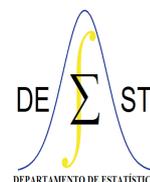




UNIVERSIDADE FEDERAL DE OURO PRETO
INSTITUTO DE CIÊNCIAS EXATAS E BIOLÓGICAS
DEPARTAMENTO DE ESTATÍSTICA
BACHARELADO EM ESTATÍSTICA



Uma abordagem pela teoria de grafos para descrever similaridades de ativos do Índice BOVESPA no período de janeiro de 2008 a junho de 2021

Daiane Jule Silva

Ouro Preto - MG
Dezembro de 2021

Daiane Jule Silva

**Uma abordagem pela teoria de grafos para
descrever similaridades de ativos do Índice
BOVESPA no período de janeiro de 2008 a junho
de 2021**

Monografia de Graduação apresentada ao
Departamento de Estatística do Instituto
de Ciências Exatas e Biológicas da Uni-
versidade Federal de Ouro Preto como re-
quisito parcial para a obtenção do grau de
bacharel em Estatística.

Orientador

Professor Dr. Rivert Paulo Braga Oliveira

UNIVERSIDADE FEDERAL DE OURO PRETO – UFOP
DEPARTAMENTO DE ESTATÍSTICA – DEEST

Ouro Preto - MG

Dezembro de 2021

SISBIN - SISTEMA DE BIBLIOTECAS E INFORMAÇÃO

S586a Silva, Daiane Jule.

Uma abordagem pela teoria de grafos para descrever similaridades de ativos do Índice BOVESPA no período de janeiro de 2008 a junho de 2021. [manuscrito] / Daiane Jule Silva. - 2021.

64 f.: il.: color., gráf., tab..

Orientador: Prof. Dr. Rivert Oliveira.

Monografia (Bacharelado). Universidade Federal de Ouro Preto. Instituto de Ciências Exatas e Biológicas. Graduação em Estatística .

1. Economia. 2. Mercado financeiro. 3. Bolsa de Valores de São Paulo. I. Oliveira, Rivert. II. Universidade Federal de Ouro Preto. III. Título.

CDU 330(81)

Bibliotecário(a) Responsável: Luciana De Oliveira - SIAPE: 1.937.800



MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DE OURO PRETO
REITORIA
INSTITUTO DE CIÊNCIAS EXATAS E BIOLÓGICAS
COLEGIADO DO CURSO DE ESTATÍSTICA



FOLHA DE APROVAÇÃO

Daiane Jule Silva

Uma abordagem pela teoria de grafos para descrever similaridades de ativos do Índice BOVESPA no período de janeiro de 2008 a junho de 2021

Monografia apresentada ao Curso de Estatística da Universidade Federal de Ouro Preto como requisito parcial para obtenção do título de Bacharel em Estatística

Aprovada em 1º de dezembro de 2021

Membros da banca

Dr. Rivert Paulo Braga Oliveira - Orientador - Universidade Federal de Ouro Preto
Dr. Anderson Ribeiro Duarte - Universidade Federal de Ouro Preto
Dr. Helgem de Souza Ribeiro Martins - Universidade Federal de Ouro Preto

Professor Dr. Rivert Paulo Braga Oliveira, orientador do trabalho, aprovou a versão final e autorizou seu depósito na Biblioteca Digital de Trabalhos de Conclusão de Curso da UFOP em 01/12/2021



Documento assinado eletronicamente por **Helgem de Souza Ribeiro Martins, CHEFE DA DIVISÃO DE ARTICULAÇÃO DE PESQUISA E PÓS-GRADUAÇÃO?**, em 07/12/2021, às 10:53, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Anderson Ribeiro Duarte, PROFESSOR DE MAGISTERIO SUPERIOR**, em 07/12/2021, às 11:09, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Rivert Paulo Braga Oliveira, PROFESSOR DE MAGISTERIO SUPERIOR**, em 07/12/2021, às 11:49, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site http://sei.ufop.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **0245436** e o código CRC **6F65E348**.

Referência: Caso responda este documento, indicar expressamente o Processo nº 23109.012026/2021-21

SEI nº 0245436

R. Diogo de Vasconcelos, 122, - Bairro Pilar Ouro Preto/MG, CEP 35400-000
Telefone: - www.ufop.br

Aos meus pais, Aldair Silva e Crislaine Silva.

Ao meu irmão, Aldair Junio.

Ao meu noivo, Thales Fernandes.

Agradecimentos

Agradeço a todos que, direta ou indiretamente, me apoiaram neste trabalho. Ao meu orientador, Dr. Rivert Paulo Braga Oliveira pela orientação e dedicação à elaboração da monografia e da Iniciação Científica, além das oportunidades e aprendizado que transpassaram o mundo acadêmico. Aos meus pais, Aldair Silva e Crislaine Silva, pelo carinho e por acreditarem em mim, essa conquista só foi possível com o apoio de vocês. Ao meu irmão, Aldair Junio, pelo companheirismo, apoio, e carinho. Ao meu noivo Thales Fernandes, pelo companheirismo, incentivo, dedicação, carinho e pelas palavras de conforto nos momentos difíceis. Aos amigos do IFMG-OP , da Estatística-UFOP, e demais amigos que fiz ao longo dos anos. Aos técnicos administrativos e demais servidores terceirizados, já que a Universidade não funciona e nem é habitável sem eles. A UFOP pelo financiamento através do Programa de Iniciação Científica da instituição. Aos professores da UFOP pelo conhecimento transmitido, em especial aos professores do DEEST. A UFOP pelo suporte e estrutura.

“Peçam, e será dado; busquem, e encontrarão; batam, e a porta será aberta. Pois todo o que pede recebe; o que busca encontra; e àquele que bate, a porta será aberta.

Mateus 7:7-8

Uma abordagem pela teoria de grafos para descrever similaridades de ativos do Índice BOVESPA no período de janeiro de 2008 a junho de 2021

Autor: Daiane Jule Silva

Orientador(a): Dr. Rivert Paulo Braga Oliveira

RESUMO

O funcionamento do mercado financeiro envolve inúmeras variáveis e é afetado por diversos fatores, como questões políticas e econômicas. Seu comportamento tem sido objeto de estudo de diversas áreas como física, estatística e ciências da computação. Vários deles buscam uma modelagem conjunta dos ativos. Atualmente o uso da teoria de grafos tem se mostrado uma ferramenta eficaz para descrição do comportamento de índices relativos à economia, em particular, de preços de ações, através de seu grau de correlação. Neste trabalho, baseado em estudos pregressos, buscou-se reproduzir sua pesquisa através da análise da variação do valor dos ativos do portfólio indexados pelo Índice BOVESPA no período compreendido entre janeiro de 2008 a junho de 2021. Observou-se, pela análise dos dados da matriz de dissimilaridades, que ativos de atividades econômicas afins se ligam em um grafo por distâncias mais curtas, enquanto aqueles que possuem dissimilaridades econômicas apresentam ligações mais longas. O método aplicado para construir este grafo foi o da árvore de ligação mínima com base na matriz de correlação da série de preços das ações durante o período estudado. Notou-se que também que,

em períodos de crise, a correlação e a distância possuem comportamento diferente em relação aos períodos que não estão em crise.

Palavras-chave: Comprimento normalizado da árvore, Taxa de sobrevivência de passo simples, Evolução temporal.

A graph theory approach to describe similarities of bovespa index assets from January 2008 to June 2021

Author: Daiane Jule Silva

Advisor: Dr. Rivert Paulo Braga Oliveira

ABSTRACT

The dynamics of the financial market involves numerous variables and it is affected by several factors, such as political and economic issues. Its behavior has been the object of study in several areas such as physics, statistics and computer science. Plenty of them seek a joint modeling of the assets. Currently, the use of graph theory has been proven to be an effective tool for describing the behavior of indexes related to the economy, in particular, of stock prices, through its degree of correlation. In this work, based on the studies of (MANTEGNA, 1999), its been sought to reproduce his research by analyzing the variation in the value of portfolio assets indexed by the BOVESPA Index from January 2008 to June 2021. Its been observed, by analysis of the dissimilarities matrix, that assets of related economic activities are connected by shorter distances, while those that have economic dissimilarities have longer connections. The method applied to construct this graph was that of the minimum span tree based on the correlation matrix of the stock price series during the studied period. It was also observed that, in times of crisis, correlation and distance have different behavior from non crisis periods.

Keywords: Normalized tree length, Single-step survival ratio, Time evolution.

Lista de figuras

1	Composição do Índice por Setor de Atuação no período de 2008 a 2021.	p. 11
2	Árvore de ligação mínima conectando as 32 ações do IBOVESPA no período de 2008 a 2021.	p. 12
3	Árvore de ligação mínima conectando as 32 ações do IBOVESPA no período de 2008 a 2015.	p. 13
4	Dendograma via ligação simples das ações do IBOVESPA, as empresas que pertencem a grupos econômicos semelhantes possuem comportamentos homogêneos, com ligações mais curtas, como a BRAP4 - VALE3(Mineração); BBDC4 - ITUB4(Bancos). Entre companhias com maior heterogeneidade, as ligações acontecem tardiamente e a distância entre elas é maior como a ligação entre o grupo BRAP4 - VALE3 com GGBR4.	p. 14
5	Correlograma das 32 ações do IBOVESPA	p. 16
6	Gráficos das funções densidades de probabilidade das distâncias, em função do tempo, para diferentes janelas de tempo.	p. 18
7	Gráficos das funções densidades de probabilidade dos coeficientes de correlação, em função do tempo, para diferentes janelas de tempo.	p. 19

8	(a): Gráfico da função densidade de probabilidade dos coeficientes de correlação em função do tempo, (b): A média, desvio padrão, assimetria e curtose dos coeficientes de correlação em função do tempo.	p. 21
9	(a): Gráfico da função densidade de probabilidade das distâncias em função do tempo, (b): A média, desvio padrão, assimetria e curtose das distâncias em função do tempo.	p. 22
10	Taxa de sobrevivência de passo unitário em função do tempo.	p. 24
11	A média, desvio padrão, assimetria e curtose em função do tempo.	p. 26

Lista de abreviaturas e siglas

B3 – Bolsa de valores do Brasil

MST – Minimum Spanning Tree / Árvore Geradora Mínima

Lista de símbolos

V (Vértice ou nós)

A (Arestas)

ρ (coeficiente de correlação)

Y_i (log-retorno)

$d(i,j)$ (matriz de distâncias)

$L(t)$ (comprimento normalizado da MST)

δT (Defasagem)

E^t (Elos)

t (tempo)

N (número de ativos)

Sumário

1	Introdução	p. 1
1.1	Objetivos	p. 2
1.2	Organização do Trabalho	p. 2
2	Literatura e Base de Dados	p. 3
2.1	Revisão de Literatura	p. 3
2.2	Material: Levantamento da Base de Dados	p. 5
3	Metodologia	p. 7
4	Resultados e Discussão	p. 11
5	Considerações Finais	p. 27
	Referências	p. 29
	Apêndice A – Código R	p. 31
	Anexo A – Abreviaturas dos Ativos	p. 63

1 Introdução

O mercado financeiro é um ambiente de negociação que envolve as operações de compra e venda de ações, fundos imobiliários, títulos públicos, entre outros. Para o funcionamento do Mercado Financeiro, existem três categorias: os Emissores, que são as instituições que emitem títulos; os Investidores, que trocam seu dinheiro pelos títulos oferecidos pelos Emissores; e os Intermediários, que são a ponte entre os Emissores e os Investidores, como por exemplo as Corretoras de Valores.

O sistema financeiro é dividido em quatro mercados, a saber: Mercado de Câmbio; Mercado Monetário; Mercado de Crédito e Mercado de Capitais. O Mercado de Capitais é formado por corretoras de valores, bolsas de valores e instituições financeiras, nos quais se negociam os principais ativos do Mercado de Capitais, as ações, Debêntures e *Commercial Papers*.

Em termos de mercado, assume-se a hipótese de um mercado financeiro eficiente, em que o preço dos ativos negociados sempre refletem inteiramente as informações disponíveis sobre os mesmos. Portanto, um mercado é eficiente quando não há maneiras de obtenção de lucros anormais mediante o uso da informação (impossibilidade de arbitragem).

Apesar do mercado financeiro ter um comportamento aleatório, existem ferramentas que permitem a modelagem e resumo da informação contida nessa aleato-

riedade. Há vários estudos em diversas áreas sobre este assunto, os quais permitem cada vez mais eficiência e descobertas de novas estratégias para auxiliar a tomada de decisão.

No contexto deste trabalho, a informação sobre o mercado financeiro é a série temporal. Entende-se por série temporal, uma sequência de observações ordenadas no tempo, a qual pode apresentar características como tendências, sazonalidade, não linearidade e pontos influentes. Mais especificamente, essa série pode ser o preço, o retorno, ou o log-retorno, todos eles no decorrer do tempo. Assim, a abordagem metodológica deste texto será desenvolvida com base nas séries de diferentes ativos que compõem o IBOVESPA da Bolsa de Valores de São Paulo (B3), através de grafos a algumas medidas de interesse.

1.1 Objetivos

Construir a Árvore de ligação Mínima que descreve o comportamento dos ativos do Índice BOVESPA, no período de janeiro de 2008 até junho de 2021, utilizando a teoria dos grafos e a matriz de coeficiente de correlação de Pearson, para estudar e compreender a dinâmica deste mercado.

1.2 Organização do Trabalho

O presente estudo está dividido da seguinte forma: Um breve referencial teórico no capítulo 2, abordando o problema do estudo. No capítulo 3 encontra-se a metodologia utilizada para alcançar os objetivos propostos. O capítulo 4 traz os resultados e discussões do trabalho. E por fim as conclusões e considerações finais estão presentes no capítulo 5.

2 Literatura e Base de Dados

2.1 Revisão de Literatura

A literatura remonta que a teoria dos grafos teve origem na cidade de Königsberg em 1736 pelo grande matemático suíço Leonhard Euler (1707-1783). No que se refere ao grafo específico deste texto, para solucionar o problema da árvore de ligação mínima BOR et al. (1926) desenvolveu o primeiro algoritmo capaz de construir tal árvore. Esse algoritmo caracteriza-se por ligar cada vértice ao seu vizinho mais próximo até que todos os vértices sejam conectados, sem formar polígonos fechados, e de modo que o tamanho das arestas seja o menor possível.

No contexto de mercado financeiro, e assumindo a hipótese de SAMUELSON (1965) de que o sistema complexo do mercado financeiro possui séries de retorno imprevisíveis, MANTEGNA (1999) aplica a teoria de grafos para estudar o comportamento do mercado financeiro através de um conjunto de ações, pertencentes aos índices Dow Jones e S&P 500 INDEX, acompanhando a variação dos valores dos seus ativos no período de julho de 1989 até outubro de 1995. Nesse artigo MANTEGNA (1999) calcula o coeficiente de correlação linear de Pearson entre os retornos dessas ações. Tais retornos são a diferença a diferença dos logaritmos dos valores dos ativos. A escala logarítmica é utilizada, por conferir aos dados propriedades matemáticas interessantes do ponto de vista estatístico.

Os coeficientes de correlação são métricas estatísticas para quantificar o grau de

associação entre variáveis. O coeficiente de correlação linear de Pearson varia entre -1 e +1. O valor +1 representa a correlação positiva perfeita entre os ativos. Já o valor -1, representa a correlação negativa perfeita entre os ativos. Por fim, o valor 0 representa a ausência completa de correlação entre os ativos.

A partir dessas medidas MANTEGNA (1999) criou uma matriz de coeficientes de correlação que foi transformada numa matriz de distâncias entre os ativos. Essa matriz de distâncias foi utilizada para determinar a árvore de ligação mínima (MST - *Minimum Spanning Tree*). Através desse grafo, pode-se compreender o comportamento do mercado financeiro.

Há também o estudo desenvolvido por ONNELA et al. (2003b) que descreve o comportamento do mercado baseado na MST. Neste artigo foi observado que em períodos de crise do mercado financeiro, a árvore tem seu comprimento encolhido.

Com base no que foi apresentado no presente texto, busca-se reproduzir o trabalho desenvolvido por MANTEGNA (1999) e ONNELA et al. (2003b) através do uso da teoria de grafos para construir a MST . Além disso, analisar o comportamento do mercado financeiro com respeito a sua topologia para o índice BOVESPA.

O Índice BOVESPA que foi criado em 1968 é o principal indicador de desempenho das ações negociadas na B3. Este índice reúne as empresas mais importantes do mercado de capitais brasileiro. O índice é reavaliado a cada quatro meses, e é o resultado de uma carteira teórica de ativos.

Até onde se tem conhecimento nenhum estudo do gênero foi realizado para o mercado financeiro brasileiro.

2.2 Material: Levantamento da Base de Dados

Inicialmente, realizou-se uma busca sobre as empresas que compõe o Índice BOVESPA. Descobriu-se que as empresas são reavaliadas a cada quatro meses e devem atender critérios descritos na sua metodologia, portanto as empresas que compõe esse Índice estão em constante mudança.

Tentou-se buscar os dados com as empresas listadas no Índice para coletar os preços dos ativos, porém, essas informações não são acessíveis livremente através da internet. Na página de índices amplos da B3 (2020) na internet encontram-se as empresas que compõem o Índice e os valores dos ativos para o dia em que for acessado, mas não o histórico completo de alterações na carteira de ativos que compõem o índice BOVESPA, o que dificultou a obtenção desses dados.

Em decorrência do problema de acesso livre à base de dados, efetuou-se a aquisição do banco de dados para o período entre 1986 e 2018 através de financiamento particular do orientador desse trabalho. As informações contidas nesse arquivo são a sigla de cada ação, o tipo, o código de especificação das ações, o nome da instituição a qual pertence, o preço de fechamento do dia, e as variações na participação de cada um dos papéis na composição total do Índice.

Após a compra dos dados, observou-se que a mesma possuía muitos valores ausentes para o período de 1986 até 2018, uma vez que a composição do Índice se alterou bastante desde 1986. Diante disso, optou-se por utilizar apenas uma parte dos dados, por um período de tempo menor(2008 a 2021). Após essa decisão, realizou-se um estudo das empresas presentes no índice em 2018 para avaliar o histórico de participação dessas empresas no Índice. Considerando esse critério, compuseram o estudo diversas empresas relatadas no Anexo A. Informações sobre substituições de algumas siglas e de empresas também são fornecidas no anexo A. O período de 2008 a 2015 foi utilizado na Iniciação Científica (Silva, D. J & Oliveira, R. P. B , 2020).

Neste trabalho foi utilizado o período de 2008 a 2021, com o intuito de observar o mercado com a pandemia de SARS-CoV-2. Os dados utilizados de 2008 a 2018 foram retirados da base de dados comprada, e os de 2019 a 2021 foram retirados dos sítios YAHOO (2021) e INVESTING.COM (2021).

3 Metodologia

Na execução do projeto utilizou-se o *software* (R Core Team, 2020) para a análise de dados. Para a construção do dendograma e da árvore de ligação mínima utilizou-se a teoria de grafos.

A teoria de grafos é usada para estudar a relação entre variáveis dentro de uma rede. Sua origem é recente na história da matemática, sendo atribuída a Euler, ao resolver o problema das pontes de Königsberg em 1736 (EULER, 1736).

Um grafo $G(V, A)$ é definido pelo par V e A , em que V é o conjunto não vazio, cujos elementos são chamados de vértices ou nós, e A é o conjunto de pares ordenados não necessariamente distintos, chamado de arestas, tal que $a=(v, w)$ em que v e $w \in V$. A existência de uma aresta indica a presença de relação ou interação entre os vértices ligados por essa aresta.

A teoria de grafos pode ser utilizada no mercado financeiro para estudar a estrutura de interação entre seus agentes, com foco nas conexões entre as ações do portfólio.

A análise é baseada no grafo da matriz de correlação, através do uso do coeficiente de correlação (ρ) associado à variação simultânea dos preços dos ativos, encontrado pelo log-retorno dos seus valores diários de mercado. Assim, gera-se uma matriz simétrica, tal que $\rho_{ij} = \rho_{ji}$, em que ρ_{ji} é dado pela equação (3.1):

$$\rho_{ij} = \frac{\langle Y_i, Y_j \rangle - \langle Y_i \rangle \langle Y_j \rangle}{\sqrt{[\langle Y_i^2 \rangle - \langle Y_i \rangle^2][\langle Y_j^2 \rangle - \langle Y_j \rangle^2]}} \quad (3.1)$$

em que, $\langle \cdot \rangle$ é um operador de médias sobre todos os elementos do seu argumento. Y_i é o log-retorno do i -ésimo ativo, $i=1, \dots, n$, dado pela equação (3.2)

$$Y_i = \ln P_i(t) - \ln P_i(t-1) \quad (3.2)$$

em que $P_i(t)$ é o preço de fechamento do ativo no dia i para o tempo t . A partir da matriz simétrica ρ é possível gerar uma matriz de distâncias, dada pela equação (3.3):

$$d(i, j) = \sqrt{2(1 - \rho)} \quad (3.3)$$

em que cada entrada da matriz d representa um valor numérico indicando a proximidade entre os ativos. A equação (3.3) atende às propriedades exigidas de uma métrica, isto é:

- se $i = j$, então $d(i, j) = 0$;
- se $i \neq j$, então $d(i, j) > 0$;
- $d(i, j) = d(j, i)$ (simetria);
- $d(i, j) + d(j, k) \geq d(i, k)$ (desigualdade triangular).

Quando um grafo possui múltiplas arestas associadas ao mesmo par de vértices e/ou quando possui uma aresta coincidente a um par de vértices não distintos, ele deixa de ser simples. Neste trabalho utiliza-se o grafo simples.

O grafo que conecta todos os nós V e tem menor comprimento possível é denominado Árvore Geradora Mínima(MST).

O algoritmo para obtenção da MST consiste em:

- verificar para cada aresta $\{v, w\}$ se $d_i = d_j, \forall i \neq j$, em que d representa os tamanhos das arestas, os pares com $d_i \neq d_j$ são removidos;
- ordenar as arestas em ordem crescente de distância;
- conectar os nós de menor distância, um de cada vez, verificar se fecham um polígono, caso isso ocorra, abandonar a conexão;
- parar quando todos os nós estiverem conectados.

O comprimento da MST é dado pela soma das distâncias nas arestas.

A partir da matriz de distâncias d é possível construir o grafo que conecta todos os ativos com base na sua medida de dissimilaridade (equação (3.3)). As conexões formadas pela MST são equivalentes àquelas formadas pelo método aglomerativo hierárquico de ligação simples (GOWER; ROSS, 1969).

O agupamento formado pode ser representado graficamente em um dendograma. Um dendograma é um diagrama de árvore que exhibe os grupos formados por agrupamento de observações em cada passo e em seus níveis de fusão. Neste trabalho, o nível de similaridade para fusão de observações é medido ao longo do eixo vertical e as diferentes observações são listadas ao longo do eixo horizontal.

Assim como apresentado por (ONNELA et al., 2003a) selecionam-se duas medidas para estudar a evolução temporal da Árvore Geradora Mínima em termos de tamanho e topologia. Para tamanho, o comprimento normalizado da MST, dado pela equação (3.4), e para a topologia, a taxa de sobrevivência de passo simples, dada pela equação (3.5), respectivamente:

$$L(t) = \frac{1}{N-1} \sum_{d_{ij} \in T^t} d_{ij}, \quad (3.4)$$

$$\sigma_t = \frac{1}{N-1} |E^t \cap E^{t-1}|, \quad (3.5)$$

em que $t, t = 1, \dots, M$, é o número de janelas de comprimento T correspondente ao número de retornos diários incluídos na mesma, de modo que as janelas consecutivas apresentam uma defasagem δT . E^t corresponde ao conjunto de arestas presentes na árvore no tempo t e \cap é o operador de interseção de conjuntos.

ONNELA et al. (2003b) utiliza a medida da equação (3.6) para obter a média do coeficiente de correlação para diferentes períodos de tempo, em que consideram-se apenas os elementos não diagonais ($i \neq j$) de uma matriz triangular superior (ou inferior).

$$\bar{\rho}(t) = \frac{1}{N(N-1)/2} \sum_{\rho_{ij} \in C^t} \rho_{ij}. \quad (3.6)$$

Na próxima seção serão aplicadas às ações componentes do IBOVESPA, o conjunto de ferramentas apresentadas até aqui. Os resultados também serão discutidos na própria seção.

4 Resultados e Discussão

Após a obtenção da base de dados, houve um interesse na informação dos Setores de Atuação dos ativos, pois é importante para interpretação das conexões. Portanto, foi gerada a Figura 1, na qual observa-se que a maioria dos ativos pertencem aos Setores de Bancos, Consumo não Cíclico e Energia Elétrica.

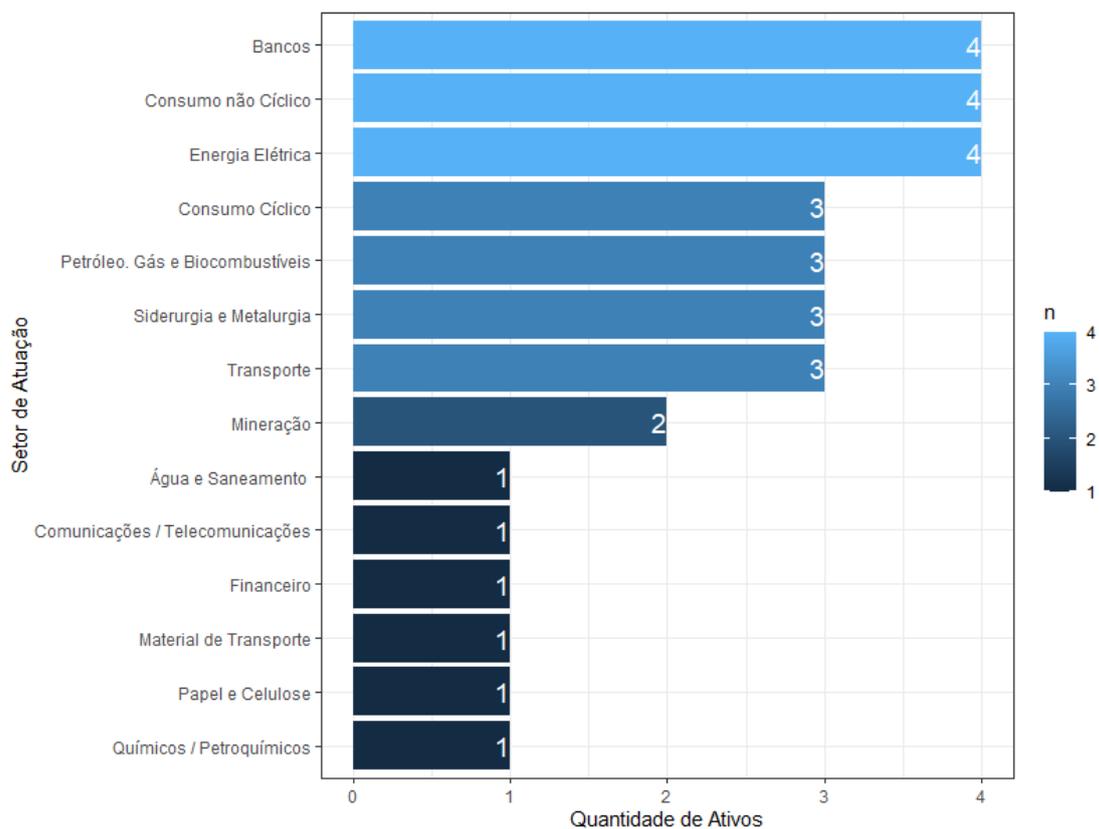


Figura 1: Composição do Índice por Setor de Atuação no período de 2008 a 2021.

Logo em seguida, calculou-se o coeficiente de correlação ρ_{ij} , para avaliar a similaridade entre os ativos do Índice BOVESPA. Essa relação é expressa pela equação (3.1), em que Y_i é o log-retorno do ativo i , conforme indicado na equação (3.2).

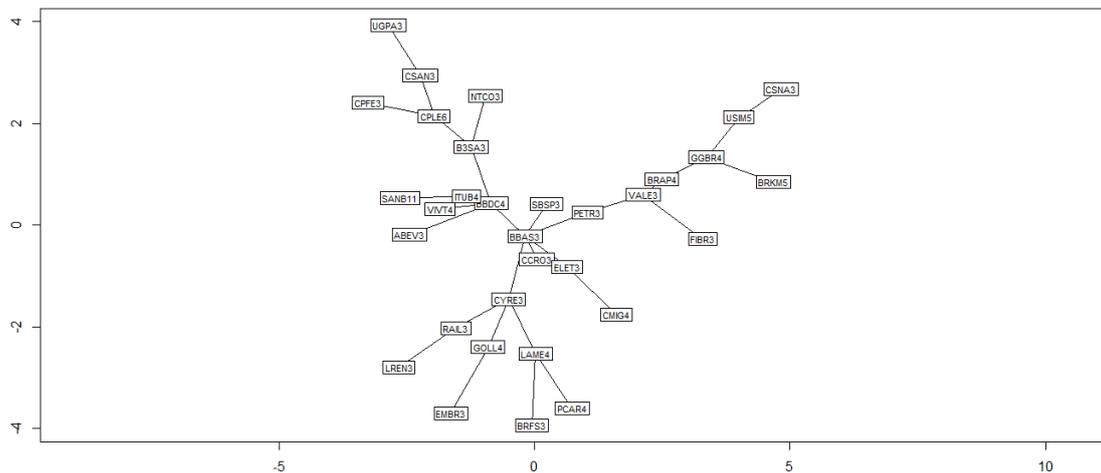


Figura 2: Árvore de ligação mínima conectando as 32 ações do IBOVESPA no período de 2008 a 2021.

Para a obtenção da matriz de distâncias, utilizou-se a relação dada na equação (3.3). A partir dessa matriz, obteve-se a MST conforme a Figura 2. Este grafo representa graficamente a ligação entre os ativos de maior similaridade. A Figura 3 foi obtida de (Silva, D. J & Oliveira, R. P. B., 2020) na qual estudou-se o Índice Bovespa no período de 2008 a 2015. Notou-se que ocorreu uma mudança na topologia da árvore, essa mudança pode ter ocorrido devido a pandemia de SARS-CoV-2.

Em seguida, construiu-se o dendograma através do método de agrupamento hierárquico via ligação simples, o qual é apresentado na Figura 4. O dendograma permite visualizar como os agrupamentos são formados em cada passo e avaliar os níveis de fusão dos agrupamentos formados. O nível de similaridade corresponde ao eixo vertical e as diferentes ações ao eixo horizontal.

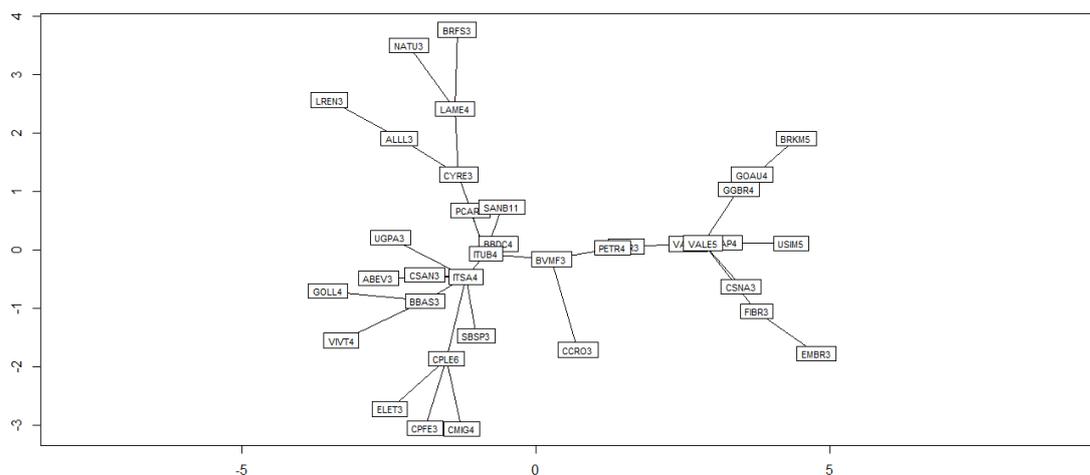


Figura 3: Árvore de ligação mínima conectando as 32 ações do IBOVESPA no período de 2008 a 2015.

Através do dendograma obtido na Figura 4 observou-se que os ativos de empresas que pertencem a grupos econômicos similares apresentam menores distâncias, pois são mais semelhantes, logo essas conexões ocorrem primeiro, como observado nas ligações entre BRAP4(BRADESCO S.A.) e VALE3(VALE S.A.) ambas são do setor de Mineração; BBDC4(BCO BRADESCO S.A.) e ITUB4(ITAU UNIBANCO HOLDING S.A.) que pertencem ao setor de Bancos. Esse tipo de agrupamento de empresas de similaridade setorial é condizente com os achados de MANTEGNA (1999). Se o agrupamento ocorre para valores pequenos de " d " (equação 3.3), isso implica que o preço dos ativos do setor varia de maneira muito similar diante de fatores econômicos, além de flutuarem de maneira similar quando fatores peculiares ao setor sofrem variações. Na MST da Figura 2, esse agrupamento para menores valores de " d " (equação 3.3) se traduz em menores comprimentos de arestas entre ativos. Assim, para aquelas conexões com arestas de maior tamanho o preço do ativo está mais ligado à peculiaridades da empresa e menos do setor, ou seja, atuam ali mais flutuações de natureza econômicas que setoriais.

Ainda com respeito ao dendograma na Figura 4, observam-se as ligações entre ELET3 (CENTRAIS ELET BRAS S.A. - ELETROBRAS) e CMIG4 (CIA ENERGETICA DE MINAS GERAIS - CEMIG) e posteriormente se ligam à CPFE3

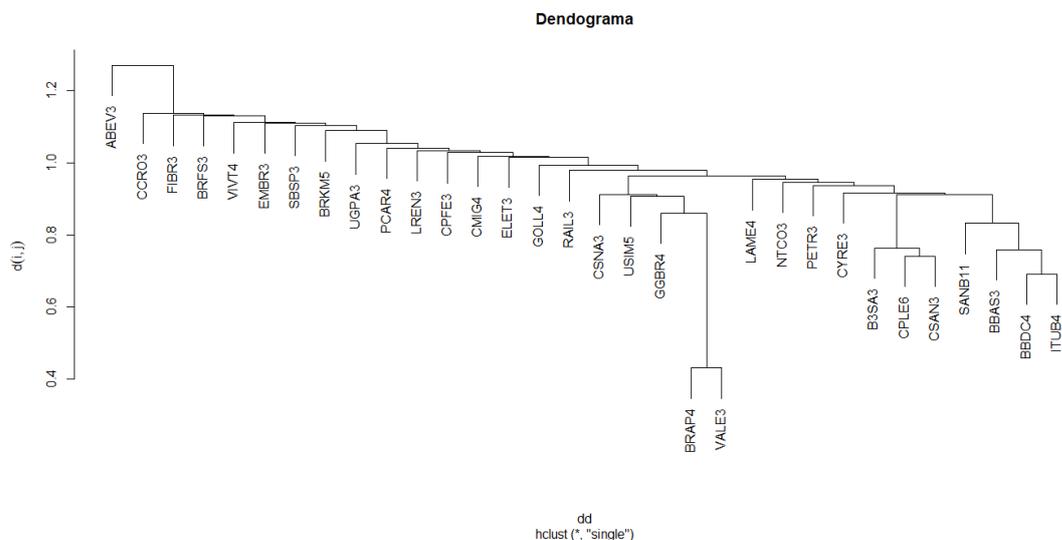


Figura 4: Dendrograma via ligação simples das ações do IBOVESPA, as empresas que pertencem a grupos econômicos semelhantes possuem comportamentos homogêneos, com ligações mais curtas, como a BRAP4 - VALE3(Mineração); BBDC4 - ITUB4(Bancos). Entre companhias com maior heterogeneidade, as ligações acontecem tardiamente e a distância entre elas é maior como a ligação entre o grupo BRAP4 - VALE3 com GGBR4.

(CPFL ENERGIA S.A.), todas essas empresas pertencem ao setor de Energia Elétrica. Ressalta-se a conexão entre VALE3(VALE S.A.) e BRAP4(BRADESPAR S.A.) e posteriormente se ligam à GGBR4(GERDAU S.A.), essas empresas são do setor de Mineração e Siderurgia e Metalurgia. Aponta-se também a conexão entre BBDC4(BCO BRADESCO S.A.) e ITUB4 (ITAU UNIBANCO HOLDING S.A.) e posteriormente se conecta à BBAS3(BCO BRASIL S.A.), que pertencem ao setor de Bancos. Os agrupamentos formados e a força de interação econômica entre ativos ficam bastante destacados na topologia da MST na Figura 2. Pode-se observar o agrupamento entre as ações LAME4(LOJAS AMERICANAS S.A.), BRFS3(BRF S.A.), PCAR4(CIA BRASILEIRA DE DISTRIBUICAO), CYRE3(CYRELA BRAZIL REALTY S.A.EMPREENDE E PART), GOLL4 (GOL LINHAS AEREAS INTELIGENTES S.A.), EMBR3(EMBRAER S.A.), RAIL3(RUMO S.A.), e LREN3(LOJAS RENNER S.A.) que são das áreas

de comércio, construção civil, alimentos, transporte aéreo e ferroviário e material de transporte. As ações PETR3 (PETROLEO BRASILEIRO S.A. PETROBRAS), VALE3 (VALE S.A.), BRAP4 (BRADESPAR S.A.), GGBR4 (GERDAU S.A.), BRKM5 (BRASKEM S.A.), USIM5 (USINAS SID DE MINAS GERAIS S.A. - USIMINAS), CSNA3 (CIA SIDERURGICA NACIONAL), FIBR3 (FIBRIA CELULOSE S.A.) pertencem as áreas de mineração, siderurgia, petroquímica, celulose, e indústria aérea. Observa-se o agrupamento entre as ações CPLE6(CIA PARANAENSE DE ENERGIA - COPEL), CPFE3(CPFL ENERGIA S.A.), CSAN3 (COSAN S.A.), e UGPA3(ULTRAPAR PARTICIPAÇÕES S.A.) que pertencem as áreas de Energia Elétrica e Petróleo, Gás e Biocombustíveis. Nota-se também as ações SANB11(BCO SANTANDER (BRASIL) S.A), BBDC4(BCO BRADESCO S.A.), B3SA3(B3 S.A. - BRASIL, BOLSA, BALCÃO), ITUB4(ITAUNIBANCO HOLDING S.A.), VIVT4(TELEFÔNICA BRASIL S.A) e BBAS3(BCO BRASIL S.A.) que são das áreas do setor financeiro, bancos e telecomunicações. Destaca-se ainda o agrupamento entre as ações ELET3(CENTRAIS ELET BRAS S.A. - ELETROBRAS), e CMIG4(CIA ENERGETICA DE MINAS GERAIS - CEMIG) que são da área de energia elétrica.

Para reforçar a apreciação da força de interação entre ativos construiu-se o correlograma das 32 ações do Índice na Figura 5. O Correlograma permite a visualização gráfica e melhor entendimento da relação dessas ações. Se a correlação é forte ou fraca, positiva ou negativa e até mesmo se não há correlação. Observou-se que todas as correlações foram positivas, o que indica que a variação na cotação de uma das ações acompanha a cotação da outra. O fato de haver somente correlações positivas não é muito bom, pois sempre uma ação estará acompanhando a outra, o que deixa o mercado previsível.

As afinidades econômicas e peculiaridades setoriais não são os únicos fatores responsáveis pela variação no valor dos ativos, as condições políticas e sociais também afetam o mercado financeiro. A instabilidade política, por exemplo, gera incertezas na economia a longo prazo, como desvalorização cambial e aumento na

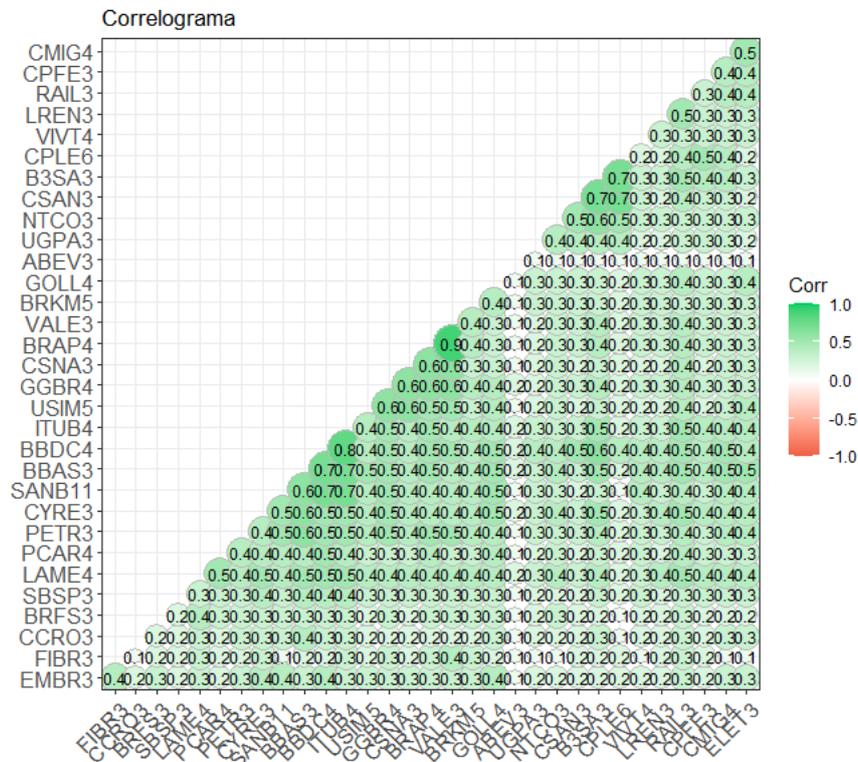


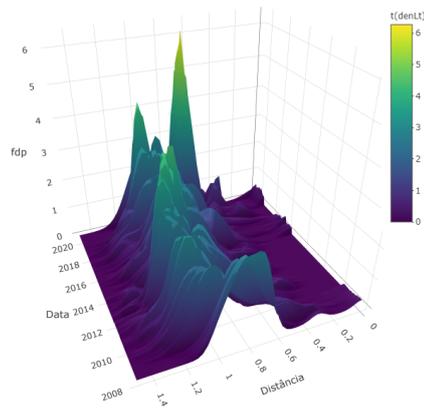
Figura 5: Correlograma das 32 ações do IBOVESPA

classificação de risco de calote. Assim como a ausência de previsão de crescimento do setor industrial tende a reduzir investimentos, diminuir a produção e reduzir a oferta de empregos. A redução no lucro afeta negativamente o valor das ações e o acontece um desequilíbrio da oferta com respeito à demanda.

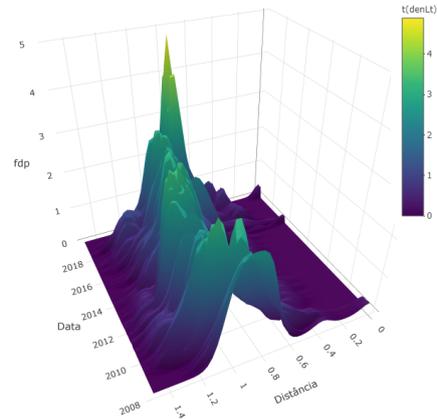
A tendência de queda cria uma espécie de "efeito manada" no mercado financeiro, o comportamento de grupos específicos afetam as decisões individuais. A redução do valor de ativos de um grupo pode simplesmente ser justificada pela decisão dos acionistas em se desfazerem de suas cotas. No entanto, a oferta crescente de um ativo produz a sensação de risco de perda de capital, e a tendência de venda aumenta. O mesmo ocorre com tendências de alta. Esse fenômeno foi explicado pelo Nobel de economia Richard Thaler, como decorrente do fator "humano" na

economia. de acordo com Thaler, ações humanas são imprevisíveis e na maioria dos casos são as que definem o aumento da compra ou venda de ativos.

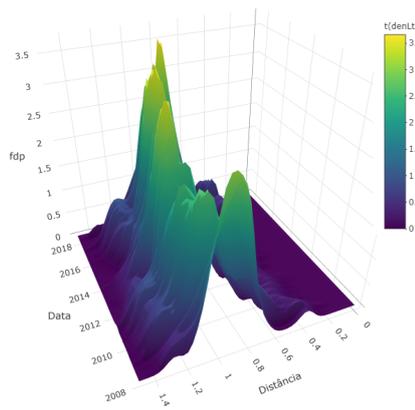
ONNELA et al. (2003b) constatou que os períodos de crise são governados por esse "efeito manada", de modo que essa característica se reflete nas correlações entre ativos e na topologia da MST. Por isso, as medidas de resumo sobre a MST podem fornecer subsídios para detecção estabilidade ou crise financeira. Por esse motivo, buscou-se estudar, através das medidas nas equações (3.4), (3.5) e (3.6), a dinâmica do mercado financeiro no decorrer de diferentes períodos históricos.



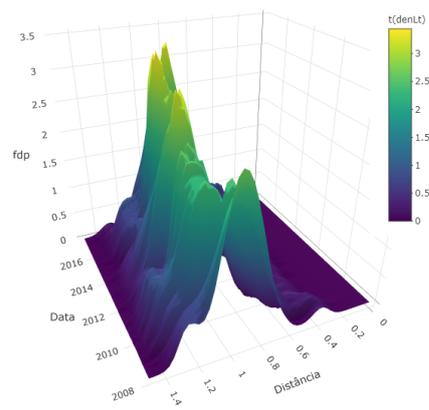
(a) Janela de 252 dias.



(b) Janela de 504 dias.



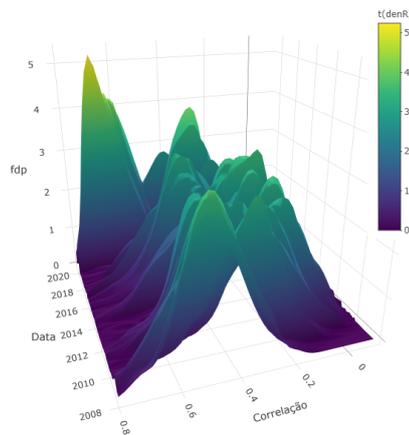
(c) Janela de 756 dias.



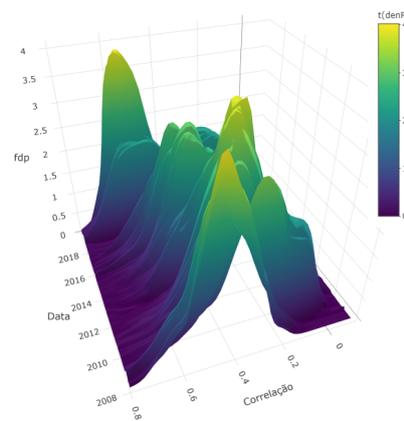
(d) Janela de 1000 dias.

Figura 6: Gráficos das funções densidades de probabilidade das distâncias, em função do tempo, para diferentes janelas de tempo.

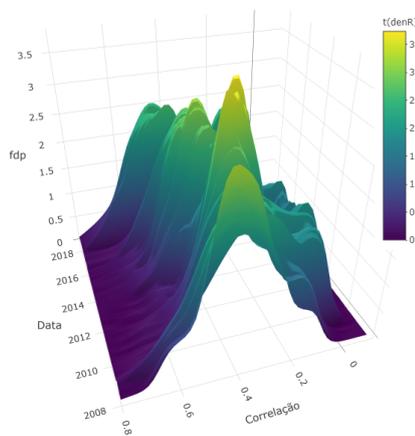
No artigo de ONNELA et al. (2003b) foi utilizado uma janela de tempo de 1000 dias para todos os cálculos. Para confirmar que a escolha de 1000 dias poderia ser transportada para a bolsa brasileira, neste trabalho foi conduzida uma análise das Figuras 6 e 7 para decidir o tamanho da janela de tempo. A Figura 6 apresenta os gráficos das funções densidades de probabilidade das distâncias, e a Figura 7 apresenta os gráficos das funções densidades de probabilidade dos coeficientes de correlação, ambas em função do tempo para diferentes janelas de tempo.



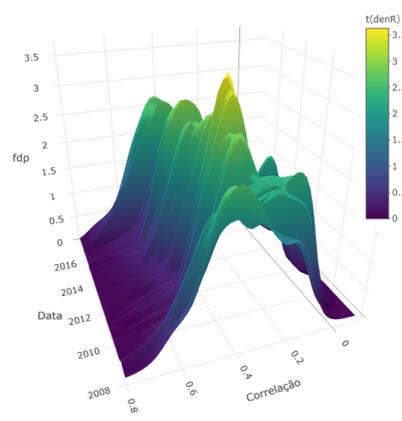
(a) Janela de 252 dias.



(b) Janela de 504 dias.



(c) Janela de 756 dias.

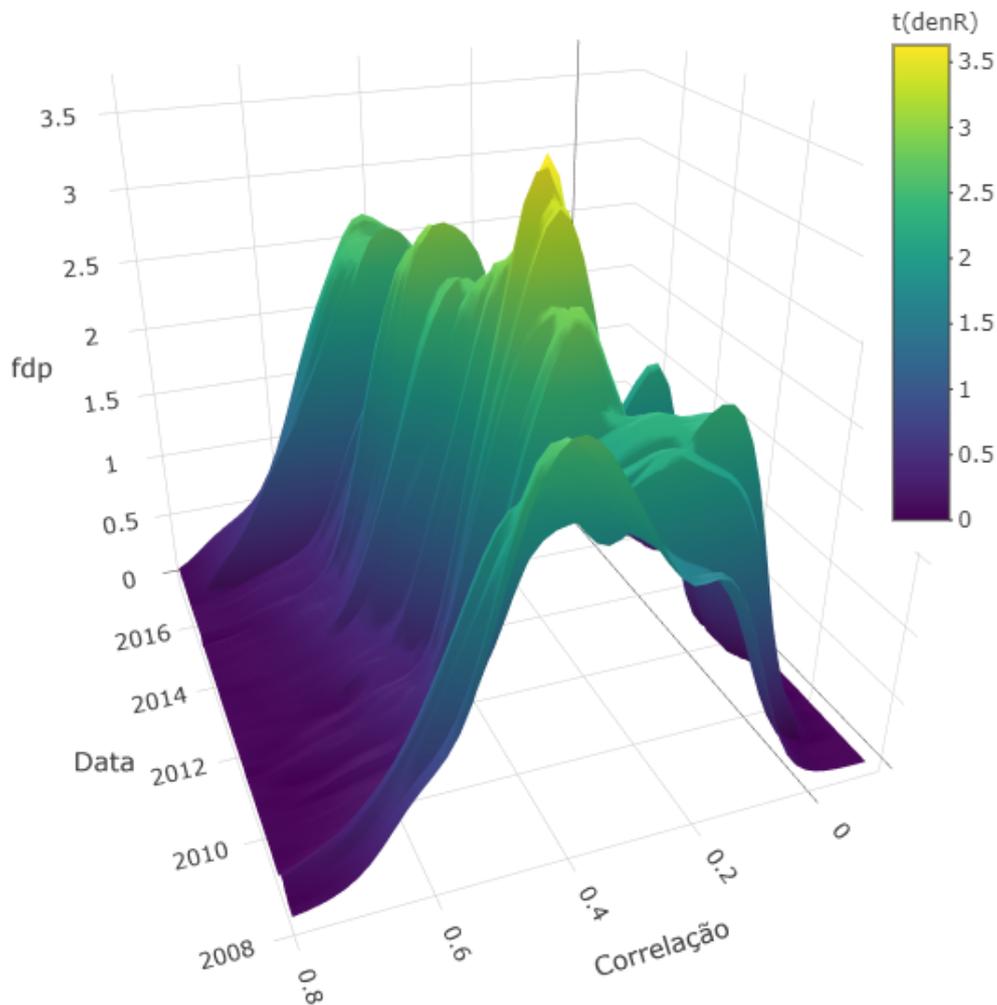


(d) Janela de 1000 dias.

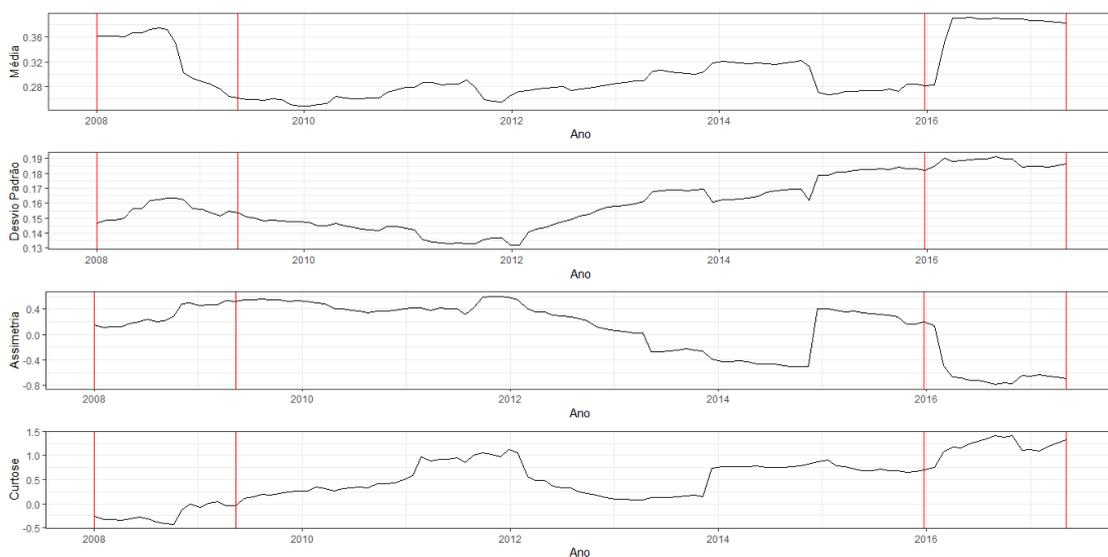
Figura 7: Gráficos das funções densidades de probabilidade dos coeficientes de correlação, em função do tempo, para diferentes janelas de tempo.

Através da análise das Figuras 6 e 7 nota-se que uma janela de tempo pequena, como nas Figuras 6(a) e 7(a), as densidades ficam muito sensíveis a mudanças instantâneas do mercado, o que pode levar a muitos alarmes falsos sobre uma possível crise. Por esse motivo optou-se pelo uso da janela de tempo de 1000 dias, pois as densidades ficam mais estáveis e não levam a muitos alarmes falsos sobre uma possível crise.

Através da Figura 8(a), construída a partir da matriz de correlações obtida da equação (3.1), percebe-se que no ano de 2008 a densidade da correlação é menor se comparada aos outros anos, já sua dispersão é maior. Agora, para a Figura 8(b), obtida pela medida na equação (3.6), percebe-se que no ano de 2008 e 2016 a média é maior e a assimetria é menor quando comparados a anos posteriores. Todas essas variações no ano de 2008 e 2016 indicam uma possível crise nestes anos, o que de fato ocorreu. A crise financeira de 2008 afetou todo o mundo, e ocorreu devido a uma bolha imobiliária nos Estados Unidos, causada por papéis imobiliários de alto risco, que foram criminosamente classificados como de baixo risco. Em 2016, houve uma crise político econômica que findou no impeachment da Presidente Dilma Rousseff. Observe ainda que, como as janelas são de 1000 dias, a crise de 2016 teve efeitos que se prorrogaram com a chegada da pandemia de SARS-CoV-2. As correlações perduraram em altos níveis, com maior variabilidade que períodos de tranquilidade político econômica.



(a)



(b)

Figura 8: (a): Gráfico da função densidade de probabilidade dos coeficientes de correlação em função do tempo, (b): A média, desvio padrão, assimetria e curtose dos coeficientes de correlação em função do tempo.

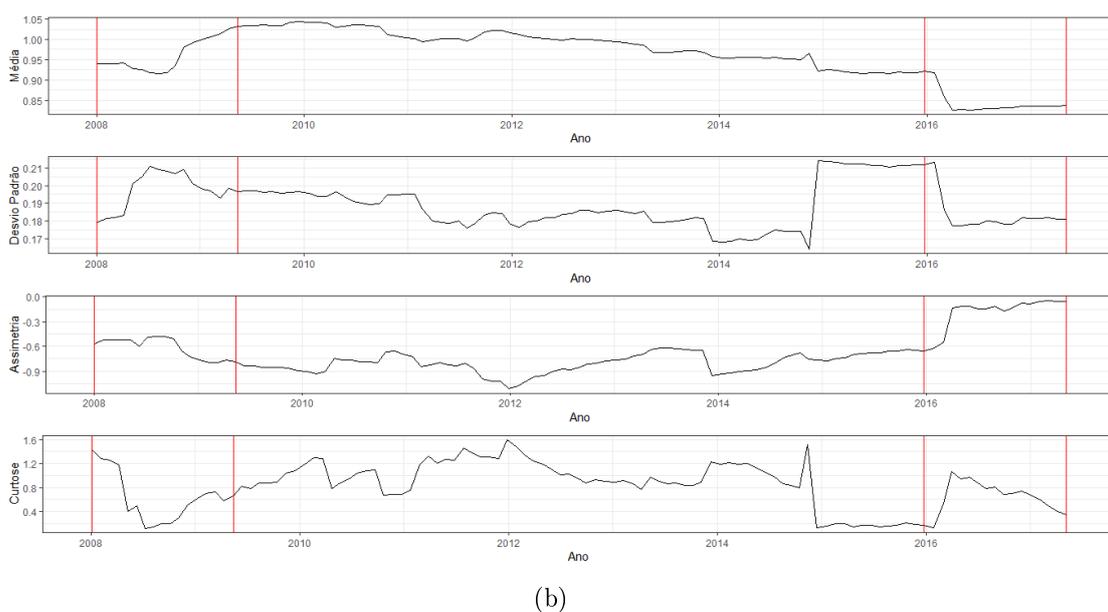
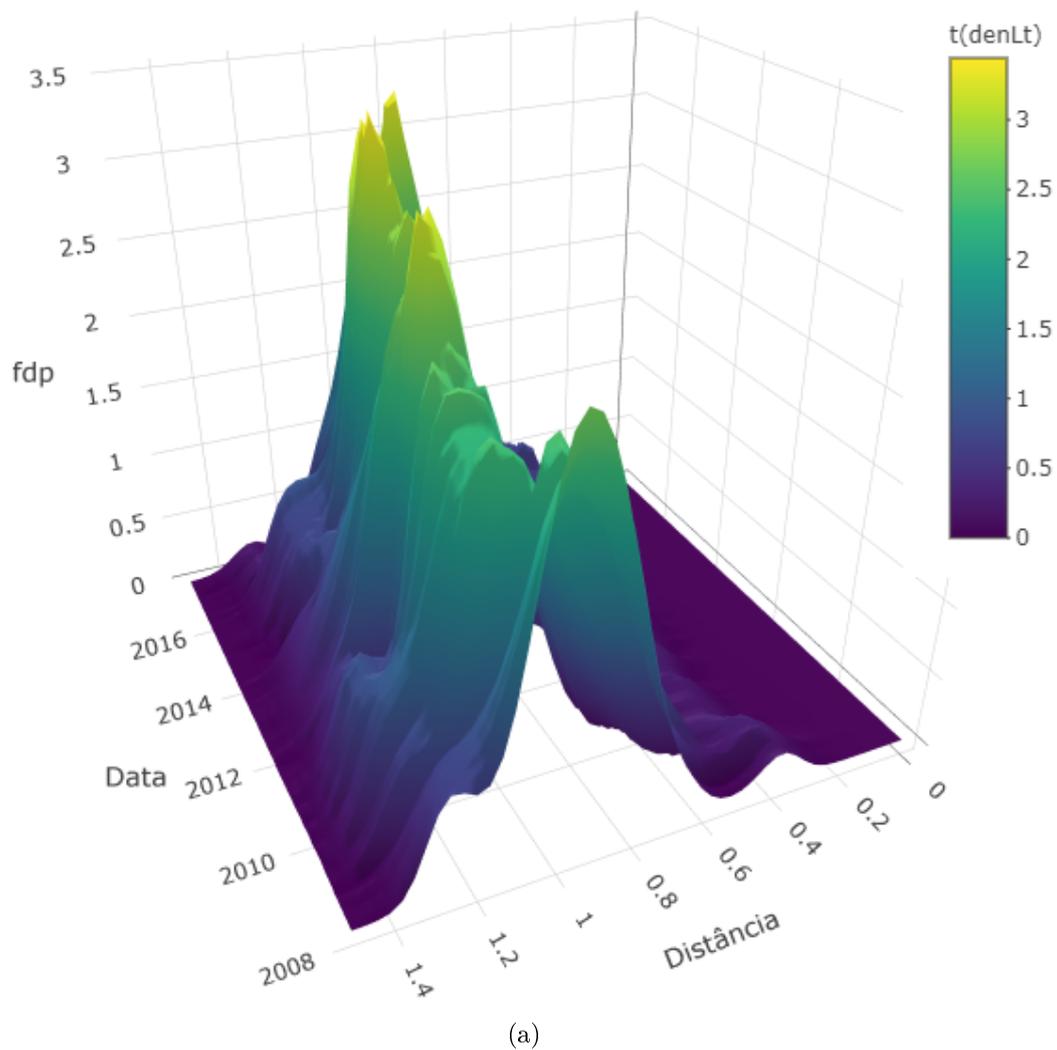
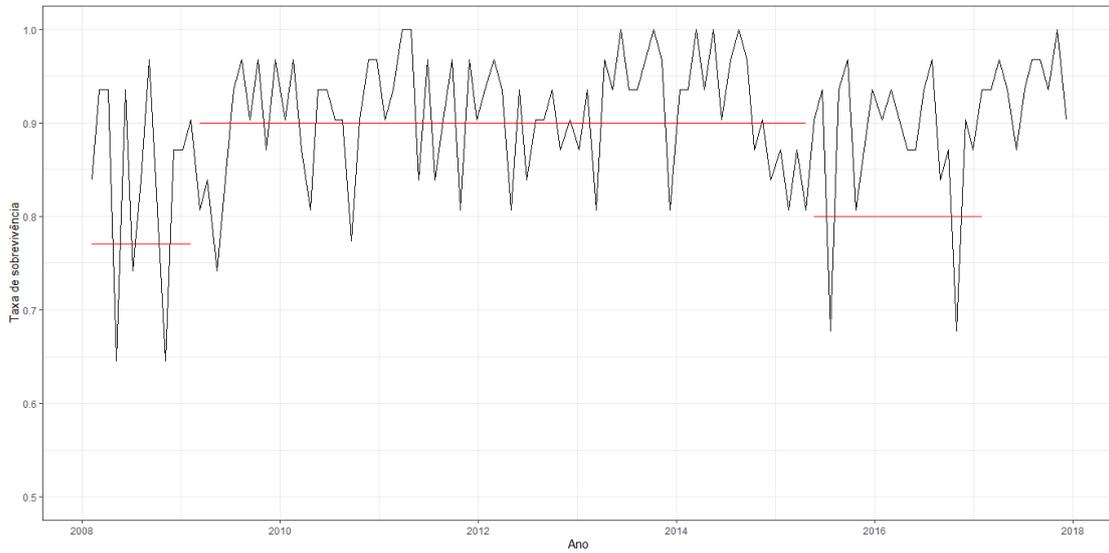


Figura 9: (a): Gráfico da função densidade de probabilidade das distâncias em função do tempo, (b): A média, desvio padrão, assimetria e curtose das distâncias em função do tempo.

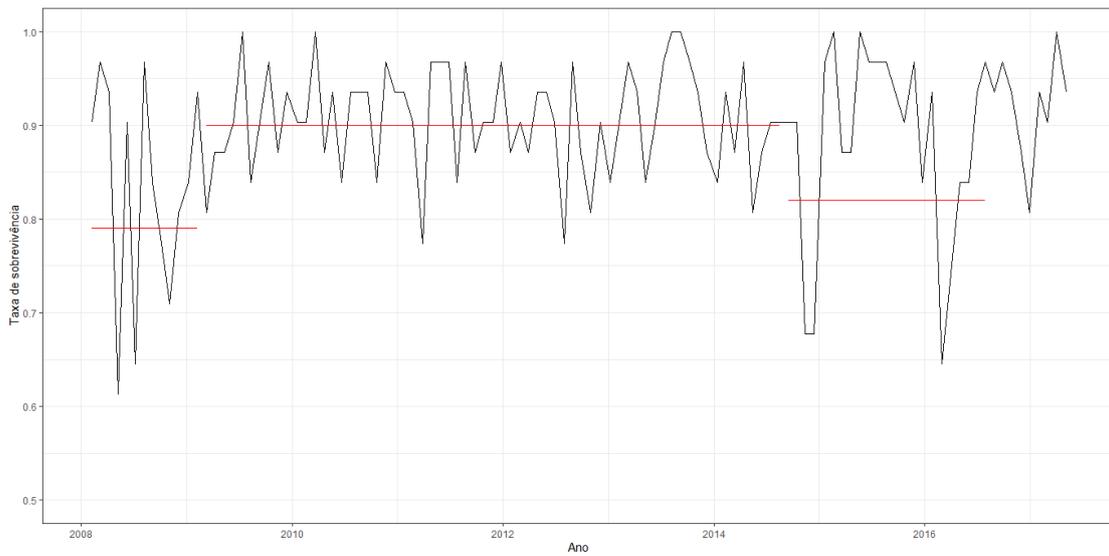
Em seguida, construíram-se os gráficos da Figura 9(a) e Figura 9(b), obtidas através das medidas de topologia da MST nas equações (3.3) e (3.4), que também permitem visualizar algum indicativo mensurável de que o mercado financeiro possa estar em crise.

Com base na Figura 9(a) percebe-se que nos anos de 2008 e 2016 a densidade da distância é menor se comparada à dos outros anos, e que a média é menor que nos outros anos, ou seja, em tempos de crise as distâncias da MST diminuem, e conseqüentemente o tamanho da MST é menor. Agora, para a Figura 9(b) percebe-se que nos anos de 2008 e 2016 o comportamento da média e curtose é menor, já a assimetria é maior em comparação a anos posteriores. Todas essas variações nos anos de 2008 e 2016 conseguem capturar o estado de crise do mercado. Nota-se que o comportamento da média e da assimetria na Figura 8(b) é o oposto ao comportamento da média e da assimetria na Figura 9(b). Posteriormente construiu-se os gráficos da Figura 10 obtidas através das medidas de topologia da MST na equação (3.5), a qual permite visualizar o momento em que o mercado entra e sai da crise, o que a torna uma medida interessante para detecção de uma crise.

Por meio da Figura 10(a) percebe-se que em 2008 e 2016 a taxa de sobrevivência de passo unitário é mais baixa, o que indica uma crise no mercado financeiro. No final de 2008 e 2016 há novamente uma queda da taxa de sobrevivência, e nos anos posteriores a taxa de sobrevivência aumenta, sinal do fim da crise no mercado financeiro. Em análise da Figura 10(b) percebe-se também uma queda da taxa de sobrevivência em 2008 e 2016 e nos anos posteriores um aumento da taxa de sobrevivência. Tal queda é mais perceptível na Figura 10(a) do que na Figura 10(b). Ademais, nota-se que com uma janela de tempo menor a taxa de sobrevivência de passo unitário é menor e sofre mais variação do que com uma janela de tempo maior.



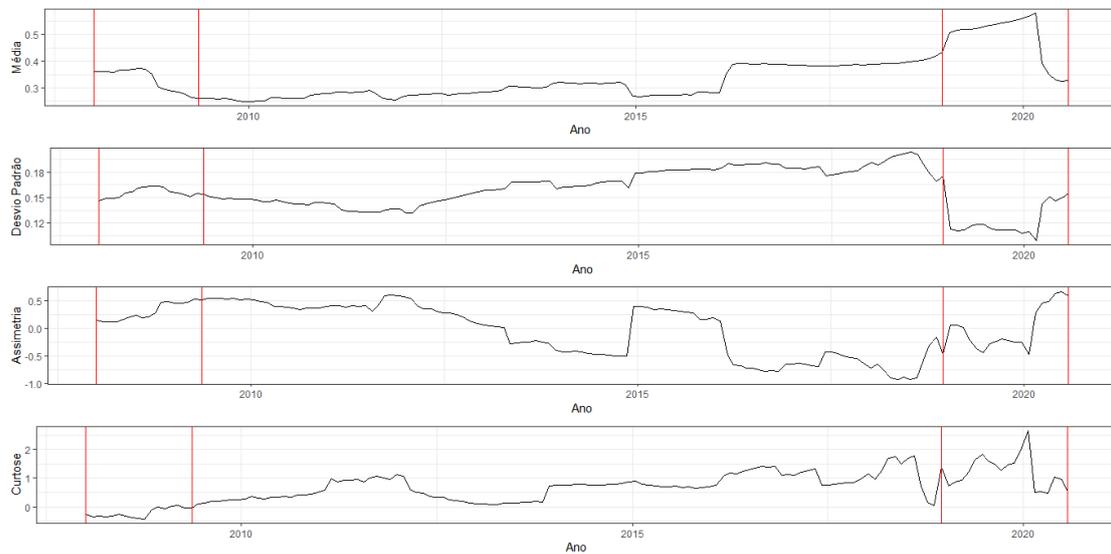
(a) Janela de 850 dias.



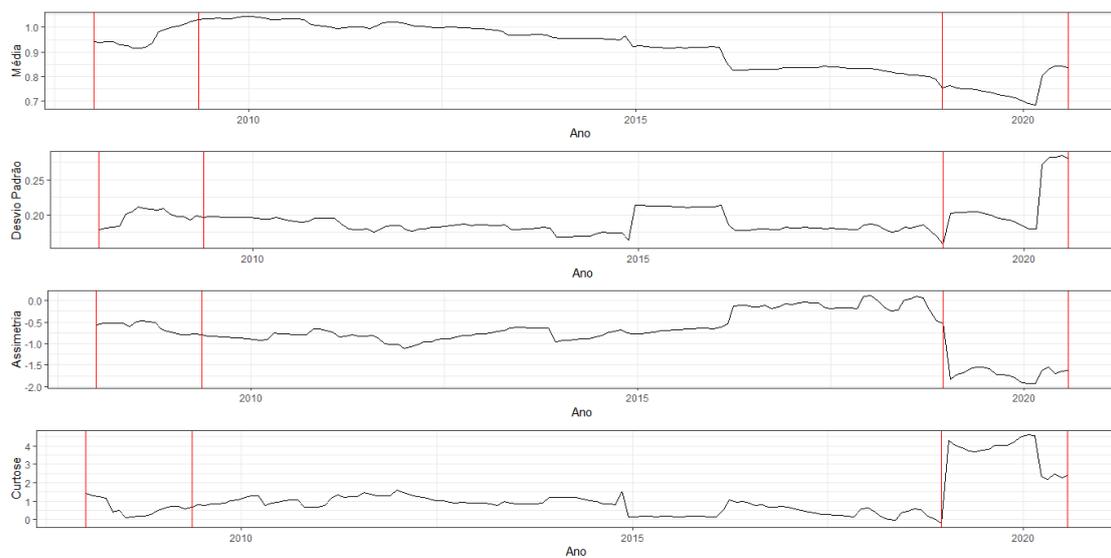
(b) Janela de 1000 dias.

Figura 10: Taxa de sobrevivência de passo unitário em função do tempo.

Através dos gráficos 9(b) e 8(b), notou-se que o eixo de datas terminava no ano de 2018 devido ao fato da janela ser de 1000 dias. De forma investigativa, implementou-se uma função para obter os dados nos gráficos até o ano de 2021 apenas como exercício de se observar os efeitos da pandemia de SARS-CoV-2 no mercado financeiro em janelas de tempo que se reduzem quando 1000 dias não estão mais disponíveis. Obteve-se as figuras 11(b) e 11(a), é possível observar com mais clareza as informações, é nítido como a crise de mercado, devido ao impacto da pandemia de SARS-CoV-2, é capturada pelas medidas utilizadas neste texto.



(a) Estatísticas dos coeficientes de correlação.



(b) Estatísticas das distâncias.

Figura 11: A média, desvio padrão, assimetria e curtose em função do tempo.

5 Considerações Finais

Neste trabalho foi possível observar a correlação entre os valores dos ativos por meio da semelhança de comportamento de seus preços frente ao aumento ou redução dos valores dos mesmos. Ativos similares guardam maior proximidade na estrutura da Árvore Geradora Mínima. Essas ligações, na grande maioria dos casos, se dão entre ativos de setores similares da economia, com destaque para os setores bancário, petrolífero, água e saneamento, e exploração de rodovias os quais ocupam posições mais centrais no grafo. A periferia do grafo é ocupada por empresas mais independentes da flutuação dos ativos do centro.

Constatou-se também que, quando o mercado financeiro está em crise, a função densidade de probabilidade das distâncias e dos coeficientes de correlação se comportam de maneira diferente de quando o mercado financeiro não está em crise. No caso da função densidade de probabilidade dos coeficientes de correlação, quando o mercado está em crise a média e o desvio padrão são maiores, e a assimetria e curtose são menores. Na função densidade de probabilidade das distâncias, a média e a curtose são menores, e a assimetria é maior, do que quando o mercado não está em crise. Assim, um mercado em crise também se reflete nessas medidas.

Outro apontamento importante diz respeito ao tamanho médio normalizado das Árvores Geradora Mínima. Nota-se que, quando o mercado financeiro está em crise, esse tamanho é menor.

Através da análise da taxa de sobrevivência de passo unitário conclui-se que a mesma também é capaz de sinalizar o momento em que o mercado financeiro entra e sai da crise. Uma janela de tempo maior para a construção do gráfico gera mais estabilidade para essa medida.

Sobre as medidas associadas à MST, os achados desse trabalho não são diferentes daqueles encontrados nos trabalhos de MANTEGNA (1999) e de ONNELA et al. (2003a, 2003b). Isso prova que a ferramenta é útil para capturar o comportamento global do mercado financeiro. No contexto brasileiro, as ferramentas apresentadas por estes autores permitiram revelar algumas estruturas interessantes para o melhor entendimento do mercado. Tais estruturas podem subsidiar gestores a tomarem decisões com base em medidas objetivas.

Referências

B3, B. B. B. *índices amplos da B3*. Fevereiro de 2020. https://www.b3.com.br/pt_br/market-data-e-indices/indices/indices-amplos/ibovespa.htm.

BOR, O. et al. O jistém problému minimálním (about a certain minimal problem). *Práce Mor. Přírodoved. Spol. v Brne III*, v. 3, 1926.

EULER, L. Solutio problematis ad geometriam situs pertinens. *Comm. Acad. Sci. Imper. Petropol.*, v. 8, p. 128–140, 1736.

GOWER, J. C.; ROSS, G. J. Minimum spanning trees and single linkage cluster analysis. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, Wiley Online Library, v. 18, n. 1, p. 54–64, 1969.

INVESTING.COM. *Investing*. Junho de 2021. <https://br.investing.com/>.

MANTEGNA, R. N. Hierarchical structure in financial markets. *The European Physical Journal B - Condensed Matter and Complex Systems*, v. 11, n. 1, p. 193–197, 1999. ISSN 1434-6036. Disponível em: <<http://dx.doi.org/10.1007/s100510050929>>.

ONNELA, J.-P. et al. Dynamic asset trees and black monday. *Physica A: Statistical Mechanics and its Applications*, v. 324, n. 1–2, p. 247 – 252, 2003. ISSN 0378-4371. Proceedings of the International Econophysics Conference. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0378437102018824>>.

ONNELA, J.-P. et al. Dynamics of market correlations: Taxonomy and portfolio analysis. *PHYSICAL REVIEW E*, v. 68, p. 056110, 2003. ISSN 1539-3755.

R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2020. Disponível em: <<https://www.R-project.org/>>.

SAMUELSON, P. A. Proof that properly anticipated prices fluctuate randomly. *Sloan Management Review*, Sloan Management Review Association., v. 6, n. 2, p. 41, 1965.

Silva, D. J & Oliveira, R. P. B . *Uma abordagem pela teoria de grafos para descrever similaridades de ativos do Índice BOVESPA. Relatório de Iniciação Científica*. Minas, Brasil, 2020.

YAHOO, F. *Yahoo*. Junho de 2021. <https://br.financas.yahoo.com/>.

APÊNDICE A – Código R

Listing A.1: Script

```
1
2 rm(list = ls())
3
4 #instalando pacotes necessarios
5
6 pkgTest <- function(x)
7 {
8   if (!require(x,character.only = TRUE))
9     {
10      install.packages(x,dep=TRUE)
11      if(!require(x,character.only = TRUE)) stop('Package not
12        found')
13    }
14 }
15 pkgTest('vegan')
16 pkgTest('e1071')
17 pkgTest('plotly')
18 pkgTest('ggplot2')
19 pkgTest('dplyr')
20 pkgTest('tidyr')
21 pkgTest('ggcorrplot')
22 pkgTest('lubridate')
```

```
22 pkgTest('grid')
23 pkgTest('corrplot')
24 pkgTest('adespatial')
25 pkgTest('spdep')
26 pkgTest('igraph')
27 require(corrplot)
28 require(grid)
29 require(ggcorrplot)
30 require(dplyr)
31 require(tidyr)
32 require(vegan)
33 require(e1071)
34 require(ggplot2)
35 require(plotly)
36 require(lubridate)
37 require(adespatial)
38 require(spdep)
39 require(igraph)
40
41 setwd('C:/Users/daian/Downloads')
42
43
44 #####PARA CALCULAR LOGRETORNOS#####
45
46 dad=read.csv('teste_monografia_2008_2021.csv', header = T,
47             sep = ';', dec = '.') #lendo a base de dados
48
49 dad$RptDt<-dmy(dad$RptDt) #transfromando a coluna de datas em
50             um objeto Date
51
52 dad2=data.frame(dad %>%
53                 select(RptDt, TckrSymb, LastPric) %>%
54                 spread(TckrSymb, LastPric))
```

```
52 #construindo uma data frame com o preco do ativo onde cada
    coluna representa uma acao e cada linha uma data
53
54 row.names(dad2)<-dad2$RptDt #nomeando as linhas do data frame
    com as datas
55 dad2$RptDt=NULL #retirando a coluna de datas
56 fix(dad2)
57
58 dia=data.frame(row.names(dad2)) #armazenando as datas
59 dia=dia[-1,] #retirando a primeira data, pois ao calcular o
    log-retorno teremos
60 #uma data a menos
61 dia=ymd(dia) #transformando em objeto Date
62 head(dia)
63
64
65 lrts<-matrix(data=NA,nrow=(dim(dad2)[1]-1),ncol=dim(dad2)[2])
    #cria matriz pra receber
66 #os logretornos totais simples
67
68 #logrts:
69 #Funcao para calcular os log-retornos
70 #Argumentos: x, representa uma coluna de um data frame com os
    precos dos ativos
71 #Retorna: vetor com os log-retornos
72 logrts<-function(x){
73   logrx<-vector()
74   for(i in 2:length(x)){
75     logrx[i-1]<-log(x[i])-log(x[i-1])}
76   logrx}
77
78 for(i in (1:dim(dad2)[2])){
```

```
79   lrts[,i]<-logrts(dad2[,i])#calcula o logretorno total
      simples por coluna
80
81 lrts<-as.data.frame(lrts)#muda a classe de vetor para data
      frame
82 names(lrts)<-names(dad2)#da nome as colunas do data frame
83 fix(lrts)
84
85
86
87 #####CALCULANDO A MATRIZ DE DISTANCIAS#####
88
89
90 rho<-cor(lrts,use = 'pairwise.complete.obs',method=c('pearson
      ')) #calcula a correlacao
91 fix(rho)
92 dados_d<-sqrt(2*(1-rho)) #tranforma a correlacao em uma
      metrica
93 dd <- as.dist(dados_d) #transforma a matriz de distancias num
      objeto 'distancia'
94 fix(dd)
95
96 #####CORRELOGRAMA#####
97
98 windows()
99 ggcorrplot(round(rho, 1), hc.order = TRUE,
100           type = 'lower',
101           lab = TRUE,
102           lab_size = 3,
103           method='circle',
104           colors = c('tomato2', 'white', 'springgreen3'),
105           title='Correlograma',
```

```
106         ggtheme=theme_bw)
107
108 corrplot(rho, order = 'hclust') #separado por grupos
109
110
111 #####ARVORE DE LIGACAO MINIMA#####
112
113 #mat<-graph_from_adjacency_matrix(dd, diag = F)
114 #g_mst<-mst(mat)
115 #plot(g_mst)
116
117 hc<-hclust(dd,method='single') #metodo de agrupamento
118 plot(hc,main='Dendograma',ylab=expression(d(i,j))) #
    construindo Dendograma
119
120 mst<-spantree(dd)
121 windows()
122 plot(mst, type = 't') #construindo a MST
123
124 #####DENSIDADE OBSERVADA#####
125
126 rhoD<-rho
127 diag(rhoD)=NA #retirando a diagonal
128 rhoD
129 hist(rhoD)
130
131 #####CALCULANDO AS MATRIZES DE CORRELACAO#####
132
133 #RHOT:
134 #Funcao para calcular a matriz de autocorrelacao para uma
    determinada janela de tempo T
135 #Argumentos: matriz de logretornos (colunas como ativos,
```

```

    linhas como logretornos),
136 #T janela de tempo em dias
137 #delta: saltos em dias (defasagem) para obter a matriz de
    logretornos
138 #data: ano de cada matriz de autocorrelacao
139 #Retorna: lista com matrizes de autocorrelacao com saltos de
    delta=21 dias, anos
140 #relacionados as matrizes de autocorrelacao
141 RHOT<-function(matriz,T,delta,data){
142     i=1; j=T; it=1; rho<-NULL
143     rho<-as.list(rho); anos<-NULL
144     while(j<=dim(matriz)[1]){
145         p=(i:j)
146         c<-cor(matriz[p,],use = 'pairwise.complete.obs',method=c(
            'pearson'))
147         c[lower.tri(c)] <- NA #transforma em matriz triangular
            superiores
148         diag(c)<-NA #retira a diagonal
149         rho[it]<-list(c)
150         anos[it]<-as.character(data[i])
151         it=it+1
152         i=i+delta
153         j=j+delta
154     }
155     rho[it]<-list(anos)
156     rho}
157
158 teste=RHOT(lrts,1000,21,dia) #invocando a funcao
159 fix(teste)
160
161 ##### GRAFICO 3D DAS CORRELACOES #####
162

```

```
163 #densRHO:
164 #Funcao para calcular a densidade das correlacoes para cada
    janela
165 #Argumentos: lista contendo matrizes de autocorrelacao,
    numero de pontos
166 #para calculo da densidade em cada data, limites do eixo de
    correlacoes
167 #que sera mostrado no grafico
168 #Retorna: matriz(cada linha corresponde a densidade daquela
    data)
169 densRHO<-function(lista,n,xlim){
170     densidade=matrix(NA,length(lista)-1,n)
171     for(i in 1:(length(lista)-1)){
172         dens<-density(as.matrix(as.data.frame(lista[i])),from=
            xlim[1],to=xlim[2],n=n,
173                 kernel='gaussian',na.rm = TRUE) #calcula a
                densidade das correlacoes
174         #da i-esima janela
175         densidade[i,<-dens$y #retorna a densidade das
            correlacoes da i-esima janela
176     }
177     densidade}
178
179 xlim=c(-0.1,0.8) # limites do eixo das correlacoes
180 n=50 # numero de pontos de correlacao para estimar a
    densidade das matrizes de correlacao
181
182 denR=densRHO(teste,n=n,xlim=xlim) #invocando a funcao
183
184 x=density(as.matrix(as.data.frame(teste[1])),from=xlim[1],to=
    xlim[2],n=n,
185         kernel='gaussian',na.rm = TRUE)$x #definindo os
```

```

                                marcadores do eixo das correlacoes
186 y=teste[length(teste)][[1]] #definindo os marcadores do eixo
    das datas
187
188 # grafico 3D das fdp's estimadas para as correlacoes
189 windows()
190 plot_ly() %>%
191     add_surface(x = ~y, y = ~x, z = ~t(denR)) %>%
192     layout(
193         scene = list(
194             xaxis = list(title = 'Data'),
195             yaxis = list(title = 'Correlacao'),
196             zaxis = list(title = 'fdp'),
197             aspectmode='cube' #aspecto mais amigavel para o
                                grafico
198             )#veja mais opcoes em https://plotly.com/r/3d-
                                surface-plots/
199     )
200
201 #####GRAFICOS DAS ESTATISTICAS DAS CORRELACOES
    #####
202
203 #est:
204 #Funcao para calcular a media, desvio padrao assimetria e
    cutose de uma lista
205 #Argumentos: lista contendo matrizes de autocorreacao
206 #Retorna: data frame com as variaveis(cada coluna uma
    variavel)
207 est<-function(lista){
208     media=vector()
209     dp=vector()
210     assimetria=vector()

```

```
211   curtose=vector()
212   for(i in 1:(length(lista)-1)){
213     media[i]<-mean(as.matrix(as.data.frame(lista[i])),na.rm =
214                   TRUE)
215     dp[i]<-sd(as.matrix(as.data.frame(lista[i])),na.rm = TRUE
216              )
217     assimetria[i]<-skewness(as.matrix(as.data.frame(lista[i])
218                              ),na.rm = TRUE)
219     curtose[i]<-kurtosis(as.matrix(as.data.frame(lista[i])),
220                          na.rm = TRUE)
221   }
222   dados=data.frame(media,dp,assimetria,curtose)
223   dados}
224
225 estatisticas=est(teste) #invocando a funcao
226 head(estatisticas)
227 da=as.data.frame(teste[length(teste)]) #recuperando as datas
228   calculadas acima
229 colnames(da)[1]<- 'dat' #nomeando a coluna
230
231 estatisticas=cbind(da,estatisticas) #criando um data frame
232 estatisticas$dat=as.Date(estatisticas$dat) #transformando a
233   data em um objeto Date
234
235 #grafico da media da correlacao
236 windows()
237 ggplot(estatisticas,aes(x=dat,y=media))+
238   geom_line(aes(x=dat,y=media))+
239   theme_bw() + labs(y='media', x='Ano')
240
241 #grafico do Desvio padrao da correlacao
242 windows()
```

```
237 ggplot(estatisticas , aes(x=dat , y=dp)) +
238   geom_line(aes(x=dat , y=dp)) +
239   theme_bw() + labs(y='Desvio padrao', x='Ano')
240
241 #grafico da Assimetria da correlacao
242 windows()
243 ggplot(estatisticas , aes(x=dat , y=assimetria)) +
244   geom_line(aes(x=dat , y=assimetria)) +
245   theme_bw() + labs(y='Assimetria', x='Ano')
246
247 #grafico da Curtose da correlacao
248 windows()
249 ggplot(estatisticas , aes(x=dat , y=curtose)) +
250   geom_line(aes(x=dat , y=curtose)) +
251   theme_bw() + labs(y='Curtose', x='Ano')
252
253 ##### GRAFICOS NA MESMA JANELA #####
254 windows()
255 pushViewport(viewport(layout = grid.layout(4, 1)))
256
257 print(ggplot(estatisticas , aes(x=dat , y=media)) +
258   geom_line(aes(x=dat , y=media)) +
259   theme_bw() + labs(y='media', x='Ano'),
260   vp = viewport(layout.pos.row = 1, layout.pos.col = 1))
261
262 print(ggplot(estatisticas , aes(x=dat , y=dp)) +
263   geom_line(aes(x=dat , y=dp)) +
264   theme_bw() + labs(y='Desvio padrao', x='Ano'),
265   vp = viewport(layout.pos.row = 2, layout.pos.col = 1))
266
267 print(ggplot(estatisticas , aes(x=dat , y=assimetria)) +
268   geom_line(aes(x=dat , y=assimetria)) +
```

```

269     theme_bw() + labs(y='Assimetria', x='Ano'),
270     vp = viewport(layout.pos.row = 3, layout.pos.col = 1))
271
272 print(ggplot(estatisticas, aes(x=dat, y=curtose))+
273       geom_line(aes(x=dat, y=curtose))+
274       theme_bw() + labs(y='Curtose', x='Ano'),
275       vp = viewport(layout.pos.row = 4, layout.pos.col = 1))
276
277 ##### GRAFICO 3D LT#####
278
279 #LISTANTL:
280 #funcao para calcular as distancias da MST para uma
      determinada janela de tempo T
281 #Argumentos: matriz de logretornos (colunas como ativos,
      linhas como logretornos),
282 #delta: saltos em dias (defasagem) para obter a matriz de
      logretornos
283 #data: ano de cada matriz de autocorrelacao
284 #T janela de tempo em dias
285 #Retorna: lista com as distancias da MST e a data(ultimo
      elemento da lista)
286 LISTANTL<-function(matriz,T,delta,data){
287   i=1; j=T; it=1; m<-NULL; m<-as.list(m); d<-ymd()
288   while(j<=dim(matriz)[1]){
289     p=(i:j)
290     c<-cor(matriz[p,],use = 'pairwise.complete.obs',method=c(
      'pearson'))
291     dados_d<-sqrt(2*(1-c))
292     dd <- as.dist(dados_d)
293     mst<-spantree(dd)
294     dmst<-mst$dist #armazena as distancia da mst
295     m[it]<-list(dmst)

```

```

296     d[it]<-as.character(data[i])
297     i=i+delta
298     j=j+delta
299     it=it+1
300   }
301   m[it]<-list(d)
302   m}
303
304 lista=LISTANTL(lrts,1000,21,dia) #invocando a funcao
305
306 xlim=c(0,1.5) #limites do eixo das distancias
307 n=50 #numero de pontos de distancias para estimar a densidade
      das distancias
308
309 denLt=densRHO(lista,n=n,xlim=xlim) #invocando a funcao
310
311 x=density(as.matrix(as.data.frame(lista[1])),from=xlim[1],to=
      xlim[2],n=n,
312           kernel='gaussian',na.rm = TRUE)$x #definindo os
      marcadores do eixo das distancias
313 y=lista[length(lista)][[1]] #definindo os marcadores do eixo
      das datas
314
315 # grafico 3D das fdp's estimadas para as correlacoes
316 plot_ly() %>%
317     add_surface(x = ~y, y = ~x,z = ~t(denLt)) %>%
318     #transposicao da matriz e troca de eixos para ter a
      mesma apresentacao do grafico de Onnela (2003)
319     layout(
320     scene = list(
321     xaxis = list(title = 'Data'),
322     yaxis = list(title = 'distancia'),

```

```

323     zaxis = list(title = 'fdp'),
324     aspectmode='cube' #aspecto mais amigavel para o
        grafico
325     )#veja mais opcoes em https://plotly.com/r/3d-surface
        -plots/
326 )
327
328 #####GRAFICO DAS ESTATISTICAS DAS DISTANCIAS#####
329
330 #NTL:
331 #funcao para calcular o tamanho normalizado da MST para uma
        determinada janela de tempo T
332 #Argumentos: matriz de logretornos (colunas como ativos,
        linhas como logretornos),
333 #delta: saltos em dias (defasagem) para obter a matriz de
        logretornos
334 #data: ano de cada matriz de autocorrelacao
335 #T janela de tempo em dias
336 #Retorna: data frame com as variaveis data(utilizando o
        primeiro ano),
337 #L_t(tamanho normalizado da arvore), assimetria(assimetria
        das distancias da MST),
338 #curtose(curtose das distancias da MST), Sd(desvio padrao das
        distancias da MST).
339 NTL<-function(matriz,T,delta,data){
340     i=1; j=T; it=1; m<-NULL
341     ass<-NULL; curt<-NULL; dpa<-NULL; d<-ymd()
342     while(j<=dim(matriz)[1]){
343         p=(i:j)
344         c<-cor(matriz[p,],use = 'pairwise.complete.obs',method=c(
            'pearson'))
345         dados_d<-sqrt(2*(1-c))

```

```

346     dd <- as.dist(dados_d)
347     mst<-spantree(dd)
348     dmst<-mst$dist
349     ass[it]<-skewness(dmst)
350     curt[it]<-kurtosis(dmst)
351     dpa[it]<-sd(dmst)
352     m[it]<-(1/(dim(matriz)[2]-1))*(sum((dmst), na.rm = TRUE))
353     d[it]<-data[i]
354     i=i+delta
355     j=j+delta
356     it=it+1
357   }
358   base=data.frame('data'=d,'L_t'=m, 'assimetria'=ass, '
359     'Curtose'=curt, 'Sd'=dpa)
360   base}
361
362 Lt=NTL(lrts,1000,21,dia) #invocando a funcao
363 head(Lt)
364
365 #grafico da media das distancias
366 windows()
367 ggplot(Lt,aes(x=data,y=L_t))+
368   geom_line(aes(x=data,y=L_t))+
369   theme_bw() + labs(y='media', x='Ano')
370
371 #grafico do Desvio padrao das distancias
372 windows()
373 ggplot(Lt,aes(x=data,y=Sd))+
374   geom_line(aes(x=data,y=Sd))+
375   theme_bw() + labs(y='Desvio padrao', x='Ano')
376

```

```

377 #grafico da Assimetria das distancias
378 windows()
379 ggplot(Lt, aes(x=data, y=assimetria))+
380   geom_line(aes(x=data, y=assimetria))+
381   theme_bw() + labs(y='Assimetria', x='Ano')
382
383 #grafico da Curtose das distancias
384 windows()
385 ggplot(Lt, aes(x=data, y=Curtose))+
386   geom_line(aes(x=data, y=Curtose))+
387   theme_bw() + labs(y='Curtose', x='Ano')
388
389 ##### GRAFICOS NA MESMA JANELA #####
390 windows()
391 pushViewport(viewport(layout = grid.layout(4, 1)))
392
393 print(ggplot(Lt, aes(x=data, y=L_t))+
394       geom_line(aes(x=data, y=L_t))+
395       theme_bw() + labs(y='media', x='Ano'),
396       vp = viewport(layout.pos.row = 1, layout.pos.col = 1))
397
398 print(ggplot(Lt, aes(x=data, y=Sd))+
399       geom_line(aes(x=data, y=Sd))+
400       theme_bw() + labs(y='Desvio padrao', x='Ano'),
401       vp = viewport(layout.pos.row = 2, layout.pos.col = 1))
402
403 print(ggplot(Lt, aes(x=data, y=assimetria))+
404       geom_line(aes(x=data, y=assimetria))+
405       theme_bw() + labs(y='Assimetria', x='Ano'),
406       vp = viewport(layout.pos.row = 3, layout.pos.col = 1))
407
408 print(ggplot(Lt, aes(x=data, y=Curtose))+

```

```

409     geom_line(aes(x=data,y=Curtose))+
410     theme_bw() + labs(y='Curtose', x='Ano'),
411     vp = viewport(layout.pos.row = 4, layout.pos.col = 1))
412
413 ##### GRAFICOS 3D RHO#####
414
415 ##### JANELA 252 DIAS#####
416
417 teste=RHOT(lrts,252,21,dia) #invocando a funcao para 252 dias
418
419 xlim=c(-0.1,0.8) # limites do eixo das correlacoes
420 n=50 # numero de pontos de correlacao para estimar a
      densidade das matrizes de correlacao
421
422 denR=densRHO(teste,n=n,xlim=xlim) #invocando a funcao
423
424 x=density(as.matrix(as.data.frame(teste[1])),from=xlim[1],to=
      xlim[2],n=n,
425           kernel='gaussian',na.rm = TRUE)$x #definindo os
      marcadores do eixo das correlacoes
426 y=teste[length(teste)][[1]] #definindo os marcadores do eixo
      das datas
427
428 # grafico 3D das fdp's estimadas para as correlacoes
429 plot_ly() %>%
430     add_surface(x = ~y, y = ~x,z = ~t(denR)) %>%
431     layout(
432     scene = list(
433     xaxis = list(title = 'Data'),
434     yaxis = list(title = 'correlacao'),
435     zaxis = list(title = 'fdp'),
436     aspectmode='cube' #aspecto mais amigavel para o

```

```

grafico
437     )#veja mais opcoes em https://plotly.com/r/3d-
        surface-plots/
438 )
439
440 ##### JANELA 504 DIAS #####
441
442 teste=RHOT(lrts,504,21,dia) #invocando a funcao para 504 dias
443
444 xlim=c(-0.1,0.8) # limites do eixo das correlacoes
445 n=50 # numero de pontos de correlacao para estimar a
        densidade das matrizes de correlacao
446
447 denR=densRHO(teste,n=n,xlim=xlim) #invocando a funcao
448
449 x=density(as.matrix(as.data.frame(teste[1])),from=xlim[1],to=
        xlim[2],n=n,
450         kernel='gaussian',na.rm = TRUE)$x #definindo os
        marcadores do eixo das correlacoes
451 y=teste[length(teste)][[1]] #definindo os marcadores do eixo
        das datas
452
453 # grafico 3D das fdp's estimadas para as correlacoes
454 plot_ly() %>%
455 add_surface(x = ~y, y = ~x,z = ~t(denR)) %>%
456 layout(
457     scene = list(
458         xaxis = list(title = 'Data'),
459         yaxis = list(title = 'correlacao'),
460         zaxis = list(title = 'fdp'),
461         aspectmode='cube' #aspecto mais amigavel para o grafico
462     )#veja mais opcoes em https://plotly.com/r/3d-surface-

```

```

plots/
463 )
464
465 ##### JANELA 756 DIAS #####
466
467 teste=RHOT(lrts,756,21,dia) #invocando a funcao para 756 dias
468
469 xlim=c(-0.1,0.8) # limites do eixo das correlacoes
470 n=50 # numero de pontos de correlacao para estimar a
densidade das matrizes de correlacao
471
472 denR=densRHO(teste,n=n,xlim=xlim) #invocando a funcao
473
474 x=density(as.matrix(as.data.frame(teste[1])),from=xlim[1],to=
xlim[2],n=n,
475 kernel='gaussian',na.rm = TRUE)$x #definindo os
marcadores do eixo das correlacoes
476 y=teste[length(teste)][[1]] #definindo os marcadores do eixo
das datas
477
478 # grafico 3D das fdp's estimadas para as correlacoes
479 plot_ly() %>%
480 add_surface(x = ~y, y = ~x,z = ~t(denR)) %>%
481 layout(
482 scene = list(
483 xaxis = list(title = 'Data'),
484 yaxis = list(title = 'correlacao'),
485 zaxis = list(title = 'fdp'),
486 aspectmode='cube' #aspecto mais amigavel para o grafico
487 )#veja mais opcoes em https://plotly.com/r/3d-surface-
plots/
488 )

```

```

489
490 ##### GRAFICOS 3D LT#####
491
492 ##### JANELA 252 DIAS#####
493
494 lista=LISTANTL(lrts,252,21,dia) #invocando a funcao para 252
      dias
495
496 xlim=c(0,1.5) # limites do eixo das distancias
497 n=50 # numero de pontos de distancias para estimar a
      densidade das distancias
498
499 denLt=densRHO(lista,n=n,xlim=xlim) #invocando a funcao
500
501 x=density(as.matrix(as.data.frame(lista[1])),from=xlim[1],to=
      xlim[2],n=n,
502           kernel='gaussian',na.rm = TRUE)$x #definindo os
      marcadores do eixo das distancias
503 y=lista[length(lista)][[1]] #definindo os marcadores do eixo
      das datas
504
505 # grafico 3D das fdp's estimadas para as correlacoes
506 plot_ly() %>%
507 add_surface(x = ~y, y = ~x,z = ~t(denLt)) %>%
508 #transposicao da matriz e troca de eixos para ter a mesma
      apresentacao do grafico de Onnela (2003)
509 layout(
510   scene = list(
511     xaxis = list(title = 'Data'),
512     yaxis = list(title = 'distancia'),
513     zaxis = list(title = 'fdp'),
514     aspectmode='cube' #aspecto mais amigavel para o grafico

```

```

515     )#veja mais opcoes em https://plotly.com/r/3d-surface-
        plots/
516 )
517
518
519 #####JANELA 504 DIAS#####
520 lista=LISTANTL(lrts,504,21,dia) #invocando a funcao para 504
        dias
521
522 xlim=c(0,1.5) # limites do eixo das distancias
523 n=50 # numero de pontos de distancias para estimar a
        densidade das distancias
524
525 denLt=densRHO(lista,n=n,xlim=xlim) #invocando a funcao
526
527 x=density(as.matrix(as.data.frame(lista[1])),from=xlim[1],to=
        xlim[2],n=n,
528           kernel='gaussian',na.rm = TRUE)$x #definindo os
           marcadores do eixo das distancias
529 y=lista[length(lista)][[1]] #definindo os marcadores do eixo
        das datas
530
531 # grafico 3D das fdp's estimadas para as correlacoes
532 plot_ly() %>%
533 add_surface(x = ~y, y = ~x,z = ~t(denLt)) %>%
534 #transposicao da matriz e troca de eixos para ter a mesma
        apresentacao do grafico de Onnela (2003)
535 layout(
536     scene = list(
537         xaxis = list(title = 'Data'),
538         yaxis = list(title = 'distancia'),
539         zaxis = list(title = 'fdp'),

```

```

540     aspectmode='cube' #aspecto mais amigavel para o grafico
541     )#veja mais opcoes em https://plotly.com/r/3d-surface-
        plots/
542 )
543
544 ##### JANELA 756 DIAS #####
545
546 lista=LISTANTL(lrts,756,21,dia) #invocando a funcao para 756
        dias
547
548 xlim=c(0,1.5) # limites do eixo das distancias
549 n=50 # numero de pontos de distancias para estimar a
        densidade das distancias
550
551 denLt=densRHO(lista,n=n,xlim=xlim) #invocando a funcao
552
553 x=density(as.matrix(as.data.frame(lista[1])),from=xlim[1],to=
        xlim[2],n=n,
554           kernel='gaussian',na.rm = TRUE)$x #definindo os
        marcadores do eixo das distancias
555 y=lista[length(lista)][[1]] #definindo os marcadores do eixo
        das datas
556
557 # grafico 3D das fdp's estimadas para as correlacoes
558 plot_ly() %>%
559   add_surface(x = ~y, y = ~x,z = ~t(denLt)) %>%
560 #transposicao da matriz e troca de eixos para ter a mesma
        apresentacao do grafico de Onnela (2003)
561 layout(
562   scene = list(
563     xaxis = list(title = 'Data'),
564     yaxis = list(title = 'distancia'),

```

```

565     zaxis = list(title = 'fdp'),
566     aspectmode='cube' #aspecto mais amigavel para o grafico
567   )#veja mais opcoes em https://plotly.com/r/3d-surface-
      plots/
568 )
569
570 #####TAXA DE SOBREVIVENCIA DE PASSO UNITARIO#####
571
572 #listaElos:
573 #funcao para calcular as matrizes de adjacencia para uma
      determinada janela de tempo T
574 #Argumentos: matriz de logretornos (colunas como ativos,
      linhas como logretornos),
575 #delta: saltos em dias (defasagem) para obter a matriz de
      logretornos
576 #data: ano de cada matriz de autocorrelacao
577 #T janela de tempo em dias
578 #Retorna: lista com as matrizes de adjacencia e a data(ultimo
      elemento da lista)
579 listaElos<-function(matriz,T,delta,data){
580   i=1; j=T; it=1; m<-NULL; m<-as.list(m); d<-ymd()
581   while(j<=dim(matriz)[1]){
582     p=(i:j)
583     c<-cor(matriz[p,],use = 'pairwise.complete.obs',method=c(
      'pearson'))
584     dados_d<-sqrt(2*(1-c))
585     dd <- as.dist(dados_d)
586     mstlistnb=mst.nb(dd) #calcula uma lista de adjacencia com
      base numa arvore de ligacao simples
587     mstadj=nb2mat(mstlistnb,zero.policy=TRUE, style='B') #
      recupera a matriz de adjacencia a partir da lista de
      adjacencia

```

```

588     mstadj[lower.tri(mstadj)] <- NA #trasnforma em uma matriz
        triangular superior
589     diag(mstadj) <- NA #retira a diagonal
590     mstadj[mstadj==0] <- NA #Substitui os 0's por NA
591     m[it]<-list(mstadj)
592     d[it]<-as.character(data[i])
593     i=i+delta
594     j=j+delta
595     it=it+1
596 }
597 m[it]<-list(d)
598 m}
599
600 N=dim(lrts)[2]#numero de ativos
601
602 #Taxa:
603 #funcao para calcular a taxa de sobrevivencia de passo
        unitario
604 #Argumentos: lista com as matrizes de adj e a data(ultimo
        elemento da lista)
605 #N: numero de ativos
606 #Retorna: vetor com as taxas de sobrevivencia de passo
        unitario
607 Taxa<-function(lista,N){
608     med=vector()
609     for (i in 1:(length(lista)-2)){
610         comp=as.matrix(as.data.frame(lista[i]))==as.matrix(as.
            data.frame(lista[i+1]))
611         med[i]=sum(comp,na.rm=TRUE)/(N-1)
612     }
613     med}
614

```

```
615 #####GRAFICO DA TAXA DE SOBREVIVENCIA PARA DIAS#####
616
617 ListaTaxa850=listaElos(lrts,850,21,dia)#invocando a funcao
        listaElos para 700 dias
618 taxafinal850=Taxa(ListaTaxa850,N)#invocando a funcao Taxa
619
620 d850=data.frame(ListaTaxa850[length(ListaTaxa850)])#
        recuperando as datas
621 d850=data.frame(d850[-1,]) #retira a primeira data
622 colnames(d850)[1]<-'data'
623
624 taxaf850=cbind(d850,taxafinal850) #junta as duas bases em uma
625
626 #grafico da Taxa de sobrevivencia
627 windows()
628 ggplot(taxaf850,aes(x=data,y=taxafinal850))+
629   geom_line(aes(x=data,y=taxafinal850))+
630   scale_y_continuous(limits = c(0.5,1))+
631   theme_bw() + labs(y='Taxa de sobrevivencia', x='Ano')
632
633 #####GRAFICO DA TAXA DE SOBREVIVENCIA 1000 DIAS#####
634
635 ListaTaxa1000=listaElos(lrts,1000,21,dia)#invocando a funcao
        listaElos para 1000 dias
636 taxafinal1000=Taxa(ListaTaxa1000,N)#invocando a funcao Taxa
637
638 d1000=data.frame(ListaTaxa1000[length(ListaTaxa1000)])#
        recuperando as datas
639 d1000=data.frame(d1000[-1,]) #retira a primeira data
640 colnames(d1000)[1]<-'data'
641
642 taxaf1000=cbind(d1000,taxafinal1000) #junta as duas bases em
```

```

    uma
643
644 #grafico da Taxa de sobrevivencia
645 windows()
646 ggplot(taxaf1000,aes(x=data,y=taxafinal1000))+
647   geom_line(aes(x=data,y=taxafinal1000))+
648   scale_y_continuous(limits = c(0.5,1))+
649   theme_bw() + labs(y='Taxa de sobrevivencia', x='Ano')
650
651 #####GRAFICOS DATA ESTENDIDA#####
652
653 #RHOT2:
654 #funcao para calcular a matriz de autocorrelacao para uma
    determinada janela de tempo T
655 #Argumentos: matriz de logretornos (colunas como ativos,
    linhas como logretornos),
656 #T janela de tempo em dias
657 #delta: saltos em dias (defasagem) para obter a matriz de
    logretornos
658 #data: ano de cada matriz de autocorrelacao
659 #Retorna: lista com matrizes de autocorrelacao com saltos de
    delta=21 dias, anos
660 #relacionados as matrizes de autocorrelacao
661 RHOT2<-function(matriz,T,delta,data){
662   i=1; j=T; it=1; rho<-NULL
663   rho<-as.list(rho); anos<-NULL
664   while(j<=dim(matriz)[1]){
665     p=(i:j)
666     c<-cor(matriz[p,],use = 'pairwise.complete.obs',method=c(
        'pearson'))
667     c[lower.tri(c)] <- NA #transforma em matriz triangular
        superiores

```

```

668     diag(c) <- NA #retira a diagonal
669     rho[it] <- list(c)
670     anos[it] <- as.character(data[i])
671     it = it + 1
672     i = i + delta
673     j = j + delta
674   }
675   k = 1
676   while(i <= (dim(matriz)[1] - (delta - 1))) {
677     p = (i : (dim(matriz)[1]))
678     c <- cor(matriz[p,], use = 'pairwise.complete.obs', method = c(
679       'pearson'))
680     diag(c) <- NA #retira a diagonal
681     rho[it] <- list(c)
682     anos[it] <- as.character(data[i])
683     it = it + 1
684     i = i + k * delta
685     k = k + 1
686   }
687   rho[it] <- list(anos)
688   rho}
689
690 teste = RHOT2(lrts, 1000, 21, dia) #invocando a funcao
691 fix(teste)
692
693 estatisticas = est(teste) #invocando a funcao
694 head(estatisticas)
695 da = as.data.frame(teste[length(teste)]) #recuperando as datas
696   calculadas acima
697   colnames(da)[1] <- 'dat' #nomeando a coluna

```

```

697
698 estatisticas=cbind(da,estatisticas) #criando um data frame
699 estatisticas$dat=as.Date(estatisticas$dat) #transformando a
      data em um objeto Date
700
701 #####GRAFICOS NA MESMA JANELA#####
702 windows()
703 pushViewport(viewport(layout = grid.layout(4, 1)))
704
705 print(ggplot(estatisticas, aes(x=dat, y=media))+
706       geom_line(aes(x=dat, y=media))+
707       theme_bw() + labs(y='media', x='Ano'),
708       vp = viewport(layout.pos.row = 1, layout.pos.col = 1))
709
710 print(ggplot(estatisticas, aes(x=dat, y=dp))+
711       geom_line(aes(x=dat, y=dp))+
712       theme_bw() + labs(y='Desvio padrao', x='Ano'),
713       vp = viewport(layout.pos.row = 2, layout.pos.col = 1))
714
715 print(ggplot(estatisticas, aes(x=dat, y=assimetria))+
716       geom_line(aes(x=dat, y=assimetria))+
717       theme_bw() + labs(y='Assimetria', x='Ano'),
718       vp = viewport(layout.pos.row = 3, layout.pos.col = 1))
719
720 print(ggplot(estatisticas, aes(x=dat, y=curtose))+
721       geom_line(aes(x=dat, y=curtose))+
722       theme_bw() + labs(y='Curtose', x='Ano'),
723       vp = viewport(layout.pos.row = 4, layout.pos.col = 1))
724
725 #LISTANTL2:
726 #funcao para calcular as distancias da MST para uma
      determinada janela de tempo T

```

```

727 #Argumentos: matriz de logretornos (colunas como ativos,
      linhas como logretornos),
728 #delta: saltos em dias (defasagem) para obter a matriz de
      logretornos
729 #data: ano de cada matriz de autocorrelacao
730 #T janela de tempo em dias
731 #Retorna: lista com as distancias da MST e a data(ultimo
      elemento da lista)
732 LISTANTL2<-function(matriz,T,delta,data){
733   i=1; j=T; it=1; m<-NULL; m<-as.list(m); d<-ymd()
734   while(j<=dim(matriz)[1]){
735     p=(i:j)
736     c<-cor(matriz[p,],use = 'pairwise.complete.obs',method=c(
      'pearson'))
737     dados_d<-sqrt(2*(1-c))
738     dd <- as.dist(dados_d)
739     mst<-spantree(dd)
740     dmst<-mst$dist #armazena as distancia da mst
741     m[it]<-list(dmst)
742     d[it]<-as.character(data[i])
743     i=i+delta
744     j=j+delta
745     it=it+1
746   }
747   k=1
748   while(i<=(dim(matriz)[1]-(delta-1))){
749     p=(i:(dim(matriz)[1]))
750     c<-cor(matriz[p,],use = 'pairwise.complete.obs',method=c(
      'pearson'))
751     dados_d<-sqrt(2*(1-c))
752     dd <- as.dist(dados_d)
753     mst<-spantree(dd)

```

```

754     dmst<-mst$dist #armazena as distancia da mst
755     m[it]<-list(dmst)
756     d[it]<-as.character(data[i])
757     it=it+1
758     i=i+k*delta
759     k=k+1
760 }
761 m[it]<-list(d)
762 m}
763
764 lista=LISTANTL2(lrts,1000,21,dia) #invocando a funcao
765
766 #####GRAFICOS DAS ESTATISTICAS DA DISTANCIA#####
767
768 #NTL2:
769 #funcao para calcular o tamanho normalizado da MST para uma
      determinada janela de tempo T
770 #Argumentos: matriz de logretornos (colunas como ativos,
      linhas como logretornos),
771 #delta: saltos em dias (defasagem) para obter a matriz de
      logretornos
772 #data: ano de cada matriz de autocorrelacao
773 #T janela de tempo em dias
774 #Retorna: data frame com as variaveis data(utilizando o
      primeiro ano),
775 #L_t(tamanho normalizado da arvore), assimetria(assimetria
      das distancias da MST),
776 #curtose(curtose das distancias da MST), Sd(desvio padrao das
      distancias da MST).
777 NTL2<-function(matriz,T,delta,data){
778     i=1; j=T; it=1; m<-NULL
779     ass<-NULL; curt<-NULL; dpa<-NULL; d<-ymd()

```

```

780 while(j<=dim(matriz)[1]){
781   p=(i:j)
782   c<-cor(matriz[p,],use = 'pairwise.complete.obs',method=c(
       'pearson'))
783   dados_d<-sqrt(2*(1-c))
784   dd <- as.dist(dados_d)
785   mst<-spantree(dd)
786   dmst<-mst$dist
787   ass[it]<-skewness(dmst)
788   curt[it]<-kurtosis(dmst)
789   dpa[it]<-sd(dmst)
790   m[it]<-(1/(dim(matriz)[2]-1))*(sum((dmst), na.rm = TRUE))
791   d[it]<-data[i]
792   i=i+delta
793   j=j+delta
794   it=it+1
795 }
796 k=1
797 while(i<=(dim(matriz)[1]-(delta-1))){
798   p=(i:(dim(matriz)[1]))
799   c<-cor(matriz[p,],use = 'pairwise.complete.obs',method=c(
       'pearson'))
800   dados_d<-sqrt(2*(1-c))
801   dd <- as.dist(dados_d)
802   mst<-spantree(dd)
803   dmst<-mst$dist
804   ass[it]<-skewness(dmst)
805   curt[it]<-kurtosis(dmst)
806   dpa[it]<-sd(dmst)
807   m[it]<-(1/(dim(matriz)[2]-1))*(sum((dmst), na.rm = TRUE))
808   d[it]<-data[i]
809   it=it+1

```

```
810     i=i+k*delta
811     k=k+1
812 }
813 base=data.frame('data'=d,'L_t'=m, 'assimetria'=ass, '
      Curtose'=curt, 'Sd'=dpa)
814 base}
815
816
817 Lt=NTL2(lrts,1000,21,dia) #invocando a funcao
818 head(Lt)
819
820 #####GRAFICOS NA MESMA JANELA#####
821 windows()
822 pushViewport(viewport(layout = grid.layout(4, 1)))
823
824 print(ggplot(Lt,aes(x=data,y=L_t))+
825       geom_line(aes(x=data,y=L_t))+
826       theme_bw() + labs(y='media', x='Ano'),
827       vp = viewport(layout.pos.row = 1, layout.pos.col = 1))
828
829 print(ggplot(Lt,aes(x=data,y=Sd))+
830       geom_line(aes(x=data,y=Sd))+
831       theme_bw() + labs(y='Desvio padrao', x='Ano'),
832       vp = viewport(layout.pos.row = 2, layout.pos.col = 1))
833
834 print(ggplot(Lt,aes(x=data,y=assimetria))+
835       geom_line(aes(x=data,y=assimetria))+
836       theme_bw() + labs(y='Assimetria', x='Ano'),
837       vp = viewport(layout.pos.row = 3, layout.pos.col = 1))
838
839 print(ggplot(Lt,aes(x=data,y=Curtose))+
840       geom_line(aes(x=data,y=Curtose))+
```

```
841     theme_bw() + labs(y='Curtose', x='Ano'),  
842     vp = viewport(layout.pos.row = 4, layout.pos.col = 1))
```

ANEXO A - Abreviaturas dos Ativos

Quadro 1: Lista de Abreviaturas - Índice BOVESPA

Código	Empresa	Setor de Atuação	Observação
ABEV3	AMBEV S.A.	Consumo não Cíclico \ Bebidas	AMBEV S.A. - Fundiu-se com CIA CERVEJARIA BRAHMA e CIA BEBIDAS DAS AMERICAS - AMBEV no ano de 2000
B3SA3	B3 S.A. - BRASIL, BOLSA, BALCÃO	Financeiro	BVFM3 substituído por B3SA3 em 2018
BBDC4	BCO BRADESCO S.A.	Financeiro \ Bancos	
BBAS3	BCO BRASIL S.A.	Financeiro \ Bancos	
SANB11	BCO SANTANDER (BRASIL) S.A	Financeiro \ Bancos	
BRAP4	BRADESPAR S.A.	Mineração	
BRKM5	BRASKEM S.A.	Químicos \ Petroquímicos	
BRFS3	BRF S.A.	Consumo não Cíclico \ Alimentos	PRGA3 substituído por BRFS3 - Em 2009, Perdigão fundiu-se com a Sadia, formando a BRF S.A.
CCRO3	CCR S.A.	Transporte \ Exploração de Rodovias	
ELET3	CENTRAIS ELET BRAS S.A. - ELETROBRAS	Energia Elétrica	

PCAR4	CIA BRASILEIRA DE DISTRIBUICAO	Consumo não Cíclico \ Comércio e Distribuição	PCAR5 substituído por PCAR4 - Em 2009, as atuais ações preferenciais "PCAR4" passam a denominar-se ações preferenciais classe A "PCAR5".
CMIG4	CIA ENERGETICA DE MINAS GERAIS - CEMIG	Energia Elétrica	
CPLE6	CIA PARANAENSE DE ENERGIA - COPEL	Energia Elétrica	
SBSP3	CIA SANEAMENTO BASICO EST SAO PAULO	Água e Saneamento	
CSNA3	CIA SIDERURGICA NACIONAL	Siderurgia e Metalurgia	
CSAN3	COSAN S.A.	Petróleo. Gás e Biocombustíveis	
CPFE3	CPFL ENERGIA S.A.	Energia Elétrica	
CYRE3	CYRELA BRAZIL REALTY S.A.EMPREENDEPART	Consumo Cíclico \ Construção Civil	
EMBR3	EMBRAER S.A.	Material de Transporte	
FIBR3	FIBRIA CELULOSE S.A.	Papel e Celulose	VCPA4 e VCPA3 substituído por FIBR3 - Em 2009, Aracruz fundiu-se com a VCP, formando a FIBRIA CELULOSE S.A.
GGBR4	GERDAU S.A.	Siderurgia e Metalurgia	
GOLL4	GOL LINHAS AEREAS INTELIGENTES S.A.	Transporte \ Transporte Aéreo	
ITUB4	ITAU UNIBANCO HOLDING S.A.	Financeiro \ Bancos	ITAU4 substituído por ITUB4 - Em 2009, as ações preferenciais "ITAU4" foram convertidas em ações preferenciais ED "ITUB4".

LAME4	LOJAS AMERICANAS S.A.	Consumo Cíclico \ Comércio	
LREN3	LOJAS RENNER S.A.	Consumo Cíclico \ Comércio	
NTCO3	NATURA COSMETICOS S.A.	Consumo não Cíclico \ Produtos de Uso Pessoal e de Limpeza	Em 2019 a Natura Cosméticos foi incorporada pela Natura&CO Holding e NATU3 foi substituído por NTCO3
PETR3	PETROLEO BRASILEIRO S.A. PETROBRAS	Petróleo. Gás e Biocombustíveis	
RAIL3	RUMO S.A.	Transporte \ Transporte Ferroviário	ALLL11 substituído por ALLL3 - Em 2010, e em 2018 foi substituído por RAIL3.
VIVT4	TELEFÔNICA BRASIL S.A	Comunicações \ Telecomunicações	TLPP4 substituído por VIVT4 - Em 2011 a Telefônica fundiu-se com a Vivo Participações, formando a TELEFÔNICA BRASIL S.A.
UGPA3	ULTRAPAR PARTICIPAÇÕES S.A.	Petróleo. Gás e Biocombustíveis	UGPA4 substituído por UGPA3 - Em 2012, as ações preferenciais "UGPA4" foram convertidas em ações ordinárias "UGPA3".
USIM5	USINAS SID DE MINAS GERAIS S.A.-USIMINAS	Siderurgia e Metalurgia	
VALE3	VALE S.A.	Mineração	