CONTABILIDADE EM TEXTO
20 anos
PPGCONT-UFRGS
Programa de Pós-Graduação em Controladoria e Contabilidade
Universidade Federal do Rio Grande do Sul

APLICAÇÃO DE CADEIAS DE MARKOV MULTIVARIADA E REGRESSÃO LINEAR MÚLTIPLA PARA INVESTIMENTO NO IBOVESPA COM O DÓLAR COMO INFLUENCIADOR

APPLICATION OF MULTIVARIATED MARKOV CHAINS AND MULTIPLE LINEAR REGRESSION FOR INVESTMENT IN IBOVESPA WITH THE DOLLAR AS INFLUENCER

RAFAEL SCARIOTE
Engenheiro de Produção. Universidade de Caxias do Sul (UCS).
E-mail: rscariote@ucs.br

LEANDRO LUÍS CORSO
Doutor e Mestre em Engenharia com foco em otimização pela Universidade Federal do Rio Grande do Sul (UFRGS). Professor na Universidade de Caxias do Sul (UCS).
E-mail: llcorso@ucs.br

Resumo: Em um mercado financeiro, que se mostra cada vez mais dinâmico e complexo, é de suma importância utilizar ferramentas que auxiliem os investidores a diminuir as perdas e maximizar os ganhos. Este artigo tem como objetivo avaliar o comportamento dos índices Ibovespa e Dólar nos anos 2019 e 2020, por meio de uma simulação de investimento no Ibovespa utilizando o Dólar como influenciador. Para isto, foram utilizados dois modelos matemáticos: Cadeias de Markov Multivariadas e Regressão Linear Múltipla. Avaliaram-se os dados dos índices, com a criação de matrizes de transição percentual entre os estados e criação da equação de regressão. Por meio de uma simulação de investimento inicial de R$ 100,00, foi possível verificar quais seriam os ganhos ou perdas utilizando os métodos. As simulações foram realizadas no segundo semestre de cada um dos anos. Para o modelo de Cadeias de Markov Multivariadas, observou-se a possibilidade de ganho de 8,1% no ano de 2019 e de 8,9% em 2020. Já para o modelo de Regressão Linear Múltipla, verificou-se uma possível perda de 8,6% no ano de 2019 e um ganho de 23,1% no ano de 2020.

Palavras-chave: Cadeias de Markov Multivariada. Regressão Linear Múltipla. Mercado Financeiro. Ibovespa. Dólar.

Abstract: *In the financial market, which is increasingly dynamic and complex, it is extremely important to use tools that help investors reduce losses and maximize gains. This article evaluates the behavior of the Ibovespa and Dollar indices in the years 2019 and 2020, through a simulation of investment in the Ibovespa using the Dollar as an influencer. For this, two mathematical models were used: Multivariate Markov Chains and Multiple Linear Regression. The index data were evaluated, with the creation of percentage transition matrices between the states and the creation of the regression equation. Through a simulation of an initial investment of R$ 100.00, it was possible to verify what the gains or losses would be using. The simulations were made in the second half of each year. For the Multivariate Markov Chains model, there was a gain possibility of 8.1% in 2019 and 8.9% in 2020. As for the*

Multiple Linear Regression model, there was a possible loss of 8.6% in the year 2019 and a gain of 23.1% in the year 2020.

***Keywords**: Multivariate Markov Chains. Multiple Linear Regression. Financial market. Ibovespa. Dollar*

1 INTRODUÇÃO

Atualmente, segundo Neto (2018) e Carrete (2019), o mercado de capitais é considerado um setor fundamental para a expansão dos países a nível mundial. Este mercado é responsável por atender as necessidades dos agentes econômicos por meio de recursos disponibilizados através de financiamentos a médio e longo prazo, para suprir capital de giro e até mesmo capital fixo. O mercado financeiro brasileiro é caracterizado por altos investimentos em tecnologia e é considerado, dentre os países emergentes e até mesmo entre países mais desenvolvidos, como referência internacional por conta disto (BRITO, 2020).

Para Virgilito (2017), os investidores deparam-se constantemente com a insegurança da volatilidade de suas aplicações financeiras, questões de como e em qual tipo de ativo investir para minimizar os riscos e maximizar os ganhos são sempre um desafio constante. Desta forma, a Pesquisa Operacional (PO) é muito empregada nos processos de tomada de decisões diante de problemas nos ambientes de negócios (RODRIGUES, 2017).

Segundo Belfiore e Fávero (2013), a PO consiste na utilização de modelos matemáticos, estatísticos e computacionais. Tais modelos permitem a possibilidade de testes e de tomadas de decisões mais assertivas antes de serem implementadas (ANDRADE, 2015). Dentro desta área, podem ser abordados os processos estocásticos, que descrevem o comportamento probabilístico de um sistema ao longo do tempo. Este processo pode ser englobado dentro do modelo de Cadeias de Markov, desde que obedeça ao critério de que a ocorrência de um estado futuro dependa somente do estado atual (HILLIER; LIEBERMAN, 2013).

Os métodos de Cadeias de Markov (CM) de primeira ordem e Cadeias de Markov Multivariadas (CMM) são utilizados como tema de estudo de trabalhos atuais. Nas aplicações, Cechin e Corso (2019) utilizaram CMM para avaliar a transição de estados entre o índice Ibovespa e *Dow Jones*. Com este estudo, foi possível constatar que a aplicação dos métodos pode ajudar na tomada de decisões para investidores. Já Bolson *et al.* (2019) aplicaram CM para prever a variação do Dólar americano. Como resultado, o método mostrou as probabilidades de variações futuras dado um estado presente.

Também na área financeira, com foco no mercado imobiliário, Araujo Filho, Oliveira e Bastos (2019) utilizaram Regressão Linear Múltipla (RLM) para relacionar variáveis macroeconômicas com as variáveis específicas do mercado imobiliário. Por meio deste estudo, chegaram à conclusão de que a porcentagem de desemprego influencia mais o mercado imobiliário do que a média salarial brasileira.

O objetivo principal deste trabalho é analisar como se comportam as variações dos dados relativos ao Ibovespa e ao Dólar, além de avaliar como se comportam os métodos de CMM e RLM para investimento no Ibovespa, tendo o Dólar como um fator de entrada. O fato de se utilizar CMM tem como premissa avaliar o comportamento do método, que trata dados estocásticos e comporta diferentes parâmetros de entrada.

2 REFERENCIAL TEÓRICO

2.1 MERCADOS FINANCEIROS

Segundo Neto (2018), a necessidade de conhecimento do sistema financeiro vem crescendo ao longo do tempo, visto que exerce papel muito importante na economia no segmento empresarial do país e pela complexidade que as suas operações apresentam. Compõem o sistema financeiro, um conjunto de instituições financeiras públicas e privadas, onde o órgão normativo máximo é o Conselho Monetário Nacional. O sistema financeiro nacional viabiliza a relação entre os agentes com necessidade de recursos e os agentes capazes de financiar o crescimento da economia.

O mercado financeiro pode ser subdividido em mercado monetário, mercado de crédito, mercado cambial e mercado de capitais. Tais divisões podem ser feitas de outra forma, visto que os agentes do mercado financeiro podem utilizar recursos de diferentes segmentos (CARRETE, 2019).

Segundo Kerr (2011), no mercado cambial, são realizadas transações cambiais, ou até mesmo compra e venda de moeda estrangeira, em que, no âmbito internacional, a moeda predominante é o Dólar. Sua principal função está atrelada a transferência de recursos de um país para outro. Carrete (2019) afirma que os mercados de câmbio, em momentos de crises ou de transições, são os mais sensíveis, pois são eles que traduzem a real situação do país sobre a economia e o valor da moeda local.

Para Neto (2018) e Carrete (2019), o mercado mais relevante no processo de desenvolvimento econômico é o mercado de capitais. Nesse mercado, os principais prestadores de serviços para as empresas com a finalidade de auxiliar a disponibilização dos títulos aos investidores são os bancos, as corretoras de valores e as distribuidoras de valores mobiliários. Os títulos de valores mobiliários têm a característica de mudar a denominação de um investidor para outro, sendo então, um ativo móvel. Desta forma, o mercado de capitais foi estruturado de tal maneira para ser a organização das negociações (PEREIRA, 2013).

Izidoro (2015) explica que o mercado de capitais é um sistema de distribuição de valores com a finalidade de oferecer liquidez aos títulos de emissão de empresas, além de viabilizar o processo de capitalização. Para Pereira (2013), diferente do mercado de crédito, onde a empresa precisa buscar crédito a custos elevados, no mercado de capitais, as empresas conseguem captar recursos a preços menores, visto que o acionista se torna sócio da empresa, e não um credor. Desta forma, este mercado atua como um impulsionador de capitais para investimentos.

2.1.1 AÇÕES E BOLSAS DE VALORES

As ações são os títulos que representam a menor fração no capital social de uma empresa, representando um investimento permanente na empresa. Com isto, o acionista é um proprietário da companhia, participando dos resultados obtidos pela empresa, de acordo com a quantidade de ações adquiridas (KERR, 2011).

Segundo Izidoro (2015), as empresas que possuem ações disponíveis para serem comercializadas nos mercados de ações são denominadas sociedades anônimas e podem ter capital aberto ou fechado. As empresas de capital aberto obrigatoriamente estão registradas na Comissão de Valores Mobiliários, portanto dão ao mercado informações sociais, econômicas e financeiras, visto que suas ações são comercializadas na bolsa de valores. Já as empresas de capital fechado, geralmente são empresas menores e de cunho familiar, contendo um número restrito de sócios com possibilidade de adquirir ações (KERR, 2011; IZIDORO, 2015).

Para Kerr (2011), bolsas de valores são onde as ações são comercializadas, ocorrendo na forma de um leilão organizado. Elas desempenham papel fundamental na sociedade, pois beneficiam os agentes econômicos e a sociedade levantando capital para aplicação em projetos de investimentos, além de permitir que pequenos investidores se tornem acionistas de grandes empresas.

Segundo Izidoro (2015), a única bolsa de valores no Brasil é a BM&FBovespa, fusão da bolsa de valores de São Paulo (Bovespa) e da Bolsa de Mercadorias e Futuros (BM&F) em 2007, após a Bovespa se transformar em sociedade anônima.

O índice da BM&FBovespa é o Ibovespa, que é o valor atual da moeda corrente de uma carteira teórica de ações. Basicamente, este índice serve como indicador do comportamento do mercado e é calculado considerando os negócios efetuados no mercado de ações de suas respectivas carteiras (PEREIRA, 2013).

O mercado de ações classifica-se em primário e secundário. O mercado é primário quando as negociações das ações são realizadas pela primeira vez, e o valor arrecadado é destinado para a empresa emitente. Por outro lado, o mercado é secundário quando estas negociações são realizadas por terceiros, onde os valores simplesmente trocam de mãos, sem financiar investimentos para a respectiva empresa (KERR, 2011; IZIDORO, 2015).

Franco *et al.* (2021) utilizaram o método de Cadeias de Markov Multivariada para avaliar a probabilidade de variação do preço de ações da Petrobrás com o dólar e com o petróleo WTI (*West Texas Intermediate*). Por meio desta aplicação do modelo, foi possível definir um intervalo onde existe maior probabilidade de se ter um estado estável e também seus respectivos tempos de recorrência.

Também na área de ações, Gururaj, Shriya e Ashwini (2019) utilizaram o modelo de regressão linear aplicado na variação dos preços das ações da Coca-Cola, para avaliar como se comportava. Como conclusão, o modelo se mostrou eficiente e possível de ser utilizado para previsões futuras.

2.2 CADEIAS DE MARKOV

As Cadeias de Markov (CM) possuem a propriedade de que as probabilidades de eventos futuros dependem somente do estado atual e independem dos eventos passados. Os modelos probabilísticos que evoluem ao longo do tempo são denominados processos estocásticos (HILLIER; LIEBERMAN, 2013).

Para Hillier e Lieberman (2013), um processo estocástico é um conjunto de variáveis aleatórias X_t, onde o índice *t* percorre certo conjunto *T*. Normalmente, *T* é um conjunto de inteiros não negativos e X_t representa uma característica mensurável no instante *t*.

Para Taha (2008), um processo estocástico é um processo de Markov se o resultado de um estado futuro depender única e exclusivamente do estado atual. Desta forma, em um processo markoviano com n resultados exclusivos em determinado tempo *t* = 0, 1, 2..., tem-se a propriedade descrita na Equação 1.

$$p_{ij} = P\{X_t = j | X_{t-1} = i\}, (ij) = 1,2, \dots, n, t = 0,1,2, \dots, T \quad (1)$$

Esta definição é denominada como a probabilidade de transição em uma etapa de passar do estado i em $t-1$ ao estado j em t. Com isso, a definição se dá conforme as Equações 2 e 3.

$$p_{ij} = P\{X_{t+1} = j \, | X_t = i\}, \quad (2)$$

$$p_{ij}^{(n)} = P\{X_{t+n} = j \, | X_t = i\}, \quad (3)$$

Desta forma, a probabilidade de transição em n-etapas $p_{ij}^{(n)}$ é a condicional de que o sistema se encontra no estado *j* após n-etapas, visto que ele inicia no estado *i* a qualquer instante *t* (HILLIER; LIEBERMAN, 2013).

Cechin e Corso (2019) utilizaram a matriz conforme Equação 4 para definir as probabilidades de transição dos estados de uma CM.

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1j} \\ p_{21} & p_{22} & \cdots & p_{2j} \\ \cdots & \cdots & \cdots & \cdots \\ p_{i1} & p_{i2} & \cdots & p_{ij} \end{bmatrix} \quad (4)$$

Segundo Taha (2008), a matriz P define CM. Nesta matriz, todas as probabilidades de transição p_{ij} são estacionárias e não dependem do tempo. Para Barbosa (2009), como todas as entradas são probabilidades, então devem ser não-negativas, além de que a soma de cada uma das linhas deve sempre ser igual a 1.

Ainda para Taha (2008) e Barbosa (2009), uma CM homogênea irredutível e aperiódica com espaço de estados finito é considerada ergódica, se atender as definições das Equações 5 e 6.

$$\pi_j = \lim_{n \to \infty} a_j^{(n)}, j = 0,1,2, \dots \quad (5)$$

$$\pi = \pi P \quad (6)$$

A Equação 6 mostra que as probabilidades π permanecem inalteradas após uma transição, desta forma representam a distribuição do estado no equilíbrio. Em alguns casos, a CM também é utilizada para extrair a função densidade de probabilidade (DAMÁSIO; MENDONÇA, 2018), já em outras situações, é utilizada para avaliar a correlação em séries temporais (TANG et al., 2019).

Para que se determinem as probabilidades de um processo voltar a um determinado estado j pela primeira vez, também conhecido como tempo médio do primeiro retorno ou tempo médio de recorrência, é calculado uma CM de n-estados conforme Equação 7.

$$\mu_{ij} = \frac{1}{\pi_j}, j = 1,2, \dots, n$$

2.2.1 CADEIAS DE MARKOV MULTIVARIADAS

Segundo Cechin e Corso (2019), as probabilidades associadas a várias mudanças de estado são denominadas de probabilidades de transição. Neste processo, uma matriz de transição descreve as probabilidades de transições de um estado para outro. Desta forma, a CMM é utilizada quando se tem uma série temporal de dados e a transição ocorre para diferentes etapas, podendo ser considerada de ordem *n*.

Para Ching, Ng e Fung (2008) e Li et al. (2019), uma CMM de primeira ordem modela o comportamento de múltiplos dados para determinada etapa no tempo. Com isso, existem *s* sequências e *m* estados possíveis, onde a probabilidade do estado *j* no tempo $t = r + 1$ depende da probabilidade de estado de todas as sequências, inclusive em $t = r$, conforme Equação 8.

$$X_{r+1}^{(j)} = \sum_{k=1}^{s} \lambda_{jk} P^{(jk)} X_r^{(k)} \; para \; j = 1,2, \dots, s. r = n-1, n, \dots \tag{8}$$

Segundo Ching, Fung e Ng (2004), a probabilidade da sequência $X_{r+1}^{(j)}$ no tempo $r + 1$ depende da média ponderada obtida entre $P_{jk} x_r^{(k)}$. A matriz P_{jk} é uma matriz de transição do estado no tempo *t* em uma determinada sequência *k*, até o estado no tempo $t + r$ na sequência *j*, e $x_r^{(k)}$ é uma distribuição de probabilidade de estados no tempo *r* na sequência *k*. Nas Equações 9 e 10, é apresentada a matriz.

$$X_{r+1} = \begin{pmatrix} x_{r+1}^{(1)} \\ x_{r+1}^{(2)} \\ \vdots \\ x_{r+1}^{(s)} \end{pmatrix} = \begin{pmatrix} \lambda_{11}P_{11} & \lambda_{12}P_{12} & \cdots & \lambda_{1s}P_{1s} \\ \lambda_{21}P_{21} & \lambda_{22}P_{22} & \cdots & \lambda_{2s}P_{2s} \\ \vdots & \vdots & \vdots & \vdots \\ \lambda_{s1}P_{s1} & \lambda_{s2}P_{s2} & \cdots & \lambda_{ss}P_{ss} \end{pmatrix} \begin{pmatrix} x_r^{(1)} \\ x_r^{(2)} \\ \vdots \\ x_r^{(s)} \end{pmatrix} \tag{9}$$

Ou

$$X_{r+1} = QX_r \tag{10}$$

2.3 REGRESSÃO LINEAR MÚLTIPLA

Segundo Devore (2018), a Regressão Linear Múltipla (RLM) é um modelo probabilístico que relaciona uma variável independente y a mais de uma variável independente. A equação do modelo de regressão múltipla é mostrada na Equação 11.

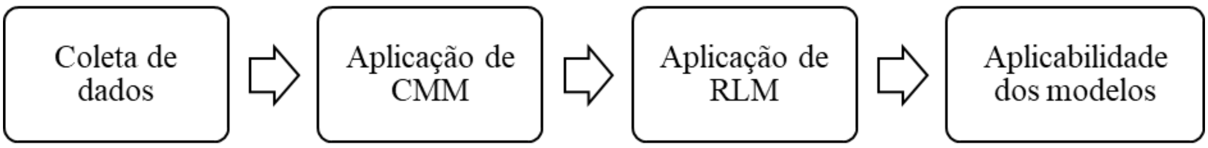
$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon \tag{11}$$

Este modelo é chamado de modelo de regressão linear múltipla com k variáveis repressoras. Na equação, os parâmetros $\beta_j, j = 0,1, \dots, k$ são chamados de coeficientes de regressão, e representam a variação esperada na resposta Y por unidade de variação unitária em x_j, desde que todos os outros regressores restantes x_i $(i \neq j)$ forem mantidos constantes (MONTGOMERY; RUNGER, 2016).

3 METODOLOGIA

Neste capítulo, é apresentada a metodologia para a aplicação de CMM e RLM na análise da variação dos dados relativos para investimentos no Ibovespa, tendo o Dólar como influenciador. A metodologia foi separada conforme mostra a Figura 1.

Figura 1 – Etapas da metodologia



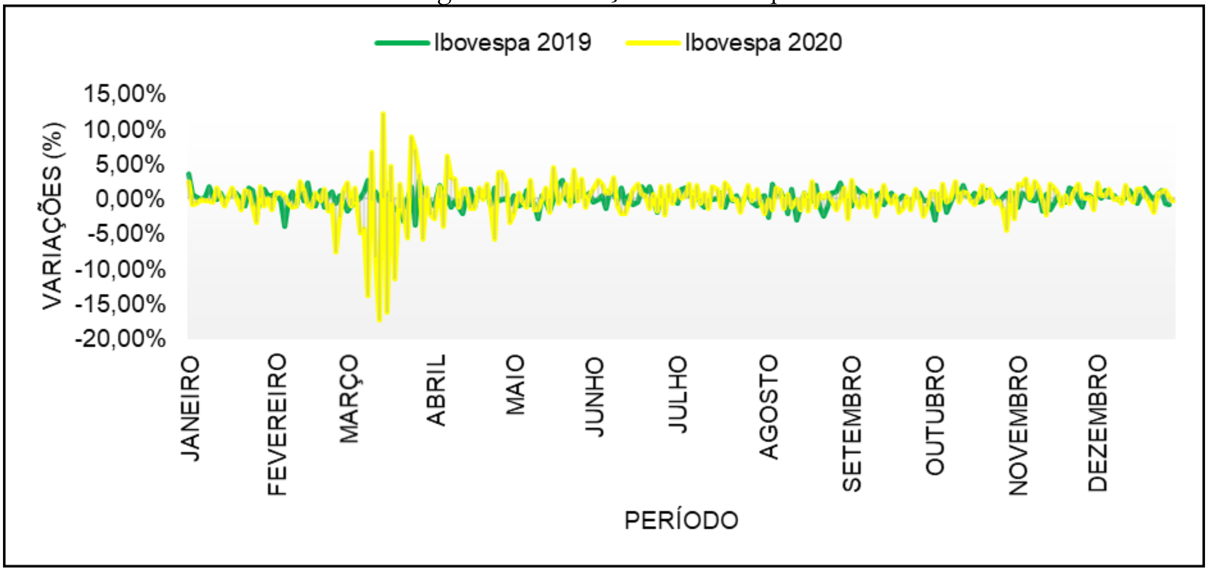
Fonte: Autores (2021)

3.1 COLETA DE DADOS

Com a coleta dos dados, o principal objetivo é avaliar a variação percentual dos valores do Dólar e do índice Ibovespa. O Dólar é a moeda oficial em diversos países, por outro lado, o índice Ibovespa é o principal indicador de desempenho médio das ações no Brasil, podendo-se dizer que este índice funciona como um termômetro do mercado de ações brasileiro.

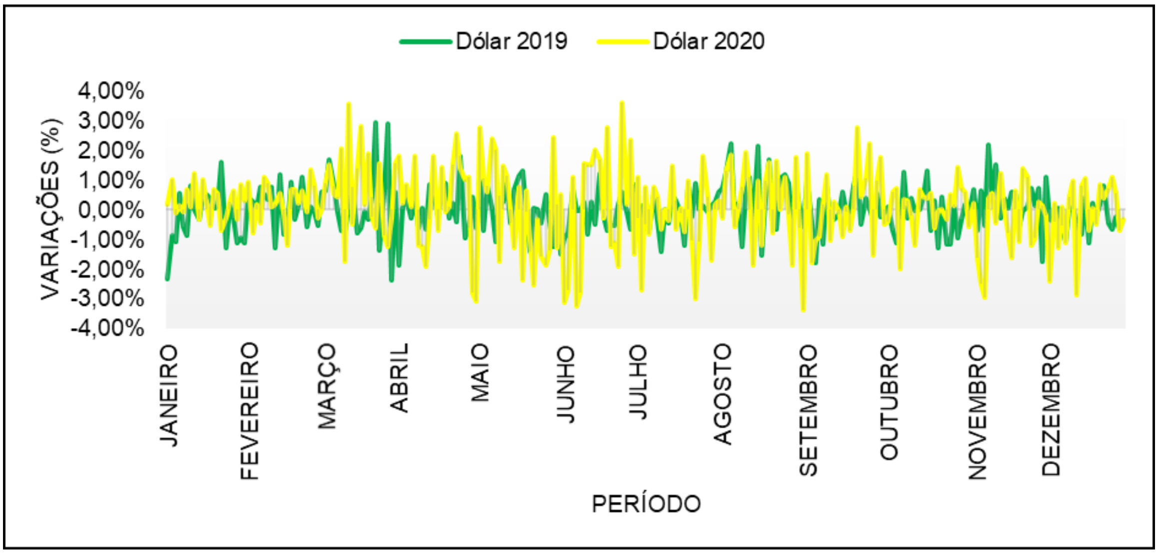
Para este trabalho, os dados da variação do Ibovespa e a variação do Dólar foram extraídos através do endereço eletrônico Investing (2020). Foram coletados dados de dois períodos diferentes, sendo um deles com início em 02 de janeiro de 2019 e término em 31 de dezembro de 2019 e outro contemplando o período de 02 de janeiro de 2020 e término em 30 de dezembro de 2020. As Figuras 2 e 3 representam a variação do Ibovespa e do dólar nos períodos de 2019 e 2020.

Figura 2 – Variação do Ibovespa



Fonte: Autores (2021)

Figura 3 –Variação Dólar



Fonte: Autores (2021)

Salienta-se que no período de março a abril do ano de 2020, existem picos de variação que destoam do padrão de variação, isso deve-se ao fato de ser o período onde iniciou-se a pandemia do Covid-19 no Brasil, mostrando de fato, como estes eventos influenciam diretamente nos índices de bolsas de valores.

3.2 APLICAÇÃO DE CADEIAS DE MARKOV MULTIVARIADAS

Utilizando Cadeias de Markov Multivariadas aplicadas a este modelo, pode-se analisar mais de uma série temporal. Desta forma, verifica-se como acontecem as transições de estados em diferentes séries. Também por meio deste modelo, é possível analisar projeções futuras, com um modelo de ordem *n* passos no tempo.

Após a extração e ajustes dos dados, é possível definir os intervalos para análise. Para cada intervalo, é atribuída uma variável normalmente indicada por uma letra. Os intervalos são definidos de acordo com a variação dos dados analisados, conforme ilustra o Quadro 1.

Com a definição das faixas, é possível elaborar a matriz de frequência, que mostra com qual frequência ocorre a transição de uma faixa para outra em um passo de tempo. Neste modelo, foi criada a matriz De Dólar/Para Ibovespa, contemplando o somatório da quantidade de vezes que ocorreu transição de uma faixa para outra.

Quadro 1 – Faixas de variação percentual

Faixa	Variável
Menor – 1,5%	A
-1,5% a -1,01%	B
-1,0% a -0,01%	C
0% a 0,99%	D
1% a 1,5%	E
Maior 1,5%	F

Fonte: Autores (2021)

Após a matriz de frequência, é possível obter a matriz de transição, efetuando a divisão do valor da quantidade de vezes que o processo muda de estado, pelo somatório de sua respectiva linha na matriz, sendo este um percentual de probabilidade de ocorrência.

3.3 APLICAÇÃO DE REGRESSÃO LINEAR MÚLTIPLA

Para a aplicação do modelo de RLM, o objetivo é prever uma variável, ajustando-a a melhor relação linear entre uma variável dependente e uma ou mais variáveis independentes.

Na aplicação deste modelo, a variável dependente a ser utilizada foi o Ibovespa no tempo ($t+1$), enquanto as variáveis independentes foram o Dólar e Ibovespa no tempo (t). Com isto, pode-se obter uma equação com a finalidade de prever eventos futuros no índice Ibovespa.

3.4 APLICABILIDADE DOS MODELOS

Para a validação dos modelos, ocorreram simulações de investimento ao longo dos períodos analisados, verificando quais seriam os retornos financeiros, caso fossem utilizados estes métodos. Para avaliar a eficácia dos modelos, foi aplicado o modelo de CMM e RLM nos 6 primeiros meses de cada ano (2019 e 2020) e verificar quais seriam os ganhos ou perdas nos próximos 6 meses respectivamente. Para ser possível comparar os modelos, foram feitas as seguintes premissas:

a) a equação da RLM foi utilizada para avaliar o comportamento do Ibovespa do dia $(t+1)$, tendo como entrada o Dólar e o Ibovespa no tempo (t). Ou seja, sempre foi considerado os dados do dia atual para prever o dia seguinte;
b) para avaliar a técnica de CMM, se considerou que no dia (t) está em uma faixa de variação do Dólar. A partir disso, foi considerado como compra em uma sugestão de investimento onde a probabilidade de alta for positiva, e venda em caso de a maior probabilidade for de queda;
c) o valor do ganho/perda a ser considerado foi o próprio valor do Ibovespa no tempo $(t+1)$; e
d) os resultados foram avaliados mês a mês, de forma a identificar se a distância dos dados de construção do modelo, apresentam maior influência em algum dos métodos.

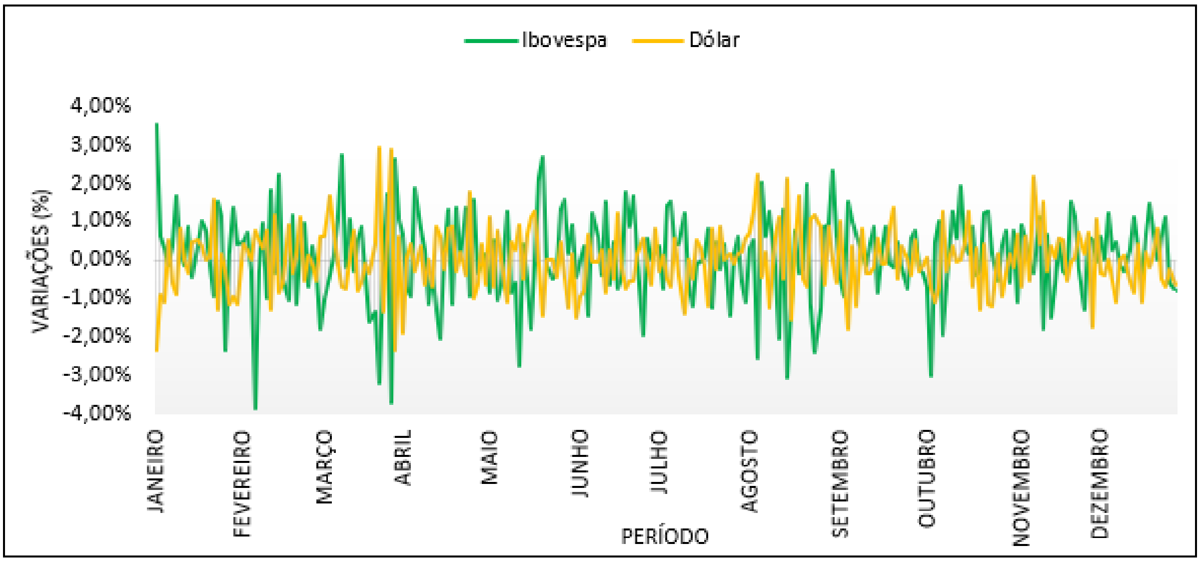
4 RESULTADOS

Neste capítulo, mostra-se os resultados obtidos na aplicação dos modelos de Cadeias de Markov Multivariadas e de Regressão Linear Múltipla tendo em vista as etapas descritas no capítulo 3.

4.1 COLETA DE DADOS

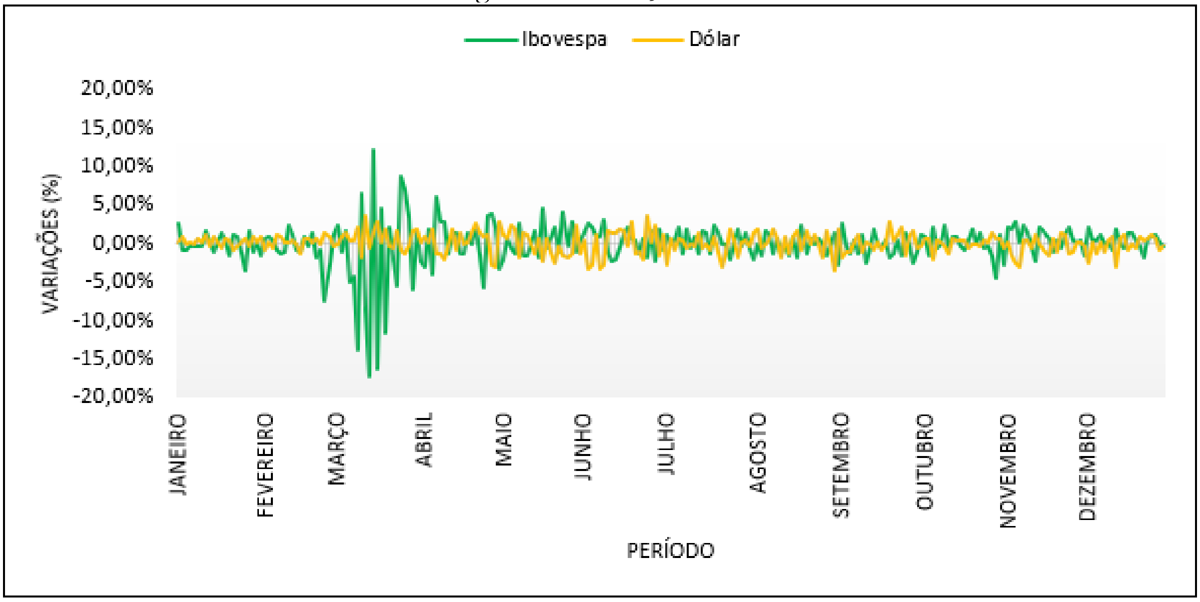
Para a análise dos índices, os dados foram retirados do endereço eletrônico Investing (2020). As Figuras 4 e 5 representam a variação do Dólar e Ibovespa nos períodos de 2019 e 2020.

Figura 4 – Variação 2019



Fonte: Autores (2021)

Figura 5 – Variação 2020



Fonte: Autores (2021)

4.2 MATRIZ DE TRANSIÇÃO

Por meio das faixas de variação, criam-se as matrizes de transição relacionando Dólar e Ibovespa, ou seja, cria-se a matriz "De Dólar/Para Ibovespa", com o intuito de prever os dados financeiros com um tempo futuro de um dia.

A matriz de transição demonstra a probabilidade percentual de mudança ou permanência na mesma faixa dentre os dados analisados. Os Quadros 2 e 3 mostram as matrizes de transição para os seis primeiros meses de 2019 e 2020 respectivamente.

Quadro 2 – Matriz de transição "De Dólar / Para Ibovespa" – 2019

DE DÓLAR/PARA IBOVESPA	A	B	C	D	E	F	SOMA
A	0,00%	0,00%	50,00%	25,00%	25,00%	0,00%	100%
B	0,00%	0,00%	44,44%	22,22%	22,22%	11,11%	100%
C	4,35%	15,22%	32,61%	30,43%	6,52%	10,87%	100%
D	15,69%	3,92%	27,45%	31,37%	13,73%	7,84%	100%
E	0,00%	0,00%	33,33%	33,33%	0,00%	33,33%	100%
F	0,00%	0,00%	20,00%	20,00%	0,00%	60,00%	100%

Fonte: Autores (2021)

Quadro 3 – Matriz de transição "De Dólar / Para Ibovespa" – 2020

DE DÓLAR/PARA IBOVESPA	A	B	C	D	E	F	SOMA
A	26,67%	13,33%	26,67%	6,67%	0,00%	26,67%	100%
B	10,00%	20,00%	0,00%	20,00%	10,00%	40,00%	100%
C	18,18%	13,64%	22,73%	18,18%	9,09%	18,18%	100%
D	27,03%	5,41%	21,62%	10,81%	13,51%	21,62%	100%
E	33,33%	13,33%	13,33%	26,67%	0,00%	13,33%	100%
F	21,74%	8,70%	4,35%	8,70%	8,70%	47,83%	100%

Fonte: Autores (2021)

4.3 REGRESSÃO LINEAR MÚLTIPLA

Na aplicação do modelo de RLM, foram utilizados os dados relativos ao Dólar e Ibovespa dos seis primeiros meses de 2019 e 2020, respectivamente, com o objetivo de prever os eventos futuros no índice Ibovespa.

Foram utilizadas duas variáveis independentes, o Dólar e o Ibovespa no tempo *(t)* e uma variável dependente no tempo *(t+1)*, sendo o Ibovespa. Após a aplicação do modelo de regressão, utilizando o *software* Excel, obtiveram-se os dados dos coeficientes das respectivas variáveis, como mostram os Quadros 4 e 5.

Quadro 4 – Coeficientes de Regressão – 2019

Interseção	0,001
IBOVESPA	0,058
DÓLAR	0,198

Fonte: Autores (2021)

Quadro 5 – Coeficientes de Regressão – 2020

Interseção	-0,002
IBOVESPA	-0,365
DÓLAR	-0,596

Fonte: Autores (2021)

4.3 AVALIAÇÃO DE APLICABILIDADE DOS MODELOS

Para a avaliação da aplicabilidade dos modelos de CMM e RLM, foram utilizados os dados relativos aos seis primeiros meses de cada ano, 2019 e 2020. Com estes dados, foi possível prever os próximos seis meses de cada ano com a utilização de simulações matemáticas. Esta avaliação tem como princípio analisar o comportamento dos métodos para este caso em específico.

É importante salientar que, tanto o modelo de CMM quanto a equação de RLM, foram construídos somente levando em conta os seis primeiros meses de cada um dos anos.

4.3.1 AVALIAÇÃO DO MODELO DE CADEIAS DE MARKOV MULTIVARIADA

Por meio da matriz de transição, foi possível elencar quais seriam as faixas de alta ou de queda do índice Ibovespa com maior probabilidade de ocorrência. As faixas A, B e C foram consideradas faixas de baixa do valor do índice, enquanto as faixas D, E e F foram consideradas de alta do mesmo. O Quadro 6 mostra quais são probabilidades das faixas mencionadas.

Quadro 6 – Probabilidades de Alta ou Queda

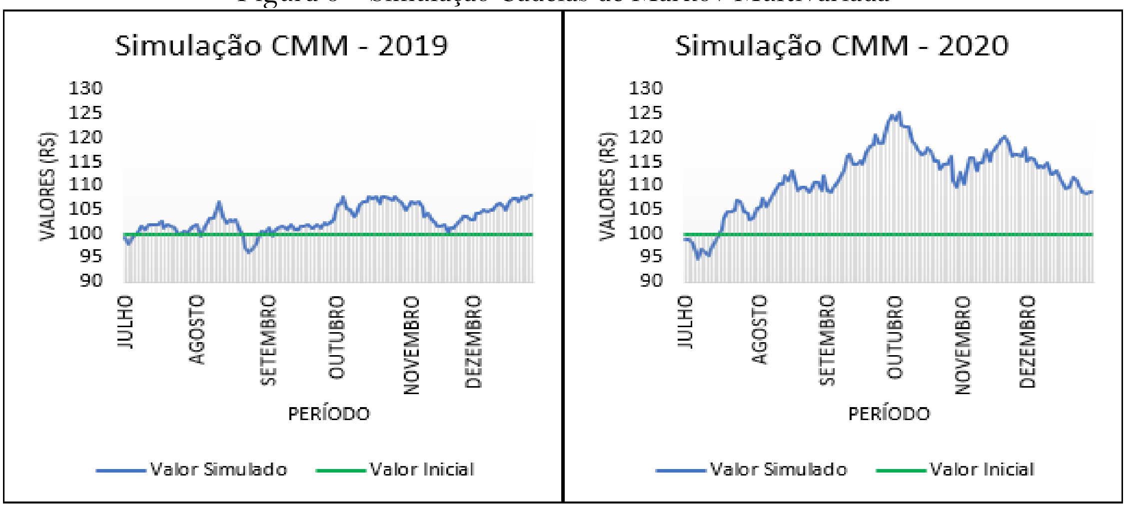
FAIXA	2019		2020	
	QUEDA	ALTA	QUEDA	ALTA
A	50,00%	50,00%	66,67%	33,33%
B	44,44%	55,56%	30,00%	70,00%
C	52,17%	47,83%	54,55%	45,45%
D	47,06%	52,94%	54,05%	45,95%
E	33,33%	66,67%	60,00%	40,00%
F	20,00%	80,00%	34,78%	65,22%

Fonte: Autores (2021)

Com estas faixas definidas, foi possível simular quais seriam os ganhos no segundo semestre de cada ano. Para isto, utilizou-se um investimento inicial no valor de R$ 100,00, valor este, escolhido de forma arbitrária. A simulação levou em conta sempre o valor do Ibovespa no tempo *(t+1)*, ou seja, somente haveria ganhos, caso o modelo matemático de CMM previsse a queda ou a alta do índice do dia seguinte. A Figura 6 ilustra o retorno do valor investido nos períodos de 2019 e 2020.

Com as simulações, pode-se observar um valor acumulado ao final de seis meses de R$ 108,10 no ano de 2019 e de R$ 108,90 no ano de 2020, resultando em um ganho de 8,1% e de 8,9% respectivamente.

Figura 6 – Simulação Cadeias de Markov Multivariada



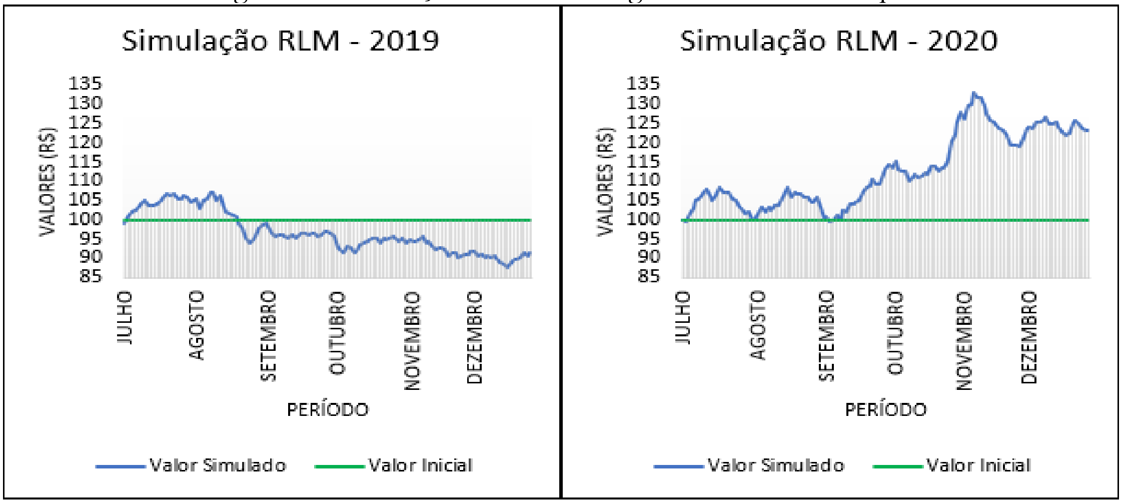
Fonte: Autores (2021)

4.3.2 AVALIAÇÃO DO MODELO DE REGRESSÃO LINEAR MÚLTIPLA

No modelo de RLM, foram utilizados os coeficientes supracitados a fim de relacionar todas as variáveis, independentes e dependentes por meio de uma equação de regressão, obtida com *software* Excel.

Com isto, foi simulado um investimento inicial também no valor arbitrário de R$ 100,00. Para a validação da equação, foram utilizados meios que comparassem o valor previsto com o valor real da variação do Ibovespa. A Figura 7 mostra os ganhos ou perdas dos respectivos anos de 2019 e 2020.

Figura 7 – Simulação utilizando Regressão Linear Múltipla



Fonte: Autores (2021)

Analisando as simulações, pode-se observar um valor acumulado ao final dos seis meses de R$ 91,40, o que significa uma perda de 8,6% no ano de 2019. Já para o ano de 2020, o valor acumulado é de R$ 123,10, resultando em um ganho 23,1%.

5 CONSIDERAÇÕES FINAIS

Em um mercado financeiro que se mostra cada vez mais complexo e dinâmico, a aplicação dos modelos matemáticos propostos se mostrou eficaz para auxílio de investidores tomarem decisões mais precisas no mercado de ações. Com este trabalho, foi possível prever variações no índice Ibovespa, utilizando o Dólar como um influenciador direto e também o próprio valor passado do Ibovespa.

Inicialmente, coletaram-se todos os dados das variações percentuais dos índices Ibovespa e Dólar, possibilitando definir faixas de variações para tais. Definidas as faixas, foi possível aplicar o modelo de CMM, avaliando a probabilidade de ocorrer mudanças do índice Ibovespa no dia seguinte. Este fato fez com que se conseguisse trabalhar com múltiplas variáveis com comportamento estocástico. Para o modelo de RLM, foram utilizados o Dólar e o Ibovespa como variáveis independentes e o Ibovespa em um tempo *(t+1)*. Desta forma, foi possível criar a equação de regressão, com a finalidade de prever eventos futuros no índice Ibovespa.

Como resultado, obteve-se um ganho de 8,1% no ano de 2019 e de 8,9% no ano de 2020, tendo como base o modelo de CMM. Já para o modelo de RLM, no ano de 2019 houve uma perda de 8,6% e no ano de 2020, houve um ganho de 23,1%. Tal perda pode ser relacionada a uma possível mudança brusca de comportamento no índice Ibovespa. Comparando os métodos à uma aplicação simples na Poupança, os ganhos no mesmo período analisado seriam de 1,67% e 0,61% para os anos de 2019 e 2020 respectivamente. Com isto, o modelo criado para este banco de dados, mostrou-se eficaz, entretanto, para novos dados é necessária uma nova avaliação. Estes resultados mostram um bom potencial de aplicação de CMM para este tipo de problema, onde percebe-se que o método é pouco explorado na literatura.

Como sugestão para estudos futuros, propõem-se o uso de Cadeias de Markov atribuído a mais índices de bolsas de valores ao mesmo tempo, juntamente com a atualização diária da matriz de transição, com a finalidade de estudar o comportamento e a influência do método em diferentes bolsas de valores com ainda mais precisão. Já para o uso de Regressão Linear Múltipla, podem ser atribuídos também mais índices de bolsas de valores como variáveis independentes, para avaliação, além de otimizar a equação e atualizá-la em um menor espaço de tempo, como, por exemplo, utilizar o passo diário de atualização de dados para prever o dia seguinte.

REFERÊNCIAS

ARAUJO FILHO, ALFREDO GONÇALVES DE; OLIVEIRA, DAVI MACHADO DE; BASTOS, EDSON VINICIUS PONTES. Regressão linear múltipla com indicadores macroeconômicos e do mercado imobiliário. *In*: ENCONTRO NACIONAL DE ENGENHARIA DE PRODUÇÃO (ENEGEP), XXXIX. **Anais...** Santos, São Paulo: 2019.

ANDRADE, EDUARDO LEOPOLDINO de. **Introdução à Pesquisa Operacional**: método e modelos para análise de decisões. 5. ed. Rio de Janeiro: LTC, 2015.

BARBOSA, HELENICE LOPES. **Métodos estatísticos em cadeias de Markov.** 2009. 49 f. Dissertação (Mestrado em Probabilidade e Estatística; Modelagem Matemática) – Universidade Federal do Rio Grande do Norte, Natal, 2009. Disponível em: <http://repositorio.ufrn.br/handle/123456789/18629>. Acesso em: 17 out. 2020.

BELFIORE, PATRÍCIA; FÁVERO, LUIS. PAULO. **Pesquisa Operacional** - Para Cursos de Engenharia. Rio de Janeiro: Elsevier, 2013.

BOLSON, MARCELO H.; BORELLA, LUCAS DE C.; TOMÉ, FERNANDA; ORLANDIN, BRUNA C.; CORSO, LEANDRO L.. Aplicação de cadeias de Markov para análise de variação do dólar americano. *In*: ENCONTRO NACIONAL DE ENGENHARIA DE PRODUÇÃO (ENEGEP), XXXIX. **Anais...** Santos, São Paulo: 2019.

BRITO, OSIAS. **Mercado financeiro**. 3. ed. São Paulo: Saraiva Educação, 2020.

CARRETE, LILIAM SANCHEZ. **Mercado Financeiro Brasileiro**. 1. ed. São Paulo: Grupo GEN, 2019.

CECHIN, RAFAELA BOEIRA; CORSO, LEANDRO LUÍS. High-order Multivariate Markov Chain applied in Dow Jones and Ibovespa indexes. **Pesquisa Operacional**, Rio de Janeiro, v. 39, n. 1, p. 205-223, 2019. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0101-74382019000100008&lng=en&nrm=iso>. Acesso em: 01 set. 2020.

CHING, WAI KI; FUNG, ERIC S.; NG, MICHAEL K. Higher-order Markov chain models for categorical data sequences. **Naval Research Logistics (NRL)**, v. 51, n. 4, p. 557-574, 2004. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/nav.20017>. Acesso em: 01 nov. 2020.

CHING, WAI-KI; NG, MICHAEL K.; FUNG, ERIC S. Higher-order multivariate Markov chains and their applications. **Linear Algebra and its Applications**, v. 428, n. 2-3, p. 492-507, 2008. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0024379507002169>. Acesso em: 01 nov. 2020.

DAMÁSIO, BRUNO; MENDONÇA, SANDRO. Modelling insurgent-incumbent dynamics: Vector autoregressions, multivariate Markov chains, and the nature of technological competition. **Applied Economics Letters**, v. 26, n. 10, p. 843-849, 2019. Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/13504851.2018.1502863>. Acesso em: 21 ago. 2021.

DEVORE, JAY L. **Probabilidade e estatística para engenharia e ciências** – Tradução da 9ª edição norte-americana. São Paulo: Cengage Learning Brasil, 2018.

FRANCO, M. M.; CECHIN, R. B.; PASOLINI, M.; MONEGAT, A. D. R.; CORSO, L. L. Cadeia de Markov Multivariada Aplicada na Bolsa de Valores Utilizando Dados da Petrobrás, Dólar e Petróleo WTI. **Contexto - Revista do Programa de Pós-Graduação em Controladoria e Contabilidade da UFRGS**, v. 21, n. 47, p. 79-94, 2021.
GURURAJ, VAISHNAVI; SHRIYA, V. R.; ASHWINI, K. Stock market prediction using linear regression and support vector machines. **International Journal of Applied Engineering Research**, v. 14, n. 8, p.

1931-1934, 2019. Disponível em: <https://www.ripublication.com/ijaer19/ijaerv14n8_24.pdf>. Acesso em: 07 nov. 2020.

HILLIER, FREDERICK S.; LIEBERMAN, GERALD J. **Introdução à pesquisa operacional**. 9. ed. Porto Alegre: Amgh, 2013.

INVESTING: Finanças, Câmbio e Bolsas de Valores. **Site institucional**. 2020. Disponível em: <https://br.investing.com/>. Acesso em: 16 nov. 2020.

IZIDORO, CLEYTON. **Mercado de capitais**. São Paulo: Pearson Education, 2015.

KERR, ROBERTO BORGES. **Mercado financeiro e de capitais**. São Paulo: Editora Pearson, 2011.

LI, WEN; KE, RIHUAN; CHING, WAI-KI; NG, MICHAEL. A C-eigenvalue problem for tensors with applications to higher-order multivariate Markov chains. **Computers & Mathematics with Applications**, v. 78, n. 3, p. 1008-1025, 2019. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0898122119301415>. Acesso em: 21 ago. 2021.

MONTGOMERY, DOUGLAS C.; RUNGER, GEORGE C. **Estatística aplicada e probabilidade para engenheiros**. 6. ed. Rio de Janeiro: LTC, 2016.

NETO, ALEXANDRE ASSAF. **Mercado Financeiro**, 14. ed. São Paulo: Grupo GEN, 2018.

PEREIRA, CLEVERSON LUIZ. **Mercado de capitais**. Curitiba: Editora Ibpex, 2013.

RODRIGUES, R. **Pesquisa Operacional**. Porto Alegre: Grupo A, 2017.

TAHA, HAMDY A. **Pesquisa operacional**: uma visão geral. 8. ed. São Paulo: Pearson Prentice Hall, 2008.

TANG, JINJUN; HU, JIN; HAO, WEI; CHEN, XINQIANG; QI, YONG. Markov Chains based route travel time estimation considering link spatio-temporal correlation. **Physica A**: Statistical Mechanics and its Applications, v. 545, p. 123759, 2020. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0378437119320941>. Acesso em: 21 ago. 2021.

VIRGILITO, SALVATORE B. **Pesquisa operacional**. São Paulo: Editora Saraiva, 2017.