandas de confiança para ruído branco. - Teste Ljung–Box:

$$Q(m) = T(T+2) \sum_{h=1}^m \frac{\hat{\rho}(h)^2}{T-h} \sim \chi_{m-(p+q)}^2.$$

- Teste ARCH nos resíduos:

$$\hat{\varepsilon}_t^2 = \alpha_0 + \sum_{i=1}^k \alpha_i \hat{\varepsilon}_{t-i}^2 + u_t,$$

verificar se coeficientes não nulos indicam heterocedasticidade remanescente. - QQ-plot: avaliar caudas; se residuais exibem caudas pesadas, pode-se modelar com distribuição t ou GED em GARCH.

2.2.7 Previsão ARIMA

- Representação infinita: para ARMA estável:

$$X_t = \mu + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}, \quad \psi_0 = 1, \quad \psi_j = \sum_{i=1}^p \phi_i \psi_{j-i} - \sum_{k=1}^q \theta_k \psi_{j-k},$$

com convenção $\psi_j = 0$ para índices negativos. - Previsão h passos à frente (no caso ARIMA, aplicando diferenciação inversa na previsão):

$$\hat{X}_{T+h|T} = \mu + \sum_{j=1}^{h-1} \psi_j \hat{\varepsilon}_{T+h-j}, \quad \text{Var}(\hat{X}_{T+h|T}) = \sigma^2 \sum_{j=0}^{h-1} \psi_j^2.$$

- Se sazonal, similar mas com coeficientes sazonais. - Em espaço de estado/Kalman Filter, a previsão e variância de previsão são obtidas diretamente pelas recursões de previsão do estado.

2.2.8 Representação em Espaço de Estado e Kalman Filter

- Representar modelo ARIMA em forma de espaço de estado:

$$\begin{cases} \mathbf{s}_t = \mathbf{F}\mathbf{s}_{t-1} + \mathbf{G}\varepsilon_t, \\ X_t = \mathbf{H}^\top \mathbf{s}_t + c_t, \end{cases}$$

onde \mathbf{s}_t é vetor de estado que contém defasagens e erros passados. - Exemplo ARIMA(1,1,1): defina $Z_t = \Delta X_t$, então

$$Z_t = \phi_1 Z_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1}.$$

Estado pode ser $\mathbf{s}_t = [Z_t, \varepsilon_t]^\top$. - Kalman filter/tracker fornece estimação do estado (quando há exógenas ou intervenção) e verossimilhança exata para parâmetros.

2.3 Modelos de Volatilidade

2.3.1 Heterocedasticidade condicional

Em finanças, retornos frequentemente exibem clusters de volatilidade. Modelos ARCH/GARCH capturam essa característica.

2.3.2 Modelo ARCH(q)

- Definição:

$$X_t = \mu + \varepsilon_t, \quad \varepsilon_t = \sigma_t z_t, \quad z_t \sim \text{i.i.d. } (0, 1),$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2,$$

com $\alpha_0 > 0$, $\alpha_i \geq 0$. - Persistência: somatório $\sum_{i=1}^q \alpha_i$. Se próximo de 1, volatilidade persiste.

2.3.3 Modelo GARCH(p,q)

- Definição (Bollerslev, 1986):

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2,$$

com $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$. - Persistência: $\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j$. Se < 1 , há estacionaridade fraca. - Verossimilhança condicional:

$$L(\theta) = \prod_{t=t_0+1}^T \frac{1}{\sqrt{2\pi\sigma_t^2(\theta)}} \exp\left(-\frac{\varepsilon_t^2}{2\sigma_t^2(\theta)}\right),$$

com $\theta = (\alpha_0, \alpha_1, \dots, \beta_1, \dots)$. Otimização numérica via Newton ou BFGS.

2.3.4 Extensões de GARCH

- **EGARCH** (Nelson, 1991): modela log-volatilidade para permitir coeficiente negativo sem violar positividade:

$$\ln(\sigma_t^2) = \omega + \sum_{i=1}^q \alpha_i \frac{\varepsilon_{t-i}}{\sigma_{t-i}} + \gamma_i \left(\frac{|\varepsilon_{t-i}|}{\sigma_{t-i}} - E\left|\frac{\varepsilon_{t-i}}{\sigma_{t-i}}\right| \right) + \sum_{j=1}^p \beta_j \ln(\sigma_{t-j}^2).$$

Aqui, termos assimétricos permitem efeito “leverage”.

- **GJR-GARCH** (Glosten-Jagannathan-Runkle, 1993): inclui efeito de choques negativos:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \gamma \varepsilon_{t-1}^2 I_{\{\varepsilon_{t-1} < 0\}} + \beta_1 \sigma_{t-1}^2.$$

- **GARCH-MIDAS**: separa componentes de curto e longo prazo. Exemplo:

$$\log(\sigma_t^2) = \omega + \underbrace{\sum_{j=1}^p \beta_j \log(\sigma_{t-j}^2)}_{\text{curto prazo}} + \underbrace{\theta \sum_{k=0}^K w_k(\vartheta) X_{t-k\Delta}}_{\text{longo prazo via variável macro } X}.$$

Onde $w_k(\vartheta)$ são pesos decrescentes (MIDAS).

- **Volatilidade estocástica (SV)**: modelo onde volatilidade segue processo estocástico latente, exigindo métodos Bayesianos ou filtros de partículas.
- **Distribuições não gaussianas nos erros**: usar t de Student, GED, Skewed t, ajustando verossimilhança com densidade apropriada:

$$f(\varepsilon_t | \sigma_t^2, \nu) \propto \left(1 + \frac{\varepsilon_t^2}{(\nu - 2)\sigma_t^2} \right)^{-\frac{\nu+1}{2}}.$$

2.3.5 Diagnóstico em modelos de volatilidade

- Resíduos padronizados: $\hat{z}_t = \varepsilon_t / \hat{\sigma}_t$. Devem ser ruído branco sem autocorrelação e sem heterocedasticidade remanescente: teste ACF de \hat{z}_t^2 e teste Ljung–Box sobre quadrados. - Persistência: se soma de coeficientes próxima de 1, volatilidade muito persistente; prever convergência lenta à variância incondicional:

$$\text{Var}_\infty = \frac{\alpha_0}{1 - \sum \alpha_i - \sum \beta_j}.$$

- Forecast de volatilidade: recursivo:

$$\hat{\sigma}_{T+h}^2 = \alpha_0 + (\alpha_1 + \beta_1)\hat{\sigma}_{T+h-1}^2, \quad h \geq 2.$$

Converge a Var_∞ se soma < 1 .

2.4 Redes Neurais Recorrentes e LSTM

2.4.1 Motivação para séries temporais

Modelos clássicos assumem linearidade e dependências de curto prazo; redes recorrentes (RNN) e LSTM capturam dependências não lineares e de longo prazo.

2.4.2 Equações internas de LSTM

Para cada passo de tempo t , dado entrada escalar/ vetor x_t , estado oculto h_{t-1} e célula c_{t-1} , definem-se:

$$\begin{aligned} i_t &= \sigma(W_i x_t + U_i h_{t-1} + b_i) \quad (\text{porta de entrada}), \\ f_t &= \sigma(W_f x_t + U_f h_{t-1} + b_f) \quad (\text{porta de esquecimento}), \\ o_t &= \sigma(W_o x_t + U_o h_{t-1} + b_o) \quad (\text{porta de saída}), \\ \tilde{c}_t &= \tanh(W_c x_t + U_c h_{t-1} + b_c) \quad (\text{candidato de célula}), \\ c_t &= f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \quad (\text{novo estado da célula}), \\ h_t &= o_t \odot \tanh(c_t) \quad (\text{novo estado oculto}). \end{aligned}$$

- $\sigma(\cdot)$ é função sigmoide; \odot produto elemento a elemento. - Matrizes W_* e U_* , vetores b_* são parâmetros treináveis. - A saída final \hat{y}_{t+1} pode ser obtida por camada densa:

$$\hat{y}_{t+1} = V^\top h_t + b_y.$$

- Função de custo usual: MSE,

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2.$$

- Retropropagação através do tempo (BPTT) computa gradientes via cadeia de derivadas das equações acima.

2.4.3 Preparação de dados para LSTM

- Série $Y_t = \log(P_t)$, normalização: z-score ou MinMax:

$$\tilde{Y}_t = \frac{Y_t - \bar{Y}_{\text{treino}}}{s_{\text{treino}}}.$$

- Janelas deslizantes de tamanho L : entrada $X^{(i)} = [\tilde{Y}_{t-L+1}, \dots, \tilde{Y}_t]$, saída $y^{(i)} = \tilde{Y}_{t+1}$.
- Divisão temporal: treino (ex.: jan2015–dez2021), validação (jan2022–dez2022), teste (jan2023–dez2024). - Formato para Keras: \mathbf{X} de dimensão $(n_{\text{amostras}}, L, 1)$.

2.4.4 Função de perda e otimização

- Usualmente MSE; otimização via gradient descent adaptativo (Adam). - Early stopping: monitorar perda em validação, interromper quando não melhorar após certa paciência. - Regularização: dropout em camadas, L2, etc.

2.4.5 Validação walk-forward

- Processo:

1. Treinar modelo até tempo T_0 .
2. Prever $T_0 + 1$.
3. Incluir observação real em série de treino e re-treinar (ou re-ajustar incrementalmente) até $T_0 + 1$.
4. Repetir até fim da amostra.

- Avaliar erros acumulados em janelas de interesse, especialmente em períodos de choque.

2.4.6 Explainability e interpretação

- Métodos como SHAP, LIME podem ser adaptados para redes temporais, mas cuidado com dependências de sequência. - Analisar importância de janelas passadas ou de variáveis exógenas em LSTM multivariado.

2.5 Modelos de Mudança de Regime

2.5.1 Modelos Markov-Switching

- Pressuposto: série muda de regimes latentes $S_t \in \{1, 2, \dots, K\}$ seguindo cadeia de Markov com matriz de transição $P = [p_{ij}]$. - Em regime i , modelo ARIMA/GARCH com parâmetros específicos $\theta^{(i)}$. Exemplo:

$$X_t = \mu_{S_t} + \sum_{j=1}^{p_{S_t}} \phi_{j,S_t} X_{t-j} + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \sigma_{S_t}^2).$$

- Estimação via algoritmo de expectation-maximization (EM) ou filtros de Hamilton. - Pode estender a GARCH de regime: volatilidade condicional depende de S_t .

2.5.2 Modelos de intervenção e dummies

- Incluir variáveis dummy D_t para choques exógenos:

$$\phi(B)(1 - B)X_t = \delta D_t + \theta(B)\varepsilon_t.$$

- $D_t = 1$ em períodos de choque (por exemplo, meses de pico de volatilidade); estima efeito médio de choque.

2.6 Avaliação de Modelos e Critérios de Seleção

2.6.1 Critérios de informação

- AIC:

$$\text{AIC} = -2 \ln \hat{L} + 2k,$$

onde k é número de parâmetros estimados. - BIC:

$$\text{BIC} = -2 \ln \hat{L} + k \ln(T).$$

- Para GARCH, semelhante, mas atenção ao número de observações efetivas pela inicialização condicional.

2.6.2 Testes de comparação de previsões

- Teste Diebold-Mariano para comparar erros de duas previsões:

$$d_t = g(e_{1,t}) - g(e_{2,t}),$$

onde $g(\cdot)$ é função de perda (ex.: quadrática) e $e_{i,t} = y_t - \hat{y}_{i,t}$. Estatística:

$$\text{DM} = \frac{\bar{d}}{\sqrt{\widehat{\text{Var}}(\bar{d})/T}} \sim N(0, 1).$$

- Testes de hipóteses sobre superioridade de um modelo.

2.6.3 Backtesting de VaR

- Se usar GARCH para VaR, checar frequência de violações:

$$\text{VaR}_{t+1}(\alpha) = \mu_t + z_\alpha \sigma_{t+1},$$

onde z_α quantil da distribuição assumida. - Backtest Kupiec (proportion of failures) e Christoffersen (independência das violações).

2.7 Referências Teóricas Ampliadas

Além das referências já listadas, aprofundar em:

- Hamilton [1994]: fundamentos de séries temporais.
- Box e Jenkins [1970]: metodologia Box-Jenkins.
- Tsay [2010]: detalhamento de econometria de séries financeiras.
- Textos sobre Kalman Filter e espaço de estado, p.ex. Durbin e Koopman [2012].
- Artigos sobre Whittle estimation e espectral.
- Artigos sobre GARCH-MIDAS e Modelos de volatilidade de longo prazo.
- Materiais sobre aprendizado de máquina para séries temporais e explainability.

Capítulo 3

Dados e Pré-processamento

Nesta seção descreve-se de forma sucinta, pois o foco deste template é a parte teórica com equações. Em Apêndice incluir-se-ão scripts Python.

3.1 Fontes de Dados

Listar as fontes, exemplos:

- U.S. Energy Information Administration: série “Europe Brent Spot Price FOB” (código DCOILBRENTU).
- Trading Economics: exportação de CSV de preços diários.
- MacroTrends: tabelas históricas mensais.
- St. Louis Fed (ALFRED) para vintages.
- Outras fontes: relatórios BP, OPEC, IEA para variáveis exógenas.

3.2 Período e Frequência

Escolhido jan/2015–dez/2024:

- Série mensal: média dos preços diários em cada mês, necessária para modelar sazonalidade 12.
- Série diária: preços de fechamento diários para análise de volatilidade e choques.

3.3 Pré-processamento

Exemplificar em texto como converter data, lidar com missing, criar log-preço e retornos:

$$Y_t = \log(P_t), \quad r_t = Y_t - Y_{t-1}.$$

Descrever testes ADF/KPSS em nível e primeira diferença, ACF/PACF, estatísticas de retornos (média, σ , skewness, kurtosis, Jarque-Bera), identificação de outliers.

Capítulo 4

Metodologia de Modelagem

4.1 Modelos ARIMA/SARIMA

Repetir brevemente:

$$\phi(B)(1-B)^d X_t = c + \theta(B)\varepsilon_t, \quad \Phi(B^s)\phi(B)(1-B)^d(1-B^s)^D X_t = c + \Theta(B^s)\theta(B)\varepsilon_t.$$

Incluir equações de Yule-Walker, função de verossimilhança condicional e exata, recursões de forecast:

$$\hat{X}_{T+h|T} = \mu + \sum_{j=1}^{h-1} \psi_j \hat{\varepsilon}_{T+h-j}, \quad \text{Var}(\hat{X}_{T+h|T}) = \sigma^2 \sum_{j=0}^{h-1} \psi_j^2.$$

Mostrar como derivar ψ_j recursivamente:

$$\psi_j = \phi_1 \psi_{j-1} + \cdots + \phi_p \psi_{j-p} - \theta_1 \psi_{j-1} - \cdots - \theta_q \psi_{j-q}, \quad \psi_0 = 1.$$

Detalhar critérios AIC/BIC:

$$\text{AIC} = -2 \ln \hat{L} + 2k, \quad \text{BIC} = -2 \ln \hat{L} + k \ln(T).$$

Descrever representação em espaço de estado:

$$\mathbf{s}_t = \mathbf{F}\mathbf{s}_{t-1} + \mathbf{G}\varepsilon_t, \quad X_t = \mathbf{H}^\top \mathbf{s}_t + c,$$

com Kalman Filter recursões:

$$\begin{aligned} \hat{\mathbf{s}}_{t|t-1} &= \mathbf{F}\hat{\mathbf{s}}_{t-1|t-1}, \quad P_{t|t-1} = \mathbf{F}P_{t-1|t-1}\mathbf{F}^\top + \mathbf{G}\mathbf{G}^\top, \\ \nu_t &= X_t - \mathbf{H}^\top \hat{\mathbf{s}}_{t|t-1}, \quad S_t = \mathbf{H}^\top P_{t|t-1}\mathbf{H} + R, \\ K_t &= P_{t|t-1}\mathbf{H}S_t^{-1}, \quad \hat{\mathbf{s}}_{t|t} = \hat{\mathbf{s}}_{t|t-1} + K_t\nu_t, \quad P_{t|t} = P_{t|t-1} - K_t\mathbf{H}^\top P_{t|t-1}. \end{aligned}$$

Explicar interpretação e uso em previsão e estimação exata de verossimilhança.

4.2 Modelos de Volatilidade

Descrever modelos:

$$X_t = \mu + \varepsilon_t, \quad \varepsilon_t = \sigma_t z_t, \quad z_t \sim D(0, 1).$$

GARCH(1,1):

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \quad \alpha_0 > 0, \alpha_1, \beta_1 \geq 0, \alpha_1 + \beta_1 < 1.$$

Função de verossimilhança condicional:

$$\ell(\theta) = -\frac{1}{2} \sum_{t=t_0+1}^T \left[\ln(2\pi) + \ln(\sigma_t^2(\theta)) + \frac{\varepsilon_t^2}{\sigma_t^2(\theta)} \right].$$

Extensões: - **EGARCH:**

$$\ln(\sigma_t^2) = \omega + \sum_{i=1}^q \alpha_i \frac{\varepsilon_{t-i}}{\sigma_{t-i}} + \sum_{i=1}^q \gamma_i \left(\frac{|\varepsilon_{t-i}|}{\sigma_{t-i}} - E \left| \frac{\varepsilon_{t-i}}{\sigma_{t-i}} \right| \right) + \sum_{j=1}^p \beta_j \ln(\sigma_{t-j}^2).$$

- **GJR-GARCH:**

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \gamma \varepsilon_{t-1}^2 I_{\{\varepsilon_{t-1} < 0\}} + \beta_1 \sigma_{t-1}^2.$$

- **GARCH-MIDAS:** separar efeitos de curto e longo prazo:

$$\ln(\sigma_t^2) = \underbrace{\omega + \beta \ln(\sigma_{t-1}^2)}_{\text{curto prazo}} + \underbrace{\theta \sum_{k=0}^K w_k(\vartheta) X_{t-k\Delta}}_{\text{longo prazo}},$$

com $w_k(\vartheta)$ pesos decrescentes. - **Distribuições não gaussianas:** erro $\varepsilon_t/\sigma_t \sim t_\nu$ ou GED. Ajuste na verossimilhança:

$$\ln f(\varepsilon_t | \sigma_t, \nu) = \text{const} - \frac{\nu + 1}{2} \ln \left(1 + \frac{\varepsilon_t^2}{(\nu - 2)\sigma_t^2} \right).$$

Descrever diagnóstico: - resíduos padronizados \hat{z}_t e análise de ACF de \hat{z}_t e \hat{z}_t^2 . - backtesting de VaR:

$$\text{VaR}_{t+1}(\alpha) = \hat{\mu}_t + z_\alpha \hat{\sigma}_{t+1},$$

testar proporção de violações conforme Kupiec:

$$LR_{POF} = -2 \ln \left[(1 - \hat{p})^{T-N} \hat{p}^N / (1 - \alpha)^{T-N} \alpha^N \right] \sim \chi_1^2,$$

onde $\hat{p} = N/T$ é proporção de violações.

4.3 Modelos de Mudança de Regime

Descrever Markov-Switching:

$$X_t = \mu_{S_t} + \sum_{j=1}^{p_{S_t}} \phi_{j,S_t} X_{t-j} + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \sigma_{S_t}^2),$$

cadeia de Markov latente $P(S_t = j | S_{t-1} = i) = p_{ij}$. Estimação EM: - E-step: obter probabilidades a posteriori $P(S_t = i | \text{dados}, \theta^{\text{old}})$ via filtro e suavizador de Hamilton. - M-step: atualizar θ (parâmetros AR e variâncias) maximizando verossimilhança ponderada pelas probabilidades de regime. Equações do filtro recursivo e suavizador podem ser descritas detalhadamente.

4.4 Redes Neurais Recorrentes e LSTM

Revisitar equações internas de LSTM (já apresentadas), função de custo, retropropagação, otimização. Incluir equações de gradiente:

$$\frac{\partial \text{MSE}}{\partial W_i} = \frac{2}{N} \sum_t (y_t - \hat{y}_t) \frac{\partial \hat{y}_t}{\partial W_i},$$

e cadeia até parâmetros internos W_i, U_i, b_i . Descrever algoritmos de otimização (Adam: momentos de primeira e segunda ordem):

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t, \quad v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2,$$

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t}, \quad \hat{v}_t = \frac{v_t}{1 - \beta_2^t}, \quad \theta_{t+1} = \theta_t - \eta \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon}.$$

Explicar early stopping, regularização e validação walk-forward.

4.5 Critérios de Comparação e Testes Estatísticos

- Comparar modelos via AIC/BIC (ARIMA, GARCH), e para LSTM usar validação cruzada temporal (walk-forward) e métricas RMSE, MAE, MAPE:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{t=1}^N (\hat{y}_t - y_t)^2}, \quad \text{MAE} = \frac{1}{N} \sum_{t=1}^N |\hat{y}_t - y_t|, \quad \text{MAPE} = \frac{100}{N} \sum_{t=1}^N \left| \frac{\hat{y}_t - y_t}{y_t} \right|.$$

- Teste Diebold-Mariano para comparar previsões de dois modelos:

$$d_t = L(e_{1,t}) - L(e_{2,t}), \quad \text{estatística DM} = \frac{\bar{d}}{\sqrt{\widehat{\text{Var}}(d)/T}} \sim N(0, 1).$$

- Backtesting de VaR para GARCH, como descrito acima.

Capítulo 5

Resultados e Estrutura de Apresentação

Nesta seção insira tabelas de resultados reais após execução dos scripts Python. Como este template foca nas equações, aqui descreve-se a estrutura de apresentação:

5.1 Teste de Estacionaridade

Tabela 5.1: Resultados de testes de raiz unitária

Série	Teste	Estatística / p-valor
LogPrice	ADF	-1.80 / 0.35
DiffLog	ADF	-3.50 / 0.01
DiffLog	KPSS	0.20 / >0.10

5.2 Ajuste ARIMA/SARIMA

5.2.1 Modelos Testados e Seleção

Tabela 5.2: Comparação de ordens ARIMA/SARIMA via AIC/BIC

Modelo	AIC	BIC
ARIMA(1,1,1)	150.2	158.7
SARIMA(1,1,1)(1,1,1)[12]	123.5	134.0
SARIMA(2,1,1)(1,1,1)[12]	124.0	136.0
SARIMA(1,1,2)(1,1,1)[12]	123.8	135.8

5.2.2 Parâmetros Estimados

Exemplo:

$$\begin{aligned}\hat{\phi}_1 &= 0.52 \text{ (p} < 0.01), & \hat{\theta}_1 &= -0.45 \text{ (p} < 0.01), \\ \hat{\Phi}_1 &= 0.31 \text{ (p} < 0.05), & \hat{\Theta}_1 &= -0.25 \text{ (p} < 0.05).\end{aligned}$$

Listar variância do erro $\hat{\sigma}^2$, estatísticas t e p-valores.

5.2.3 Diagnóstico de Resíduos

Tabelas:

- Ljung–Box: lag 12, lag 24, estatística Q e p-valores.
- Teste ARCH nos resíduos: estatística e p-valor.
- QQ-plot (mencionar qualitativamente se há caudas pesadas).

5.2.4 Forecast ARIMA/SARIMA

Apresentar em texto: previsão de h passos (ex.: 12 meses adiante), intervalos de confiança, fórmulas usadas. Incluir equações de forecast já mostradas anteriormente.

5.3 Ajuste de Modelos de Volatilidade

5.3.1 Parâmetros GARCH

Exemplo de resultados:

$$\hat{\alpha}_0 = 0.0006, \quad \hat{\alpha}_1 = 0.12, \quad \hat{\beta}_1 = 0.83, \quad \alpha_1 + \beta_1 = 0.95.$$

Descrever se usou distribuição t e qual número de graus de liberdade ν .

5.3.2 Diagnóstico

Listar:

- Teste ACF de resíduos padronizados e resíduos padronizados ao quadrado.
- Teste Ljung–Box nos resíduos ao quadrado.
- Backtesting de VaR: proporção de violações, estatística Kupiec, etc.

5.3.3 Forecast de Volatilidade

Descrever recursão e convergência à variância incondicional. Mencionar aplicação em VaR.

5.4 Resultados de LSTM

5.4.1 Treinamento e Validação

- Curva de perda em treino e validação (descrever comportamento, por ex. convergência em época X). - Equações de otimização mostradas previamente.

5.4.2 Desempenho Fora da Amostra

Tabela:

Tabela 5.3: Comparação de métricas (jan/2023–dez/2024)

Modelo	RMSE (log)	MAE (log)	MAPE (%)
SARIMA(1,1,1)(1,1,1)[12]	0.09	0.07	4.0
ARIMA(1,1,1)	0.11	0.08	5.2
LSTM	0.08	0.06	3.5

Descrever como os valores foram obtidos via equações de métricas.

5.4.3 Validação Walk-Forward

Descrever recursão de re-treinamento e cálculo de erro acumulado:

$$\text{erro acumulado} = \sqrt{\frac{1}{T'} \sum_{t=T_0+1}^T (\hat{y}_{t|t-1} - y_t)^2}.$$

Comparar para ARIMA e LSTM.

5.5 Modelos de Regime e Intervenção

Apresentar resultados de modelo Markov-Switching (se aplicável): probabilidades de regime, parâmetros em cada regime, testes de mudança de regime. Incluir equações de verossimilhança EM (já descritas em Revisão).

Capítulo 6

Discussão Geral

Nesta seção analisa-se comparativamente:

- Precisão de previsão (métricas) de ARIMA vs. LSTM vs. híbridos.
- Robustez em choques: performance em janelas de crise.
- Comportamento de volatilidade: persistência e aplicabilidade a hedge e VaR.
- Interpretabilidade: ARIMA/GARCH são mais interpretáveis; LSTM “caixa preta”.
- Custo computacional e necessidade de pipeline automatizado.
- Benefícios de incluir variáveis exógenas e ensembles.

Discussão qualitativa referenciando equações: por exemplo, se persistência de volatilidade alta ($\alpha_1 + \beta_1 \approx 1$), previsão de volatilidade tende a permanecer elevada, afetando decisões de hedge.

Capítulo 7

Conclusão e Trabalhos Futuros

Resumo das principais conclusões, remeter a equações centrais explicadas anteriormente (e.g., importância de verificar raízes de $\phi(B)$ para estacionaridade, importância de diagnóstico de resíduos para heterocedasticidade, utilidade de LSTM mas necessidade de validação).

Trabalhos Futuros

- Incluir variáveis exógenas: estoques, indicadores macro (PIB, inflação, taxa de câmbio), sentimento de notícias, processados via embeddings em LSTM multivariado.
- Modelos ARIMAX e GARCH-X: equações:

$$\phi(B)(1 - B)^d X_t = c + \sum_{k=1}^K \beta_k Z_{k,t} + \theta(B)\varepsilon_t,$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + \sum_{k=1}^K \gamma_k W_{k,t-1}.$$

- Modelos de regime (Markov-Switching ARIMA/GARCH): equações já descritas, estimar via EM.
- GARCH-MIDAS: incorporar fatores de longo prazo.
- Volatilidade estocástica: equações de modelo SV e métodos de estimação via MCMC ou filtros de partículas.
- Ensembles: combinar previsões:

$$\hat{y}_t = \sum_{m=1}^M w_m \hat{y}_t^{(m)}, \quad \sum w_m = 1,$$

ou stacking com meta-modelo.

- Validação robusta: backtesting de VaR, teste Diebold-Mariano periódico, monitorar performance em tempo real.

- Explainability em LSTM: SHAP adaptado para séries, análise de sensibilidade de entrada.
- Pipeline automatizado: scripts para coleta, pré-processamento, modelagem, geração automática de relatório LaTeX periódico.

Referências

Referências Bibliográficas

- Box, G. E. P.; Jenkins, G. M.; Reinsel, G. C. *Time Series Analysis: Forecasting and Control*. Holden-Day, 1970.
- Tsay, R. S. *Analysis of Financial Time Series*. Wiley, 2010.
- Hamilton, J. D. *Time Series Analysis*. Princeton University Press, 1994.
- Durbin, J.; Koopman, S. J. *Time Series Analysis by State Space Methods*. Oxford University Press, 2012.
- Engle, R. F. “Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation.” *Econometrica*, 1982.
- Bollerslev, T. “Generalized Autoregressive Conditional Heteroskedasticity.” *Journal of Econometrics*, 1986.
- Nelson, D. B. “Conditional Heteroskedasticity in Asset Returns: A New Approach.” *Econometrica*, 1991.
- Glosten, L. R.; Jagannathan, R.; Runkle, D. “On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks.” *Journal of Finance*, 1993.
- Engle, R.; Ghysels, E.; Sohn, B. “Stock Market Volatility and Macroeconomic Fundamentals.” *Review of Economics and Statistics*, 2013.
- Zhao, Y.; Hu, B.; Wang, S. “Prediction of Brent crude oil price based on LSTM model under the background of low-carbon transition.” arXiv:2409.12376, 2024.
- Pasula, P. “Real World Time Series Benchmark Datasets with Distribution Shifts: Global Crude Oil Price and Volatility.” arXiv:2308.10846, 2023.
- Albulescu, C. “Coronavirus and Oil Price Crash.” arXiv:2003.06184, 2020.
- Dai, P.-F.; Xiong, X.; Zhou, W.-X. “The Role of Global Economic Policy Uncertainty in Predicting Crude Oil Futures Volatility: Evidence from a Two-Factor GARCH-MIDAS Model.” arXiv:2007.12838, 2020.
- Kang, S. H.; Yoon, S.-M. “Hybrid models for energy price forecasting.” *Energy Economics*, 2015.
- Zhang, X.; et al. “Combining statistical and machine learning methods for commodity price forecasting.” *Journal of Forecasting*, 2018.

- Estudo comparativo de ARIMA, SVR e LSTM na previsão de preços energéticos. *Energy Economics*, 2019.
- U.S. Energy Information Administration. “Europe Brent Spot Price FOB (DCOILBRENTEU).” Disponível em: <https://www.eia.gov/dnav/pet/hist/rbrtem.htm>, acessado em junho de 2025.
- Trading Economics. “Brent Crude Oil Price.” Disponível em: <https://tradingeconomics.com/commodity/brent-crude-oil>, acessado em junho de 2025.
- MacroTrends LLC. “Brent Crude Oil Prices (histórico mensal 2015–2024).” Disponível em site MacroTrends, acessado em junho de 2025.
- BP. “BP Statistical Review of World Energy 2024.” Disponível online, acesso em 2025.
- OPEC. “Monthly Oil Market Report,” vários volumes, 2015–2024.
- International Energy Agency. “Oil Market Report,” 2015–2024.
- Hamilton, J. D. “Specification tests and inference in Markov-switching models.” *Journal of Econometrics*, 1996.
- Clements, M. P.; Hendry, D. F. “Forecasting Economic and Financial Time Series.” Cambridge University Press, 2005.
- Ng, E. “Exploring GARCH-MIDAS models for commodity volatility forecasting.” *International Journal of Forecasting*, 2011.
- Durbin, J.; Koopman, S. J. *Time Series Analysis by State Space Methods*. Oxford University Press, 2012.
- Wikipedia. “Box–Jenkins method.” Última revisão em 2025. Disponível em: https://en.wikipedia.org/wiki/Box%E2%80%93Jenkins_method.

Apêndice A

Código em Python

A.1 Coleta e Pré-processamento

Aqui incluir em apêndice o script completo em Python para:

- Obter dados da API EIA ou ler CSV.
- Converter para série mensal e diária.
- Criar log-preço Y_t e retornos r_t .
- Testes de estacionaridade ADF/KPSS.
- Calcular ACF/PACF.

A.2 Ajuste ARIMA/SARIMA

Script para estimar modelos, extrair parâmetros e diagnósticos via statsmodels.

A.3 Ajuste GARCH e extensões

Script para ajustar GARCH(1,1), EGARCH, GJR-GARCH, GARCH-MIDAS via pacote ‘arch’.

A.4 Implementação LSTM

Script para preparar janelas, treinar LSTM em Keras, avaliar métricas e gerar CSV com previsões.

A.5 Testes Estatísticos Detalhados

Gerar tabelas ADF/KPSS, Ljung–Box, ARCH, backtesting de VaR.

Apêndice B

Derivações Matemáticas Complementares

B.1 Equações de Yule-Walker

Derive sistema de equações para $AR(p)$, relacionando autocovariâncias e coeficientes.

B.2 Representação em Espaço de Estado

Detalhar como construir matrizes F, G, H para ARIMA e SARIMA.

B.3 Kalman Filter

Derivar recursões de predição e atualização de estado.

B.4 Função de Verossimilhança

Mostrar derivação da verossimilhança condicional e exata para ARMA e GARCH.

B.5 Whittle Estimation

Descrever derivação de log-likelihood aproximada no domínio espectral.