



UNIVERSIDADE ESTADUAL PAULISTA

"JÚLIO DE MESQUITA FILHO"

FACULDADE DE ENGENHARIA DE BAURU

Departamento de Engenharia de Produção



PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO

DEBORA BARBOSA AIRES

**ANÁLISE DE CRISES FINANCEIRAS BRASILEIRAS USANDO TEORIA
DAS MATRIZES ALEATÓRIAS**

BAURU

2021

DEBORA BARBOSA AIRES

**ANÁLISE DE CRISES FINANCEIRAS BRASILEIRAS USANDO TEORIA
DAS MATRIZES ALEATÓRIAS**

Tese apresentada como requisito para obtenção do título de Doutor em Engenharia de Produção pela Faculdade de Engenharia de Bauru da Universidade Estadual Paulista "Júlio de Mesquita Filho".

Área de concentração:

Gestão de Sistemas Produtivos

Orientador:

Prof. Dr. Antonio Fernando Crepaldi

BAURU

2021

Aires, Debora Barbosa.

Análise de crises financeiras brasileiras
usando teoria das matrizes aleatórias / Debora
Barbosa Aires, 2021

149 p. : il.

Orientador: Antonio Fernando Crepaldi

Tese (Doutorado)-Universidade Estadual
Paulista. Faculdade de Engenharia, Bauru, 2021

1. Crise econômica. 2. Crises financeiras. 3.
Correlação (Estatística). 4. Crise financeira
global 2008-2009. 5. Análise multivariada. I.
Universidade Estadual Paulista. Faculdade de
Engenharia. II. Título.

ATA DA DEFESA PÚBLICA DA TESE DE DOUTORADO DE DEBORA BARBOSA AIRES, DISCENTE DO PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO, DA FACULDADE DE ENGENHARIA - CÂMPUS DE BAURU.

Aos 05 dias do mês de março do ano de 2021, às 09:00 horas, por meio de Videoconferência, realizou-se a defesa de TESE DE DOUTORADO de DEBORA BARBOSA AIRES, intitulada **ANÁLISE DE CORRELAÇÃO DE AÇÕES, NO PERÍODO DA CRISE DO SUBPRIME E VIZINHANÇAS TEMPORAIS, POR MEIO DA TEORIA DE MATRIZES ALEATÓRIAS**. A Comissão Examinadora foi constituída pelos seguintes membros: Prof. Dr. ANTONIO FERNANDO CREPALDI (Orientador(a) - Participação Virtual) do(a) Departamento de Engenharia de Produção / Faculdade de Engenharia de Bauru - UNESP, Prof. Dr. ENZO BARBERIO MARIANO (Participação Virtual) do(a) Departamento de Engenharia de Produção / Faculdade de Engenharia de Bauru - UNESP, Prof. Dr. JOSE DE SOUZA RODRIGUES (Participação Virtual) do(a) Departamento de Engenharia de Produção / Faculdade de Engenharia de Bauru - UNESP, Prof. Dr. FERNANDO FAGUNDES FERREIRA (Participação Virtual) do(a) Escola de Artes, Ciências e Humanidades (EACH) / Universidade de São Paulo (USP), Professora Associada DAISY APARECIDA DO NASCIMENTO REBELATTO (Participação Virtual) do(a) Departamento de Engenharia de Produção / Universidade de São Paulo / São Carlos. Após a exposição pela doutoranda e arguição pelos membros da Comissão Examinadora que participaram do ato, de forma presencial e/ou virtual, a discente recebeu o conceito final: aprovada. Nada mais havendo, foi lavrada a presente ata, que após lida e aprovada, foi assinada pelo(a) Presidente(a) da Comissão Examinadora.

Prof. Dr. ANTONIO FERNANDO CREPALDI



UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
Campus de Bauru

PROPOSTA DE ALTERAÇÃO DO TÍTULO

A COMISSÃO EXAMINADORA PROPÕE A ALTERAÇÃO DO TÍTULO DO TRABALHO DA ALUNA: **DEBORA BARBOSA AIRES**

DE: "ANÁLISE DE CORRELAÇÃO DE AÇÕES, NO PERÍODO DA CRISE DO SUBPRIME E VIZINHANÇAS TEMPORAIS, POR MEIO DA TEORIA DE MATRIZES ALEATÓRIAS"

PARA:

Análise de crises financeiras brasileiras usando Teoria das Matrizes

Aleatórias

Bauru, 05 de março de 2021.

Prof. Dr. Antonio Fernando Crepaldi
Orientador



Faculdade de Engenharia de Bauru – Pós-graduação
Av. Eng. Luiz Edmundo Carrijo Coube, 14-01 17033-360 Bauru - SP
tel. (14) 3103-6108 spg.feb@unesp.br www.feb.unesp.br

Dedico este trabalho a meus pais, Cláudio Aires e Maria Tereza Barbosa Aires, por sempre me incentivarem a ir mais além, por me ensinarem o valor da formação acadêmica e por serem exemplos de honestidade, dedicação e esforço.

Recebam esta conquista como prova do meu amor por vocês. Parabéns.

Agradecimentos

Agradeço pelos frutos deste trabalho a Deus todo poderoso por nunca me desamparar, principalmente nas horas mais difíceis. Sem Ele nada seria possível.

Agradeço também aos meus pais, irmão, familiares e ao meu companheiro, Leandro Pires Maciel, por todo apoio, incentivo, compreensão, ajuda em todos os sentidos e pelo imenso amor que recebo diariamente.

Agradeço a meus amigos, desde os mais próximos até os mais distantes, pela companhia, amizade, companheirismo, paciência e pela alegria que me transmitem, me ajudando, assim, a seguir em frente.

Agradeço a Universidade Estadual Paulista “Júlio de Mesquita Filho” (Unesp) por disponibilizar toda a infraestrutura para que este trabalho fosse realizado, a meus professores de graduação e pós-graduação da Faculdade de Ciência e da Faculdade de Engenharia da Unesp do campus de Bauru, por me acompanharem durante todo meu processo de desenvolvimento pessoal e intelectual, em especial ao meu orientador Prof. Dr. Antonio Fernando Crepaldi, aos professores Prof. Dr. Enzo Barberio Mariano, Prof. Dr. José de Souza Rodrigues, Prof. Dr. Fernando Fagundes Ferreira e Prof^a. Dr^a. Daisy Aparecida do Nascimento Rebelatto que compuseram minha banca de qualificação e/ou defesa e contribuíram de forma substancial para a qualidade deste trabalho. Agradeço também aos funcionários da Seção Técnica de Pós-graduação da Faculdade de Engenharia de Bauru por todo apoio e orientação quanto aos trâmites legais e institucionais relativos a esta tese.

“Não foram razões humanas que nos moveram, mas a graça de Deus”. 2Cor 1, 12b

Resumo

Diversas abordagens e conceitos da física, como Teoria de Matrizes Aleatórias, têm sido utilizadas para investigar séries temporais financeiras, dada sua complexidade. Neste estudo será analisado o espectro de correlação de ações da bolsa de valores brasileira utilizando a Teoria das Matrizes Aleatórias (TMA), juntamente com algumas técnicas de análise multivariada, nos períodos da crise financeira do *subprime de 2008* e crise financeira asiática além de suas vizinhanças temporais. Os resultados mostram uma alteração substancial e consistente do comportamento das matrizes de correlação, dos autovalores e dos autovetores em períodos de crise. Também mostram que o período que antecede uma crise apresenta sintomas que poderiam predizer tal crise. Mostram ainda e que no período posterior à crise há um movimento do mercado na direção de uma recuperação da crise. No entanto, essa recuperação não é total, visto que se percebem nos dados analisados características persistentes da crise superada. Esse efeito seria uma marca das crises nas correlações das séries temporais que é observável pela aplicação da TMA. O estudo também apresenta a sugestão de uma carteira de investimentos equilibrada, mesmo em períodos de crise financeira, por meio de Análises de Componentes Principais e métodos de agrupamento.

Palavras-chave: Crise econômica; Crise financeira global, 2008-2009; Ações (Finanças); Mercado financeiro; Correlação (Estatística); Análise multivariada; Análise por agrupamento; Crises financeiras; Análise de componentes principais; Investimentos bancários; Ações (Finanças) Preços.

Abstract

Several approaches and concepts of physics, such as Random Matrix Theory, have been used to investigate financial time series, given their complexity. In this study, the spectrum of stock correlation in the Brazilian stock market will be analyzed using Random Matrix Theory (RMT) along with some multivariate analysis techniques in the 2008 subprime crisis and Asian financial crisis periods beyond its temporal neighborhoods. Results show a substantial and consistent change in the behavior of the correlation matrices, eigenvalues and eigenvectors in periods of crisis. They also show that post-crisis period has symptoms that could predict such a crisis. They show as well that in the post-crisis period there is a market movement towards recovery from the crisis. However, this recovery is not complete since the analyzed data show persistent characteristics of the overcome crisis. This effect would be a mark of the crises in the correlations of the time series, which is observable by the application of the RMT. The study also presents the suggestion of a balanced investment portfolio, even in periods of financial crisis, through Principal Component Analysis and grouping methods.

Keywords: *Stocks; Cluster analysis; Financial crises; Principal components analysis; Bank investments; Stocks Prices.*

Lista de Abreviaturas e Siglas

ACP – Análise de Componentes Principais;

B3 - Brasil Bolsa Balcão;

CAPM - *Capital Asset Pricing Model*

FMI - Fundo Monetário Internacional

i.i.d. - independentes e identicamente distribuídos;

Ibovespa - Índice Bovespa;

PCA - *Principal Component Analysis*

SIM - *Single Index Model*;

TMA - Teoria de Matrizes Aleatórias;

Índice de Figuras

Figura 1 - A Pesquisa Operacional e Engenharia Econômica e suas subáreas.	29
Figura 2 - Distribuição de Wigner: Histograma de Autovalores de um Núcleo Pesado.....	49
Figura 3 - Passos da Metodologia da TMA.....	55
Figura 4 - Diagrama esquemático do processo de unfolding.....	57
Figura 5 - Diagrama da dinâmica de comparação-filtragem das propriedades universais sobre a correlação empírica.	58
Figura 6 - Distribuição de Wigner: Histograma da distribuição do espaçamento entre vizinhos próximos.	65
Figura 7 - Divisão do Período I em subperíodos analisados sobre a variação do Ibovespa.	76
Figura 8 - Início e fim da crise do subprime sobre a variação trimestral do PIB brasileiro.....	77
Figura 9 - Início e fim da crise do subprime sobre índices internacionais.....	78
Figura 10 - Divisão do Período II em subperíodos analisados sobre a variação do Ibovespa.	79
Figura 11 - Início e fim da crise de contaminação asiática sobre índices internacionais.	80
Figura 12 - Densidade de probabilidade das correlações empíricas do Período I: (a) para dados empíricos originais; (b) para dados empíricos adicionada a cotação do dólar em reais simulando uma ação anticorrelacionada.....	82
Figura 13 - Densidade de probabilidade das correlações empíricas do Período II.	83
Figura 14 - Distribuição de probabilidade dos autovalores empíricos vs. comportamento universal TMA para o Período I.....	85
Figura 15 - Distribuição de probabilidade dos autovalores empíricos vs. comportamento universal TMA para o Período II.....	86
Figura 16 - Diagrama de Caixa para os autovalores empíricos para o Período I.	88
Figura 17 - Diagrama de Caixa para os autovalores empíricos para o Período II.....	89
Figura 18 - Distribuição de probabilidade dos autovalores empíricos do Período I adicionado a variação do dólar vs. comportamento universal TMA.	90
Figura 19 - Distribuição dos autovalores unfolded para o Período I.....	91
Figura 20 - Distribuição dos autovalores unfolded para o Período II.	91
Figura 21 - Distribuição dos vizinhos próximos dos autovalores unfolded vs. distribuição de Wigner para o Período I.	92
Figura 22 - Distribuição dos vizinhos próximos dos autovalores unfolded vs. distribuição de Wigner para o Período II.	92
Figura 23 - Alcance dos autovalores unfolded vs. alcance de autovalores teóricos para o Período I.	94
Figura 24 - Alcance dos autovalores unfolded vs. alcance de autovalores teóricos para o Período II.	94
Figura 25 - Densidade dos componentes dos autovetores v' 100normalizados para o Período I.	96

Figura 26 - Densidade dos componentes dos autovetores v'_{61} normalizados para o Período II.....	96
Figura 27 - Densidade dos componentes dos autovetores empíricos v'_{101} normalizados do Período I gerados juntamente com a variação da cotação do dólar em reais.....	98
Figura 28 - Assimetria da distribuição dos componentes dos autovetores empíricos v_i para o Período I.....	99
Figura 29 - Assimetria da distribuição dos componentes dos autovetores empíricos v_i para o Período II.....	99
Figura 30 - Assimetria da distribuição dos componentes dos autovetores empíricos v_i do Período I gerados juntamente com a variação da cotação do dólar em reais.....	100
Figura 31 - Medida de não achatamento padrão da distribuição dos componentes dos autovetores empíricos v_i para o Período I.....	101
Figura 32 - Medida de não achatamento padrão da distribuição dos componentes dos autovetores empíricos v_i para o Período II.....	101
Figura 33 - Diagrama de Caixa para os autovetores empíricos divergentes do Período I.....	102
Figura 34 - Diagrama de Caixa para os autovetores empíricos divergentes do Período II.....	103
Figura 35 - Taxa de Participação Inversa dos autovetores empíricos v_i do Período I.....	104
Figura 36 - Taxa de Participação Inversa dos autovetores empíricos v_i do Período II.....	104
Figura 37 - Taxa de Participação dos autovetores empíricos v_i do Período I.....	105
Figura 38 - Taxa de Participação dos autovetores empíricos v_i do Período II.....	105
Figura 39 - Composição dos autovetores v_{100} considerando as ações com participação significativa para o Período I.....	107
Figura 40 - Composição dos autovetores v_{61} considerando as ações com participação significativa para o Período II.....	107
Figura 41 - Regressão das séries temporais $rIbovespa(t)$ contra $r_{100}(t)$ para o Período I.....	109
Figura 42 - Regressão das séries temporais $rIbovespa(t)$ contra $r_{61}(t)$ para o Período II.....	109
Figura 43 - Composição dos autovetores v_{99} considerando as ações com participação significativa para o Período I.....	110
Figura 44 - Composição dos autovetores v_{60} considerando as ações com participação significativa para o Período II.....	111
Figura 45 - Composição dos autovetores v_{98} considerando as ações com participação significativa para o Período I.....	111
Figura 46 - Composição dos autovetores v_{59} considerando as ações com participação significativa para o Período II.....	112
Figura 47 - Diagrama de Pareto para os dez primeiros componentes principais do Período I.....	113

Figura 48 - Diagrama de Pareto para os dez primeiros componentes principais do Período II.....	113
Figura 49 - Escores de cada dia para ACP dos retornos diários sobre os dois primeiros componentes principais para o Período I.	115
Figura 50 - Escores de cada dia para ACP dos retornos diários sobre os dois primeiros componentes principais para o Período II.....	116
Figura 51 - Escores de cada dia para ACP dos retornos diários sobre os três primeiros componentes principais para o Período I.	117
Figura 52 - Escores de cada dia para ACP dos retornos diários sobre os três primeiros componentes principais para o Período II.....	117
Figura 53 - Coeficientes de cada ação para ACP com agrupamento por setores de mercado e por grupos k- Means sobre os dois primeiros componentes principais para p Período I.	118
Figura 54 - Coeficientes de cada ação para ACP com agrupamento por setores de mercado e por grupos k- Means sobre os dois primeiros componentes principais para p Período II.	119
Figura 55 - Dendrograma da AHCT das Matrizes de Correlação Empírica para o Período I.	123
Figura 56 - Dendrograma da AHCT das Matrizes de Correlação Empírica para o Período II.....	124
Figura 57 - Densidade de probabilidade de $C100$ para o Período I.....	128
Figura 58 - Densidade de probabilidade de $C1bovespa$ para o Período I.	128
Figura 59 - Densidade de probabilidade de $C1bovespa$ para o Período II.....	129
Figura 60 - Diagrama de Caixa para os autovetores $vilbovespa$ divergentes para o Período I.	131
Figura 61 - Diagrama de Caixa para os autovetores $vilbovespa$ divergentes para o Período II.....	131
Figura 62 - Composição dos autovetores $v99Ibovespa$ considerando as ações com participação significativa para o Período I.	133
Figura 63 - Composição dos autovetores $v60Ibovespa$ considerando as ações com participação significativa para o Período II.....	133
Figura 64 - Composição dos autovetores $v98Ibovespa$ considerando as ações com participação significativa para o Período I.	135
Figura 65 - Composição dos autovetores $v59Ibovespa$ considerando as ações com participação significativa para o Período II.....	135
Figura 66 - Composição dos autovetores $v97Ibovespa$ considerando as ações com participação significativa para o Período I.	136
Figura 67 - Composição dos autovetores $v58Ibovespa$ considerando as ações com participação significativa para o Período II.....	137
Figura 68 - Composição dos autovetores $v96Ibovespa$ considerando as ações com participação significativa para o Período I.	138
Figura 69 - Composição dos autovetores $v57Ibovespa$ considerando as ações com participação significativa para o Período II.....	138

Índice de Quadros

Quadro 1 - Ações utilizadas na análise do Período I.....	69
Quadro 2 - Ações utilizadas na análise do Período II.....	73

Sumário

1. Introdução	25
1.1 Justificativa	28
1.2 Objetivo	31
1.3 Características da Pesquisa	32
1.4 Estrutura da Tese	32
2. Revisão da Literatura	33
2.1 Mercado de Ações no Brasil	33
2.2 Crise Financeira	34
2.2.1 Crise do Subprime.....	37
2.2.2 Crise Asiática e Russa.....	38
2.3 Minimização de Riscos de Portifólio	41
2.4 Noções Básicas de Conceitos Fundamentais	43
2.4.1 Matriz Hamiltoniana.....	43
2.4.2 Autovalores e Autovetores.....	44
2.4.3 Análise de Componentes Principais.....	45
2.5 Teoria das Matrizes Aleatórias	48
2.6 Utilização da TMA na análise de séries temporais financeiras	50
3. Metodologia	55
3.1 Determinação da Série Temporal Financeira	60
3.2 Determinação da Matriz de Correlação Empírica	61
3.3 Determinação da Matriz de Correlação Aleatória	62
3.4 Função Densidade de Probabilidade	62
3.5 Processo Unfolding	62
3.6 Propriedades Universais da TMA	64
3.6.1 Distribuição dos espaçamentos dos vizinhos próximos - distribuição de Wigner 64	
3.6.2 Distribuição dos espaçamentos de longo alcance - alcance de autovalores	65
3.7 Análise dos Autovetores	66
4. Coleta e Análise de Dados	68
4.1 Procedimentos de Coleta de Dados	68
4.2 Procedimentos de Análise de Dados	75

5. Resultados e Análises	81
5.1 Análise das Matrizes de Correlação	81
5.2 Análise dos Autovalores	85
5.3 Autovalores Empíricos vs. Propriedades Universais da TMA	90
5.4 Análise dos Autovetores Divergentes	95
5.5 ACP para Correlação Empírica de Retornos	112
5.6 Remoção da Influência do Mercado	127
6. Conclusões	140
Referências Bibliográficas	144

1. Introdução

O comportamento das ações no mercado financeiro é uma análise de difícil tratamento devido ao comportamento não linear desses sistemas, o que gera a necessidade de utilização de ferramentas capazes de tratar esse tipo de problema de forma eficiente (SILVA; SAPATTI; FLAUZINO, 2010). Esta análise é importante não só para entender o comportamento coletivo de um sistema complexo, mas também para alocação de ativos e estimar o risco de um portfólio (BOUCHAUD; POTTERS, 2009; MARKOWITZ, 1952; OH, 2011; PARTOVI; CAPUTO, 2004; SENSOY; YUKSEL; ERTURK, 2013). Assim sendo, o avanço tecnológico elevou o grau de sofisticação das operações oferecidas pelo mercado financeiro, o que eleva as possibilidades de ganhos e redução de riscos nas atividades de investimento em ações (MARTINS; METTE; MACEDO, 2008).

É do conhecimento comum que quaisquer empresas na economia estão correlacionadas. Mesmo empresas pertencentes a diferentes setores de uma indústria podem ser correlacionadas por causa de correlações "indiretas" (PLEROU *et al.*, 2000). Devido à complexidade do sistema financeiro, várias abordagens e conceitos advindos da física, como Teoria de Matrizes Aleatórias (TMA), têm sido utilizadas para investigar os mecanismos determinísticos nas tendências dos preços nas séries temporais financeiras por meio de análise da correlação (COELHO *et al.* 2008; KUMAR; DEO, 2015). Dessa forma o desenvolvimento de métodos que lidem com problemas de correlações é de fundamental importância (LIVAN, REBECCHI, 2012).

A TMA vem auxiliar na compreensão das complexas interações entre elementos de diferentes sistemas, buscando destacar a possível aleatoriedade existente nessas interações (PLEROU *et al.*, 2000). Pode-se dizer que a aplicação do método TMA serve de filtro para as correlações, a fim de promover uma separação entre ruído (gerado pela aleatoriedade) e correlações realmente significativas, através da análise espectral dos autovalores das matrizes de correlação (BOUCHAUD; POTTERS, 2009; COELHO *et al.* 2008; DYSON, 1962; GUHR; KULKARNI; DEO, 2007; LIVAN;

REBECCHI, 2012; MÜLLER–GROELING; WEIDENMÜLLER, 1998; PLEROU *et al.*, 2000; PLEROU *et al.*, 2002).

A TMA originou-se da física nuclear na década de 1950 através de estudos sobre as mudanças de níveis energéticos de núcleos atômicos de elementos com grande massa atômica (GUHR; MÜLLER–GROELING; WEIDENMÜLLER, 1998). Na década de 1990, começaram a surgir, com frequência significativa, trabalhos publicados que empregavam a TMA ao estudo do comportamento de matrizes de correlação de dados financeiros (BOUCHAUD; POTTERS, 2009; CONLON; RUSKIN; CRANE, 2009; SENSOY; YUKSEL; ERTURK, 2013).

Resultados de estudos apresentam frequentemente que desvios do comportamento de matrizes de correlação geradas a partir séries históricas financeiras, comparadas a matrizes de correlação aleatórias, fornecem informações genuínas sobre o comportamento do mercado (FENN *et al.*, 2011; KULKARNI; DEO, 2007; KUMAR; DEO, 2015; KUMAR; DEO, 2012; LIM *et al.*, 2009; LIM *et al.*, 2014; MEIYING; BIN, 2011; PLEROU *et al.*, 1999; PLEROU *et al.*, 2000; PLEROU *et al.*, 2001; PLEROU *et al.*, 2002; SANDOVAL JUNIOR; BORTOLUZZO; VENEZUELA, 2014; SANDOVAL JUNIOR; FRANCA, 2012; SENSOY; YUKSEL; ERTURK, 2013; SUPARNO *et al.*, 2016; UTSUGI; INO; OSHIKAWA, 2004).

Resultados de estudos mostram ainda que para períodos de maior volatilidade do mercado, estes desvios são mais eminentes, apresentando correlações com maiores magnitudes e um comportamento sincronizado do mercado (PLEROU *et al.*, 2002; KULKARNI; DEO, 2007; DALKIR, 2009; SANDOVAL JUNIOR; FRANCA, 2012). Sandoval Junior e Franca (2012) analisaram a correlação dos mercados financeiros mundiais com relação a crises que afetaram o mercado financeiro mundial ocorridas nas últimas três décadas (segunda-feira negra de 1987, crise russa de 1998, estouro da bolha ponto com em 2001 e 11 de setembro de 2001 e a crise das hipotecas *subprime* de 2008).

A crise financeira de 2008, também conhecida como crise do *subprime*, teve origem nos Estados Unidos da América e se alastrou por

quase todos os mercados financeiros do mundo (KRUGMAN, 2009; SENSOY; YUKSEL; ERTURK, 2013), motivando estudos sobre a estrutura de correlação entre mercados financeiros em todo o mundo (SENSOY; YUKSEL; ERTURK, 2013).

A crise asiática, por sua vez, foi uma crise que se iniciou em alguns países da Ásia a partir de meados de 1997 e contaminou vários mercados financeiros pelo mundo todo, afetando com maior força a países emergentes como o Brasil e Rússia por exemplo, os quais sofreram fortes quedas em suas bolsas e viveram um período de recessão econômica (REIS, 2019; ZAHAR, 2005).

A crise asiática, então, apesar de ter atingido proporções mundiais, afetou de forma mais impetuosa a países emergentes, diferente do que aconteceu com a crise do *subprime*, a qual teve início em um país desenvolvido, Estados Unidos da América, e contaminou todas as outras bolsas mundiais de uma mesma forma. Não foi possível encontrar na literatura uma análise do mercado brasileiro durante a crise asiática a partir da aplicação da metodologia TMA.

Sandoval Junior, Bortoluzzo e Venezuela (2014) teve como proposta analisar os efeitos de dois métodos de filtragem (TMA e Modelo de Índice Único (do inglês *Single-Index Model* - SIM)) da matriz de correlação de retornos de ações do mercado brasileiro sobre a previsão de riscos de um portfólio dentro e fora do período da crise do *subprime*. Os autores concluíram que o método de filtragem baseada na TMA falhou ao prever riscos principalmente em anos de crise. No entanto, tais autores utilizaram uma abordagem de filtragem a partir da TMA diferente da que é aplicada no presente estudo. No trabalho aqui apresentado utilizou-se a metodologia desenvolvido por Plerou, Gopikrishnan, Rosenow, Amaral, Guhr e Stanley (PLEROU *et al.*, 2002) a qual será apresentada detalhadamente seção 3 deste trabalho. Desta forma, não foi possível encontrar na literatura a aplicação da metodologia aqui utilizada para o mercado de ações brasileiro.

Neste trabalho, os conceitos e métodos da TMA, juntamente com algumas técnicas de análise multivariada, serão aplicados na análise de

correlações entre ativos no mercado financeiro brasileiro no contexto da crise do *subprime* e crise asiática e suas vizinhanças temporais. Optou-se por verificar a aplicação da TMA em dois casos de crise financeira que afetaram o mercado financeiro brasileiro para evitar que os resultados encontrados em um único período possam ser um “falso-positivo”, uma vez que será possível comparar os resultados obtidos para dois períodos distintos. Desta forma, as informações extraídas das correlações empíricas poderão ser comparadas às propriedades de matrizes de correlação aleatórias para dois períodos diferentes contendo crises com diferentes características que afetaram o Brasil, levando a conclusões mais generalizadas sobre a aplicação deste método sobre períodos de crises de âmbitos internacionais que afetaram o mercado de ações brasileiro. Pretende-se, desta forma, verificar se a TMA é potente o suficiente para capturar alterações no mercado antes mesmo de uma crise propriamente dita ocorrer e para capturar a movimento de retomada do mercado a um comportamento rotineiro após uma crise.

Os resultados deste estudo mostram uma alteração substancial e consistente do comportamento das matrizes de correlação, dos autovalores e dos autovetores em períodos de crise. Também mostram que o período que antecede uma crise apresenta sintomas que poderiam predizer tal crise. Mostram ainda que no período posterior à crise há um movimento do mercado na direção de uma recuperação da crise. No entanto, essa recuperação não é total, visto que se percebem nos dados analisados características persistentes da crise superada. Esse efeito seria uma marca das crises nas correlações das séries temporais que é observável pela aplicação da TMA. O estudo também apresenta a sugestão de uma carteira de investimentos equilibrada, mesmo em períodos de crise financeira, por meio de Análises de Componentes Principais e métodos de agrupamento.

1.1 Justificativa

Duas das grandes áreas de conhecimento da Engenharia de Produção são a Pesquisa Operacional e a Engenharia Econômica (ABEPRO, 2008) as quais são apresentadas na Figura 1 e relacionadas às respectivas subáreas

abrangentes neste trabalho.

Figura 1 - A Pesquisa Operacional e Engenharia Econômica e suas subáreas.

Pesquisa Operacional	Engenharia Econômica
<ul style="list-style-type: none"> • Programação Matemática • Processos Decisórios • Processos Estocásticos • Inteligência Computacional 	<ul style="list-style-type: none"> • Gestão Econômica • Gestão de Investimentos • Gestão de Riscos

Fonte: Fonte: Autor. Baseado em ABEPRO (2008).

Pelo exposto na Figura 1, esta pesquisa enquadra-se entre a Pesquisa Operacional e a Engenharia Econômica e propõe analisar o espectro de correlação de ações no mercado brasileiro utilizando a Teoria de Matrizes Aleatórias no período da crise do *subprime*, crise asiática e suas vizinhanças temporais.

As características do mercado de ações brasileiro são interessantes uma vez que este é um país em desenvolvimento (ISI, 2018), em 2019 foi o sexto maior destinatário de Investimentos Estrangeiros Diretos no mundo (IEDI, 2020) e a bolsa de valores brasileira, B3 (Brasil Bolsa Balcão), é uma das principais empresas de infraestrutura do mercado financeiro do mundo, uma das maiores em valor de mercado com posição global de destaque no setor de bolsa de valores (B3, 2017).

As crises no mercado de ações podem ser extremamente prejudiciais e contagiosas aos outros processos que se desenvolvem em uma economia (KUMAR; DEO, 2015; LAEVEN; VALENCIA, 2008). Por exemplo, as contaminações de variáveis reais pela crise financeira do *subprime* ocorrida a partir do ano de 2008 pôde ser observada em escala mundial e foi a pior crise financeira desde a grande depressão dos anos 1930 (KRUGMAN, 2009; KUMAR; DEO, 2015). Assim sendo, muitos dos estudos publicados após 2010 que aplicaram o método da TMA, abordaram a análise do espectro da crise mundial de 2008 (FENN *et al.* 2011; KUMAR; DEO, 2012; KUMAR; DEO, 2015; LIM *et al.*, 2014; SANDOVAL JUNIOR; BORTOLUZZO; VENEZUELA, 2014; SANDOVAL JUNIOR; FRANCA, 2012; SENSOY; YUKSEL; ERTURK, 2013). Em particular, o trabalho de Kumar e Deo (2012) mostrou

que houve mudanças substanciais na estrutura de correlação entre mercados financeiros pelo mundo durante a crise financeira de 2008. Assim, é interessante para a sociedade evitar grandes prejuízos a partir de modelos que possam indicar alterações significativas do comportamento do mercado acionário.

O mercado financeiro é tido como um sistema complexo, não existindo um algoritmo padronizado para calcular a correlação entre duas empresas i e j por basicamente três razões (PLEROU *et al.*, 2000; PLEROU *et al.*, 2002; SANDOVAL JUNIOR; BORTOLUZZO; VENEZUELA, 2014):

- mesmo demonstrando correlação direta ou indireta, a interação entre elas é de natureza não-trivial, isso porque as correlações entre as empresas não ocorrem apenas entre pares, mas podem ocorrer entre grupos, sendo que nem sempre estes grupos são fixos, podendo uma nova empresa entrar ou alguma sair destes grupos ou pares de correlações com o passar do tempo;
- a disponibilidade de registros de cada empresa i é finita (no nosso caso, cotação dos preços das ações) e com uma taxa de dados históricos diários relativamente baixa, o que dificulta a quantificação das correlações entre as empresas;
- notícias, mercados externos, crises e os próprios preços das ações interagem para orientar o preço de uma ação.

Todas estas razões citadas geram ruídos nas estatísticas sobre correlações entre empresas, o que tornam a análise de séries temporais financeiras um desafio não trivial e não determinístico, ou seja, as interações são geralmente muito complexas para serem acomodadas por qualquer modelo econométrico (SANDOVAL JUNIOR; BORTOLUZZO; VENEZUELA, 2014).

Assim sendo, é previsível que a matriz de correlação gerada a partir de ações do mercado brasileiro tenha contribuições de origem aleatória. Portanto, se faz necessário um método que separe a correlação aleatória da correlação genuína, o qual carrega informações acerca da natureza do relacionamento entre os diferentes papéis analisados, mesmo durante

períodos de crises financeiras. O método TMA pode ser aplicado para investigar a estrutura da matriz de correlação em séries temporais financeiras globais antes durante e após uma crise financeira (KUMAR; DEO, 2015). Dessa forma, os conceitos e métodos da TMA serão aplicados durante a análise das correlações para dois períodos diferentes estudados neste trabalho, dentre os quais estão contempladas duas crises financeiras diferentes as quais tiveram impactos na bolsa de valores brasileira, possibilitando conclusões mais generalizadas sobre a aplicação do método da TMA sobre períodos de crises de âmbitos internacionais que afetaram o mercado de ações brasileiro.

1.2 Objetivo

Este estudo tem como objetivo a análise do espectro de correlação de ações da bolsa de valores brasileira utilizando a TMA, juntamente com algumas técnicas de análise multivariada, no período da crise do *subprime*, da crise asiática e de ambas as vizinhanças temporais.

Objetiva-se especificadamente:

- comparar dados empírica gerados a partir de séries temporais financeiras de ações do mercado brasileiro com as propriedades universais da TMA conhecidas para matrizes aleatórias normalmente distribuídas;
- verificar a possibilidade de dados empíricos gerados a partir de séries temporais financeiras de ações do mercado brasileiro representarem índices representativos no mercado brasileiro;
- fornecer uma sugestão de composição de portfólio equilibrado através da análise dos dados utilizando ferramentas de análise multivariada.

Para tanto a ferramenta Matlab será utilizada para o desenvolvimento de algoritmos baseados nestes objetivos, visando uma aplicação prática.

Será verificado não somente o período do *subprime* no contexto brasileiro entre 15/09/2008 e 31/12/2009 e o período da crise asiática no contexto brasileiro entre 23/10/1997 e 31/12/1999, como também suas vizinhanças temporais.

1.3 Características da Pesquisa

Este estudo pode ser classificado como aplicado, quantitativo e experimental (DA SILVA, 2005).

A base conceitual desta pesquisa foi estruturada a partir de documentos obtidos das seguintes bases de dados científicas: *Web of Science; Scopus; Google Scholar*. As principais palavras-chave utilizadas foram: *Random Matrix Theory; Cross-Correlations; Econophysics; Temporary Series; Financial; Subprime; Crisis; Crises; BM&FBovespa; B3; Brazil; Ibovespa*. Além de consultas em livros, dissertações, teses e notícias jornalísticas que puderam ser acessadas.

1.4 Estrutura da Tese

Este trabalho compreende 6 Capítulos. O Capítulo 1 introduz a importância de um método alternativo para tratamento de séries temporais financeiras, apresentando os motivos que justificam a aplicação da TMA nesta análise e os objetivos e características deste trabalho. O Capítulo 2 apresenta a revisão da literatura, introduzindo os principais conceitos que serão tratados neste trabalho, como o mercado de ações no Brasil, a crise do *subprime*, a crise asiática e russa, minimização de riscos de portfólio, noções de matemática e análise multivariada e a Teoria de Matrizes Aleatórias. No Capítulo 3 descreve-se a metodologia que foi aplicada para a análise dos dados. O Capítulo 4 traz os detalhes sobre os dados que foram utilizados para esta análise bem como os tratamentos prévios aplicados a estes. No Capítulo 5 são apresentados os resultados obtidos e, por fim, no Capítulo 6 apresenta-se as conclusões.

2. Revisão da Literatura

Este capítulo apresenta a fundamentação teórica para os principais assuntos abordados neste trabalho. A seguir será introduzido conceitos sobre o Mercado de Ações no Brasil, a Crise do *Subprime*, a Crise Asiática e Russa, Minimização de Riscos de Portifólio, Noções de conceitos básicos sobre a matemática análise multivariada aplicadas e sobre a Teoria de Matrizes Aleatórias.

2.1 Mercado de Ações no Brasil

Em março de 2017 as empresas BM&FBOVESPA S.A. (Bolsa de Valores, Mercadorias e Futuros) e Cetip S.A. (Mercados Organizados) se fundiram, resultando na B3 S.A. (Brasil, Bolsa, Balcão), que é uma das maiores empresas provedoras de serviços financeiros do Brasil (B3, 2017; BM&FBOVESPA, 2016). Em 2019 o Brasil foi o sexto maior destinatário de Investimentos Estrangeiros Diretos no mundo (IEDI, 2020). A B3 realiza todas as atividades de negociação de ativos no Brasil e está entre as maiores bolsas do mundo em valor de mercado, negociando ativos de aproximadamente 450 empresas as quais integram os índices Ibovespa, IBrX-50, IBrX e Itag, entre outros (BM&FBOVESPA, 2016). Bolsas de valores são “[...] ambientes físicos ou virtuais organizados para compra e venda de títulos e valores mobiliários.” (B3, 2017).

A compra e venda de ações são investimentos de renda variável que podem ser realizadas por qualquer pessoa. Porém este tipo de investimento envolve riscos de flutuação no valor investido e seu potencial de alta rentabilidade se dá no longo prazo (BM&FBOVESPA, 2016). “Ações são pequenas partes de uma empresa. Quando uma Cia se torna aberta, o patrimônio dela é dividido em várias cotas, que são distribuídas para os investidores.” (BM&FBOVESPA, 2016).

O possuidor de ações pode compra-las ou vende-las a qualquer momento, porém, para lucrar neste tipo de mercado recomenda-se seguir algumas precauções, como conhecer a empresa que receberá o investimento, conhecer o mercado onde ela atua, certificar-se da credibilidade desta e avaliar a possibilidade de lucratividade no futuro

(BM&FBOVESPA, 2016). Uma das formas de obter lucro no mercado de ações é com a alta das ações adquiridas, sendo que o preço da ação é formado pelo fluxo de oferta e procura de cada papel (BM&FBOVESPA, 2016).

Existem dois modelos principais que regem a compra e venda de ações no mercado financeiro: modelo fundamentalista e modelo técnico. O modelo fundamentalista analisa aspectos da empresa em questão e aspectos macroeconômicos, já o modelo técnico se fundamenta em dados históricos a respeito da empresa, tentando, assim, inferir o futuro a partir de comportamentos passados (SILVA; SAPATTI; FLAUZINO, 2010). O histórico de preços de uma determinada ação, portanto, produz uma série temporal não-linear, “[...] cuja previsão de comportamento é de fundamental importância para negociações futuras” (SILVA; SAPATTI; FLAUZINO, 2010, p. 302).

O Índice da Bolsa de Valores de São Paulo (Ibovespa), criado em 1968, é o indicador do desempenho (variação) médio diário das ações de maior negociabilidade e representatividade do mercado brasileiro negociadas na B3, resultado de uma carteira teórica de ativos composto por, aproximadamente, 60 papéis de empresas (BM&FBOVESPA, 2016; SANDRONI, 1999). A negociabilidade é o principal fator para a utilização do Ibovespa como índice de mercado (LEITE; SANVICENTE, 1994) uma vez que este índice “[...] representa a tendência geral dos preços das ações negociadas na Bolsa” (SANDRONI, 1999, p. 288).

2.2 Crise Financeira

As crises no mercado de ações são extremamente prejudiciais aos outros processos que se desenvolvem em uma economia (KUMAR; DEO, 2015). Crises financeiras podem ser prejudiciais e contagiosas, espalhando-se rapidamente, e historicamente levam as economias afetadas a recessões profundas (LAEVEN; VALENCIA, 2008).

Crises financeiras são muito diversas em suas origens e nas suas consequências (KRUGMAN, 2009). Entre as muitas causas das crises financeiras está uma combinação de políticas macroeconômicas

insustentáveis (incluindo grandes déficits em conta corrente e dívida pública insustentável), booms excessivos de crédito, grandes entradas de capital e fragilidades de balanço, combinadas com paralisia política devido a uma variedade de políticas e restrições econômicas (LAEVEN; VALENCIA, 2008).

Segundo Kenton e Scott (2020) em uma crise financeira, os preços dos ativos sofrem uma queda acentuada no valor, as empresas e os consumidores não conseguem pagar suas dívidas e as instituições financeiras enfrentam escassez de liquidez. Ainda segundo estes autores, uma crise financeira costuma estar associada a um pânico ou corrida a um banco durante a qual os investidores vendem ativos ou retiram dinheiro de contas de poupança por temerem que o valor desses ativos caia se eles permanecerem em uma instituição financeira. Outras situações que podem ser rotuladas de crise financeira incluem o estouro de uma bolha financeira especulativa, um *crash* do mercado de ações, um padrão soberano ou uma crise cambial, podendo ser limitada a bancos ou espalhar-se por uma única economia regional ou, ainda, espalhar-se por economias em todo o mundo (KENTON; SCOTT, 2020).

Recessão é o “[...] declínio da atividade econômica, caracterizada por queda da produção, aumento do desemprego, diminuição da taxa de lucros e crescimento dos índices de falências e concordatas” (SANDRONI, 1999, p. 512) a qual, quando se estende de forma prolongada, configura uma crise econômica ou uma depressão. Uma crise econômica, por sua vez, está atrelada a um desequilíbrio entre produção e consumo em setores isolados da produção (SANDRONI, 1999). Já depressão econômica está atrelada a um declínio acelerado da produção, gerando queda nos lucros, perda do poder aquisitivo da população e desemprego (SANDRONI, 1999).

“As crises são classificadas em endógenas (crises de superprodução, venda, crédito e especulação) e exógenas (de causas não-econômicas, como guerras, desastres naturais e epidemias)” (SANDRONI, 1999, p. 142). Na economia capitalista são mais comuns crises relacionadas a superprodução, venda, crédito e especulação (SANDRONI, 1999).

Durante um período de crise, problemas macroeconômicos se

espalharam de forma quase homogênea entre as empresas dos mercados afetados, o que leva a um comportamento sincronizado do mercado. Isto porque é sabido que em períodos de maior volatilidade do mercado, a tendência é ter correlações proeminentemente mais fortes quando comparado a períodos calmos, ou seja, o mercado tende a comportar-se como um só durante períodos de crise (PLEROU *et al.*, 2002; KULKARNI; DEO, 2007; DALKIR, 2009; SANDOVAL JUNIOR; FRANCA, 2012).

Recessão técnica é um termo utilizado para caracterizar o período de dois trimestres consecutivos de queda no PIB, calculados contra o trimestre imediatamente anterior (GUIMARÃES, 2020). Apesar de não ser um termo de cunho científico, não caracterizando necessariamente uma recessão real, tal situação definida por este termo serve, no entanto, como um “sinal de alerta” para uma economia sendo um termo muito utilizado no meio econômico (G1, 2014).

Após uma crise, é esperado que se tenha uma recuperação. Tratando-se de análise de gráficos de períodos de crise, espera-se observar a formações de diferentes padrões, como por exemplo (GUIMARÃES, 2020):

- Recuperação em V: caracteriza uma recuperação rápida da economia depois de uma queda brusca;
- Recuperação em W: indica uma recuperação após uma retração dos índices econômicos, seguida de uma segunda queda;
- Recuperação em L: quando ocorre uma retração e em seguida uma estagnação da economia;
- Recuperação em \surd : retifica um retorno gradual após um declínio econômico.
- Recuperação em $\sqrt{\quad}$: situação em que ocorre uma rápida recuperação após uma retração e, após, estabiliza-se um platô.

No entanto, o comportamento de um mercado após uma crise muitas vezes não é retomado imediatamente ao seu “ritmo rotineiro”. Este comportamento de um período que segue após uma crise e que continua apresentando resquícios de um comportamento de alta volatilidade vem ao

encontro do que já foi apresentado na literatura e pode ser explicada pelo comportamento dos investidores no mercado financeiro, os quais, após uma crise, passam a enxergar o mercado todo como se existissem altas correlações entre diferentes mercados, o que não é necessariamente verdadeiro, porém acaba por se concretizar por meio do comportamento destes investidores no mercado financeiro, os quais enxergam a volatilidade de um outro mercado como um prelúdio do que virá a acontecer em seu próprio mercado de atuação (SENSOY; YUKSEL; ERTURK, 2013; DALKIR, 2009).

2.2.1 Crise do Subprime

A crise financeira internacional de 2008, evidenciada a partir da crise no mercado imobiliário norte-americano em 2007, conhecida como a crise de 2008 ou crise do crédito hipotecário *subprime* ou ainda simplesmente por crise do *subprime*, foi resultante de uma profunda fragilização na confiança econômica, sentimento este que se alastrou pelos diversos mercados mundiais (KRUGMAN, 2009; LIMA; DEUS, 2013). Esta foi a pior crise financeira desde a grande depressão dos anos 1930 (KRUGMAN, 2009; KUMAR; DEO, 2015).

Nos anos anteriores à crise, o mercado americano era estável, com taxas de juros baixas (KRUGMAN, 2009; LIMA; DEUS, 2013). Por volta de 1998, os bancos dos Estados Unidos da América começaram a conceder empréstimos a muitas pessoas que não apresentarem garantias, mesmo pessoas sem comprovante de renda ou histórico de crédito desfavorável conseguiam obter aprovação de financiamento pelos bancos, pois era possível hipotecar a própria casa como garantia para múltiplos empréstimos. Esse tipo de crédito era conhecido como *subprimes* (IEDI, 2008; LIMA; DEUS, 2013).

O crédito fácil e a disseminação pelo mundo todo de um investimento sem lastro, estavam na raiz da crise financeira de 2008. Esta situação foi originalmente decorrente de uma cadeia de empréstimos imobiliários baseados em devedores insolventes (LIMA; DEUS, 2013). Os investimentos eram baseados na ideia de que quando a dívida da hipoteca fosse paga, o

dinheiro iria para o possuidor do pacote de investimento, o qual receberia o valor com juros (UOL, 2016).

A expansão do crédito imobiliário fácil, sem necessidade de comprovação de renda, atraiu os consumidores, o que causou a valorização dos imóveis e aumento da taxa de juros, desproporcionalmente ao aumento da renda da população que, por sua vez, não conseguiram honrar com suas dívidas de hipoteca, descapitalizando bancos e instituições financeiras (KRUGMAN, 2009; LIMA; DEUS, 2013). Esta situação gerou a maior bolha financeira da história, que quanto estorou, causou a queda de diversos títulos bancários, uma vez que houve uma profunda fragilização na confiança econômica levando investidores a preferirem liquidez. Tal situação tornou diversas instituições financeiras insolventes, levando-as a quebra, e gerando o aumento do desemprego no auge da crise econômica nos Estados Unidos. (BORÇA JUNIOR, TORRES FILHO, 2008; KRUGMAN, 2009; LIMA; DEUS, 2013). Como estes títulos estavam nas mãos de bancos e fundos de investimento do mundo todo, houve um efeito dominó no mercado (IEDI, 2008).

Um dos marcos da crise foi a chamada segunda-feira negra em 15 de setembro de 2008, quando um dos bancos de investimentos mais tradicionais dos Estados Unidos, o Lehman Brothers, foi à falência. As bolsas do mundo todo despencaram e o Ibovespa apresentou a maior desvalorização percentual diária desde 11 de setembro de 2001 (MATUCHESKI; CLEMENTE; SANDRINI, 2009; UOL, 2016).

2.2.2 Crise Asiática e Russa

A crise financeira asiática de 1997 foi um período de recessão econômica que atingiu grande parte dos países da Ásia (inicialmente afetando a Tailândia, Malásia, Taiwan, Indonésia, Filipinas entre outros), e em especial os chamados tigres asiáticos (Hong Kong, Coreia do Sul, Cingapura e Taiwan), o conjunto de países emergentes do Sudeste Asiático que, até meados dos anos 1990, se destacavam pelo seu desenvolvimento econômico acelerado, baseado no crescimento da exportação de bens de

consumo, inclusive bens tecnológicos, investimentos de capital estrangeiro e aproveitamento de mão-de-obra barata (REIS, 2019; REMDE, 2013; ZAHAR, 2005).

No entanto, a partir de julho de 1997, com o crescimento dos salários locais, o endividamento excessivo do setor privado e a incapacidade do setor financeiro de recuperar boa parte dos empréstimos concedidos geraram uma crise de confiança e investidores passaram a enxergar melhores oportunidades em outros países, resultando em uma especulação de investidores, fuga de capital e deflação de ativos financeiros das economias da região, o que os forçou a desvalorizar as suas moedas devido a pressões especulativas, revelando assim a fragilidade do sistema financeiro da Ásia (REMDE, 2013; ZAHAR, 2005). “os países emergentes asiáticos gozavam de credibilidade internacional [...] A abrupta alteração da expectativa internacional quanto ao desempenho de suas economias foi fundamental para a crise ocorrer” (REMDE, 2013).

Para os países mais afetados na Ásia, a crise levou a severas consequências como a forte desvalorização de suas moedas, queda de Bolsas de Valores, interrupção do crescimento econômico, quedas de governos, aumento de problemas sociais devido à alta taxa de desemprego, grandes bancos e empresas em geral decretando falência, companhias financeiras tiveram atividades suspensas por falta de liquidez e governos aumentaram as taxas de juros, o que os levou, a partir de 1998, a pedirem ajuda ao FMI (Fundo Monetário Internacional), e ao G7, grupo que reúne os sete países mais industrializados do planeta (REIS, 2019).

A crise asiática rapidamente se espalhou para outros países, afetando inclusive as principais potências econômicas mundiais de forma indireta o que provocaram o efeito dominó nos mercados do mundo todo e passou, assim, a ser chamada de Contágio Asiático, uma vez que ganhou escala global (FOLHA DE SÃO PAULO, 1997). Esta foi a primeira grande crise da era da economia globalizada (FOLHA DE SÃO PAULO, 2006; REIS, 2019).

Segundo noticiários da época, “A crise econômica foi noticiada em 23 de outubro de 1997, quando a bolsa de Hong Kong sofreu a sua pior baixa.”

(ZAHAR, 2005), levando “[...] ao fechamento das operações em Nova York uma hora antes do normal e provocou a quarta maior baixa na história da Bolsa de Valores de São Paulo [...], o pior desempenho entre todas as Bolsas do mundo.” (FOLHA DE SÃO PAULO, 1997). As baixas nas bolsas de valores da Ásia também afetavam os mercados da Europa, Estados Unidos, América Latina e o Japão, o que levou investidores do mundo todo a buscar investimentos menos arriscados, uma vez que foi considerado que ações de grandes empresas estavam supervalorizadas (ZAHAR, 2005).

A crise financeira na Ásia afetou também as finanças da Rússia. Apesar da desvalorização de várias moedas frente ao dólar no mundo todo durante a crise asiática, o Estado russo manteve o rublo em paridade com o dólar, o que aumentou ainda mais as despesas para manter essa política cambial, uma vez que as despesas para manter a máquina pública já eram altas, herança dos gastos do antigo regime soviético (PINTO, 2021; REMDE, 2013). “Ao mesmo tempo, houve uma queda do preço do petróleo, uma das principais fontes de receita de exportação da economia russa.” (PINTO, 2021). O rublo russo entrou em colapso, houve dificuldades na manutenção de serviços públicos estatais, dificuldades de pagamento de dívidas internas e externas, quebra de bancos e a economia Russa também entrou em recessão, deflagrando na conhecida crise Russa em 17 de agosto de 1998, resultado das fugas de capitais que haviam sido investidos no país (PINTO, 2021; REIS, 2019). A crise resultou em mais desconfiança dos investidores que optaram por trocar rublos por dólar, e o crescente aumento do desemprego no país, o que fez com que o estado, por fim, desvalorizasse sua moeda, fazendo com que investidores voltassem ao país (PINTO, 2021).

Apesar de haver contaminação global durante as crises asiática e russa, houve uma disseminação mais evidente sobre países emergentes como Rússia e Brasil, uma vez que havia um temor de que os demais países emergentes seriam os próximos afetados pela crise asiática, desta forma os investidores retiravam o capital de países que apresentavam semelhanças aos tigres asiáticos (GOLDFAJN, 2000; REMDE, 2013). Desta forma, a crise asiática teve um desenrolar diferente daquele ocorrido durante a crise de

2008, a qual originou-se nos EUA e afetou bolsas do mundo todo sem distinção.

Além do fato de ser um país emergente, o Brasil tinha uma grande dependência de capitais estrangeiros, resultado de uma “abertura” indiscriminada da economia brasileira nos anos 1990, que deixaram entrar crescentes quantias de capital estrangeiro sem controle, o que, associada à sobrevalorização do real frente ao dólar, aumentando o endividamento externo do país e deixando-o vulnerável a crises internacionais (REIS, 2019). Da mesma forma que ocorreu nos países emergentes asiáticos afetados pela crise, o Brasil também sofreu ataques especulativos, que levaram a um aumento das taxas de juros, do desemprego e a um período de recessão econômica (GOLDFAJN, 2000; REMDE, 2013). O Banco Central interveio durante as duas crises, gastando reais para tentar defender o câmbio. Esta estratégia, a qual resultava no aumento na taxa básica de juros e promessa de redução de gastos pelo Brasil, funcionou durante a crise asiática, porém já estavam desgastadas na época da crise russa e somente agravou a situação fiscal do país (REMDE, 2013). Desta forma, o Banco Central do Brasil foi forçado a desvalorizar o real em janeiro de 1999 (REMDE, 2013). As consequências das crises asiática e russa fizeram com que a economia mundial ficasse contida nos anos de 1998 e 1999 (REIS, 2019).

2.3 Minimização de Riscos de Portifólio

Um portfólio, ou cesta ou carteira de investimentos é o conjunto de ativos financeiros de um investidor, podendo ser formado por diferentes tipos de investimentos, como por exemplo ações comercializadas no mercado financeiro, objetivando um retorno financeiro positivo.

Para a teoria financeira de risco é muito importante o conhecimento das correlações entre as diversas ações que potencialmente farão parte de uma carteira. A partir do trabalho de Markowitz (1952) se desenvolveu toda a teoria de otimização de carteiras buscando calibrar o peso de cada ação na composição da carteira com a finalidade de minimizar o risco para um dado nível geral de rentabilidade (BOUCHAUD; POTTERS, 2009; OH, 2011;

SENSOY; YUKSEL; ERTURK, 2013).

A construção de carteiras é baseada na correlação entre os papéis, de forma a minimizar os riscos ao agregar papéis que tendem a se comportar de forma oposta, ou seja, anticorrelacionados, pois assim a oposição do movimento dos preços se compensaria, em certa proporção, reduzindo a volatilidade (BOUCHAUD; POTTERS, 2009). Em outras palavras, quando uma ação ou grupo de ações de um portfólio diversificado sofre uma desvalorização, deseja-se que exista uma segunda ação ou grupo de ações de comportamento diametralmente oposto (sofrendo valorização) para, então, compensar as possíveis perdas sobre tal portfólio.

O risco de um portfólio é avaliado pela volatilidade de cada ação que compõe a carteira e pelas correlações entre pares de ações (SANDOVAL JUNIOR; BORTOLUZZO; VENEZUELA, 2014). Tal perspectiva surgiu a partir do trabalho de Markowitz (1952).

No entanto, apesar de ser um conceito aparentemente intuitivo, existem várias barreiras que torna a composição de um portfólio diversificado um desafio em um contexto prático. Algumas delas são (SANDOVAL JUNIOR; BORTOLUZZO; VENEZUELA, 2014):

- A previsão de risco de uma carteira geralmente é diferente do risco real, uma vez que os dados utilizados para os cálculos de previsão são os antigos retornos das ações (históricos), os quais refletem o mercado como ele era, e não como será, uma vez que eventos futuros não necessariamente ocorrerão da mesma forma como no passado;
- Para minimizar os efeitos de eventos que podem mudar o comportamento de um mercado, não se pode usar dados muito distantes no tempo, desta forma há apenas uma quantidade limitada de dados, o que gera desvios associados ao efeito de amostra finita;
- Uma carteira eficiente deve ser construída a partir de muitas e diversas ações, devendo-se, dessa forma, manter um escopo bastante recente de dados históricos, ampliando ainda mais o problema de finitude de dados devido ao número de retornos diários ser pequeno quando comparado a quantidade de empresas no

portifólio, uma vez que quanto maior a razão entre quantidade de retornos diários e quantidade de empresas, mais precisa será a modelagem;

- As interações entre os muitos elementos de um mercado de ações são muito complexas para serem abrangidas por um modelo econômico. Notícias, mercados externos, crises e os próprios preços das ações interagem para orientar o preço de uma ação.

É de conhecimento comum que no mercado de ações pode haver grandes quedas de índices de mercado e desvalorizações de ações. Entre as diversas maneiras que um investidor tem para se proteger de eventos como estes, principalmente durante períodos de declínio do mercado, é utilizar o mecanismo de venda a descoberto (VAD). Venda a descoberto (do inglês *short selling*) é a atividade de vender ações que o investidor tomou emprestado (geralmente mediante a cobrança de uma comissão), esperando que o preço caia antes de comprá-las de volta e devolvê-las ao dono (geralmente uma corretora financeira), objetivando o lucro ou mesmo a proteção de seu capital por meio da especulação na queda dos mercados financeiros (CAMBRIDGE DICTIONARY, 2020; CHEN, 2020). Ou seja, ao lançar mão de uma VAD, o investidor está apostando na queda do preço de da ação em questão, da mesma forma que ao comprar uma ação este investidor está apostando na valorização desta.

2.4 Noções Básicas de Conceitos Fundamentais

Esta seção é dedicada a destacar alguns pontos importantes de conceitos básicos de matemática e estatística multivariada os quais serão utilizados durante o desenvolvimento deste trabalho.

2.4.1 Matriz Hamiltoniana

Uma matriz H de dimensões $2n \times 2n$ (de polinômio característico par) na forma $H = \begin{pmatrix} E & F \\ G & -E^T \end{pmatrix}$ é chamada de Matriz Hamiltoniana se $E, F, G \in M_n(\mathbb{C})$, sendo $M_n(\mathbb{C})$ o conjunto das matrizes complexas de dimensões $n \times n$, com $F^T = F$ e $G^T = G$ (FABBENDER *et al.*, 1999; IKRAMOV, 2001) onde A^T denota a transposta de uma matriz A .

Assim sendo, compreende-se facilmente as seguintes propriedades para matrizes hamiltonianas: se A e B são matrizes hamiltonianas, então as seguintes matrizes também serão matrizes do tipo hamiltoniana (MEYER; HALL; OFFIN, 2009):

- A^T ;
- αA ($\alpha \in \mathbb{C}$);
- $A \pm B$;

Da mesma forma, pode-se caracterizar o conjunto \mathcal{H} de todas as matrizes hamiltonianas de dimensões $2n \times 2n$ por $\mathcal{H} = \{H \in M_{2n}(\mathbb{C}) \mid (JH)^T = JH\}$, onde $J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$ é ortogonal e antissimétrica, ou seja, $J^{-1} = J^T = -J$ e I_n é a matriz identidade de dimensões $n \times n$ (FABBENDER *et al.*, 1999; IKRAMOV, 2001; MEYER; HALL; OFFIN, 2009). O conjunto \mathcal{H} também pode ser denotado por $sp(2n, \mathbb{C})$ (IKRAMOV, 2001; MEYER; HALL; OFFIN, 2009).

2.4.2 Autovalores e Autovetores

Autovalores e autovetores são conceitos fundamentais para a análise estatística multivariada (JOHNSON; WICHERN, 2002). Considerando uma matriz A simétrica e quadrada de dimensões $k \times k$, esta terá k pares de autovalores λ_i ($i = 1, 2, \dots, k$) e correspondentes autovetores $v_i \neq 0$, se $AV = \lambda V$, onde λ é um vetor de autovalores λ_i de tamanho k e V é uma matriz de autovetores v_i de dimensões $k \times k$ (JOHNSON; WICHERN, 2002). A escolha dos autovetores pode ser tal que sejam perpendiculares entre si e satisfaça $1 = v_1'v_1 = \dots = v_k'v_k$. Os autovetores serão únicos a não ser que exista dois ou mais autovalores iguais entre si (JOHNSON; WICHERN, 2002).

Um autovetor de uma matriz de correlação de retornos de ações do mercado financeiro pode ter diferentes utilidades. Alguns autores trataram este tipo de autovetor como índice de mercado ou índice de setores de mercado (considerando apenas aqueles relacionados aos maiores autovalores que ultrapassaram o espectro da correlação aleatória) (COELHO *et al.* 2008; JOHNSON; WICHERN, 2002; PLEROU *et al.*, 2002; SENSOY; YUKSEL; ERTURK, 2013). Já Sandoval Junior, Bortoluzzo e Venezuela (2014) utilizaram estes autovetores como pesos para o cálculo

de retorno e de risco de um dado portfólio, ou seja, cada componente vai dizer o quanto uma dada ação está influenciando em um dado portfólio (o percentual de papéis pertencentes a uma certa ação de um portfólio).

Dessa forma, sinais opostos entre componentes destes autovetores podem ter diferentes interpretações. Segundo Johnson e Wichern (2002), dois componentes de sinais opostos em um autovetor pode significar o contraste de comportamento entre dois setores do mercado financeiro. Já para Sandoval Junior, Bortoluzzo e Venezuela (2014), pode significar que a ação referente a um componente positivo do autovetor deve ser comprada para compor o portfólio, ou seja, espera-se que a ação sofra uma valorização. Já o sinal negativo, segundo estes mesmos autores, significa que esta ação é um alvo provável para VAD, ou seja, existe a expectativa que a ação sofra uma desvalorização.

2.4.3 Análise de Componentes Principais

Uma das dificuldades inerentes à estatística multivariada é o problema de visualização e, conseqüentemente, de interpretação de dados que possuem muitas variáveis. Felizmente, em conjuntos de dados com muitas variáveis, grupos de variáveis costumam se mover juntos. Uma razão para isso é que mais de uma variável pode estar medindo a mesma dimensão inerente comum (fatores) que governa o comportamento de um dado sistema (HAIR *et al.*, 2009; MATHWORKS, 2020). Assim sendo, pode-se simplificar um certo problema substituindo várias variáveis originais por um novo conjunto menor de variáveis estatísticas (fatores) com uma perda mínima de informação (HAIR *et al.*, 2009; MATHWORKS, 2020).

A Análise de Componentes Principais (ACP) (do inglês *Principal Component Analysis* – PCA), é um método quantitativamente rigoroso para alcançar essa simplificação. Este método gera um novo conjunto de variáveis, chamadas de componentes principais, os quais são uma combinação linear das variáveis originais, ortogonais entre si (não correlacionadas), o que evita a redundância de informações, sendo que o tamanho máximo do conjunto de componentes principais é o mesmo tamanho do conjunto original de variáveis (BOUCHAUD; POTTERS, 2009;

HAIR *et al.*, 2009; JOHNSON; WICHERN, 2002; MATHWORKS, 2020). Segundo Johnson e Wichern (2002), a análise dos componentes principais frequentemente revela relacionamentos que não eram suspeitados anteriormente e, portanto, permite interpretações mais apuradas.

Algebricamente, cada componente principal ortogonal PC_i é um fator da matriz de correlação das variáveis normalizadas originais, os quais resultam de combinações lineares de modo que $PC_i = a_{i1}z_1 + a_{i2}z_2 + \dots + a_{iN}z_N$, onde a_{ij} são os coeficientes para o fator i , ou seja, as cargas fatoriais ou os coeficientes do i -ésimo componente principal, também conhecido como *loadings*, de modo que os a_{ij} são os componentes do autovetor v_i da matriz de correlação das variáveis normalizadas originais. Estes coeficientes multiplicam as variáveis normalizadas originais z_j , as quais são vetores com T observações (HAIR *et al.*, 2009; JOHNSON; WICHERN, 2002).

Geometricamente, existe originalmente uma base não-ortogonal (formado pelas variáveis originais as quais podem estar correlacionados entre si) sendo que cada eixo representa uma variável do problema em questão. Pode-se, desta forma, projetar cada observação sobre estes eixos, de forma que uma nuvem de T pontos é formada sobre o espaço N -dimensional. A ACP rotaciona os eixos originais no espaço N -dimensional para a nova posições (componentes principais). Assim sendo, os componentes principais como um todo formam uma base ortogonal (não correlacionados) para o espaço dos dados, sendo que cada componente principal é um eixo neste espaço. Esses novos eixos representam as direções de maior variação na nuvem N -dimensional de T pontos, sendo que são ordenados de tal forma que o primeiro componente principal apresenta a direção com a máxima variabilidade (ou seja, a dimensão com maior dispersão), o segundo componente aponta para próxima direção com maior variabilidade e assim por diante (HAIR *et al.*, 2009; JOHNSON; WICHERN, 2002; MATHWORKS, 2020). A variância indicada pelo primeiro componente principal é dada pelo maior autovalor associado a matriz de correlação das variáveis originais normalizadas e a variância percentual

explicada por cada componente principal i é calculada dividindo-se o i -ésimo autovalor pela quantidade total de componentes principais existentes (JOHNSON; WICHERN, 2002).

Pode-se, então, representar cada observação no novo espaço ortogonal formado pelos componentes principais (realizando uma combinação linear entre as observações originais normalizadas e os coeficientes dos componentes principais), de forma que uma nova nuvem de T pontos é formada sobre o novo espaço ortogonal. Esta nova nuvem, onde cada ponto T é chamado de *escore*, irá representar a variação entre as observações. O gráfico de *escores* com duas dimensões, por exemplo, representa os *escores* do primeiro e segundo componentes principais (MINITAB, 2020).

Da mesma forma, pode-se transpor as N variáveis para este novo espaço ortogonal, as quais são representadas pelos coeficientes equivalentes, chamados de *cargas fatoriais*, podendo assim, também avaliar a variação entre as variáveis (JOHNSON; WICHERN, 2002; MATHWORKS, 2020). O gráfico de *cargas fatoriais* com duas dimensões, por exemplo, ilustra as *cargas fatoriais* dos primeiros dois componentes principais (MINITAB, 2020).

Assim sendo, ao tomar-se os primeiros k componentes, eles definem um hiperplano k -dimensional que “melhor se ajusta” à nuvem de pontos e fornece uma descrição mais simples da estrutura de uma matriz de correlação (HAIR *et al.*, 2009; JOHNSON; WICHERN, 2002), os quais, juntamente com técnicas de agrupamento, ajudam a visualizar candidatos verdadeiramente anticorrelacionados para a composição de um portfólio equilibrado, o que influencia na redução do risco deste (BOUCHAUD; POTTERS, 2009; MARKOWITZ, 1952; PARTOVI; CAPUTO, 2004).

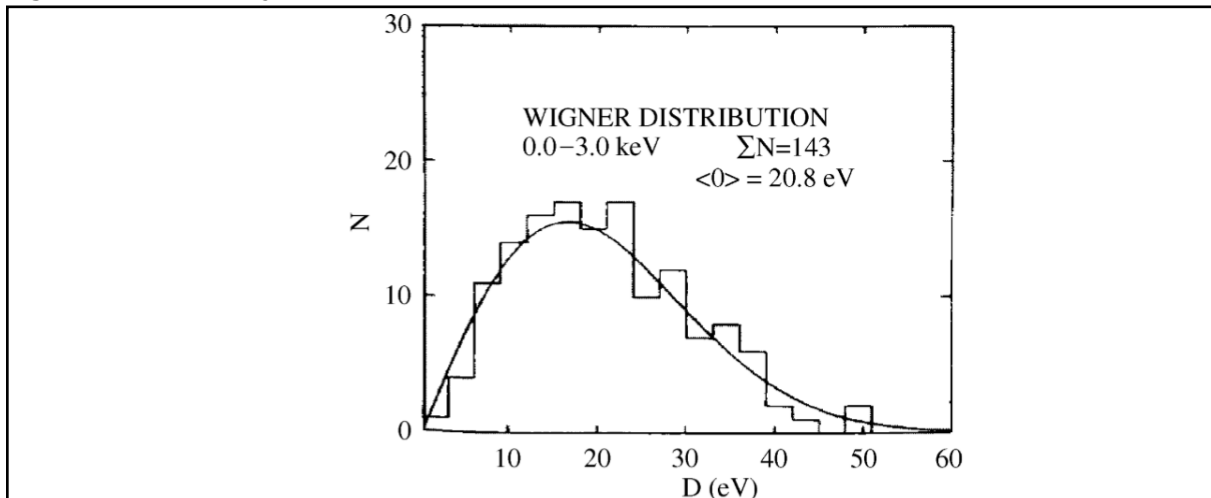
No caso deste estudo, as N variáveis em questão são as ações aqui considerada durante a análise e as T observações são os dias com retornos considerados para cada período estudado, ou seja, a quantidade de retornos diários de uma dada ação.

2.5 Teoria das Matrizes Aleatórias

A TMA foi introduzida em 1928 na matemática estatística por Wishart e ganhou destaque na década de 1950 dentro do contexto da física nuclear com Wigner através de estudos sobre as mudanças de níveis energéticos de núcleos atômicos pesados, onde a natureza precisa da interação entre subunidades é desconhecida (GUHR; MÜLLER-GROELING; WEIDENMÜLLER, 1998; PLEROU, 2002; WIGNER, 1955). Neste cenário, os modelos matemáticos até então existentes não conseguiam explicar a estatística dos níveis energéticos de sistemas quânticos complexos (PLEROU, 2002). O físico Eugene Paul Wigner assumiu, então, como hipótese, que as interações entre tais subunidades eram tão complexas que poderiam ser modeladas como uma aleatoriedade (DYSON, 1962; GUHR; MÜLLER-GROELING; WEIDENMÜLLER, 1998; LIOU; CAMARDA; RAHN, 1972). Neste contexto, os pesquisadores Wigner, Dyson, Mehta, entre outros, definiram uma distribuição teórica através de uma matriz hamiltoniana aleatória, com elementos independentes e identicamente distribuídos (i.i.d.), que era capaz de se ajustar aos dados empíricos, possibilitando predição do comportamento da distribuição de energia de núcleos pesados (GUHR; MÜLLER-GROELING; WEIDENMÜLLER, 1998). Tal distribuição teórica foi batizada de distribuição de Wigner e deu origem à TMA (LIOU; CAMARDA; RAHN, 1972; PLEROU *et al.*, 2000; PLEROU *et al.*, 2002).

A distribuição de Wigner aplicada à área de Física Nuclear é apresentada na Figura 2.

Figura 2 - Distribuição de Wigner: Histograma de Autovalores de um Núcleo Pesado



Fonte: PLEROU *et al.* (2000).

A Figura 2 apresenta um histograma dos níveis de energia (autovalores) em elétron-volts de um núcleo pesado. A linha contínua mostra a distribuição de Wigner (PLEROU *et al.*, 2000).

Para matrizes do tipo hamiltonianas aleatórias, com elementos i.i.d., normalmente distribuídos, de dimensões $M \times M$, no limite de $M \rightarrow \infty$ e reescaladas para espaçamento de autovalores em termos de média local, sabe-se que existe um comportamento estatístico de propriedades universais (PLEROU *et al.*, 2002).

Assim, a aplicação da TMA envolve a comparação entre o comportamento dos autovalores de uma matriz de correlação calculada a partir de dados empíricos e as propriedades universais das matrizes aleatórias (referente a TMA) (PLEROU *et al.*, 2002). Entre essas propriedades universais tem-se a distribuição dos espaçamentos dos vizinhos mais próximos (distribuição de Wigner) e uma outra distribuição destinada aos espaçamentos de longo alcance, determinada pela variância do número de autovalores em uma determinada janela móvel (alcance de autovalores) (PLEROU *et al.*, 2002).

A comparação entre o comportamento dos autovalores empíricos e o comportamento universal da TMA fornece informações sobre o sistema analisado (PLEROU *et al.*, 2000; PLEROU *et al.*, 2002; SENSOY; YUKSEL; ERTURK, 2013). Assim, esta técnica pode ser pensada como um teste de hipótese em que a H_0 representa o comportamento i.i.d. (independentes e

identicamente distribuídas) e a hipótese alternativa é a existência de uma informação significativa nos dados empíricos (PLEROU *et al.*, 2002). Pode-se dizer que esta técnica serve como um filtro para as correlações, de modo a promover uma separação entre ruído e as correlações realmente significativas, por meio da análise espectral dos autovalores das matrizes (BOUCHAUD; POTTERS, 2009; DYSON, 1962; GUHR; MÜLLER-GROELING; WEIDENMÜLLER, 1998; LIVAN; REBECCHI, 2012; KULKARNI; DEO, 2007; PLEROU *et al.*, 2000; PLEROU *et al.*, 2002).

Desta forma, a TMA vem auxiliar na compreensão das complexas interações entre elementos de diferentes sistemas, buscando prever a possível aleatoriedade (resultado de correlações não estacionárias e séries temporais finitas) existente nessas interações (PLEROU *et al.*, 2000).

2.6 Utilização da TMA na análise de séries temporais financeiras

Em diversos modelos do comportamento econômico-financeiro existe o interesse em determinar as relações entre variáveis, seja para construção de teorias explicativas, realização de previsões, otimização etc. Porém, muitas das situações colocadas apresentam comportamentos não lineares, em que técnicas tradicionais não são indicadas, pois captam apenas correlações lineares.

Além do aspecto de não linearidade, o comportamento mais frequente observado em séries temporais econômico-financeiras é a não estacionariedade (PLEROU *et al.*, 2002). Novamente, existem diversas técnicas em econometria que são capazes apenas de abordar com sucesso séries estacionárias. Para analisar séries temporais com características de não estacionariedade e não linearidade, se faz necessária a aplicação de técnicas alternativas.

Uma das principais dificuldades nesses estudos é a natureza complexa e desconhecida das interações subjacentes ao mercado financeiro (PLEROU *et al.*, 2002). A análise das correlações entre ações não é feita apenas de par em par, mas envolvem grupos de ações. O desafio se dá pelas constantes mudanças nas interações entre estes grupos de ações ao passar do tempo, sendo que nem sempre estes grupos são fixos, podendo uma

nova empresa entrar ou alguma sair destes grupos, além da interferência do tamanho finito da série temporal disponível, de notícias, dos mercados externos, de crises e dos próprios preços das ações, todos estes efeitos interagindo para orientar o preço de uma ação (PLEROU *et al.*, 2000; SANDOVAL JUNIOR; BORTOLUZZO; VENEZUELA, 2014). Devido a tal complexidade, dificilmente tem-se um modelo econométrico poderoso o suficiente para prever com precisão o comportamento das diversas ações no mercado financeiro, o que gera “medições sujas” nos resultados (PLEROU *et al.*, 2002; SANDOVAL JUNIOR; BORTOLUZZO; VENEZUELA, 2014). Sendo assim, mesmo que fosse possível utilizar uma série temporal longa o suficiente, de forma que o problema causado pelo tamanho finito da série fosse irrelevante, ainda haveria problemas com a não-estacionariedade da correlação (PLEROU *et al.*, 2002).

Assim sendo, a geração de matriz de correlação gerada a partir de dados empíricos incorpora todos os efeitos colaterais de uma aplicação prática já discutidos. Desta forma, parte dessas correlações serão ruídos provenientes de medições sujas ou resultantes de aleatoriedades, o que dificulta a análise e inviabiliza conclusões sobre as implicações dos coeficientes da matriz de correlação empírica (PLEROU *et al.*, 2002). Isto posto, a remoção de alguns destes efeitos da matriz de correlação trará maior clareza sobre a análise do comportamento das ações no mercado financeiro (SANDOVAL JUNIOR; BORTOLUZZO; VENEZUELA, 2014). Alguns métodos de filtragem da matriz de correlação, com o intuito de separar informações genuínas de ruídos, foram propostos pela literatura, porém uma abordagem mais efetiva foi o uso da metodologia proposta pela TMA (LIVAN; REBECCHI, 2012).

Na década de 1990, começaram a surgir, com frequência significativa, trabalhos publicados que empregavam a TMA ao estudo do comportamento de matrizes de correlação de dados financeiros (BOUCHAUD; POTTERS, 2009; CONLON; RUSKIN; CRANE, 2009; SENSOY; YUKSEL; ERTURK, 2013). Os físicos estavam interessados no mecanismo que gerava invariância de escala em séries financeiras, e para isso era necessário

entender a correlação entre as diferentes unidades que compõem esse sistema (PLEROU *et al.*, 2000).

No contexto da análise de dados financeiros, a TMA tem sido útil na análise de propriedades universais de correlações entre diferentes ações (KULKARNI; DEO, 2007). Resultados de estudos apresentam frequentemente que desvios do comportamento de matrizes de correlação geradas a partir séries históricas financeiras, comparadas a matrizes de correlação aleatórias, fornecem informações genuínas sobre o comportamento do mercado (FENN *et al.* 2011; KULKARNI; DEO, 2007; KUMAR; DEO, 2015; KUMAR; DEO, 2012; LIM *et al.*, 2009; LIM *et al.*, 2014; MEIYING; BIN, 2011; PLEROU *et al.*, 1999; PLEROU *et al.*, 2000; PLEROU *et al.*, 2001; PLEROU *et al.*, 2002; SANDOVAL JUNIOR; BORTOLUZZO; VENEZUELA, 2014; SANDOVAL JUNIOR; FRANCA, 2012; SENSOY; YUKSEL; ERTURK, 2013; SUPARNO *et al.*, 2016; UTSUGI; INO; OSHIKAWA, 2004). Resultados de estudos mostram ainda que para períodos de maior volatilidade do mercado, estes desvios são mais eminentes, apresentando correlações com maiores magnitudes e um comportamento sincronizado do mercado (PLEROU *et al.*, 2002; KULKARNI; DEO, 2007; DALKIR, 2009; SANDOVAL JUNIOR; FRANCA, 2012).

Alguns dos autores mais ativos nesta linha de pesquisa são Lim *et al.* (LIM *et al.* 2009; LIM *et al.* 2014), Sandoval Junior (SANDOVAL JUNIOR; FRANCA, 2012; SANDOVAL JUNIOR; BORTOLUZZO; VENEZUELA, 2014), Kumar e Deo (KUMAR; DEO, 2015; KUMAR; DEO, 2012) e principalmente Plerou que aplicaram os conceitos da TMA para filtrar as correlações, removendo ruídos (PLEROU *et al.*, 1999; PLEROU *et al.*, 2000; PLEROU *et al.*, 2001; PLEROU *et al.*, 2002).

Para análise da matriz de correlação, estudos anteriores apresentaram, juntamente com a aplicação da TMA, vários métodos estatísticos, como Análise de Componentes Principais (FENN *et al.* 2011; KUMAR; DEO, 2012; KUMAR; DEO, 2015; LIM *et al.*, 2014; MEIYING; BIN, 2011; PENG *et al.*, 2019; PLEROU *et al.*, 1999; PLEROU *et al.*, 2000; PLEROU *et al.*, 2002; SENSOY; YUKSEL; ERTURK, 2013; SUPARNO *et al.*,

2016), Regressão (EL ALAOUI; BOURI; AZOURY, 2020; SANDOVAL JUNIOR; BORTOLUZZO; VENEZUELA, 2014), Topologia de Rede (LIM *et al.*, 2014), Multifractalidade (KUMAR; DEO, 2015), Análise de Agrupamento (KUMAR; DEO, 2012; SENSOY; YUKSEL; ERTURK, 2013), Teoria de Redes (LIM *et al.*, 2009), Modelo de Regime de Comutação e Algoritmo de Gradiente Descendente (EL ALAOUI; BOURI; AZOURY, 2020), Matriz de Correlação Ampliada (SANDOVAL JUNIOR, 2014) e Modelo de Alocação de Portfólio de Markowitz (PENG *et al.*, 2019; SANDOVAL JUNIOR; BORTOLUZZO; VENEZUELA, 2014; SANDOVAL JUNIOR; FRANCA, 2012).

García (2016) usou a abordagem da TMA para analisar a estrutura da matriz de correlação de uma coleção de tweets públicos (do serviço de rede social Twitter) e a série temporal de retorno correspondente associada a índices financeiros globais.

Sandoval Junior (2014), Sensoy, Yuksel e Erturk (2013), Kumar e Deo (2012) e Peng *et al.* (2019) realizaram uma análise da correlação entre índices internacionais de mercado aplicando a TMA em períodos que inclui a crise de 2008. Sandoval Junior e Franca (2012) analisaram a correlação dos mercados financeiros mundiais com relação a crises que afetaram o mercado financeiro mundial ocorridas nas últimas três décadas (segunda-feira negra de 1987, crise russa de 1998, estouro da bolha ponto com em 2001 e 11 de setembro de 2001 e a crise das hipotecas *subprime* de 2008).

O artigo de Sandoval Junior, Bortoluzzo e Venezuela (2014) teve como proposta analisar os efeitos de dois métodos de filtragem da matriz de correlação de retornos de ações do mercado brasileiro sobre a previsão de riscos de um portfólio dentro e fora do período da crise do *subprime*. Os métodos foram chamados de Limpeza da Matriz de Correlação por TMA e Modelo de Índice Único (do inglês *Single-Index Model - SIM*). Porém os autores concluíram que o método de filtragem baseada na TMA, o qual foi utilizado para limpar e remover ruídos da matriz de correlação, falhou ao prever riscos principalmente em anos de crise e que, mesmo a estratégia com melhor resultado sendo a combinação de ambas as filtragens, o método de limpeza por TMA não surtiu efeitos efetivos, sendo que a

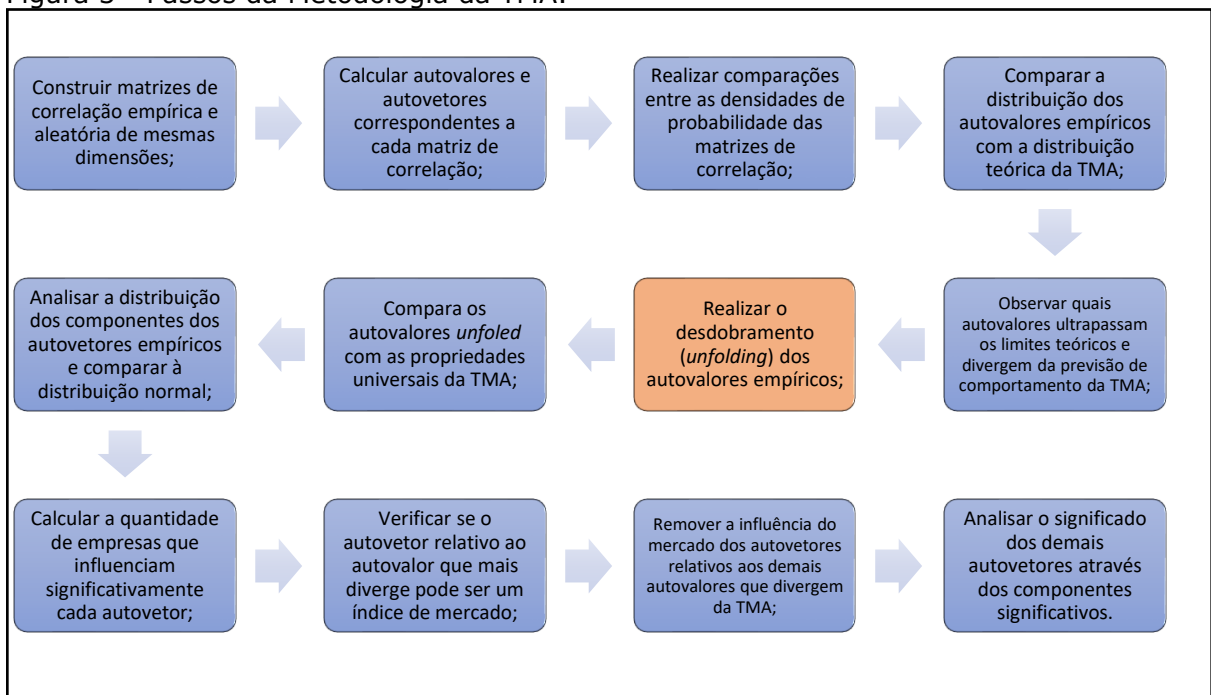
aplicação da filtragem pelo método SIM já era o suficiente para evidenciar melhorias na previsão de riscos de portfólio. No entanto, tais autores utilizaram uma abordagem diferente de filtragem a partir da TMA, construindo uma nova matriz de correlação filtrada a partir de uma matriz diagonal onde os elementos da diagonal são os autovalores da matriz de correlação original, mas agora com todos os autovalores correspondentes ao ruído (dentro da região de Wishart) substituídos por sua média. No trabalho aqui apresentado utilizou-se a metodologia de TMA desenvolvido por Plerou, Gopikrishnan, Rosenow, Amaral, Guhr e Stanley (PLEROU *et al.*, 2002) o qual também aplica o método SIM como filtragem da matriz de correlação e, como será visto nos resultados, traz uma maior clareza para a análise. Esta metodologia será detalhada passo-a-passo na seção 3 deste trabalho.

3. Metodologia

Os conceitos e métodos de TMA desenvolvidos por Plerou, Gopikrishnan, Rosenow, Amaral, Guhr e Stanley (PLEROU *et al.*, 2002) foram utilizados na análise dos dados no presente trabalho, além de algumas técnicas de análise multivariada. Esta seção será dedicada a descrever o passo a passo e esclarecer como a metodologia foi aplicada neste trabalho.

Este método da TMA de Plerou *et al.* (2002), o qual pode ser utilizado para análise de séries temporais financeiras, segue alguns passos, conforme pode ser observado de forma simplificada na Figura 3 e os quais serão detalhados a seguir. O passo da metodologia que está destacados de laranja foi realizado de forma alternativa uma vez que não foi possível encontrar na literatura uma descrição matemática aplicada a dados empíricos e, por tanto, este representa uma contribuição originais deste trabalho a metodologia da TMA já conhecida, a qual será detalhadamente apresentada na Seção 3.5.

Figura 3 - Passos da Metodologia da TMA.



Fonte: Autor.

Primeiramente deve-se construir as matrizes empíricas de correlação para os períodos que serão analisados, calculando-se, na sequência, os

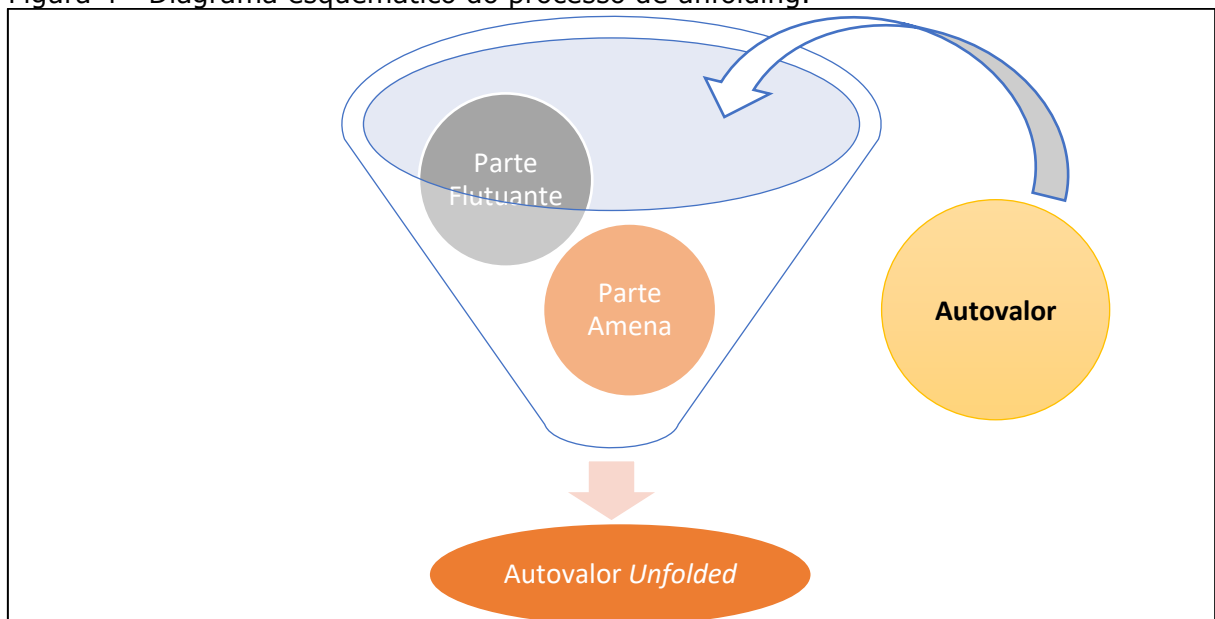
autovalores e autovetores correspondentes. Analogamente, constrói-se matrizes de correlação aleatória a partir de matrizes com valores aleatórios i.i.d. com as mesmas dimensões das matrizes de retornos empíricas, e calcula-se seus respectivos autovalores e autovetores. Assim, é possível fazer comparações entre os diferentes períodos a partir da distribuição de suas matrizes de correlação empírica além de comparar com o comportamento de uma matriz de correlação aleatória.

A TMA descreve uma distribuição $\rho_{rm}(\lambda)$ que é esperada para autovalores aleatórios. Dessa forma é possível realizar uma comparação entre as distribuições dos autovalores empíricos e o comportamento aleatório esperado, uma vez que o comportamento dos autovalores tem implicações na identificação da natureza das correlações empíricas (KULKARNI; DEO, 2007). Nesta etapa tem-se o interesse em destacar a parte da distribuição empírica que se desvia do comportamento aleatório, a qual carrega informações genuínas sobre a correlação segundo vários estudos (FENN *et al.* 2011; KULKARNI; DEO, 2007; KUMAR; DEO, 2015; KUMAR; DEO, 2012; LIM *et al.*, 2009; LIM *et al.*, 2014; MEIYING; BIN, 2011; PLEROU *et al.*, 1999; PLEROU *et al.*, 2000; PLEROU *et al.*, 2001; PLEROU *et al.*, 2002; SANDOVAL JUNIOR; BORTOLUZZO; VENEZUELA, 2014; SANDOVAL JUNIOR; FRANCA, 2012; SENSOY; YUKSEL; ERTURK, 2013; SUPARNO *et al.*, 2016; UTSUGI; INO; OSHIKAWA, 2004). No entanto, esta comparação por si só não é o bastante para suportar quaisquer conclusões (PLEROU *et al.*, 2002).

Portanto, é de interesse a investigação do comportamento dos autovalores empíricos para verificar o quanto eles se ajustam às propriedades universais da TMA, ou seja, a distribuição dos espaçamentos dos vizinhos mais próximos e a estatística de alcance de autovalores (PLEROU *et al.*, 2002). No entanto, para que esta análise seja possível, faz-se necessário ordenar e realizar um procedimento conhecido como *unfolding* (ou em português, desdobramento) o qual mapeia cada autovalor λ_i para um novo autovalor *unfolded* ξ_i , reescalando cada autovalor em termos de unidades de espaçamento local, o que facilita a comparação com

resultados analíticos, isto porque o objetivo desse procedimento é remover a parte flutuante da distribuição da densidade acumulada de autovalores, tornando-a muito insensível aos detalhes da distribuição e mantendo apenas a parte “amena” da densidade de λ_i , de forma que a distribuição de densidade dos autovalores aleatórios *unfolded* será uniforme e a média das diferenças entre vizinhos próximos será unitária (BRODY *et al.*, 1981; GUHR; MÜLLER-GROELING; WEIDENMÜLLER, 1998; MEYER; D'AURIAC; BRUUS, 1996; PLEROU *et al.*, 2002; SENSOY; YUKSEL; ERTURK, 2013). Desta forma, alteração de comportamento após o procedimento de *unfolding* nos dados empíricos pode indicar características genuínas da distribuição. Este procedimento funciona como um filtro inicial dos dados, filtrando o espectro da distribuição dos autovalores de forma a manter apenas a parte menos sensível às características próprias da distribuição. Na prática, no entanto, a separação de um espectro em partes amena e flutuante pode ser uma tarefa não trivial (GUHR; MÜLLER-GROELING; WEIDENMÜLLER, 1998). A Figura 4 representa uma abstração esquemática do objetivo do processo de *unfolding*.

Figura 4 - Diagrama esquemático do processo de unfolding.



Fonte: Autor.

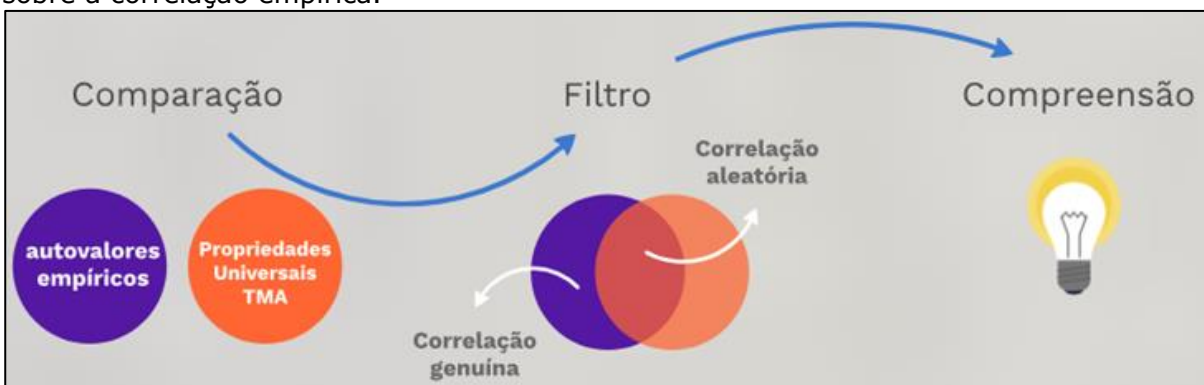
Desta forma, então, pode-se verificar o ajuste dos autovalores *unfolded* empíricos às propriedades universais da TMA (PLEROU *et al.*,

2002).

Ao comparar propriedades universais previstas pela TMA com o comportamento apresentado pelos autovalores *unfolded* empíricos, é esperado que os maiores autovalores ultrapassem os limites teóricos apresentados pela TMA. Sendo assim, os autovalores que ultrapassem os limites teóricos estão relacionados às correlações “puras e verdadeiras”, uma vez que os autovalores que estão no espectro do comportamento previsto pela TMA são apenas ruídos, resultantes de correlações aleatórias, não sendo, dessa forma, uma correlação genuína (PLEROU *et al.*, 2002). O maior autovalor em especial (o qual demonstra-se ser o mais distante do limite teórico da TMA de acordo com a literatura) refletirá o comportamento global do mercado (KULKARNI; DEO, 2007; LIVAN; REBECCHI, 2012; PLEROU *et al.*, 2002). Os demais autovalores que ultrapassem os limites teóricos, carregam informações sobre setores de mercado (PLEROU *et al.*, 2002; KULKARNI; DEO, 2007).

Neste estágio, as propriedades universais funcionaram como um filtro, separando a correlação resultante da aleatoriedade da correlação realmente genuína, possibilitando, desta forma, uma melhor compreensão da natureza da correlação entre as ações analisadas. A Figura 5 esquematiza e esclarece o objetivo desta fase da metodologia.

Figura 5 - Diagrama da dinâmica de comparação-filtragem das propriedades universais sobre a correlação empírica.



Fonte: Autor.

Uma investigação mais aprofundada sobre as características das correlações empíricas envolve a análise de seus autovetores. A análise da distribuição dos componentes de um autovetor fornece informações

importantes sobre a natureza da correlação analisada. Ao normalizar os componentes dos autovetores, é possível comparar a distribuição de seus componentes à distribuição Gaussiana, a qual é a distribuição esperada para os componentes de autovetores estocásticos, além da verificação da medida de não achatamento padrão e assimetria, as quais fornecem mais informações acerca da distribuição de componentes dos autovetores (PLEROU *et al.*, 2002).

Também é possível calcular a quantidade de empresas que influenciam significativamente cada autovetor (Taxa de Participação - TP), podendo dessa forma, em uma análise mais profunda, saber quais componentes estão influenciando cada índice uma vez que autovetores referentes aos autovalores que se destacaram dos limites teóricos podem ser tratados como índices de mercado (PLEROU *et al.*, 2002).

O autovetor associado ao autovalor mais distante, de acordo com as análises de Plerou *et al.* (2002), Coelho *et al.* (2008), Sensoy, Yuksel e Erturk (2013), Sandoval Junior, Bortoluzzo e Venezuela (2014) e Sandoval Junior (2014) pode representar um resultado coletivo das interações de todo o mercado, ou seja, um índice de mercado. Para verificar essa possibilidade de forma mais profunda pode-se comparar a projeção (produto escalar) da série temporal de retornos diários sobre este autovetor com a medida de desempenho padrão do mercado de ações brasileiro, o Ibovespa. Caso tal suspeita seja verificada, deve-se, então, remover a influência do mercado sobre os retornos das ações analisadas, eliminando as variações comuns de todas as ações devido aos movimentos do mercado através do modelo SIM (SANDOVAL JUNIOR; BORTOLUZZO; VENEZUELA, 2014). Este procedimento objetiva tornar a análise dos dados mais precisa, podendo enxergar de forma mais clara o comportamento de índices do mercado. Em seguida, deve-se realizar uma nova análise sobre os dados filtrados os quais espera-se que tragam conclusões mais assertivas (SANDOVAL JUNIOR; BORTOLUZZO; VENEZUELA, 2014; PLEROU *et al.*, 2002).

Além disso, pode-se realizar uma Análise de Componentes Principais

(ACP), a fim de simplificar a análise de correlações entre as várias empresas.

Para um melhor entendimento do comportamento das diferentes empresas, pode-se agrupá-las usando o método *k-Means Clustering* (JOHNSON; WICHERN, 2002), que atribui cada empresa a um dos k clusters definidos pelos centroides a partir da distância euclidiana quadrada. Dessa forma, é possível identificar visualmente quais grupos diferentes de empresas possuem comportamentos opostos e, assim, selecionar empresas com comportamentos opostos para uma carteira equilibrada.

O agrupamento também pode ser alcançado através do método de agrupamento hierárquico que utiliza uma determinada métrica de distância para comparar o comportamento entre as empresas e assim vincular aquelas que têm comportamento mais semelhante entre si (distância menor) de forma ascendente e hierárquica (HAIR *et al.*, 2009; JOHNSON; WICHERN, 2002; SENSOY; YUKSEL; ERTURK, 2013).

Os cálculos para aplicação da metodologia da TMA na análise dos dados financeiros são descritos nas subseções a seguir.

A metodologia utilizada neste trabalho está baseada nos trabalhos de Kumar e Deo (2015), Meyer, D'auriac e Bruus (1996), Plerou *et al.* (2002), Plerou *et al.* (2000) e Sensoy, Yuksel e Erturk (2013). Esta seção está subdividida para melhor compreensão da metodologia, sendo que a seção 3.1 aborda como a série temporal financeira foi obtida; as Seções 3.2 e 3.3 demonstram os métodos utilizados para obtenção das matrizes de correlação empírica e aleatória, respectivamente; na Seção 3.4 é apresentada a função densidade de probabilidade dos autovalores; a Seção 3.5 demonstra o processo pelo qual os autovalores sofrem *unfolding*. A Seção 3.6 apresenta as propriedades universais da TMA; por fim, a Seção 3.7 trata sobre algumas abordagens de análise dos autovetores.

3.1 Determinação da Série Temporal Financeira

Considere as séries temporais de N ações por um tempo T . Seja $P_i(t)$ o preço do fechamento diário da ação i no tempo t ($i = 1, 2, \dots, N$; $t = 1, 2, \dots, T$), o retorno $r_i(t)$ de uma ação em determinado período será dado por

(DENG *et al.* 2011; KULKARNI; DEO, 2007; KUMAR; DEO, 2015; SENSOY; YUKSEL; ERTURK, 2013; SANDOVAL JUNIOR; BORTOLUZZO; VENEZUELA, 2014)

$$r_i(t) = \ln P_i(t + \Delta t) - \ln P_i(t), \quad (1)$$

onde, no caso deste estudo, $\Delta t = 1$ dia. Normalizando os retornos, tem-se (KULKARNI; DEO, 2007; KUMAR; DEO, 2015; SENSOY; YUKSEL; ERTURK, 2013)

$$r'_i(t) = \frac{r_i(t) - \langle r_i \rangle}{\sigma_i}, \quad (2)$$

onde o símbolo $\langle r_i \rangle$ representa a média dos retornos da ação i ao longo do tempo T , e σ_i o respectivo desvio-padrão. Com os valores obtidos na normalização dos retornos ($r'_i(t)$), constrói-se uma matriz R de tamanho $T \times N$.

3.2 Determinação da Matriz de Correlação Empírica

A matriz de correlação C (construída através dos retornos normalizados empíricos) expressa em notação matricial será (COELHO *et al.*, 2008; LIVAN; REBECCHI, 2012; PLEROU *et al.*, 2000; PLEROU *et al.*, 2002; SENSOY; YUKSEL; ERTURK, 2013)

$$C = \frac{1}{T} R^t R, \quad (3)$$

onde R^t é a matriz transposta de R . Portanto, cada elemento da matriz C pode ser expresso como (KULKARNI; DEO, 2007; PLEROU *et al.*, 2000; PLEROU *et al.*, 2002; SENSOY; YUKSEL; ERTURK, 2013)

$$c_{i,j} = \langle r'_i r'_j \rangle. \quad (4)$$

Os elementos de $c_{i,j}$ estão limitados ao domínio $-1 \leq c_{i,j} \leq 1$, onde $c_{i,j} = 1$ define correlações positivas perfeitas, $c_{i,j} = -1$ corresponde a correlações negativas perfeitas e $c_{i,j} = 0$ corresponde a não correlação (KUMAR; DEO, 2015; KULKARNI; DEO, 2007). Uma vez que cada elemento $c_{i,j}$ representa a correlação entre um par de empresas, espera-se que a

diagonal principal seja unitária (perfeitamente correlacionada), pois demonstra a correlação entre uma mesma empresa. Também é esperado que os elementos acima da diagonal principal sejam iguais aos elementos abaixo, pois representam a correlação de um mesmo par de empresas.

3.3 Determinação da Matriz de Correlação Aleatória

Toma-se N amostras de um processo estocástico i.i.d., de modo que $y_i(t)$ representa os valores assumidos pela variável aleatória correspondente a esse processo no tempo t . A matriz aleatória A , de tamanho $T \times N$, terá como componentes $a_{t,i} = y_i(t)$. A matriz de correlação W (construída através de dados aleatórios) expressa em notação matricial é dada por (OH, 2011; PLEROU *et al.*, 2000; PLEROU *et al.*, 2002; SENSOY; YUKSEL; ERTURK, 2013)

$$W = \frac{1}{T} A^t A. \quad (5)$$

3.4 Função Densidade de Probabilidade

Supondo que $T \rightarrow \infty$ e $N \rightarrow \infty$ de tal forma que $Q \equiv \frac{T}{N} \geq 1$, e considerando que os valores obtidos por um processo estocástico siga uma distribuição normal ($y_i(t) \sim N(0,1)$), a função densidade de probabilidade $\rho_{rm}(\lambda)$ dos autovalores λ , referente a uma matriz W será (COELHO *et al.*, 2008; KUMAR; DEO, 2015; PLEROU *et al.*, 2000; PLEROU *et al.*, 2002; SENSOY; YUKSEL; ERTURK, 2013)

$$\rho_{rm}(\lambda) = \frac{Q}{2\pi} \frac{\sqrt{(\lambda_{max} - \lambda)(\lambda - \lambda_{min})}}{\lambda}, \quad (6)$$

onde $\lambda_{max} = 1 + \frac{1}{Q} + 2\sqrt{\frac{1}{Q}}$ e $\lambda_{min} = 1 + \frac{1}{Q} - 2\sqrt{\frac{1}{Q}}$ são os limites teóricos para os autovalores e $\lambda_i < \lambda_{i+1}$ (autovalores estão ordenados) os quais delimitam a chamada região de Wishart (SANDOVAL JUNIOR; BORTOLUZZO; VENEZUELA, 2014).

3.5 Processo Unfolding

A função de distribuição acumulada dos autovalores, a qual conta a quantidade de autovalores $\lambda_i \leq \lambda$, é definida da seguinte forma (PLEROU *et al.*, 2002)

$$F(\lambda) = N \int_{-\infty}^{\lambda} \rho(x) dx, \quad (7)$$

onde N é a quantidade de autovalores e $\rho(x)$ denota a densidade de probabilidade dos autovalores. Desconsiderando-se a parte flutuante de $F(\lambda)$, a função de distribuição cumulativa do autovalor em média é $F_{av.}(\lambda)$. Assim sendo, o autovalor adimensional *unfolded* ξ_i é dado por $\xi_i = F_{av.}(\lambda_i)$ (PLEROU *et al.*, 2002). Desta forma, uma aproximação para $F_{av.}(\lambda)$ pode ser dada pela seguinte expressão analítica (PLEROU *et al.*, 2002)

$$F_{rm}(\lambda) = N \int_{-\infty}^{\lambda} \rho_{rm}(x) dx, \quad (8)$$

onde $\rho_{rm}(x)$ é a densidade de probabilidade dos autovalores aleatórios da Equação (6).

Dentre os diferentes métodos de aproximação do valor *unfolded* ξ_i para dados empíricos (PLEROU *et al.*, 2002), um possível é o ajuste polinomial da densidade acumulada de autovalores, o qual busca ser uma aproximação para $F_{av.}(\lambda_i)$. Este procedimento basicamente aproxima o valor da densidade acumulada por um polinômio, gerando o novo autovalor *unfolded* ξ_i calculado sobre o antigo autovalor λ_i (VIVO, 2009). Para este procedimento iremos propor aqui uma nova maneira de abordar o processo de *unfolding* dos autovalores da matriz de correlação dos dados empíricos. Na plataforma MATLAB, pode-se usar uma adaptação do código de Vivo (2009), sendo que as principais linhas de codificação deste método são as seguintes

Algoritmo 1

```

E = cumtrapz(AutoValores);
for i = 1:n
    CumDist(i) = length(find(E <= F(i)));
end

p = polyfit(F, CumDist, Deg);
Unfold = polyval(p, E);

```

onde E é o vetor de autovalores empíricos acumulado, n é o número de pontos entre 0 e o maior autovalor (ou seja, a quantidade de autovalores), F é a função $F_{rm}(\lambda)$ definida da Eq. (8), Deg é o grau do polinômio e $Unfold$ é o vetor de autovalores *unfolded* ξ_i resultante. Desta forma, o algoritmo aqui apresentado encontra a quantidade de autovalores acumulados menores ou iguais a $F_{rm}(\lambda)$ (expressão analítica de distribuição acumulada dos autovalores aleatórios). Esta função é, então, aproximada por um polinômio o qual é computado sobre os autovalores empíricos acumulados E , resultando nos novos autovalores *unfolded* ξ_i .

3.6 Propriedades Universais da TMA

A seguir, as propriedades dos autovalores *unfolded* ξ_i são comparadas às propriedades universais da TMA.

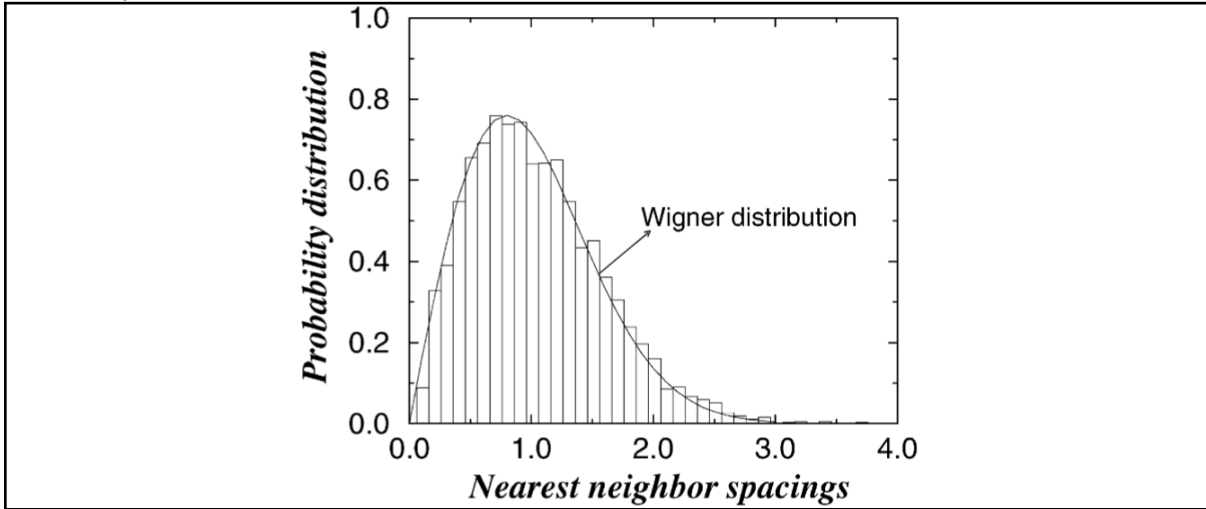
3.6.1 Distribuição dos espaçamentos dos vizinhos próximos - distribuição de Wigner

A distribuição de probabilidade do espaçamento s dos autovalores dos vizinhos mais próximos de uma matriz W é dada pela distribuição de Wigner (MEYER; D'AURIAC; BRUUS, 1996; PLEROU *et al.*, 2000; PLEROU *et al.*, 2002; SENSOY; YUKSEL; ERTURK, 2013)

$$\rho_{Wvp}(s_i) = \frac{\pi s_i}{2} e^{(-\frac{\pi}{4}s_i^2)}, \quad (9)$$

onde $s_i = \xi_{i+1} - \xi_i$, ou seja, os valores gerados pela diferença entre os autovalores *unfolded* mais próximos. Um exemplo da distribuição Wigner aplicada a dados do mercado financeiro é apresentado na Figura 6.

Figura 6 - Distribuição de Wigner: Histograma da distribuição do espaçamento entre vizinhos próximos.



Fonte: PLEROU *et al.* (2000).

3.6.2 Distribuição dos espaçamentos de longo alcance - alcance de autovalores

Diferentemente da distribuição $\rho_{Wvp}(s_i)$ (Equação 9) que apresenta propriedades para os espaçamentos dos vizinhos mais próximos, a distribuição dos espaçamentos de longo alcance Σ^2 , conhecida como alcance de autovalores, a qual determina a variação do número de autovalores *unfolded* ξ_i em cada janela móvel de comprimento L centrado em cada ξ_i , apresenta propriedades para distâncias maiores de espaçamentos entre vizinhos mais distantes. Esta distribuição é definida como (PLEROU *et al.*, 2002; SENSOY; YUKSEL; ERTURK, 2013)

$$\Sigma^2(L) = \langle [n(\xi, L) - L]^2 \rangle_{\xi}, \quad (10)$$

onde $n(\xi, L)$ é o número de autovalores *unfolded* no intervalo $[\xi - \frac{L}{2}, \xi + \frac{L}{2}]$ e $\langle \dots \rangle_{\xi}$ denota uma média em todos os ξ . Para altos valores de L , o alcance de autovalores para uma matriz W tem um comportamento $\Sigma^2 \sim \ln L$ e caso os autovalores não sejam correlacionados, o alcance de autovalores tem um comportamento $\Sigma^2 \sim L$ (PLEROU *et al.*, 2002).

O alcance de autovalores definido pela TMA, para uma matriz W , pode ser expressado por (PLEROU *et al.*, 2002)

$$\Sigma^2(L) = L - 2 \int_0^L (L - x) Y(x) dx. \quad (11)$$

Para nosso propósito, é suficiente saber que para o caso da distribuição normal, $Y(x)$ é dado por (PLEROU *et al.*, 2002)

$$Y(x) = s^2(x) + \frac{ds}{dx} \int_x^{\infty} s(x') dx', \quad (12)$$

onde

$$s(x) = \frac{\sin(\pi x)}{\pi x}. \quad (13)$$

3.7 Análise dos Autovetores

O autovetor empírico v_i pode ser normalizado ($v_i \rightarrow v_i'$) de modo que $\sum_{j=1}^N (v'_{i,j})^2 = N$, onde i corresponde aos componentes do autovetor v_i' e N corresponde a quantidade de ações analisadas (SENSOY; YUKSEL; ERTURK, 2013).

Para o autovetor normalizado v_i' associado ao autovalor λ_i gerado a partir de uma matriz de correlação aleatória do tipo W (equação (5)), seus componentes seguem uma distribuição Gaussiana com média zero e variância unitária descrita da seguinte maneira ($\sim N(0,1)$) (SENSOY; YUKSEL; ERTURK, 2013)

$$\rho(v') = \frac{1}{\sqrt{2\pi}} e^{-\frac{v'^2}{2}}. \quad (14)$$

A Taxa de Participação (TP) é a taxa de participações significativas em um autovetor e pode ser definida pela Taxa de Participação Invertida (TPI) (PLEROU *et al.*, 2002; SENSOY; YUKSEL; ERTURK, 2013)

$$TPI_i = \sum_{j=1}^N [v_{i,j}]^4. \quad (15)$$

Na TMA é esperado que $TPI = 3/N$, uma vez que a medida de não achatamento padrão para uma distribuição normalmente distribuída ($\sim N(0,1)$) é 3 (PLEROU *et al.*, 2002; SENSOY; YUKSEL; ERTURK, 2013). Dessa forma, TP é a inversa da TPI e representa a quantidade de participações significativas que compõem o autovetor (PLEROU *et al.*, 2002; SENSOY; YUKSEL; ERTURK, 2013).

O autovetor v_{100} associado ao autovalor mais distante λ_{100} , de acordo com as análises de Johnson e Wichern (2002) Plerou *et al.* (2002), Coelho *et al.* (2008) e Sensoy *et al.* (2013), pode representar um resultado coletivo das interações de todo o mercado, ou seja, um índice de mercado. Para verificar essa possibilidade de forma mais profunda pode-se comparar a projeção $r_{100}(t)$ (produto escalar) da série temporal de retornos diários $r_i(t)$ (Eq. 1) sobre este autovetor (o mais divergente) com a medida de desempenho padrão do mercado de ações brasileiro, o Ibovespa. Desta forma $r_{100}(t)$ é dado por (PLEROU *et al.*, 2002; SENSOY *et al.*, 2013)

$$r_{100}(t) = \sum_{i=1}^{100} v_{100,i} r_i(t). \quad (16)$$

Uma forma de modelar estatisticamente uma influência que é comum a todos os estoques é a utilização do Modelo de Índice Único (do inglês *Single Index Model* - SIM) o qual expressa o retorno $r_i(t)$ da ação i em termos da função de primeiro grau de um dado índice de mercado $M(t)$ adicionando-se um erro (SANDOVAL JUNIOR; BORTOLUZZO; VENEZUELA, 2014; PLEROU *et al.*, 2002)

$$r_i(t) = \alpha_i + \beta_i M(t) + \epsilon_i, \quad (17)$$

onde $M(t)$ é um termo aditivo comum a todas as ações, o qual pode ser aproximado pelo índice do mercado ou pela projeção $r_{100}(t)$ por exemplo, ϵ_i é o resíduo ou erro, α_i e β_i são os coeficientes específicos das ações, podendo ser estimados por uma regressão de mínimos quadrados. Este procedimento é padrão em muitos modelos em finanças, principalmente para o Modelo de Precificação de Ativos Financeiros (do inglês *Capital Asset Pricing Model* - CAPM) (SANDOVAL JUNIOR; BORTOLUZZO; VENEZUELA, 2014; PLEROU *et al.*, 2002). Busca-se desta forma, separa os resíduos, os quais representam o quanto cada retorno se distancia do comportamento do mercado. Assim, o resíduo irá representar os retornos de cada ação sem a influência do mercado (PLEROU *et al.*, 2002).

4. Coleta e Análise de Dados

Foram analisadas duas séries históricas de ações negociadas no mercado financeiro brasileiro:

- Período I: De 2007 a 2011, período que contempla a crise do *subprime* de 2008 e suas vizinhanças temporais. Os dados que suportam as conclusões deste estudo para o período em questão foram extraídos da base de dados do Yahoo Finanças (2020). Esta base de dados foi escolhida por se tratar de uma base gratuita e por apresentar o preço de fechamento diário ajustado tanto para dividendos quanto para *splits*.
- Período II: De 1996 a 2000, período que inclui as crises asiática e russa, de 1997 e 1998 respectivamente, e suas vizinhanças temporais. Os dados para análise deste período foram extraídos da base de dados da B3 (2020), uma vez que este banco de dados é gratuito e apresenta uma base de dados mais completa para o período em questão.

A seguir serão detalhados os procedimentos de coleta e análise destes dados.

4.1 Procedimentos de Coleta de Dados

Das séries temporais históricas de preços de ações das empresas brasileiras encontradas durante a fase de coleta de dados, para o Período I, foram excluídos os papéis de ações que apresentaram cinco ou mais dias (uma semana comercial) corridos sem registro de negociações diária, resultando no total de 100 papéis de ações selecionadas no mercado financeiro brasileiro.

Já para o Período II, foram removidos os papéis de ações que apresentaram dez dias corridos ou mais (o que equivale a duas semanas comerciais) sem registros de cotação diária, resultando em 61 papéis selecionados, os quais atuaram no mercado de ações brasileiro durante o período. Assim sendo, percebe-se que para o período II usou-se uma filtragem menos rígida para delimitação dos papeis escolhidos, uma vez que

durante este período existia uma quantidade menor de empresas de capital aberto no mercado financeiro brasileiro e, desta forma, um filtro rigoroso iria resultar em uma quantidade muito menor de empresas a serem analisadas, o que prejudicaria a aplicação do método da TMA, gerando ruídos nos dados (PLEROU *et al.*, 2002; SANDOVAL JUNIOR; FRANCA, 2012).

Ressalta-se que os registros de cotação diária dos papéis no mercado de ação brasileiro são disponibilizados somente em dias úteis com a abertura da bolsa.

As ações selecionadas para este estudo para o Período I estão apresentadas no Quadro 1, em ordem alfabética dos Código de Negociação, destacando-se o setor e segmento de atuação de cada empresa de acordo com as informações constantes na própria base do Yahoo Finanças (2020).

Quadro 1 - Ações utilizadas na análise do Período I.

Índ.	Código	Empresa	Setor	Segmento
1	ABEV3	Ambev S.A.	Consumo Não Cíclico	Cervejas e Refrigerantes
2	ALPA3	Alpargatas S.A.	Consumo Cíclico	Calçados
3	BAZA3	Banco Amazônia S.A.	Financeiro	Bancos
4	BBAS3	Banco Brasil S.A.	Financeiro	Bancos
5	BEES3	Banestes S.A.	Financeiro	Bancos
6	BMKS3	Bicicletas Monark S.A.	Consumo Cíclico	Bicicletas
7	BRAP4	Bradespar S.A.	Financeiro	Gestão de ativos
8	BRKM3	Braskem S.A.	Materiais Básicos	Produtos químicos
9	BRKM5	Braskem S.A.	Materiais Básicos	Produtos químicos
10	BTOW3	B2W - Cia Digital	Consumo Cíclico	Varejo na Internet
11	CARD3	CSU Cardsystem S.A.	Financeiro	Serviços de Crédito
12	CBEE3	Ampla Energia E Serviços S.A.	Utilidade Pública	Energia Elétrica
13	CCPR3	Cyrela Commercial Propert S.A.	Imobiliária	Exploração de Imóveis
14	CCRO3	CCR S.A.	Bens Industriais	Exploração de Rodovias
15	CESP3	CESP - Cia Energética de São Paulo	Utilidade Pública	Energia Elétrica
16	CESP6	CESP - Cia Energética de São Paulo	Utilidade Pública	Energia Elétrica
17	CGAS3	COMGAS - Cia Gás de São Paulo	Utilidade Pública	Gás

18	CGAS5	COMGAS - Cia Gás de São Paulo	Utilidade Pública	Gás
19	CMIG3	CEMIG - Cia Energética De Minas Gerais	Utilidade Pública	Energia Elétrica
20	CMIG4	CEMIG - Cia Energética De Minas Gerais	Utilidade Pública	Energia Elétrica
21	COCE5	COELCE - Cia Energética Do Ceara	Utilidade Pública	Energia Elétrica
22	CPFE3	CPFL Energia S.A.	Utilidade Pública	Energia Elétrica
23	CPL3	COPEL - Cia Paranaense De Energia	Utilidade Pública	Energia Elétrica
24	CPL6	COPEL - Cia Paranaense De Energia	Utilidade Pública	Energia Elétrica
25	CRIV4	Financeira Alfa S.A.	Financeiro	Conglomerados Financeiros
26	CSAN3	COSAN S.A.	Petróleo. Gás e Biocombustíveis	Exploração. Refino e Distribuição
27	CSMG3	COPASA - Cia Saneamento De Minas Gerais	Utilidade Pública	Água e Saneamento
28	CSNA3	Cia Siderúrgica Nacional	Materiais Básicos	Siderurgia
29	CTNM4	Coteminas - Cia Tecidos Norte De Minas	Consumo Cíclico	Fios e Tecidos
30	CTSA3	Cia Tecidos Santanense	Consumo Cíclico	Fios e Tecidos
31	CYRE3	Cyrela Brazil Realty S.A.	Consumo Cíclico	Incorporações
32	DTCY3	DTCOM - Direct To Company S.A.	Serviços de comunicação	Serviços de telecomunicações
33	EALT4	Electro Aço Altona S.A.	Bens Industriais	Máq. e Equip. Industriais
34	ELET3	Eletrobras - Centrais Elétricas Brasileiras S.A.	Utilidade Pública	Energia Elétrica
35	EMAE4	Empresa Metropolitana de Águas e Energia S.A.	Utilidade Pública	Energia Elétrica
36	EMBR3	Embraer S.A.	Bens Industriais	Material Aeronáutico e de Defesa
37	ENBR3	EDP - Energias Do Brasil S.A.	Utilidade Pública	Energia Elétrica
38	ETER3	Eternit S.A.	Bens Industriais	Produtos para Construção
39	FESA4	Ferbasa - Cia Ferro Ligas Da Bahia	Materiais Básicos	Siderurgia
40	GFA3	GAFISA S.A.	Consumo Cíclico	Incorporações
41	GGBR3	Gerdau S.A.	Materiais Básicos	Siderurgia
42	GOAU3	Metalúrgica Gerdau S.A.	Materiais Básicos	Siderurgia

43	GOAU4	Metalúrgica Gerdau S.A.	Materiais Básicos	Siderurgia
44	GOLL4	GOL Linhas Aéreas Inteligentes S.A.	Bens Industriais	Transporte Aéreo
45	GPCP3	GPC Participações S.A.	Materiais Básicos	Produtos químicos
46	GPIV33	GP Investments	Financeiro	Gestão de ativos
47	GRND3	Grendene S.A.	Consumo Cíclico	Calçados
48	GUAR3	Guararapes Confecções S.A.	Consumo Cíclico	Tecidos, Vestuário e Calçados
49	HAGA4	Haga S.A.	Bens Industriais	Produtos para Construção
50	HGTX3	Cia Hering	Consumo Cíclico	Vestuário
51	IDNT3	Ideiasnet S.A.	Financeiro	Gestão de ativos
52	IGBR3	Gradiente	Imobiliária	Exploração de Imóveis
53	INEP4	Inepar S.A.	Bens Industriais	Máq. e Equip. Industriais
54	ITSA4	Itausa S.A.	Financeiro	Bancos
55	ITUB3	Itaú Unibanco Holding S.A.	Financeiro	Bancos
56	ITUB4	Itaú Unibanco Holding S.A.	Financeiro	Bancos
57	JBDU3	Indústrias J B Duarte S.A.	Serviços de comunicação	Serviços de telecomunicações
58	JBDU4	Indústrias J B Duarte S.A.	Serviços de comunicação	Serviços de telecomunicações
59	JFEN3	João Fortes Engenharia S.A.	Imobiliária	Incorporações
60	KEPL3	Kepler Weber S.A.	Bens Industriais	Máq. e Equip. Industriais
61	LAME3	Lojas Americanas S.A.	Consumo Cíclico	Lojas de departamento
62	LAME4	Lojas Americanas S.A.	Consumo Cíclico	Lojas de Departamento
63	LIGT3	Light S.A.	Utilidade Pública	Energia Elétrica
64	LPSB3	LPS Brasil - Consultoria De Imoveis S.A.	Imobiliária	Intermediação Imobiliária
65	LREN3	Lojas Renner S.A.	Consumo Cíclico	Lojas de Departamento
66	MDIA3	M.Dias Branco S.A.	Consumo Não Cíclico	Alimentos Diversos
67	MGEL4	Mangels Industrial S.A.	Materiais Básicos	Artefatos de Ferro e Aço
68	MMXM3	MMX Mineracao e Metalicos S.A.	Materiais Básicos	Minerais Metálicos
69	MNPR3	Minupar Participacoes S.A.	Consumo Não Cíclico	Carnes e Derivados
70	MYPK3	Iochpe Maxion S.A.	Consumo Cíclico	Automóveis e Motocicletas
71	NATU3	Natura	Consumo Não Cíclico	Produtos domésticos e pessoais
72	ODPV3	Odontoprev S.A.	Saúde	Planos de Saúde
73	PETR3	Petrobras - Petroleo Brasileiro S.A.	Petróleo. Gás e Biocombustíveis	Exploração. Refino e Distribuição

74	PFRM3	Profarma Distribuidora de Produtos Farmacêuticos S.A.	Saúde	Distribuição de Medicamentos
75	PMAM3	Paranapanema S.A.	Materiais Básicos	Artefatos de Cobre
76	POMO4	Marcopolo S.A.	Consumo Cíclico	Fabricantes de automóveis
77	POSI3	Positivo Tecnologia S.A.	Tecnologia da Informação	Computadores e Equipamentos
78	PSSA3	Porto Seguro S.A.	Financeiro	Seguradoras
79	PTBL3	Portobello S.A.	Bens Industriais	Produtos para Construção
80	RADL3	Raia Drogasil S.A.	Saúde	Varejistas Farmacêuticos
81	RAPT4	Randon S.A.	Bens Industriais	Material Rodoviário
82	RCSL4	Recrusul S.A.	Bens Industriais	Material Rodoviário
83	RENT3	Localiza Rent a Car S.A.	Bens Industriais	Aluguel de carros
84	ROMI3	Indústrias Romi S.A.	Bens Industriais	Máq. e Equip. Industriais
85	RSID3	Rossi Residencial S.A.	Consumo Cíclico	Incorporações
86	SANB3	Banco SANTANDER S.A.	Financeiro	Bancos
87	SANB4	Banco SANTANDER S.A.	Financeiro	Bancos
88	SBSP3	SABESP - Cia Saneamento Básico Est São Paulo	Utilidade Pública	Água e Saneamento
89	SCAR3	São Carlos Empreend E Participações S.A.	Imobiliária	Exploração de Imóveis
90	SHUL4	Schulz S.A.	Consumo Cíclico	Autopeças
91	SLED4	Saraiva Livreiros S.A.	Consumo Cíclico	Produtos Diversos
92	SNSY5	Sansuy S.A.	Consumo Cíclico	Autopeças
93	SPRI3	Springer	Consumo Cíclico	Móveis, utensílios e eletrodomésticos
94	TELB3	Telebras - Telec Brasileiras S.A.	Serviços de comunicação	Telecomunicações
95	TOTS3	TOTVS S.A.	Tecnologia da Informação	Programas e Serviços
96	TRPL4	CTEEP - Cia Transmissão Energia Elétrica Paulista	Utilidade Pública	Energia Elétrica
97	UNIP6	UNIPAR Carbocloro S.A.	Materiais Básicos	Produtos químicos
98	USIM3	USIMINAS - Usinas Sid De Minas Gerais S.A.	Materiais Básicos	Siderurgia
99	VALE3	Vale S.A.	Materiais Básicos	Minerais Metálicos
100	WHRL4	Whirlpool S.A.	Tecnologia da Informação	Eletrodomésticos

Fonte: Autor. Baseado em dados de Yahoo Finanças (2020).

As ações utilizadas para análise do Período II estão apresentadas no Quadro 2, em ordem alfabética dos Código de Negociação, destacando-se

o setor e segmento de atuação de cada empresa de acordo com as informações constantes na própria base da B3 (2020) e do Yahoo Finanças (2020).

Quadro 2 - Ações utilizadas na análise do Período II

Índ.	Código	Empresa	Setor	Segmento
1	ACES3	Acesita S.A.	Materiais Básicos	Siderurgia
2	ACES4	Acesita S.A.	Materiais Básicos	Siderurgia
3	ARCZ6	ARACRUZ CELULOSE S.A.	Materiais Básicos	Papel e Celulose
4	AVPL3	Avipal S.A.	Consumo Não Cíclico	Alimentos e bebidas
5	BBAS3	Banco do Brasil S.A.	Financeiro	Bancos
6	BBDC3	Banco Bradesco S.A.	Financeiro	Banco
7	BBDC4	Banco Bradesco S.A.	Financeiro	Banco
8	BELG3	Cia Siderurgica Belgo Mineira	Materiais Básicos	Siderurgia
9	BELG4	Cia Siderurgica Belgo Mineira	Materiais Básicos	Siderurgia
10	BESP4	BANESPA - Banco do Estado de São Paulo S.A.	Financeiro	Bancos
11	BFIT3	Banco Sudameris Brasil S.A.	Financeiro	Bancos
12	BRDT4	PETROBRAS BR - Petrobras Distribuidora S.A.	Petróleo. Gás e Biocombustíveis	Exploração. Refino e Distribuição
13	BRIV4	ALFA INVEST	Financeiro	Intermediários Financeiros
14	CBEE3	Ampla Energia e Serviços S.A.	Utilidade Pública	Energia Elétrica
15	CESP3	CESP - Companhia Energética de São Paulo	Utilidade Pública	Energia Elétrica
16	CESP4	CESP - Companhia Energética de São Paulo	Utilidade Pública	Energia Elétrica
17	CEVA4	Ceval Alimentos S.A.	Consumo Não Cíclico	Alimentos
18	CMET4	Caemi Mineração e Metalurgia S.A.	Materiais Básicos	Minerais Metálicos
19	CMIG3	Companhia Energética de Minas Gerais	Utilidade Pública	Energia Elétrica
20	CMIG4	Companhia Energética de Minas Gerais	Utilidade Pública	Energia Elétrica
21	CNFB4	Confab Industrial S.A.	Materiais Básicos	Siderurgia e Metalurgia
22	COGU4	Gerdau S.A.	Materiais Básicos	Siderurgia
23	CPLE3	COPEL - Companhia Paranaense de Energia	Utilidade Pública	Energia Elétrica
24	CPNE5	COPENE PETROQUIMICA DO NORDESTE S.A.	Materiais Básicos	Produtos químicos

25	CPSL3	Copesul - Companhia Petroquímica do Sul	Materiais Básicos	Produtos químicos
26	CRUZ3	Souza Cruz S.A.	Consumo Não Cíclico	Tobaco
27	CSNA3	Cosan S.A.	Petróleo. Gás e Biocombustíveis	Exploração. Refino e Distribuição
28	CTNM4	COTEMINAS - Companhia de Tecidos Norte de Minas -	Consumo Cíclico	Fios e Tecidos
29	DURA4	DURATEX S.A.	Materiais Básicos	Madeira e Papel
30	ELET3	Eletrobras - Centrais Elétricas Brasileiras S.A.	Utilidade Pública	Energia Elétrica
31	ELET6	Eletrobras - Centrais Elétricas Brasileiras S.A.	Utilidade Pública	Energia Elétrica
32	FFTL4	Fosfertil - Fertilizantes Fosfatados S.A.	Materiais Básicos	Químicos
33	FJTA4	Forjas Taurus S.A.	Bens Industriais	Máquinas e Equipamentos
34	GOAU4	Metalurgica Gerdau S.A.	Materiais Básicos	Siderurgia
35	ICPI4	Companhia Cimento Portland Itaú	Consumo Cíclico	Construção Civil
36	INEP4	Inepar S.A. Industria e Construcoes	Bens Industriais	Máq. e Equip. Industriais
37	ITAU3	Itau Unibanco Holding S.A.	Financeiro	Bancos
38	ITAU4	Itau Unibanco Holding S.A.	Financeiro	Bancos
39	ITSA4	Itausa - Investimentos Itau S.A.	Financeiro	Bancos
40	KLAB4	KLABIN S.A.	Materiais Básicos	Papel e Celulose
41	LEVE4	Mahle-Metal Leve S.A.	Consumo Cíclico	Automóveis e Motocicletas
42	LIGH3	Light S.A.	Utilidade Pública	Energia Elétrica
43	MAHS4	Manah S.A.	Materiais Básicos	Químicos
44	PALF3	CPFL - Cia Paulista De Forca E Luz	Utilidade Pública	Energia Elétrica
45	PCAR4	P.Acucar-Cbd - Cia Brasileira De Distribuicao	Consumo Não Cíclico	Alimentos
46	PETR3	Petrobras - Petroleo Brasileiro S.A.	Petróleo. Gás e Biocombustíveis	Exploração. Refino e Distribuição
47	PETR4	Petrobras - Petroleo Brasileiro S.A.	Petróleo. Gás e Biocombustíveis	Exploração. Refino e Distribuição
48	PTIP4	Cia Brasileira De Petróleo Ipiranga	Petróleo. Gás e Biocombustíveis	Exploração. Refino e Distribuição
49	RHDS3	Rhodia-Ster S.A.	Materiais Básicos	Químicos

50	SUZA4	Suzano S.A.	Materiais Básicos	Papel e Celulose
51	TEKA4	Teka-Tecelagem Kuehnrich S.A.	Consumo Cíclico	Fios e Tecidos
52	TELB3	TELEBRAS - Telecomunicacoes Brasileiras S.A.	Serviços de Comunicação	Telecomunicações
53	TELB4	TELEBRAS - Telecomunicacoes Brasileiras S.A.	Serviços de Comunicação	Telecomunicações
54	TEPR4	TELEPAR - Telecomunicações do Paraná S.A.	Serviços de Comunicação	Telecomunicações
55	TERJ3	TELERJ - Telecomunicações do Rio de Janeiro S.A.	Serviços de Comunicação	Telecomunicações
56	TERJ4	TELERJ - Telecomunicações do Rio de Janeiro S.A.	Serviços de Comunicação	Telecomunicações
57	TMGR3	Telemig Celular Participações S.A.	Serviços de Comunicação	Telecomunicações
58	TMGR6	Telemig Celular Participações S.A.	Serviços de Comunicação	Telecomunicações
59	UBBR4	Unibanco S.A.	Financeiro	Bancos
60	UNIP6	Unipar Carbocloro S.A.	Materiais Básicos	Produtos químicos
61	VALE3	Vale S.A.	Materiais Básicos	Minerais Metálicos

Fonte: Autor. Baseado em dados de Yahoo Finanças (2020).

4.2 Procedimentos de Análise de Dados

Para ambos os períodos, os dados foram tratados, ordenados cronologicamente e as datas que apresentavam mais de um quarto de empresas sem atividades registradas nessa data específica foram removidas. Além disso, foram incluídos nas análises do Período I a cotação diária do dólar em reais (USD / BRL) (INVESTING.COM, 2020) simulando um estoque anticorrelacionado adicional como fator de perturbação nos dados com o objetivo de avaliar seu efeito sobre o comportamento do TMA.

Por meio das séries temporais financeiras foram calculados os valores de retorno diário de cada ação, conforme equação (2).

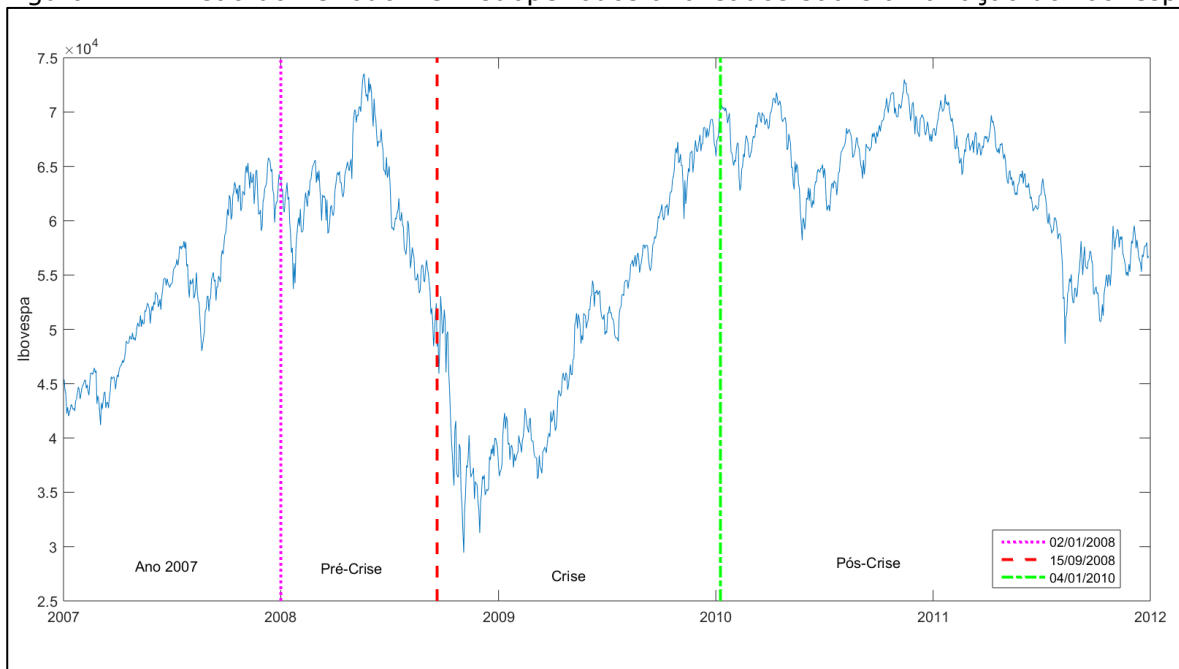
O Período I foi dividido em quatro subperíodos para análise das regiões temporais:

- Ano 2007: de 02/01/2007 a 28/12/2007 (com 245 retornos)
- Pré-Crise: de 02/01/2008 a 12/09/2008 (com 175 retornos);

- Crise: de 15/09/2008 a 28/12/2009 (com 318 retornos);
- Pós-Crise: de 04/01/2010 a 29/12/2011 (com 486 retornos).

O dia 15 de setembro de 2008 foi escolhido como marco representativo do início da crise do *subprime* pois é a data de falência de um dos bancos de investimentos mais tradicionais dos Estados Unidos, o Lehman Brothers. A Figura 7 ilustra a divisão entre os períodos que serão analisados, incluindo o início (linha vermelha) e fim (linha verde) da crise do *subprime* sobre a variação do Ibovespa. Pode-se observar a nítida queda que se inicia no segundo semestre de 2008 e se alastra até o fim daquele ano.

Figura 7 – Divisão do Período I em subperíodos analisados sobre a variação do Ibovespa.



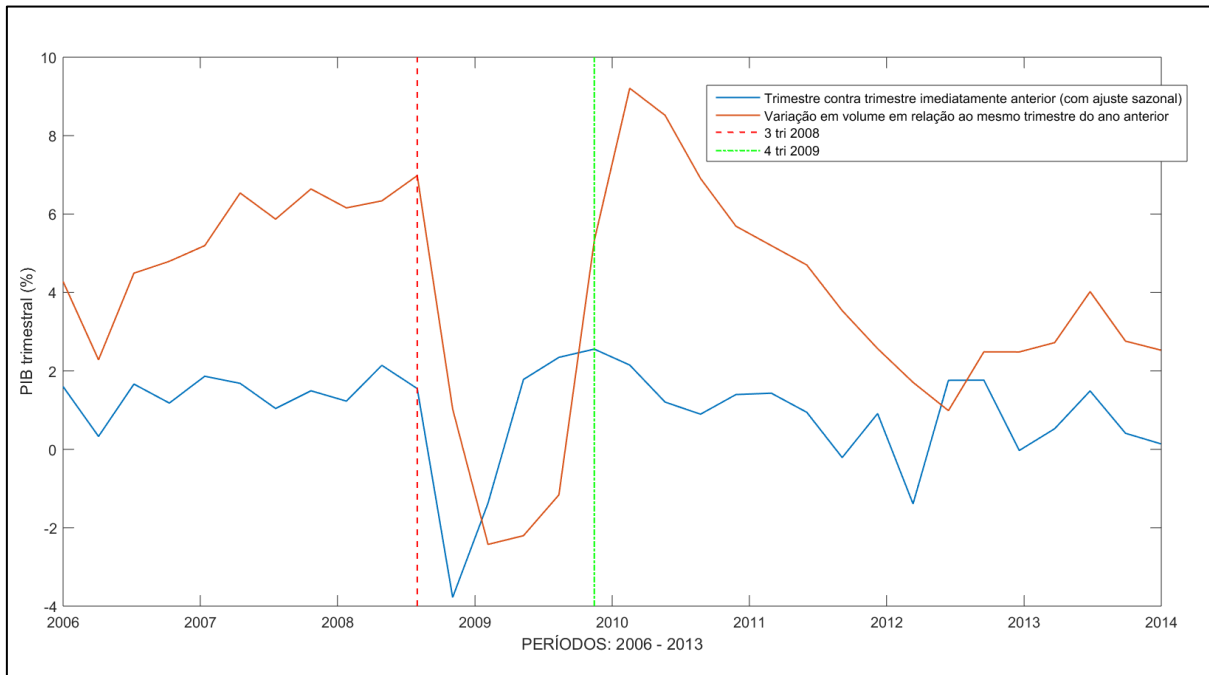
Fonte: Autor. Baseado em dados de Yahoo Finanças (2020).

O final do ano de 2009 foi escolhido como a data para o fim do período de crise visto que nesta data o mercado brasileiro já havia se recuperado, conforme pode ser observado na Figura 7, onde a formação de um V pode ser evidenciado durante o período de crise.

Pode-se, ainda, observar esse movimento de recuperação do mercado brasileiro durante este mesmo período por meio do comportamento do PIB na Figura 8. Percebe-se claramente um formato de V no gráfico dentro do período denominado de crise, o que indica uma retomada do

comportamento econômico (GUIMARÃES, 2020).

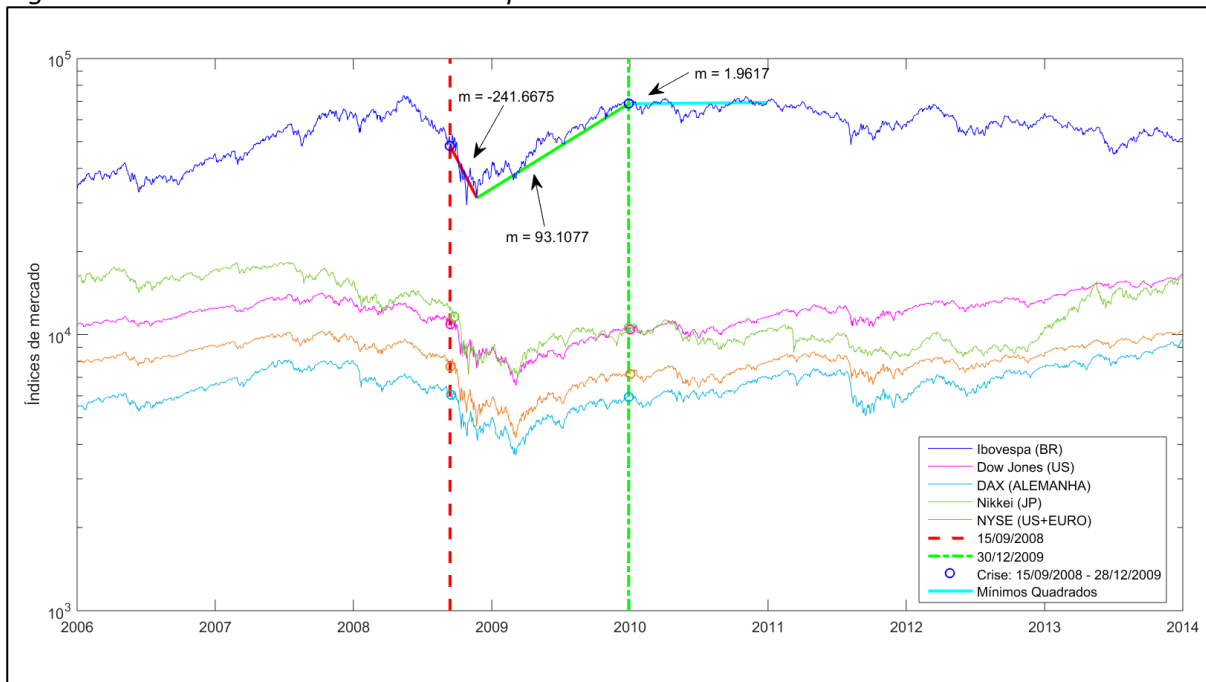
Figura 8 - Início e fim da crise do *subprime* sobre a variação trimestral do PIB brasileiro.



Fonte: Autor. Baseado em dados de IBGE (2020).

Além disso, observa-se que o PIB trimestral em comparação ao mês imediatamente anterior (linha azul), manteve-se em queda durante dois trimestres seguidos (quarto trimestre de 2008 e primeiro trimestre de 2009), o que caracteriza uma recessão técnica (GUIMARÃES, 2020).

Observa-se ainda que em 2010 muitos mercados ao redor do mundo já haviam se recuperado da crise financeira internacional de 2008, como pode ser observado pelo comportamento dos índices de diferentes mercados apresentados na Figura 9. Para o mercado acionário brasileiro, pelo método dos mínimos quadrados, observa-se uma tendência negativa a partir do início da crise (coeficiente angular $m \cong -241.7$), seguida de uma tendência positiva até o final do período de crise (coeficiente angular $m \cong 93$). Durante o ano de 2010, a inclinação é próxima de zero, atingindo-se um platô, o que caracteriza uma recuperação em $\sqrt{\quad}$ (GUIMARÃES, 2020), indicando que a crise já havia sido superada no contexto brasileiro.

Figura 9 - Início e fim da crise do *subprime* sobre índices internacionais.

Fonte: Autor. Baseado em dados de Yahoo Finanças (2020).

O Período II também foi dividido em quatro subperíodos para análise das regiões temporais:

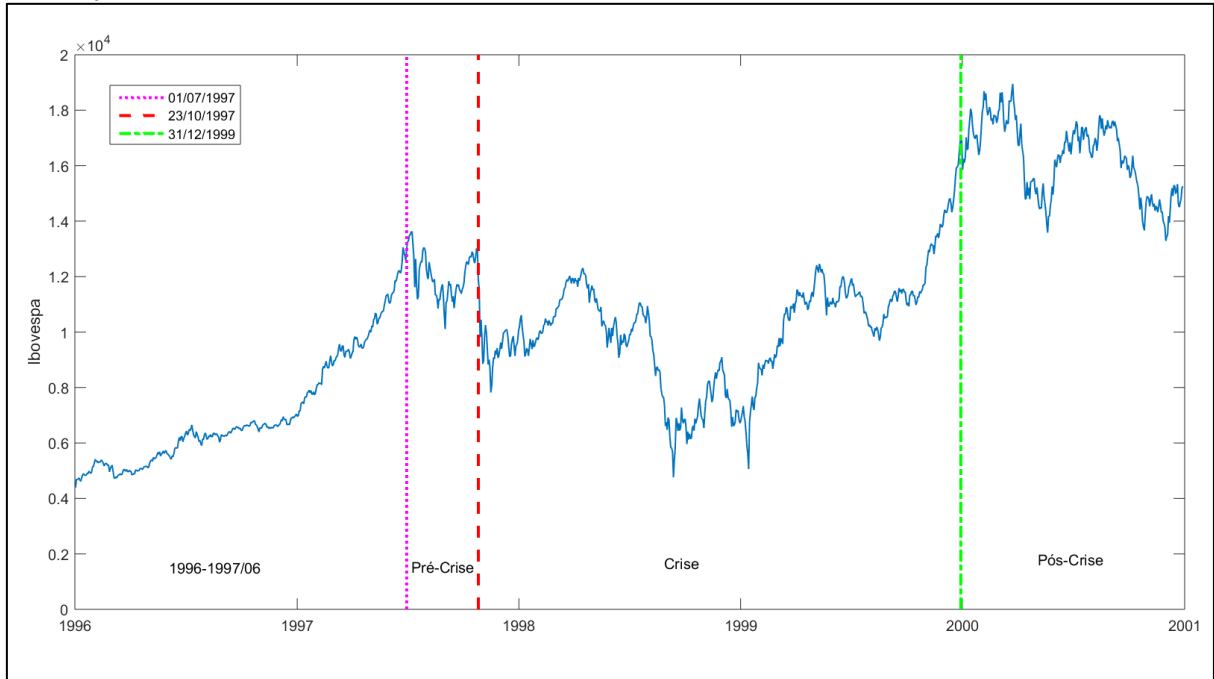
- 1996-1997/06: de 02/01/1996 a 30/06/1997 (com 370 retornos)
- Pré-Crise: de 01/07/1997 a 22/10/1997 (com 80 retornos);
- Crise: de 23/10/1997 a 30/12/1999 (com 539 retornos);
- Pós-Crise: de 03/01/2000 a 28/12/2000 (com 247 retornos).

Apesar da crise asiática ter sido iniciada a partir de julho de 1997 em alguns países asiáticos (LIMA; DEUS, 2013; ZAHAR, 2005), o dia 23 de outubro de 1997 foi definido como o início da contaminação da crise asiática em proporções mundiais pois neste dia foi noticiado pela mídia que a bolsa de Hong Kong sofreu a sua pior baixa (ZAHAR, 2005), o que levou a um efeito cascata causando um crash global. Inclusive esta contaminação foi sentida no Brasil, resultando na quarta maior baixa na história da Bolsa de Valores de São Paulo, com um recuo de 14,97%, o pior desempenho entre todas as Bolsas do mundo na época (FOLHA DE SÃO PAULO, 1997).

A Figura 10 ilustra a divisão entre os períodos que serão analisados, incluindo o início (linha vermelha) e fim (linha verde) da contaminação crise

asiática no contexto do mercado brasileiro sobre a variação do Ibovespa. Pode-se observar a nítida queda que se inicia no fim de 1997 e se alastra até o fim de 1998.

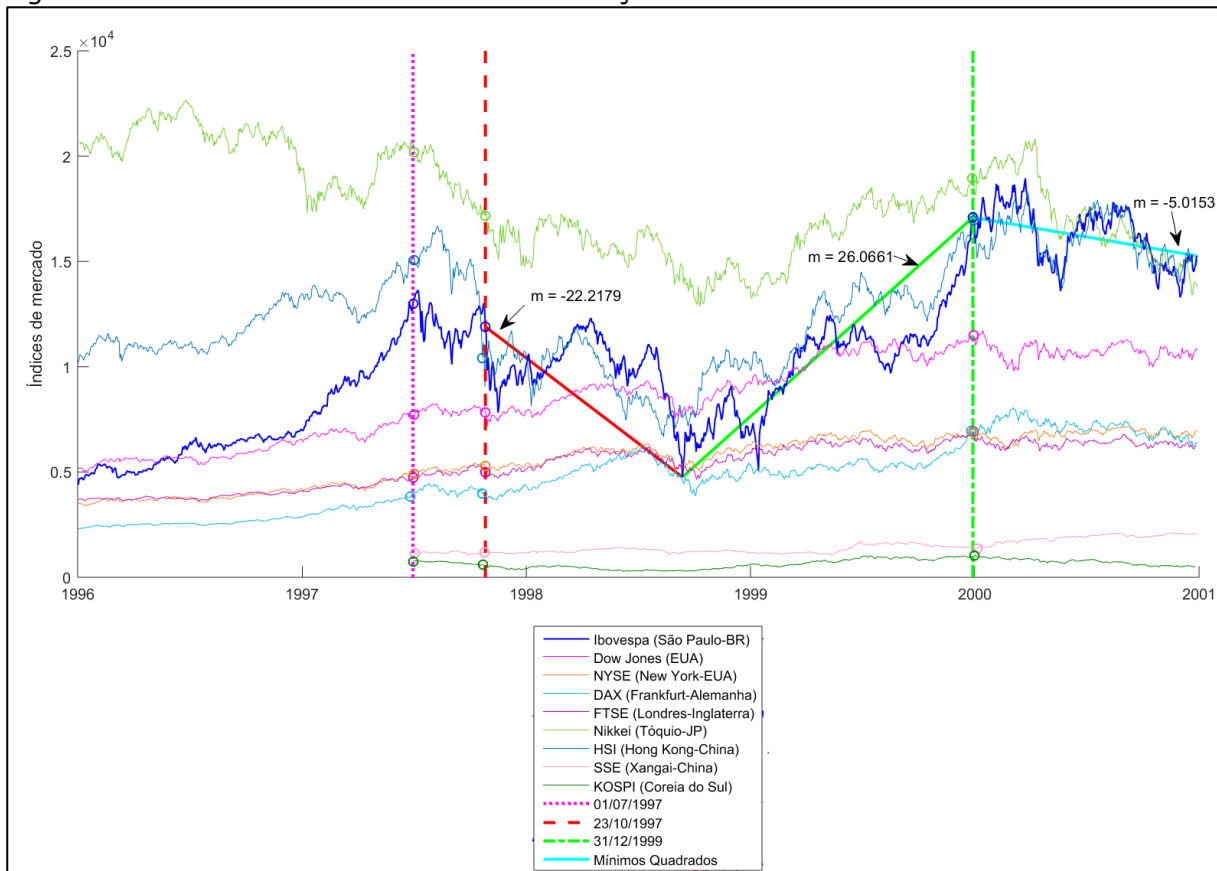
Figura 10 - Divisão do Período II em subperíodos analisados sobre a variação do Ibovespa.



Fonte: Autor. Baseado em dados de Yahoo Finanças (2020).

O final do ano de 1999 foi escolhido como a data para o fim do período de crise no contexto brasileiro visto que o país sofreu uma sucessão de impactos em sua economia em decorrência de eventos advindos a partir da crise asiática de 1997, como a crise russa de 1998 e a crise de desvalorização do real em 1999 (GOLDFAJN, 2000; REIS, 2019; REMDE, 2013). Desta forma, uma estabilidade na economia do país foi vista a partir de 2000, também acompanhado da estabilidade de outras economias de países pelo mundo, conforme pode ser observado na Figura 11, onde a formação de alguns *Ws* poder ser evidenciada durante o período de crise.

Figura 11 - Início e fim da crise de contaminação asiática sobre índices internacionais.



Fonte: Autor. Baseado em dados de Yahoo Finanças (2020).

Observa-se que em 2000 muitos mercados ao redor do mundo já haviam se recuperado da crise financeira asiática, como se pode observar através do comportamento dos índices de diferentes bolsas de valores apresentados na Figura 11. Para o mercado acionário brasileiro, pelo método dos mínimos quadrados, observa-se uma tendência negativa a partir do início da crise (coeficiente angular $m \cong -22.2179$), seguida de uma tendência positiva até o final do período de crise (coeficiente angular $m \cong 26.0661$). Durante o ano de 2000, a inclinação é pequena, atingindo-se um platô, o que caracteriza uma recuperação em $\sqrt{\quad}$ (GUIMARÃES, 2020), indicando que a crise já havia sido superada no contexto brasileiro.

5. Resultados e Análises

A seguir são apresentados os resultados da aplicação da TMA sobre os dados do mercado brasileiro juntamente com as análises baseadas na literatura e suportada por algumas técnicas de análise multivariada.

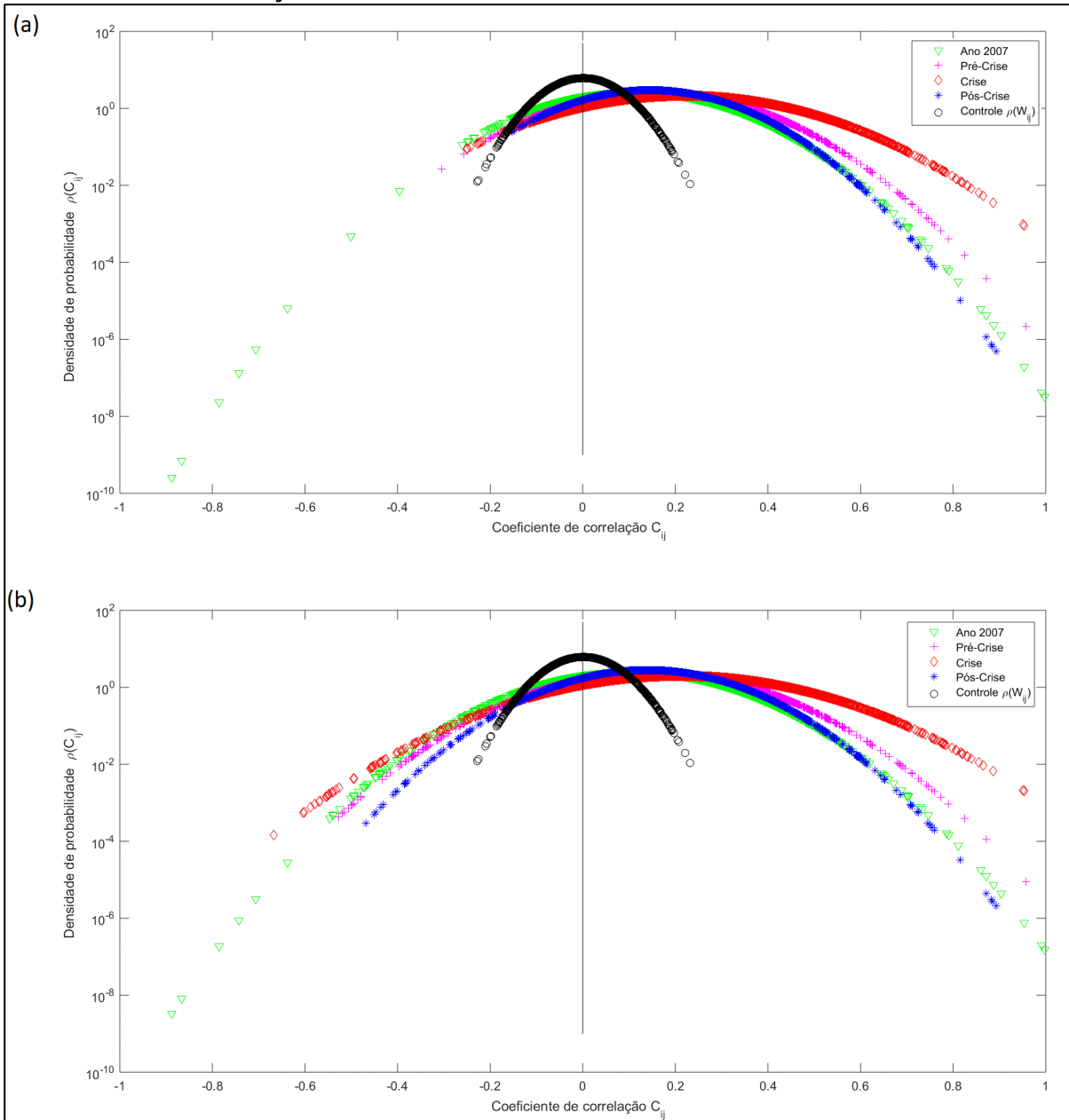
5.1 Análise das Matrizes de Correlação

As matrizes de retorno normalizado (Equação (2)), correspondentes a cada subperíodo analisado, foram geradas e deram origem as matrizes de correlação correspondentes (Equação (3)).

Como cada subperíodo analisado possui dimensões diferentes, foram geradas quatro matrizes aleatórias, cada uma com dimensões correspondentes às matrizes de retorno. A partir dessas, foram geradas as matrizes de correlação aleatória conforme Equação (5).

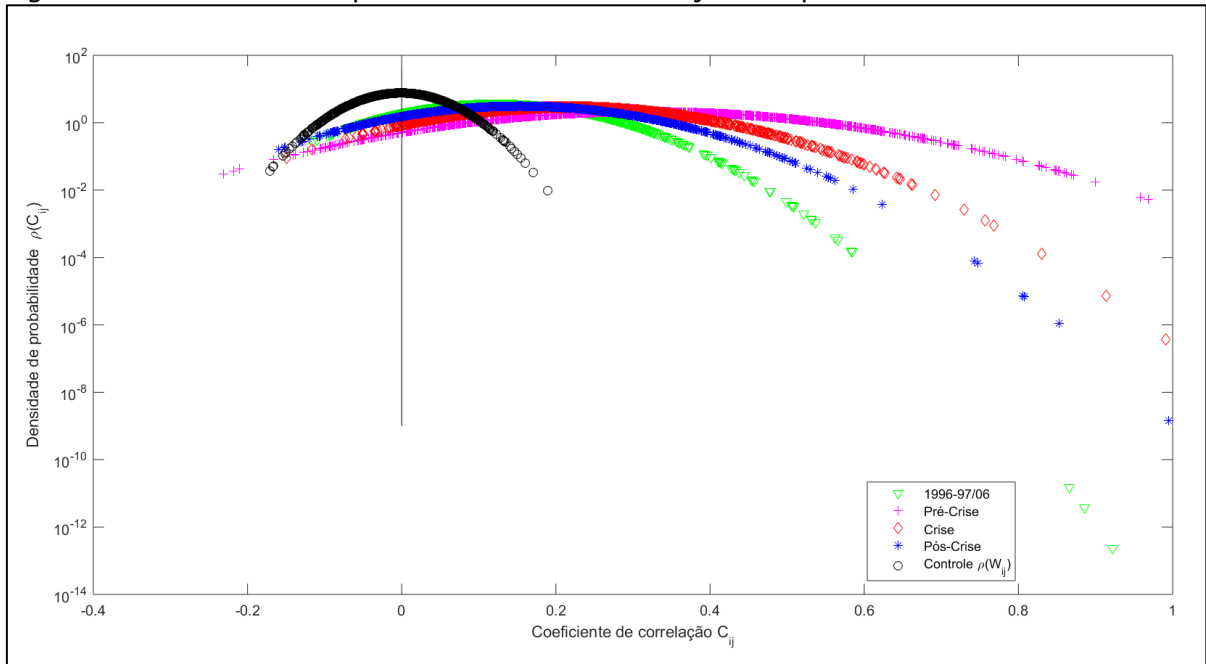
As distribuições de probabilidade $\rho(c_{i,j})$ dos elementos $\{c_{i,j}; i \neq j\}$ das matrizes de correlação empírica foram analisadas e podem ser observadas nas Figuras 12 e 13, juntamente com a curva de controle $\rho(w_{i,j})$ correspondente aos elementos $\{w_{i,j}; i \neq j\}$ de uma matriz de correlação aleatória W . Este mesmo procedimento foi realizado para a matriz de retornos normalizados adicionada a suposta ação anticorrelacionada representada pela cotação do dólar em reais para o Período I e pode ser observado na Figura 12(b).

Figura 12 - Densidade de probabilidade das correlações empíricas do Período I: (a) para dados empíricos originais; (b) para dados empíricos adicionada a cotação do dólar em reais simulando uma ação anticorrelacionada.



Fonte: Autor.

Figura 13 - Densidade de probabilidade das correlações empíricas do Período II.



Fonte: Autor.

As Figuras 12 e 13 mostram que $\rho(w_{i,j})$ (círculos pretos) é consistente com uma distribuição Gaussiana (Equação (14)) com média 0, sendo um comparativo para o comportamento das demais distribuições empíricas $\rho(c_{i,j})$. Notou-se que a $\rho(c_{i,j})$ para todos os subperíodos estudados é positivamente enviesada, centrada em torno de um valor médio positivo ($\langle c_{i,j} \rangle > 0$), o que implica que a correlação positiva é prevalente em relação ao comportamento negativamente correlacionado (anticorrelacionado). Além disso, a distribuição referente ao comportamento anticorrelacionado na Figura 12(a) para os períodos de pré-crise, crise e pós-crise, e na Figura 13 para os períodos de 1996-1997/06, crise e pós-crise, está localizada sobre a curva de controle $\rho(w_{i,j})$, podendo indicar que se trata apenas de um efeito da aleatoriedade, o que acorda com a literatura (PLEROU *et al.*, 2002). No entanto, existem correlações negativas principalmente para o subperíodo de 2007 do Período I que ultrapassam o limite da curva de controle $\rho(w_{i,j})$, demonstrando que para este subperíodo provavelmente houve anticorrelacionado genuíno.

A Figura 12(b) apresenta a mesma análise de densidades de probabilidade, mas considera a adoção da cotação do dólar em reais como

um fator de perturbação dos dados. De fato, sua inclusão influenciou $\rho(c_{i,j})$, tornando a parte negativa da correlação mais representativa para todos os subperíodos analisados neste caso.

Observou-se também que a média e variância das correlações são sensíveis ao tempo, ou seja, para o Período I tem-se:

- 2007: média = 0,11047 e variância = 0,02161;
- pré-crise: média = 0,15514 e variância = 0,02303;
- crise: média = 0,21870 e variância = 0,03482;
- pós-crise: média = 0,14532 e variância = 0,01786;

Para o Período II tem-se:

- 1996-1997/06: média = 0.12117 e variância = 0.01055;
- pré-crise: média = 0.31648 e variância = 0.03288;
- crise: média = 0.21381 e variância = 0.01903;
- pós-crise: média = 0.15301 e variância = 0.01646;

Tais comportamentos podem indicar uma tendência de readequação do mercado, retomando um comportamento semelhante àquele anterior à crise, porém mantendo alguns resquícios da crise superada. Este comportamento vem ao encontro com o que é apresentado na literatura (SENSOY; YUKSEL; ERTURK, 2013; DALKIR, 2009). Observa-se, ainda, que durante os subperíodos que antecedem à uma crise, o mercado apresentou um movimento de mudança em seu comportamento, o qual pode ser encarado como um "sinal de alerta" de que algo estava diferente do comportamento rotineiro.

Desta forma, neste trabalho será chamar de "efeito mola deformada" esse movimento de mudança no comportamento do mercado durante um período anterior à crise, já sinalizando esta, e, após uma crise, a tendência do mercado de regredir a um comportamento característico de um período sem volatilidade, porém ainda apresentando indícios da crise superada, uma vez que uma mola deformada não volta totalmente a seu formato inicial.

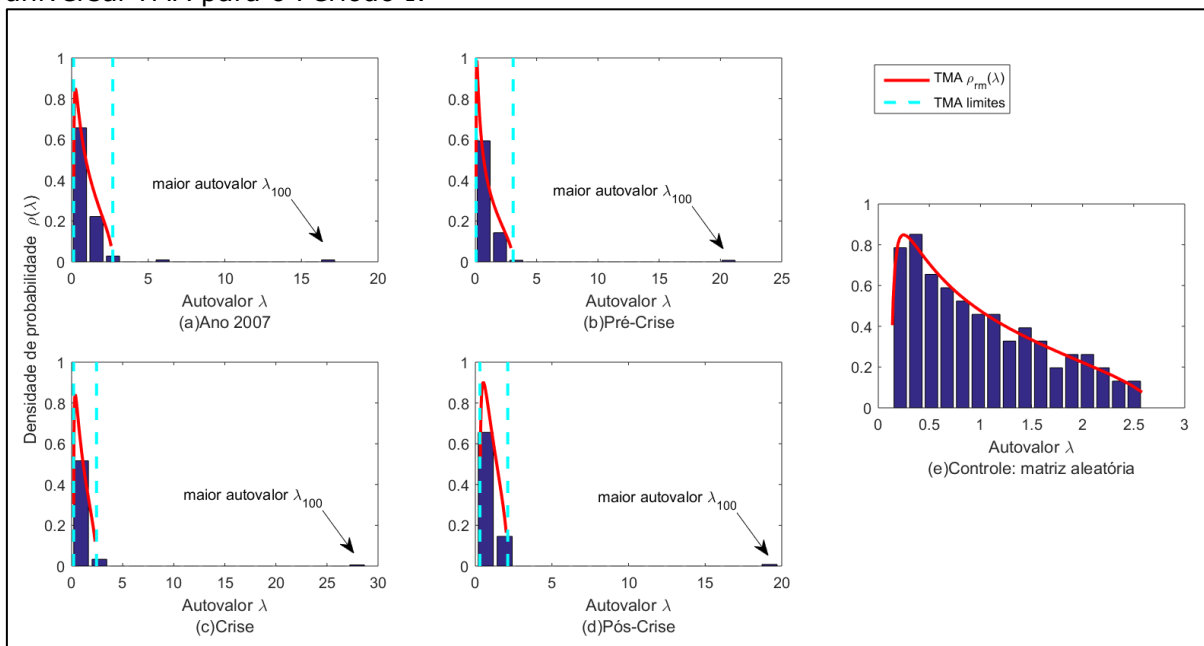
O maior grau de magnitude da média dos elementos da matriz de

correlação para o período de crise em relação um período anterior não contaminado pela crise, reflete uma probabilidade maior de existir fortes correlações positivas durante o período da crise. Assim sendo, em uma análise preliminar, os dados sugerem um maior sincronismo do mercado durante o período de crise, uma vez que, em média, cada empresa teve maior influência no comportamento das demais, o que acorda com a literatura (PLEROU *et al.*, 2002; KULKARNI; DEO, 2007; DALKIR, 2009; SANDOVAL JUNIOR; FRANCA, 2012).

5.2 Análise dos Autovalores

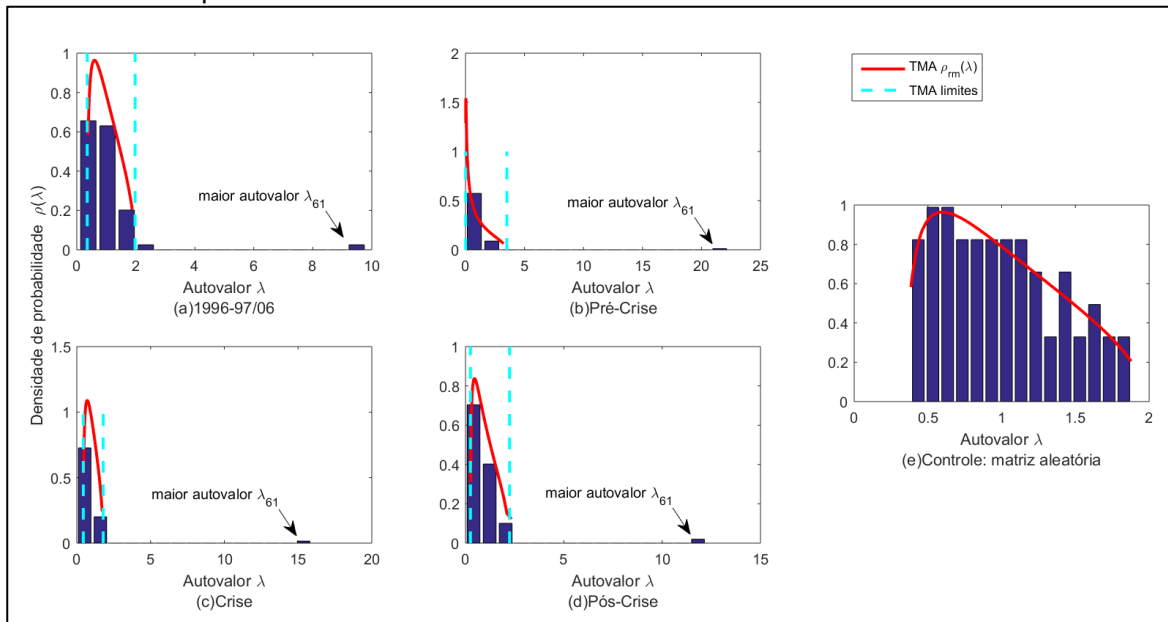
Os histogramas das Figuras 14 e 15 comparam a distribuição de probabilidade dos autovalores empíricos $\rho(\lambda)$ (barras azuis) referente aos períodos analisados, com as distribuições teóricas (linhas vermelhas), calculadas conforme Equação (6), com valores de Q para o Período I iguais a 2,45; 1,75; 3,18 e 4,83 para os subperíodos de 2007, pré-crise, crise e pós-crise respectivamente e com valores de Q para o Período II iguais a 6,07; 1,31; 8,8 e 4 para os subperíodos de 1996-1997/06, pré-crise, crise e pós-crise respectivamente. Pode-se perceber um bom ajuste entre as distribuições teórica e empírica.

Figura 14 – Distribuição de probabilidade dos autovalores empíricos vs. comportamento universal TMA para o Período I.



Fonte: Autor.

Figura 15 - Distribuição de probabilidade dos autovalores empíricos vs. comportamento universal TMA para o Período II.



Fonte: Autor.

Observou-se que a maioria dos autovalores para ambos os períodos estão dentro da região de Wishart ou próximos aos limites teóricos $[\lambda_{min}, \lambda_{max}]$ da Equação (6). Também foram observados desvios dos maiores e menores autovalores. Isto nos permite separar a correlação genuína (referente aos autovalores que ultrapassam o limite teórico da TMA) da correlação aleatória (referente aos autovalores que respeitam os limites teóricos da TMA). Desta forma, a análise dos dados nos sugere que a maior parte das correlações são aleatórias, exceto pelas correlações referente aos autovalores que se desviam.

Em particular, o maior autovalor de cada subperíodo são os que mais se distanciam dos limites teóricos. Para o Período I temos:

- 2007: $\lambda_{100} \approx 17,28$ é aproximadamente seis vezes e meia maior que o limite teórico superior ($\lambda_{max} \approx 2,6$);
- pré-crise: $\lambda_{100} \approx 21,33$ é aproximadamente sete vezes maior que o limite teórico superior ($\lambda_{max} \approx 3$);
- crise: $\lambda_{100} \approx 28,85$ o qual é aproximadamente 12 vezes maior que o limite teórico superior ($\lambda_{max} \approx 2,44$) sendo que este não só se desvia dos limites teóricos, mas apresenta uma magnitude muito maior que os demais períodos;

- pós-crise: $\lambda_{100} \approx 19,82$ é aproximadamente nove vezes maior que seu limite teórico superior ($\lambda_{max} \approx 2,12$).

Para o Período II temos:

- 1996-1997/06: $\lambda_{61} \approx 9,81$ é aproximadamente cinco vezes maior que o limite teórico superior ($\lambda_{max} \approx 1,98$);
- pré-crise: $\lambda_{61} \approx 22,29$ é aproximadamente seis vezes e meia maior que o limite teórico superior ($\lambda_{max} \approx 3,5$);
- crise: $\lambda_{61} \approx 15,9$ o qual é aproximadamente nove vezes maior que o limite teórico superior ($\lambda_{max} \approx 1,8$) sendo que este não só se desvia dos limites teóricos, mas apresenta uma magnitude muito maior que o limite superior;
- pós-crise: $\lambda_{61} \approx 12,29$ é aproximadamente cinco vezes e meia maior que seu limite teórico superior ($\lambda_{max} \approx 2,24$).

Para averiguar que os desvios encontrados nas Figuras 14 e 15 não são um efeito da limitação do conjunto de dados (PLEROU *et al.*, 2002; SANDOVAL JUNIOR; FRANCA, 2012), comparou-se as distribuições teóricas da Equação (6) com as distribuições de probabilidade dos autovalores das matrizes de correlação aleatória da Equação (5) através das Figuras 14(e) e 15(e). Encontrou-se um bom ajuste, mostrando, assim, que os desvios da TMA encontrados nas Figuras 14 e 15 para os maiores autovalores não são resultantes do fato de que T (tamanho a série temporal) e N (quantidade de ações analisadas) são finitos, uma vez que os dados das Figuras 14(e) e 15(e) tiveram origens de matrizes aleatórias de mesmas dimensões $T \times N$.

Pode-se ainda analisar a quantidade de autovalores empíricos que ultrapassam os limites teóricos esperados pela TMA. Tal comparação nos permite distinguir o espectro de autovalores que se ajustam a TMA (correlações aleatórias) dos desvios (correlações genuínas). Para o Período I temos:

- 2007: apresenta nove autovalores menores que o limite inferior λ_{min} e quatro autovalores que ultrapassam o limite superior

λ_{max} .

- pré-crise: quatro autovalores são menores que o limite teórico inferior e dois são maiores que o limite teórico superior.
- crise: apresenta 19 autovalores que são menores que λ_{min} e três são maiores que λ_{max} .
- pós-crise: apresenta 15 autovalores menores que o limite teórico inferior e quatro valores que ultrapassaram o limite λ_{max} .

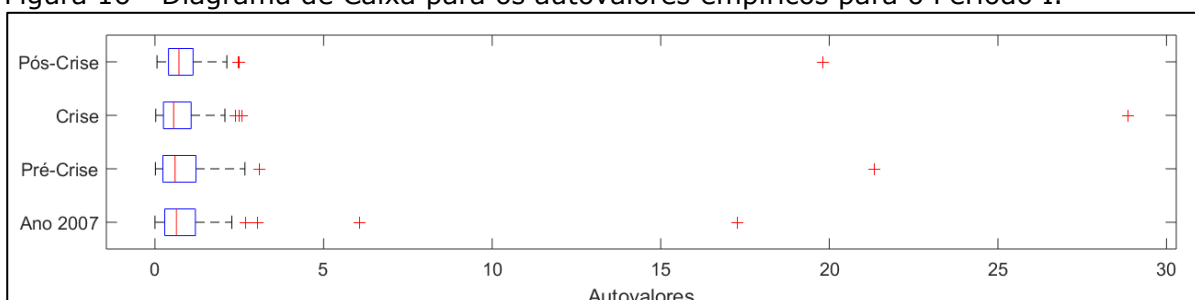
Para o Período II temos:

- 1996-1997/06: apresenta oito autovalores menores que o limite interior λ_{min} e dois autovalores que ultrapassam o limite superior λ_{max} .
- pré-crise: três autovalores são menores que o limite teórico inferior e um é maior que o limite teórico superior.
- crise: apresenta 15 autovalores que são menores que λ_{min} e três são maiores que λ_{max} .
- pós-crise: apresenta oito autovalores menores que o limite teórico inferior e dois valores que ultrapassaram o limite λ_{max} .

Os dados apresentados demonstram novamente um comportamento característico do “efeito mola deformada” já descrito previamente na seção 5.1, apresentando uma dependência temporal ainda mesmo antes do período de crise, sendo, porém, mais proeminente durante a crise propriamente dita, e tendendo a regredir após este período.

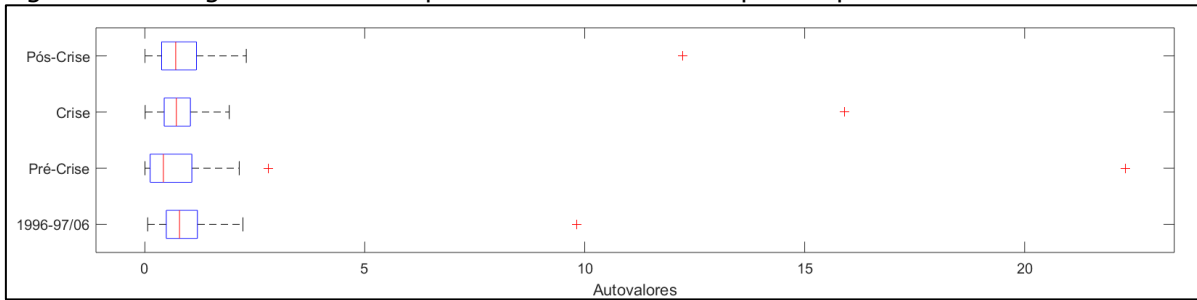
As Figuras 16 e 17 apresenta um Diagrama de Caixa para os autovalores referentes as matrizes de correlação empíricas para cada período analisado.

Figura 16 - Diagrama de Caixa para os autovalores empíricos para o Período I.



Fonte: Autor.

Figura 17 - Diagrama de Caixa para os autovalores empíricos para o Período II.

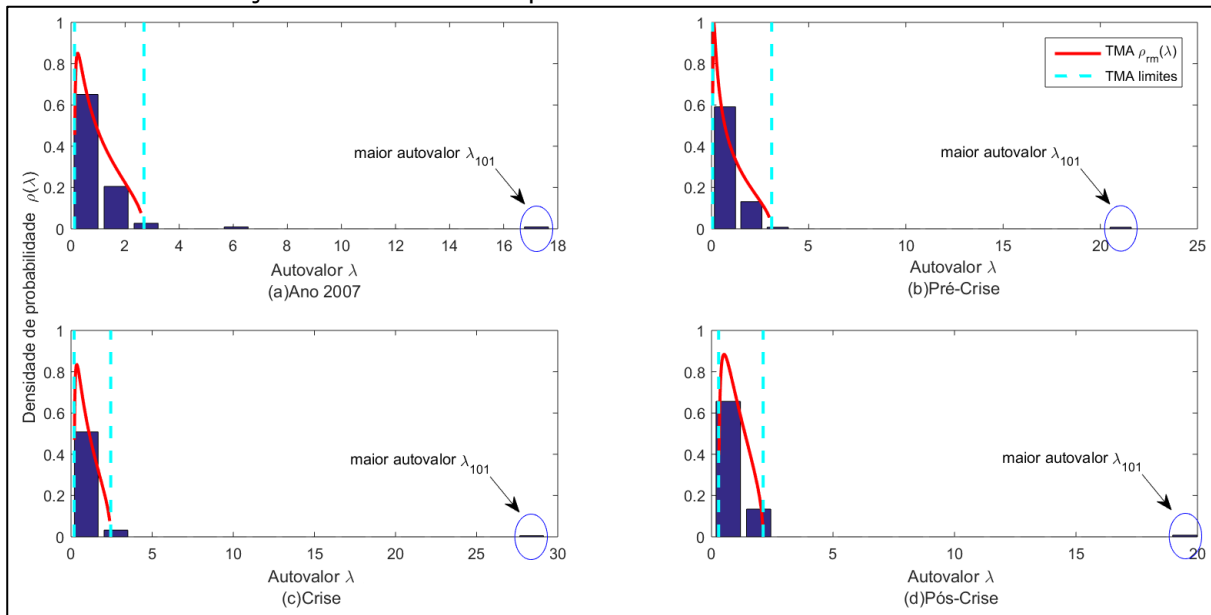


Fonte: Autor.

Percebe-se que, apesar de existir muitos autovalores que são menores que o limite inferior λ_{min} para todos os subperíodos, estes estão muito próximos da mediana de cada distribuição, sendo que os *outliers* apresentados nas Figuras 16 e 17 são referentes aos autovalores que ultrapassam superiormente os limites teóricos da TMA. Considerando que as magnitudes dos menores autovalores estão mais próximas do limite teórico inferior, é possível que as informações mais relevantes estejam representadas nos maiores valores próprios que ultrapassaram os limites teóricos.

A Figura 18 apresenta a mesma análise de densidades de probabilidade, mas considera a adoção da cotação do dólar em reais como um fator perturbador dos dados. Percebe-se que a interferência ocasionada pela participação do preço do dólar não foi percebida por esta análise o que nos leva a interpretar que uma única ação anticorrelacionada não tem muita influência nos resultados trazidos pelo TMA, por se tratar de uma técnica que avalia os resultados dos dados de forma global.

Figura 18 - Distribuição de probabilidade dos autovalores empíricos do Período I adicionado a variação do dólar vs. comportamento universal TMA.



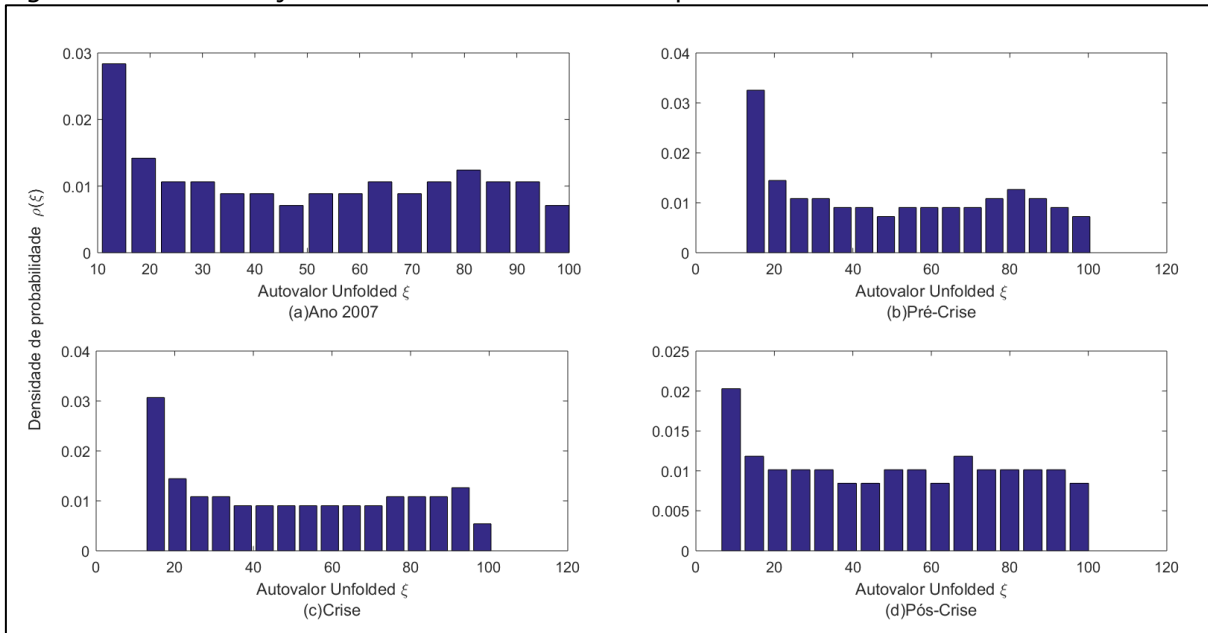
Fonte: Autor.

5.3 Autovalores Empíricos vs. Propriedades Universais da TMA

É necessário, no entanto, aprofundar a análise dos autovalores para averiguar se o que foi observado sobre estes por meio das Figuras 14, 15, 16 e 17 pode ser validado. Esta validação será feita a seguir através de comparações entre as propriedades universais da TMA apresentadas na Seção 3.6 e os dados empíricos. Esta investigação sobre o comportamento dos autovalores, no entanto, incluiu o mapeamento destes para os valores *unfolded* ξ_i . Foi aplicada, para tanto, uma nova abordagem para o processo de *unfolding*, o qual foi proposta e apresentado durante a metodologia na Seção 3.5. Os seguintes valores de variáveis foram utilizados para o processo de *unfolding* referente aos períodos analisados: $N = 100$ e $Deg = 4$.

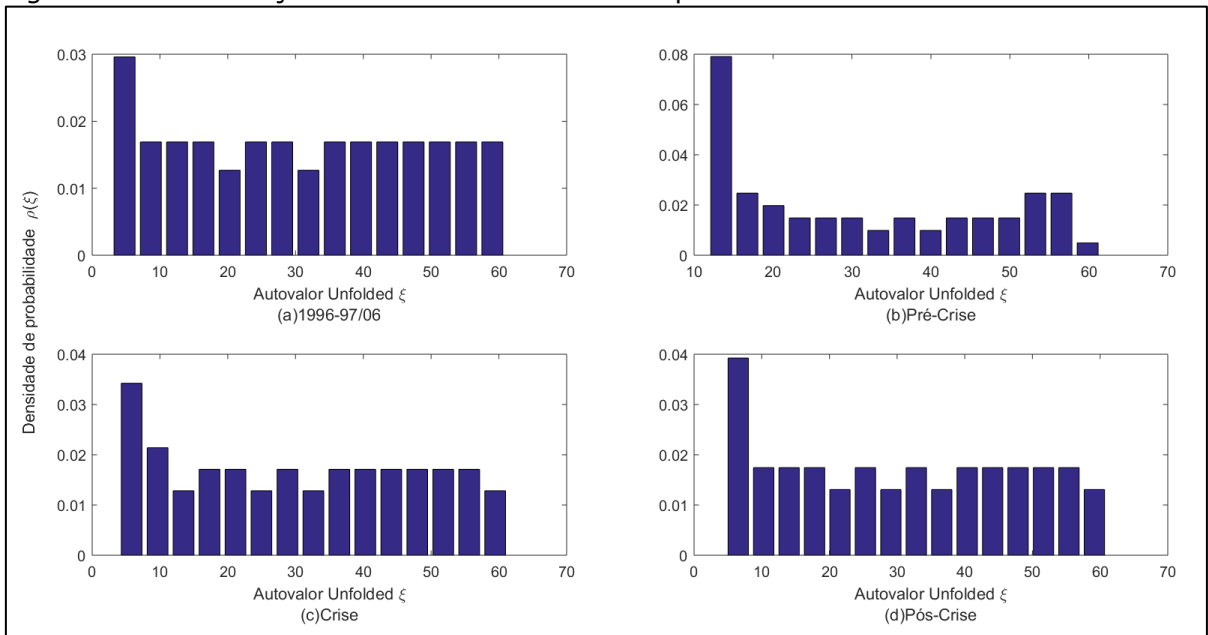
De acordo com a literatura (PLEROU *et al.*, 2002), a distribuição dos valores resultantes do processo de *unfolding* deve ser uniforme. Assim sendo, as Figuras 19 e 20 demonstram que os valores resultantes para ξ_i apresentam uma distribuição uniforme aproximada para os autovalores empíricos.

Figura 19 - Distribuição dos autovalores *unfolded* para o Período I.



Fonte: Autor.

Figura 20 - Distribuição dos autovalores *unfolded* para o Período II.

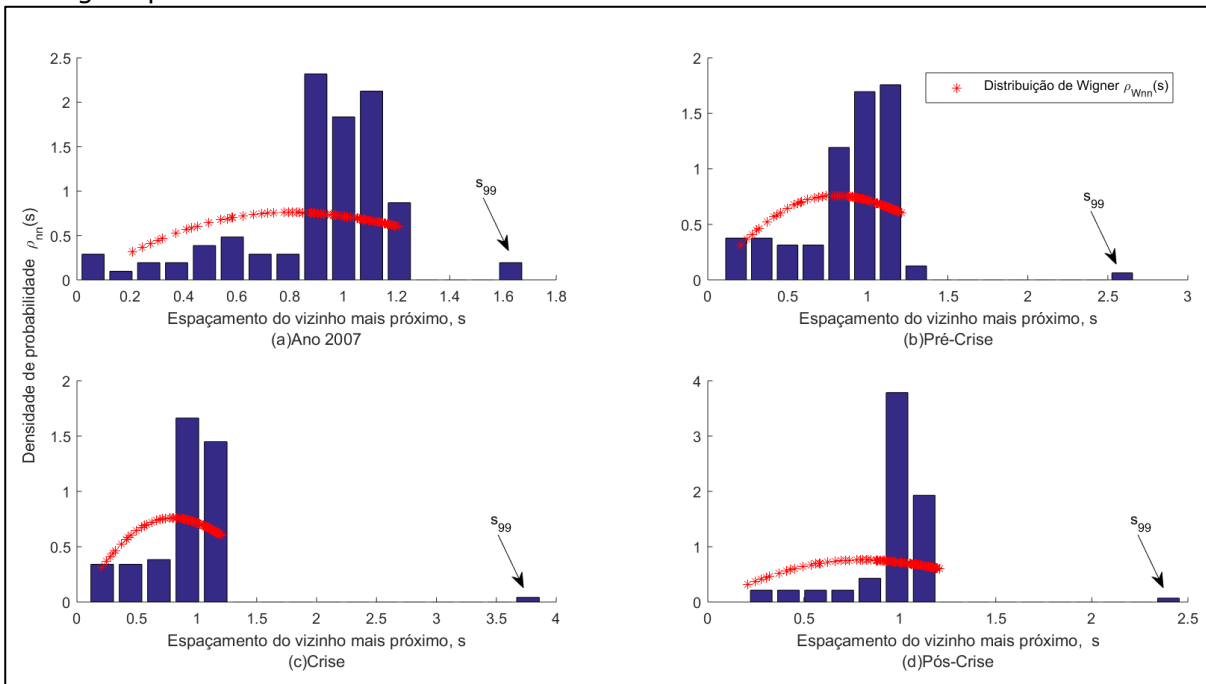


Fonte: Autor.

A seguir, os autovalores *unfolded* ξ_i são testados para as propriedades universais conhecidas para a TMA.

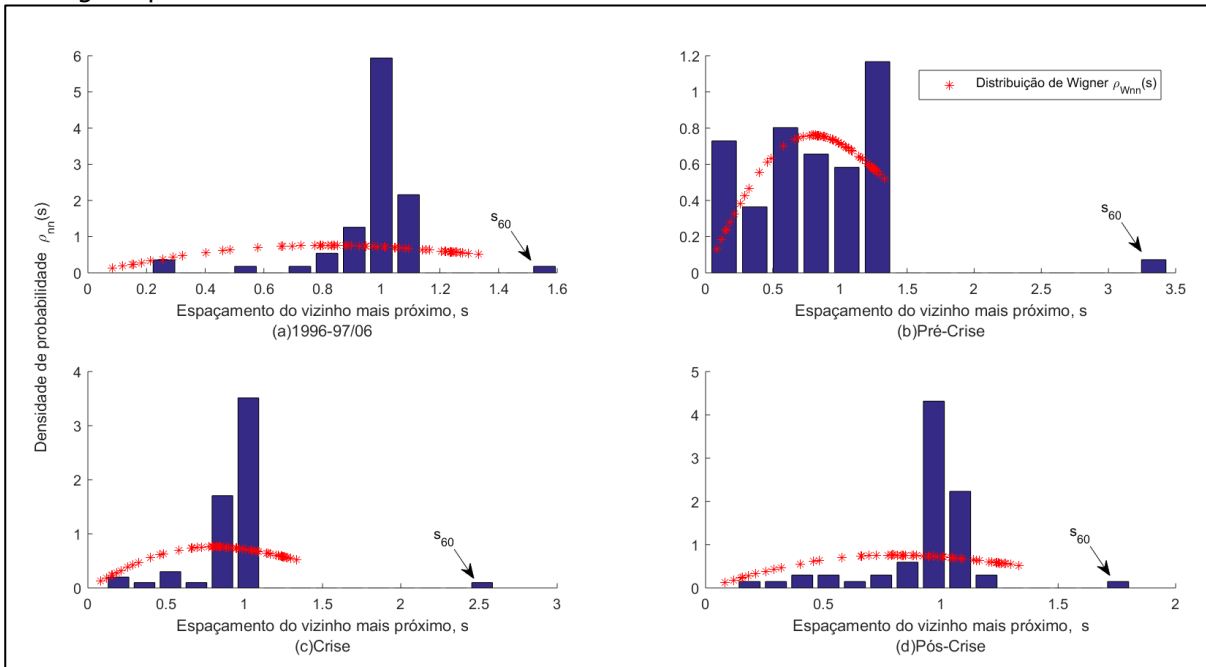
Os histogramas das Figuras 21 e 22 apresentam a distribuição dos espaçamentos entre vizinhos próximos $s = \xi_{i+1} - \xi_i$ para os períodos estudados. Esta distribuição é comparada a distribuição teórica de Wigner $P_{Wvp}(s)$, conforme Equação (9).

Figura 21 - Distribuição dos vizinhos próximos dos autovalores *unfolded* vs. distribuição de Wigner para o Período I.



Fonte: Autor.

Figura 22 - Distribuição dos vizinhos próximos dos autovalores *unfolded* vs. distribuição de Wigner para o Período II.



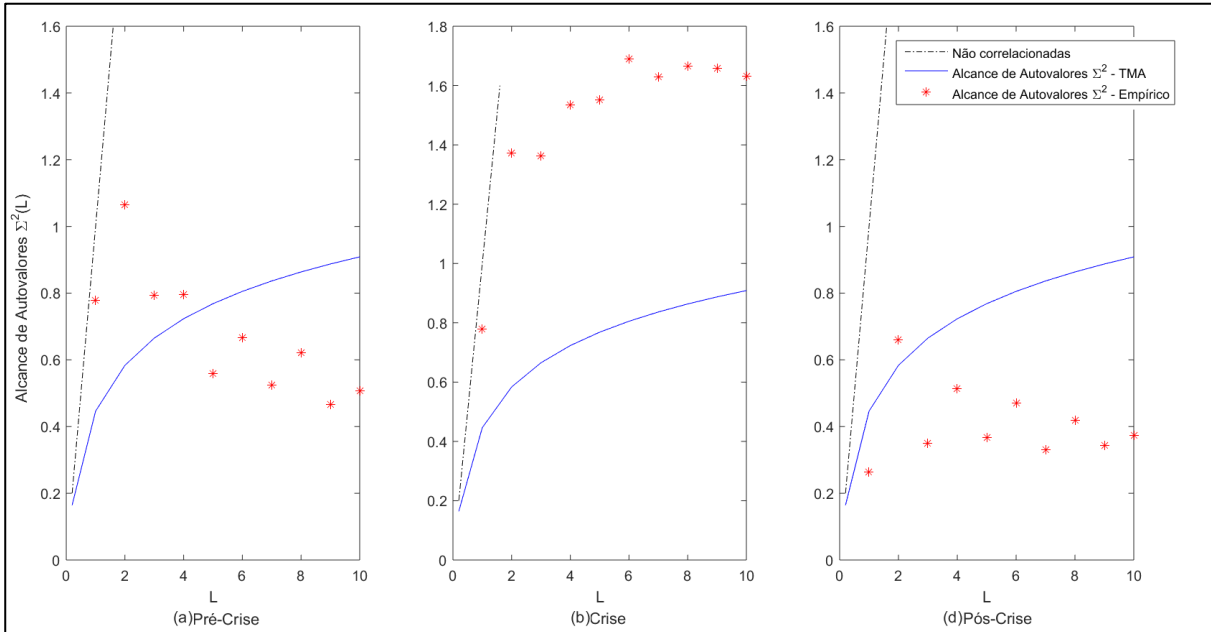
Fonte: Autor.

Vê-se um bom ajuste entre essas distribuições para todos os subperíodos analisados, o que ajuda a fortalecer o argumento de que grande parte da correlação empírica é resultante de aleatoriedade. Porém, percebe-se que as distâncias s_{99} relativa aos ξ_{100} do Período I e s_{61} relativa

aos ξ_{61} do Período II de cada subperíodo são *outliers* quando comparadas às demais distâncias. Sendo assim, estes estão relacionados às correlações “puras e verdadeiras”, e é possível perceber que aqueles autovalores que ultrapassaram o limite inferior λ_{min} na seção 5.2 perdem a importância ao retirar a parte sensível da distribuição durante o processo de *unfolding*, uma vez que os autovalores que estão no espectro do comportamento previsto pela TMA são apenas ruídos, resultantes de correlações aleatórias, não sendo, dessa forma, uma correlação genuína (PLEROU *et al.*, 2002). Esse comportamento é mais evidente no período de crise com relação ao Período I. Para o período pré-crise, as distâncias s_{99} do Período I e s_{61} do Período II cresce em relação ao ano anterior e pode novamente representar uma indicação de uma possível crise que se aproxima. O período pós-crise também apresenta um comportamento singular com tendência de retorno ao comportamento apresentado no período anterior à crise, ressaltando mais uma vez o chamado “efeito mola deformada” do mercado, quando o mercado tende a retornar ao comportamento “padrão” existente antes da crise, porém, ainda mostrando sinais deixados pela crise em seu comportamento.

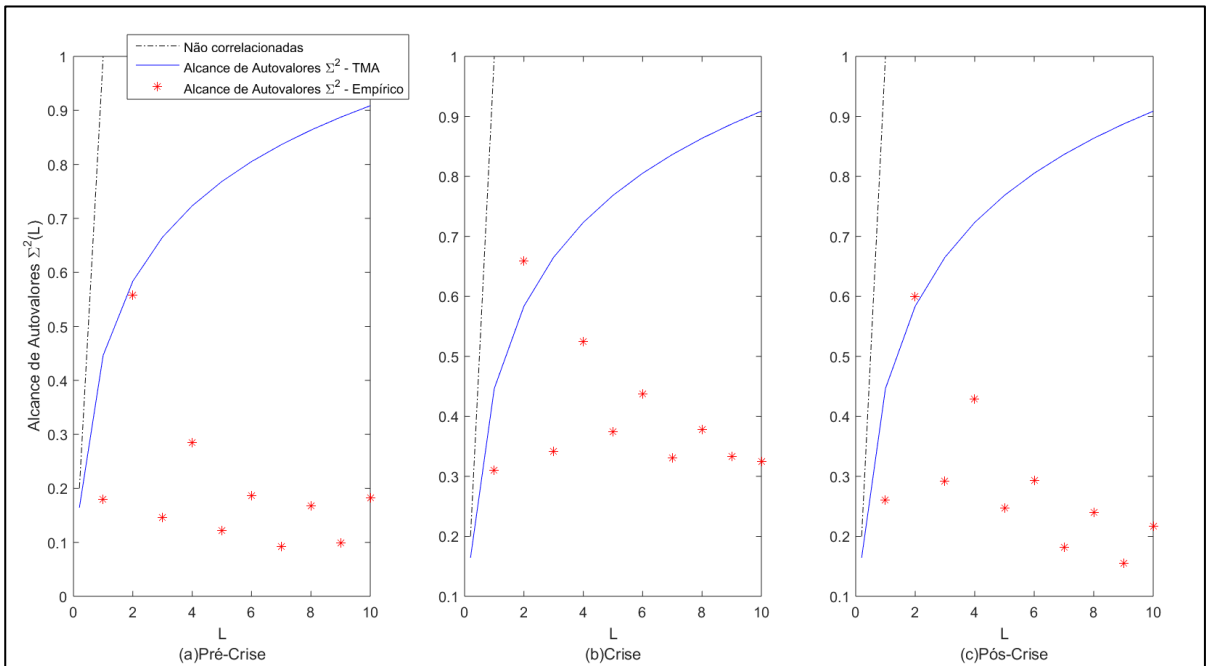
As Figuras 23 e 24 apresentam o alcance de autovalores empíricos Σ^2 da Equação (10) para as matrizes de correlação empíricas comparados ao alcance de autovalores teóricos Σ^2 da Equação (11). Vê-se um bom ajuste entre elas, principalmente quando comparadas a distribuição não correlacionada $\Sigma^2 \sim I$.

Figura 23 - Alcance dos autovalores *unfolded* vs. alcance de autovalores teóricos para o Período I.



Fonte: Autor.

Figura 24 - Alcance dos autovalores *unfolded* vs. alcance de autovalores teóricos para o Período II.



Fonte: Autor.

Para esta análise em particular, foram utilizados três períodos ao em vez de quatro, com o período de pré-crise sendo formado pela junção deste com o ano anterior, uma vez que esta divisão configurou uma melhor análise para o desfecho da crise, ou seja, um movimento de subida do Σ^2 empírico para ambos os período durante a época de crise e sua posterior

descida no pós-crise, realizando um movimento de onda com pico no subperíodo de crise, o que destaca o comportamento característico de ambas as crises. Para o Período I, o comportamento do subperíodo de crise vai além do que era esperado pela teoria TMA. Em uma análise mais profunda, a distância euclidiana entre o alcance de autovalores empírico Σ^2 da Equação (10) e o alcance de autovalores teórico Σ^2 da Equação (11), referente ao Período I, é igual a:

- pré-crise: 0,91271;
- crise: 2,82020;
- pós-crise: 1,09248.

Com relação ao Período II, é igual a:

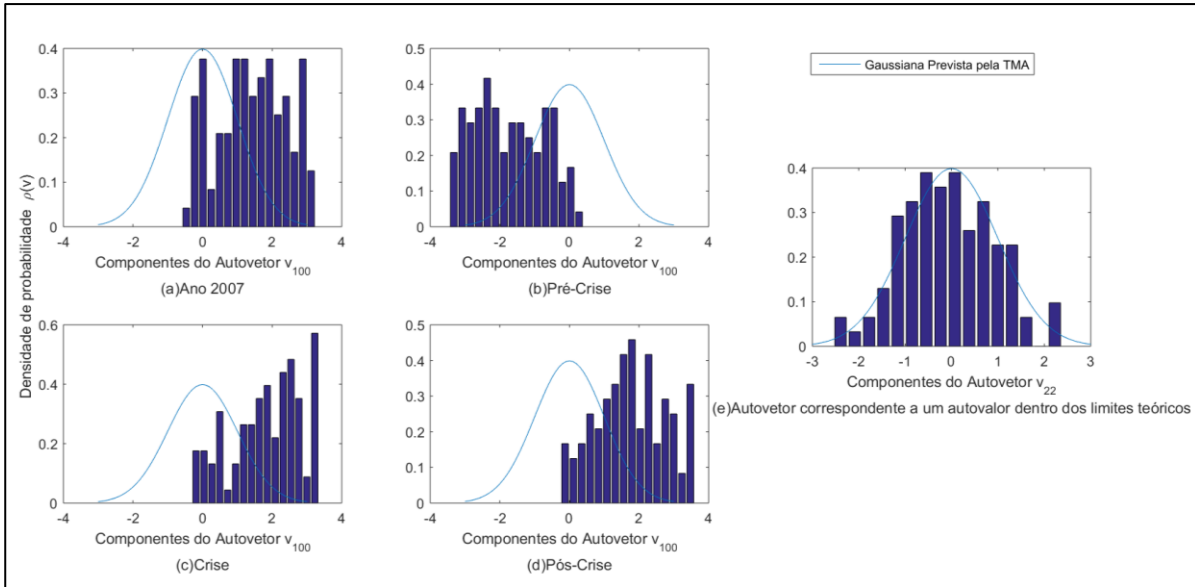
- pré-crise: 1.87666;
- crise: 1.26450;
- pós-crise: 1.62122.

Assim, novamente, torna-se evidente que o comportamento das correlações no período pós-crise tende a regredir ao comportamento existente antes da crise, porém não totalmente (efeito mola deformada).

5.4 Análise dos Autovetores Divergentes

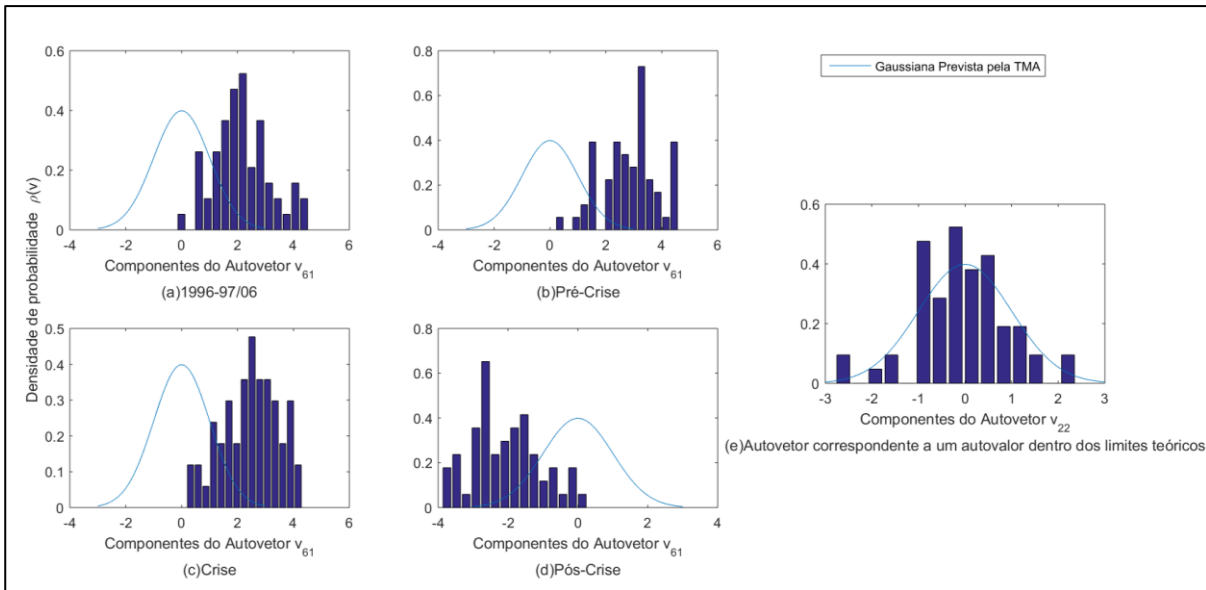
Para uma análise ainda mais profunda, os autovetores devem ser investigados quantitativamente. Assim, nas Figuras 25 e 26 é feita uma comparação entre as distribuições de componentes normalizados de autovetores relacionados aos maiores autovalores $\rho(v'_{100})$ para o Período I e $\rho(v'_{61})$ para o Período II das matrizes de correlação empíricas de cada período e a distribuição Gaussiana apresentada na Equação (14), representada pela linha sólida.

Figura 25 – Densidade dos componentes dos autovetores v'_{100} normalizados para o Período I.



Fonte: Autor.

Figura 26 - Densidade dos componentes dos autovetores v'_{61} normalizados para o Período II.



Fonte: Autor.

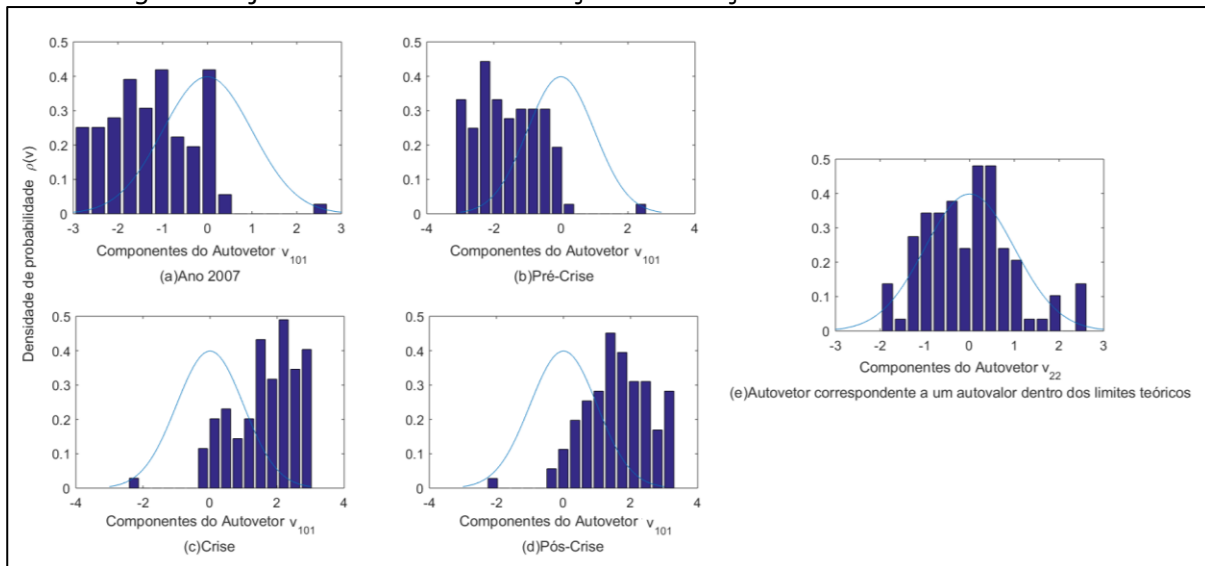
As Figuras 25(e) e 26(e) apresentam a distribuição $\rho(v'_{22})$ de um típico autovetor relativo a um autovalor pertencente ao intervalo teórico $[\lambda_{min}, \lambda_{max}]$. A distribuição dos componentes de tal autovetor se ajustou bem ao comportamento previsto pela TMA representado pela curva sólida (Equação (14)). Os outros histogramas apresentam as distribuições dos autovetores v'_{100} para o Período I e v'_{61} para o Período II relacionados ao maior autovalor λ_{100} e λ_{61} respectivamente para todos os subperíodos analisados que se

desviaram significativamente do comportamento normalmente distribuído. É perceptível que quase todos os componentes destes autovetores têm o mesmo sinal, o que desloca a distribuição para um dos lados, e são aproximadamente uniformes (sobretudo o subperíodo de crise do Período I) o que pode indicar que as empresas que participam significativamente de v_{100} no Período I e de v_{61} no Período II de cada subperíodo têm componentes comuns que afetam a todas elas e sugerem que a maioria das ações participam desse autovetor.

Percebe-se ainda que os componentes do autovetor relativo ao subperíodo de pré-crise do Período I e os componentes do subperíodo de pós-crise do Período II estão deslocados para a esquerda (lado negativo), opostamente do que ocorre com os autovetores dos demais períodos. Para uma interpretação mais assertiva deste comportamento precisa-se prosseguir com as análises. Porém, segundo os estudos de Sandoval Junior, Bortoluzzo e Venezuela (2014), um componente negativo de um autovetor pode significar uma aposta na desvalorização do ativo referente a tal componente. Desta forma, pode-se supor que para o período de pré-crise a TMA previu a desvalorização de muitos ativos, que foi exatamente o que aconteceu no período da crise de 2008.

A seguir será feita uma análise semelhante na Figura 27, porém com a interferência do suposto componente anticorrelacionado representado pela variação da taxa de câmbio do dólar em reais.

Figura 27 - Densidade dos componentes dos autovetores empíricos v'_{101} normalizados do Período I gerados juntamente com a variação da cotação do dólar em reais.



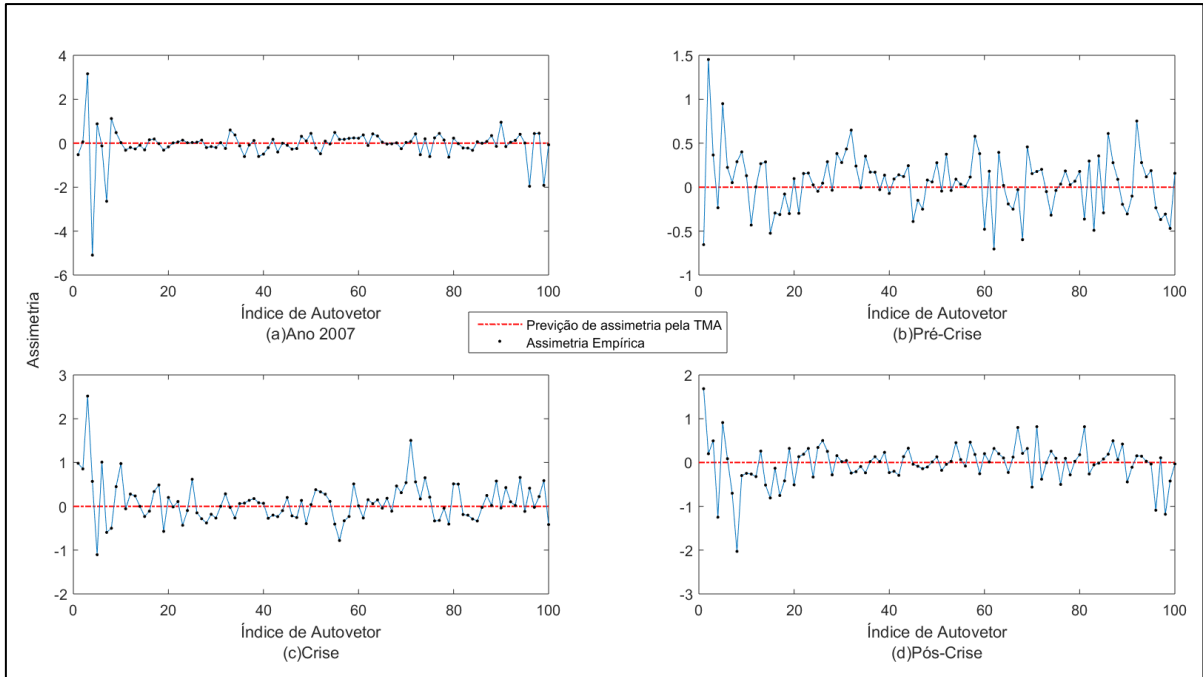
Fonte: Autor.

Uma análise semelhante a que foi realizada sobre a Figura 25 pode ser feita também para a Figura 27, porém agora a adoção da variação do dólar causa um pequeno distúrbio na distribuição dos componentes, de modo que um novo componente aparece em oposição aos demais. Além disso, a adição deste novo elemento foi o suficiente para deslocar os componentes do autovetor referente ao período de 2007 para a esquerda.

O ajuste da distribuição $\rho(v_i)$ à distribuição Gaussiana também pode ser verificado por algumas estatísticas descritivas como a assimetria e a medida de não achatamento padrão, as quais para a distribuição Gaussiana possuem valores padrões iguais a 0 e 3 respectivamente (PLEROU *et al.*, 2002; SENSOY; YUKSEL; ERTURK, 2013).

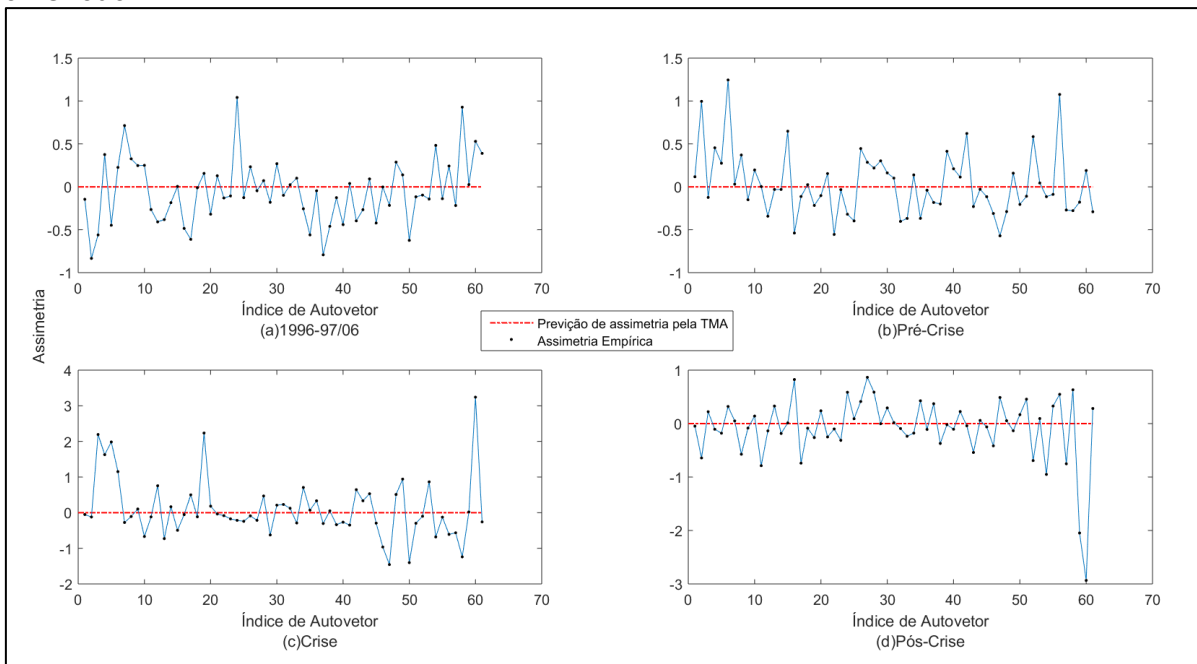
As Figuras 28 e 29 apresentam as assimetrias calculadas para as distribuições dos componentes de cada autovetor empírico v_i para todos os períodos analisados.

Figura 28 – Assimetria da distribuição dos componentes dos autovetores empíricos v_i para o Período I.



Fonte: Autor.

Figura 29 - Assimetria da distribuição dos componentes dos autovetores empíricos v_i para o Período II.

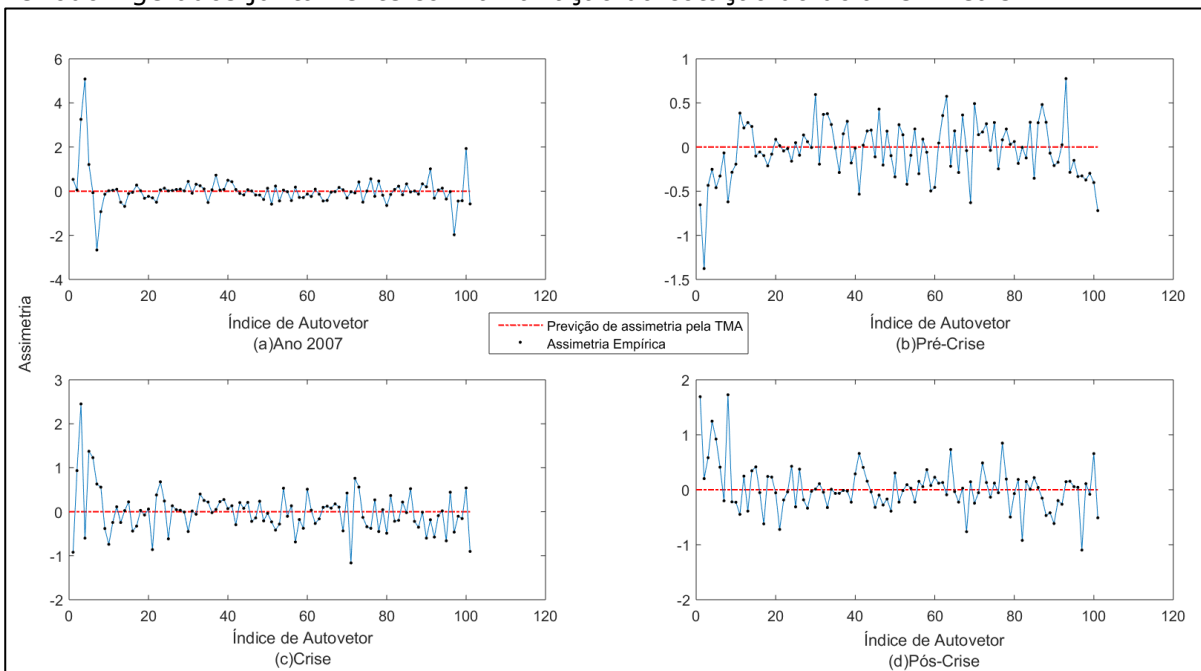


Fonte: Autor.

Percebe-se que para todos os períodos estudados as assimetrias dos componentes dos autovetores variam em torno de 0, o que é consistente com o comportamento normalmente distribuído (SENSOY; YUKSEL; ERTURK, 2013).

Comparando a Figura 28 com a Figura 30, pode-se notar outro efeito causado pela adoção da taxa de câmbio do dólar, um distúrbio, que corrobora com a análise anterior dos componentes do autovetor apresentados na Figura 27.

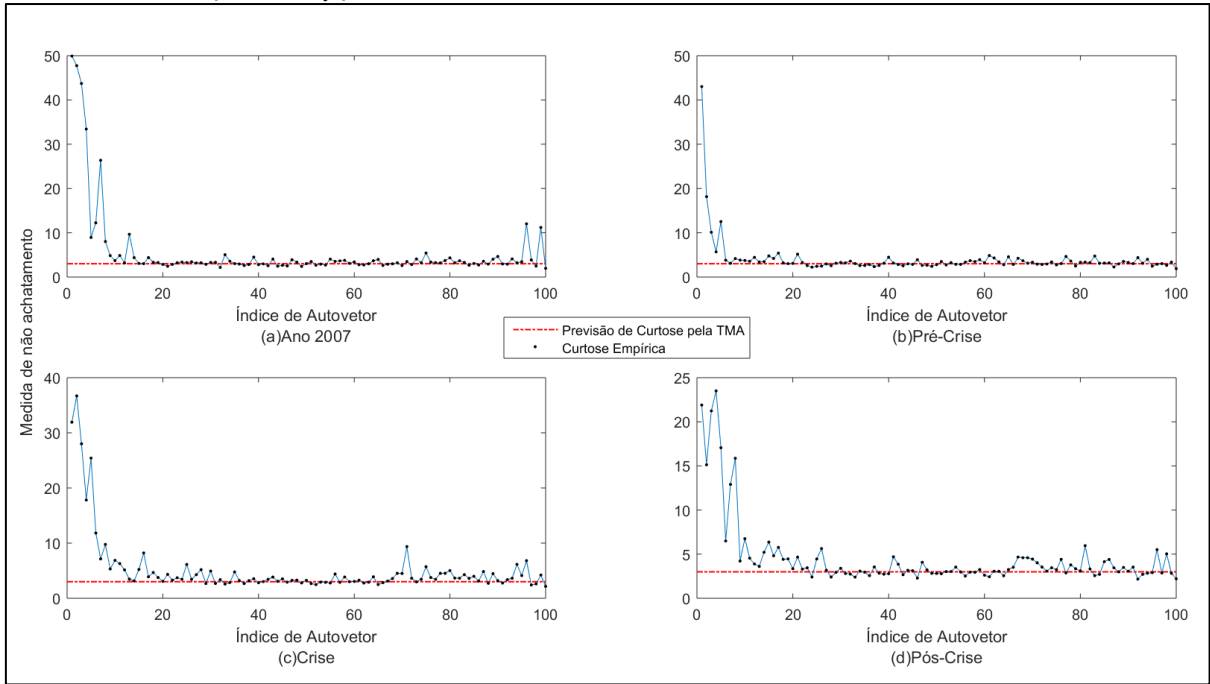
Figura 30 - Assimetria da distribuição dos componentes dos autovetores empíricos v_i do Período I gerados juntamente com a variação da cotação do dólar em reais.



Fonte: Autor.

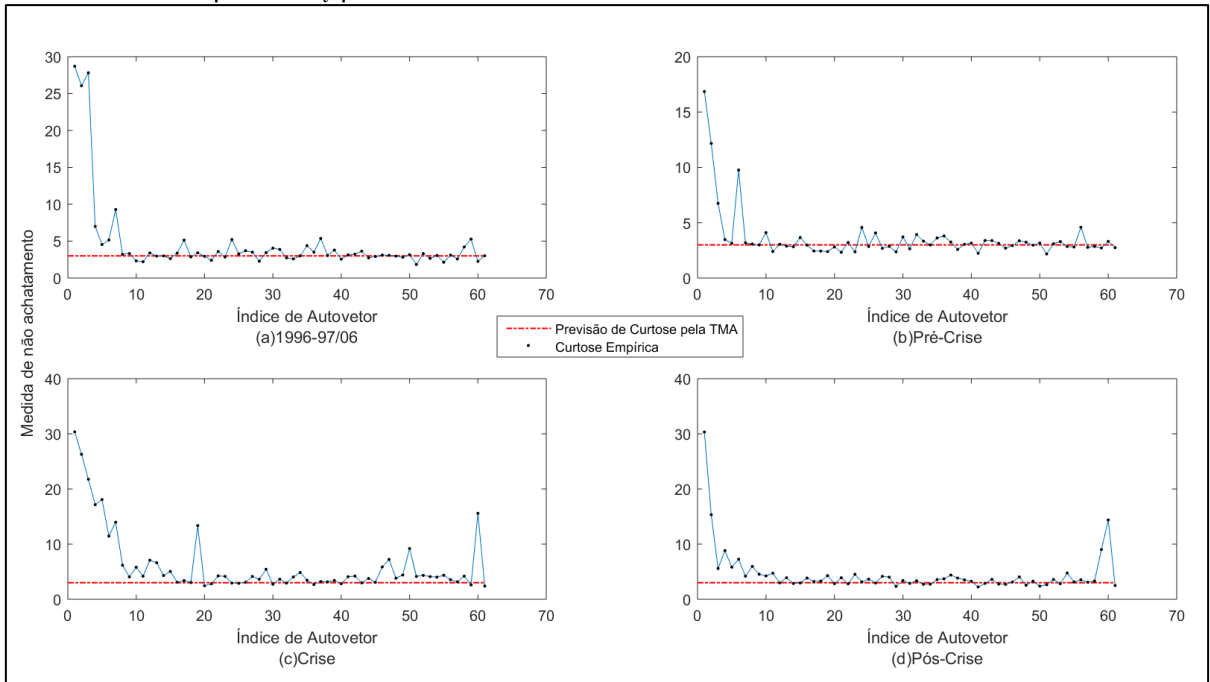
Em relação as medidas de não achatamento padrão apresentadas nas Figuras 31 e 32, em geral, os autovetores correspondentes aos menores e maiores autovalores desviam-se sistematicamente do comportamento normalmente distribuído.

Figura 31 – Medida de não achatamento padrão da distribuição dos componentes dos autovetores empíricos v_i para o Período I.



Fonte: Autor.

Figura 32 - Medida de não achatamento padrão da distribuição dos componentes dos autovetores empíricos v_i para o Período II.

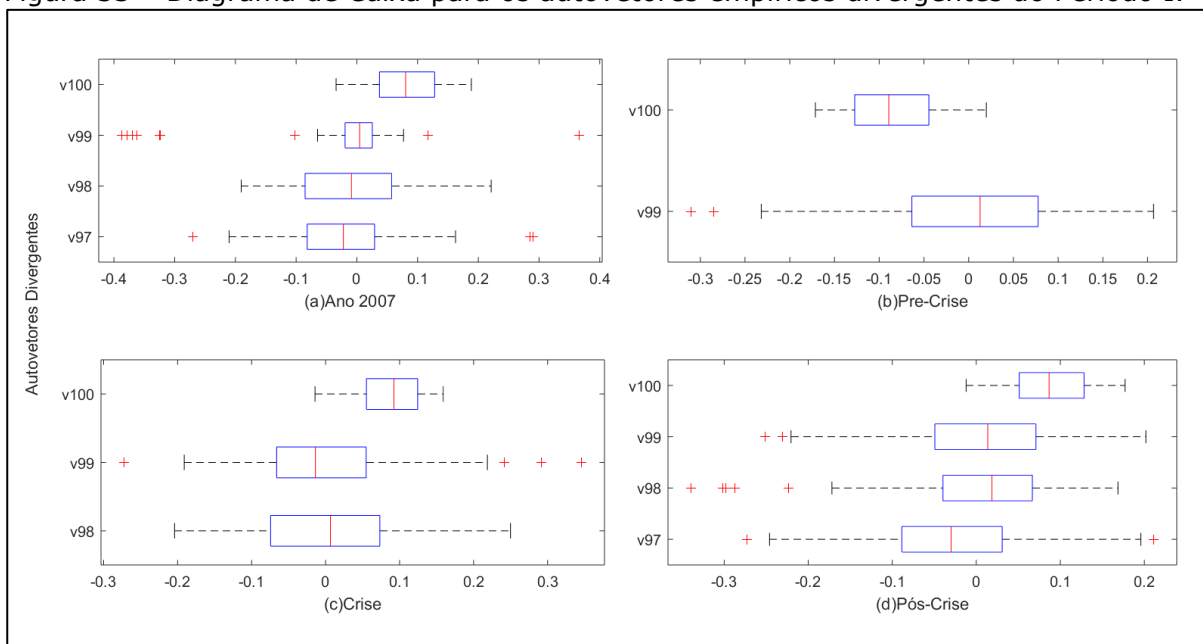


Fonte: Autor.

Os autovetores correspondentes aos maiores autovalores fora do limite do TMA, chamados de "autovetores divergentes", podem ser observados através dos diagramas de caixa nas Figuras 33 e 34, os quais ajudam a visualizar a mediana, a dispersão, os valores extremos e a

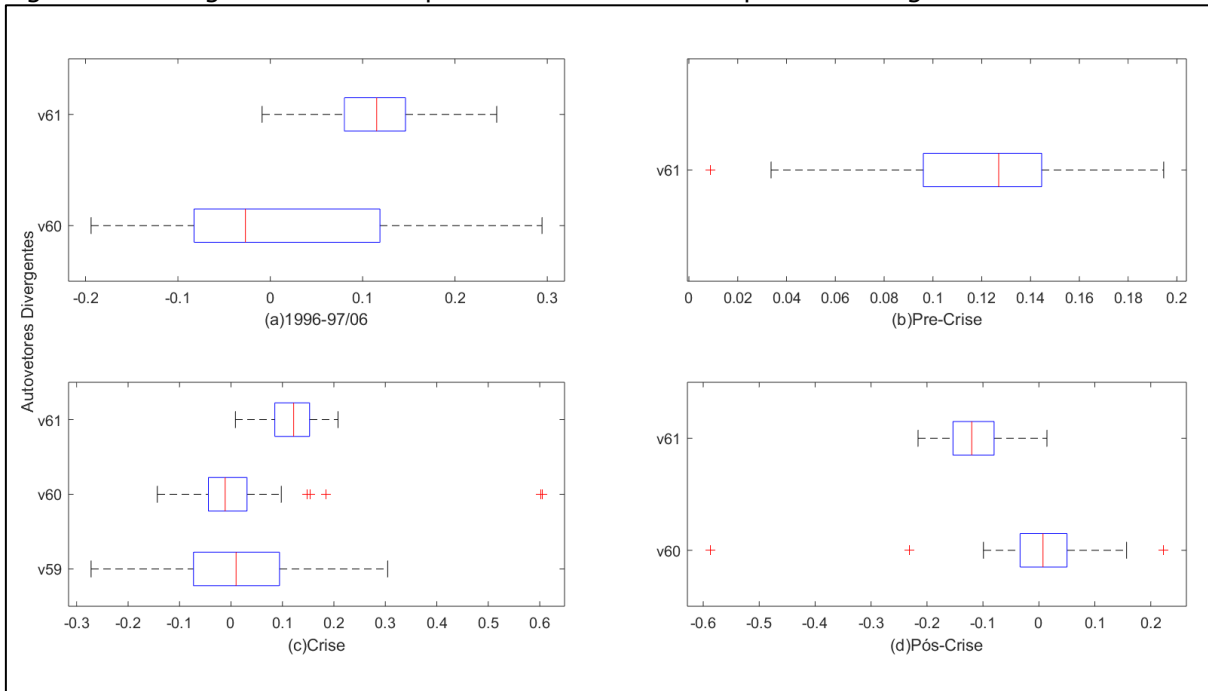
assimetria de um conjunto de dados, sendo que a caixa em cada gráfico apresenta a localização dos 50% valores mais prováveis (segundo e terceiro quartis). Sendo assim, quanto menor a caixa, mais equivalentes são as contribuições da maioria das ações. Já a presença de um *outlier* significa que uma certa ação é muito mais ou muito menos influente neste autovetor que as demais. Além disso, o deslocamento de uma caixa para o lado positivo ou negativo está relacionado com o sinal da maioria dos componentes de um autovetor.

Figura 33 – Diagrama de Caixa para os autovetores empíricos divergentes do Período I.



Fonte: Autor.

Figura 34 - Diagrama de Caixa para os autovetores empíricos divergentes do Período II.

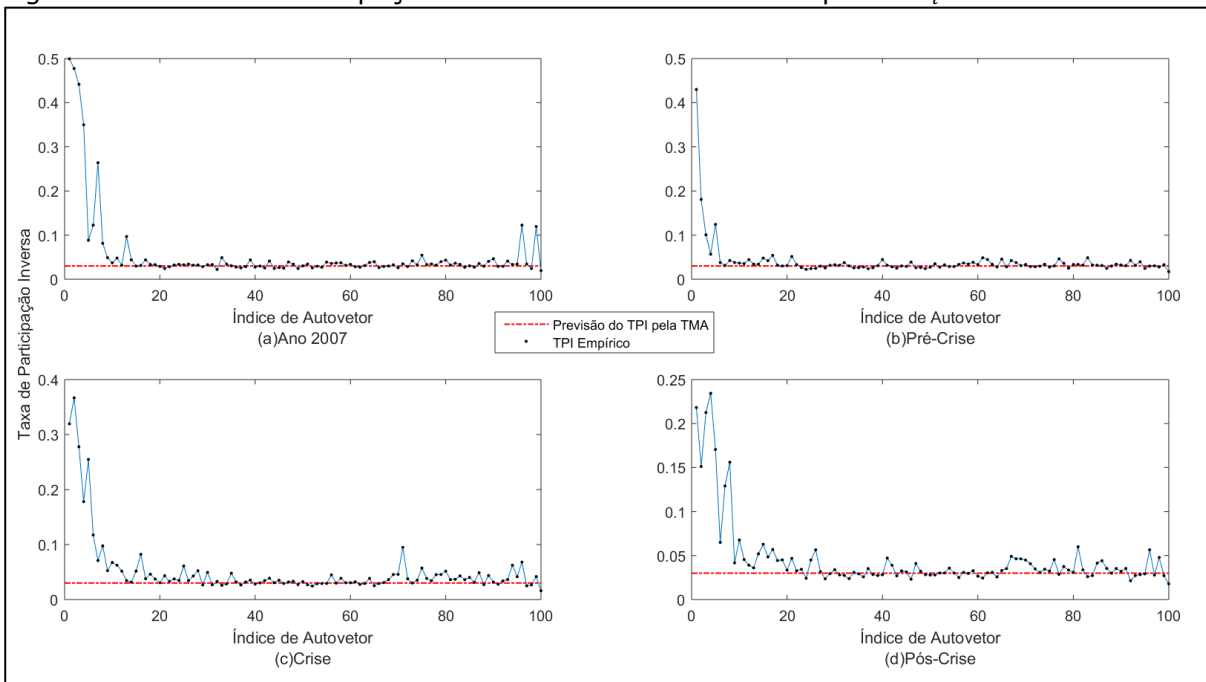


Fonte: Autor.

Para ambos os períodos é possível observar que, em geral, os componentes dos autovetores mais divergentes, v_{100} para o Período I e v_{61} para o Período II, apresentam poucos ou mesmo não apresentam *outliers*, o que aponta novamente que a maioria dos papéis participa significativamente. Também é possível observar que geralmente estes autovetores apresentam uma assimetria positivamente deslocada e com uma menor dispersão com relação a amplitude interquartil, o que sugere mais uma vez que há um componente comum que afeta a maioria dos participantes significativos com o mesmo viés.

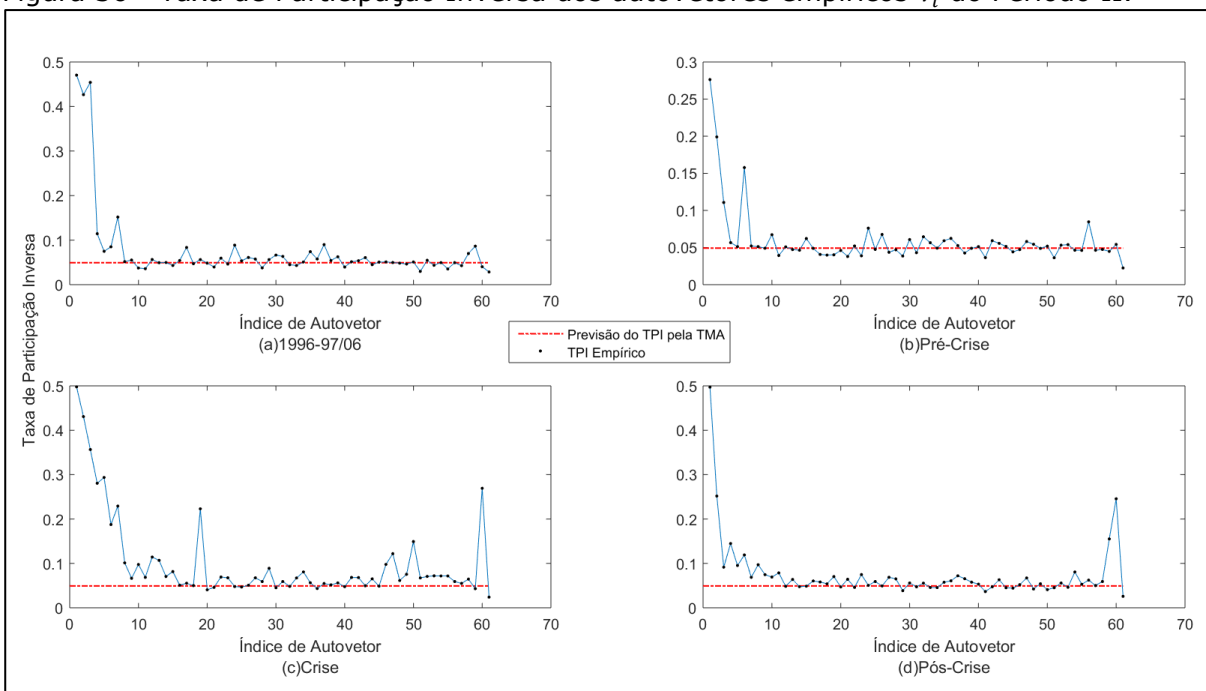
O número de ações com participação significativa pode ser quantificado pela taxa de participação (TP), conforme Equação (15). As Figuras 35 e 36 apresenta as medições TPI e as Figuras 37 e 38 apresentam as medições TP em função do índice dos autovetores (i) para os períodos analisados.

Figura 35 - Taxa de Participação Inversa dos autovetores empíricos v_i do Período I.



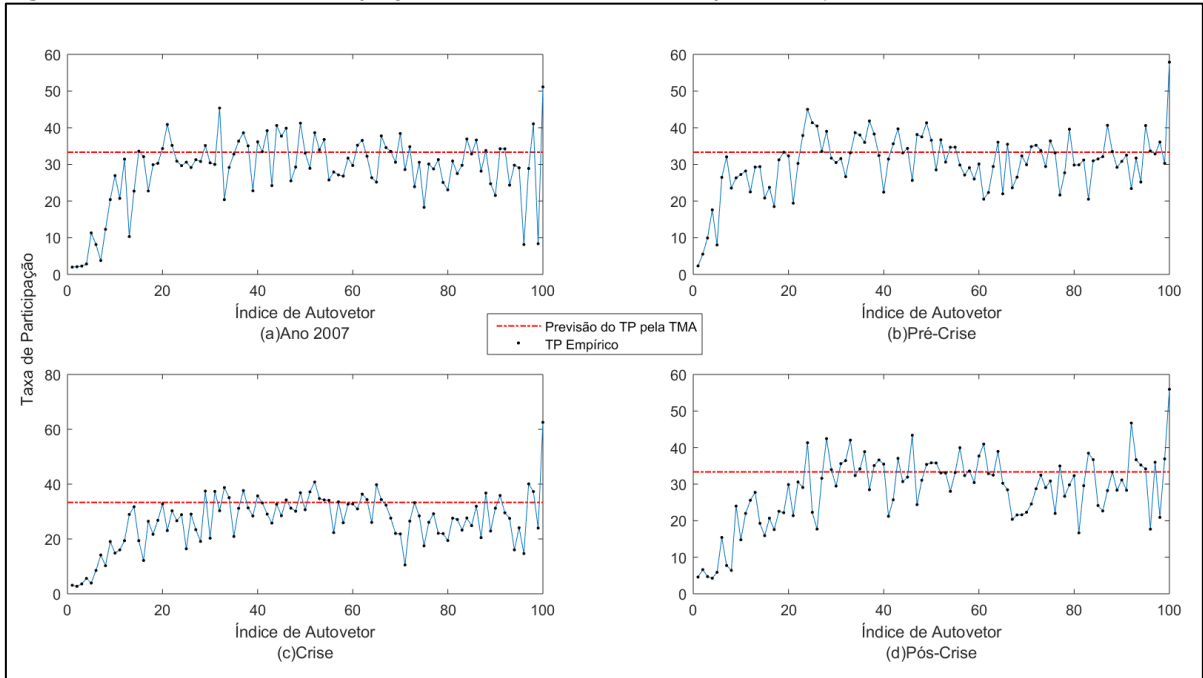
Fonte: Autor.

Figura 36 - Taxa de Participação Inversa dos autovetores empíricos v_i do Período II.



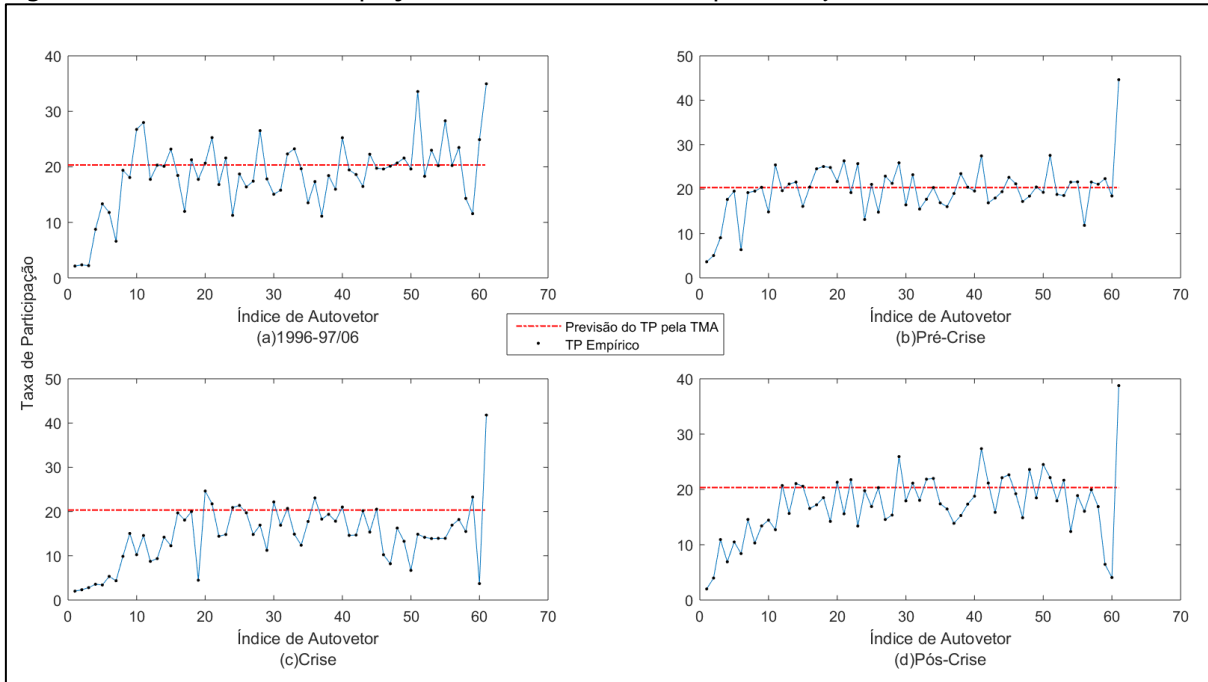
Fonte: Autor.

Figura 37 - Taxa de Participação dos autovetores empíricos v_i do Período I.



Fonte: Autor.

Figura 38 - Taxa de Participação dos autovetores empíricos v_i do Período II.



Fonte: Autor.

Conforme pode ser visto nas Figuras 37 e 38, autovetores relacionados a autovalores dentro dos limites teóricos $[\lambda_{min}, \lambda_{max}]$ tem TP próxima do valor previsto pela TMA ($N/3$). No entanto, os autovetores v_{100} do Período I e autovetores v_{61} do Período II apresentam quase que o dobro do valor esperado. Para o Período I tem-se aproximadamente:

- 2007: 51 participantes significativos;
- pré-crise: 58 participantes significativos;
- crise: 63 participantes significativos;
- pós-crise: 56 participantes significativos.

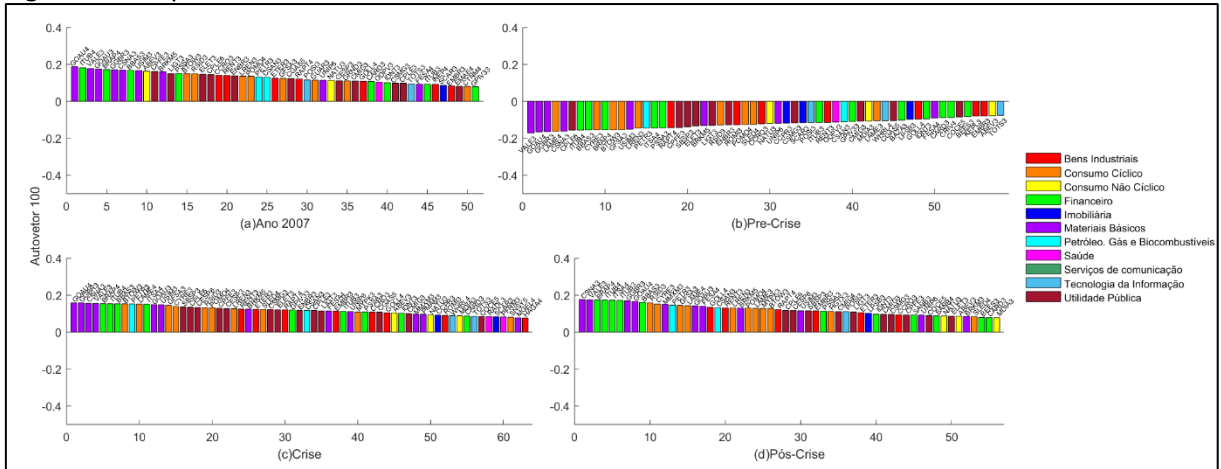
Para o Período II tem-se aproximadamente:

- 1996-1997/06: 35 participantes significativos;
- pré-crise: 45 participantes significativos;
- crise: 42 participantes significativos;
- pós-crise: 39 participantes significativos.

Podemos dizer, então, que comparativamente com o número de participações que era previsto pela TMA, existem muitas empresas que participam significativamente da composição do autovetor que mais diverge, o que consiste com a teoria de que este seria um índice de mercado (PLEROU *et al.*, 2002; SANDOVAL JUNIOR; BORTOLUZZO; VENEZUELA, 2014). Novamente, nesta análise percebeu-se o “efeito mola deformada” com uma mudança no comportamento do mercado percebida no período pré-crise, (que se intensifica durante a crise no caso do Período I) e tende a retornar ao seu comportamento original após a crise, porém ainda com resquícios. Também é possível perceber que os autovetores correspondentes aos menores autovalores têm poucos participantes significativos.

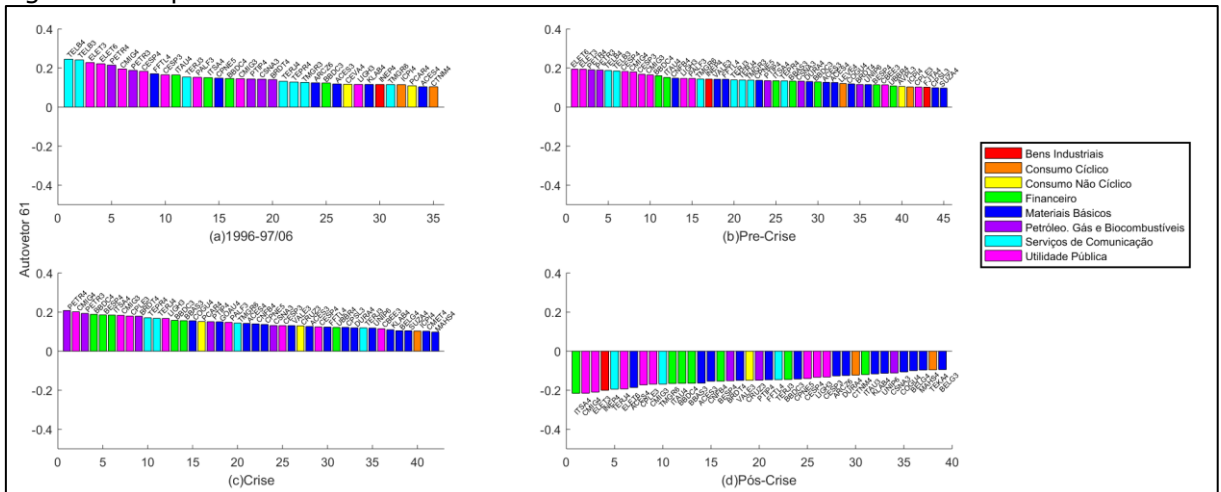
As Figuras 39 e 40 trazem uma visão detalhada dos componentes que participam significativamente dos autovetores v_{100} do Período I e v_{61} do Período II de cada período estudado.

Figura 39 – Composição dos autovetores v_{100} considerando as ações com participação significativa para o Período I.



Fonte: Autor.

Figura 40 - Composição dos autovetores v_{61} considerando as ações com participação significativa para o Período II.



Fonte: Autor.

Percebe-se que para ambos os períodos a cada subperíodo há representantes de quase todos os setores de mercado considerados na análise. Percebe-se ainda que para o Período II alguns setores de mercado ainda não eram presentes ou ainda não apareciam com força representativa no mercado brasileiro, indo surgir somente no Período I, como os setores de Imobiliária, Saúde e Tecnologia da Informação.

Para o Período I, no entanto, pode-se perceber que o setor de Materiais Básicos sempre apresenta grande influência na composição do autovetor v_{100} , encabeçando a lista de contribuições significativas, seguido do setor Financeiro, Consumo Cíclico e Utilidade Pública. Observa-se ainda que os componentes significativos de v_{100} do subperíodo de pré-crise são

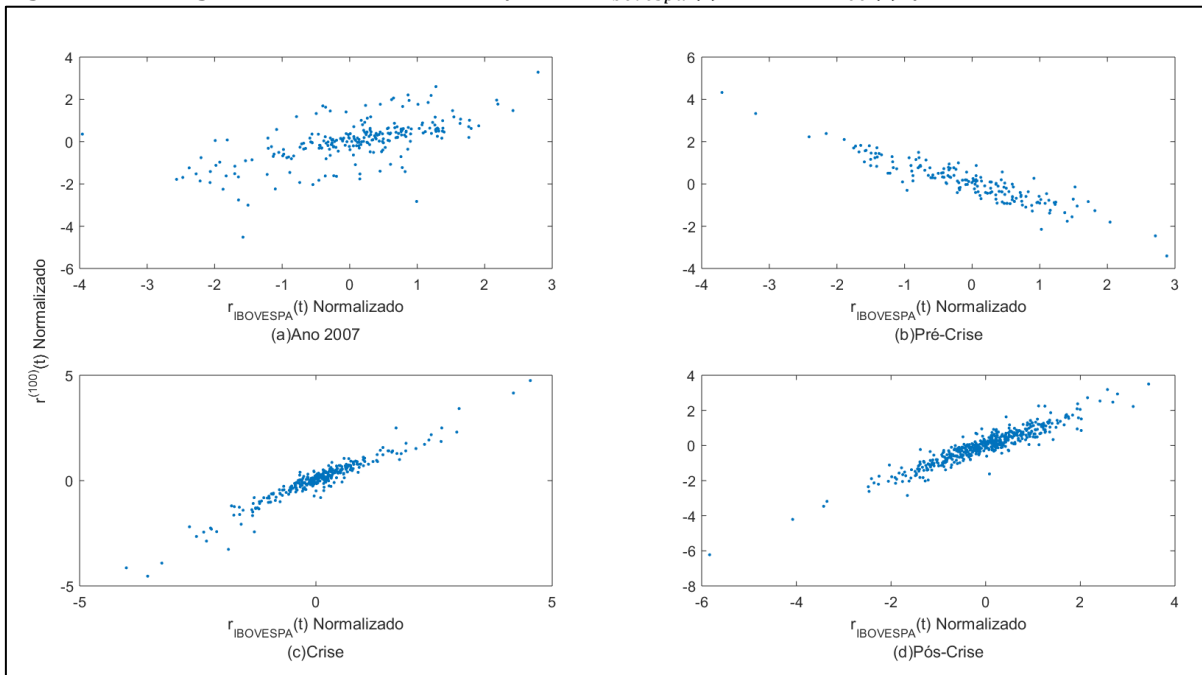
negativos, o que é coerente com o que foi apresentado e discutido sobre a Figura 25.

Para o Período II, pode-se perceber que o setor de Utilidade Pública e Petróleo, Gás e Biocombustível é sempre de grande peso na composição do v_{61} . O setor de Serviços de Comunicação, no entanto, era o setor mais forte no subperíodo de 1996-1997/06. Já no subperíodo pós-crise, o setor de Petróleo, Gás e Biocombustível praticamente perdeu toda a influência de mercado, muito provavelmente afetado pela crise da Rússia, que teve como um de seus estopins a subida do preço do petróleo. Observa-se ainda que todos os componentes com influência significativa para o autovetor v_{61} do subperíodo de pós-crise são negativos, o que acorda com o que foi discutido sobre a Figura 26.

Dessa forma, o comportamento de tais autovetores sugerem que estes representem um resultado coletivo das interações de todo o mercado, ou seja, um índice de mercado, o que está de acordo com pesquisas anteriores (COELHO *et al.* 2008; JOHNSON; WICHERN, 2002; PLEROU *et al.*, 2002; SENSOY; YUKSEL; ERTURK, 2013). Para verificar essa possibilidade pode-se comparar a projeção $r_{100}(t)$ para o Período I e $r_{61}(t)$ para o Período II da Equação (16) com a medida de desempenho padrão do mercado de ações brasileiro, o Ibovespa, representado pela série temporal dos retornos diários do Ibovespa $r_{Ibovespa}(t)$ gerada conforme a Equação (1).

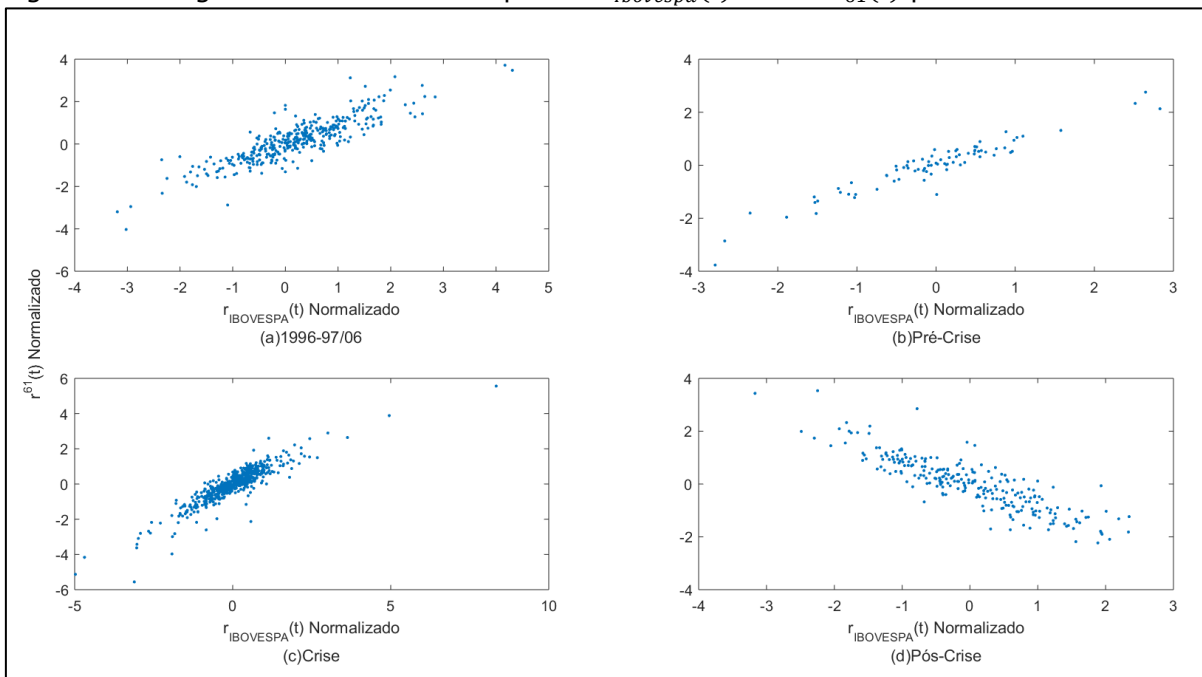
As Figuras 41 e 42 mostram essa regressão entre as medições. Ambos os eixos são reescalados por seus respectivos desvios padrão.

Figura 41 - Regressão das séries temporais $r_{Ibovespa}(t)$ contra $r_{100}(t)$ para o Período I.



Fonte: Autor.

Figura 42 - Regressão das séries temporais $r_{Ibovespa}(t)$ contra $r_{61}(t)$ para o Período II.



Fonte: Autor.

Percebe-se um comportamento extremamente correlacionado entre ambas as séries temporais para todos os períodos, com coeficiente de correlação para o Período I aproximadamente igual:

- 2007: 0,62;
- pré-crise: -0,92;

- crise: 0,96;
- pós-crise: 0,94.

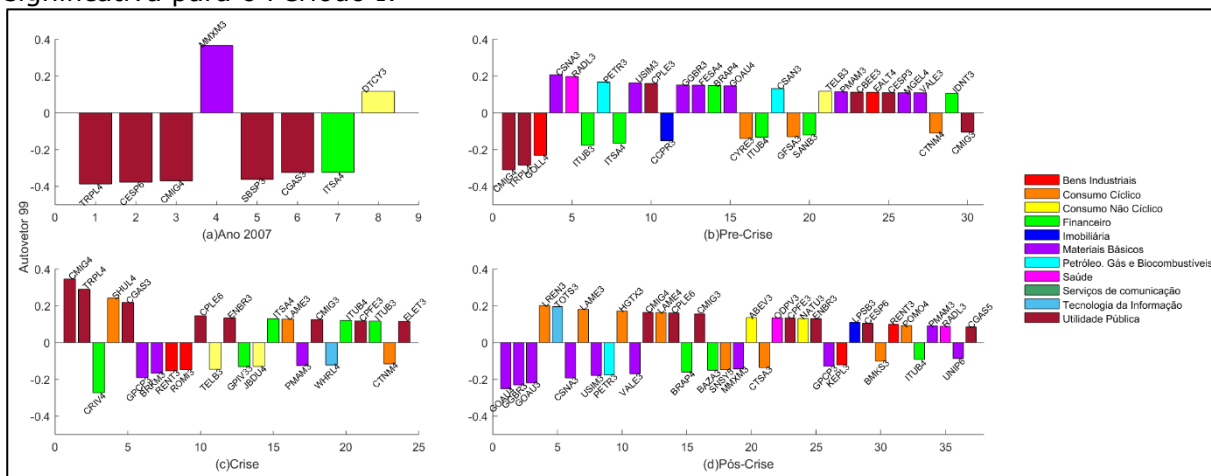
E com coeficiente de correlação para o Período II aproximadamente igual:

- 1996-1997/06: 0,87;
- pré-crise: 0,95;
- crise: 0,9;
- pós-crise: -0,87.

Desta forma, a correlação entre a projeção do subperíodo de 2007 do Período I e o Ibovespa não se mostrou tão perfeita quanto as demais projeções. Já a projeção para o subperíodo de pré-crise do Período I é quase que perfeitamente anticorrelacionada com o Ibovespa, da mesma forma que o subperíodo pós-crise do Período II. Isto se dá por ocorrência dos componentes deste autovetor serem negativamente deslocados. Desta forma, pode-se dizer que o autovetor que mais diverge para cada um dos subperíodos é um índice de mercado, o qual é praticamente equivalente a outro índice de mercado já conhecido, o Ibovespa.

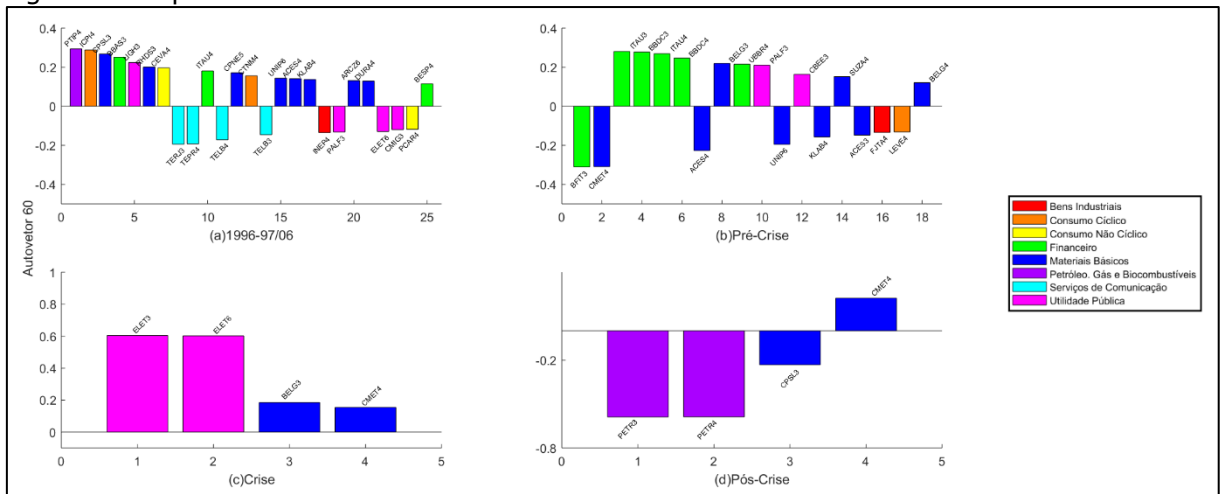
As Figuras 43 e 44 apresentam a composição de ações com participações significativas para o próximo autovetores divergentes (v_{99} do Período I e v_{60} do Período II respectivamente).

Figura 43 - Composição dos autovetores v_{99} considerando as ações com participação significativa para o Período I.



Fonte: Autor.

Figura 44 - Composição dos autovetores v_{60} considerando as ações com participação significativa para o Período II.



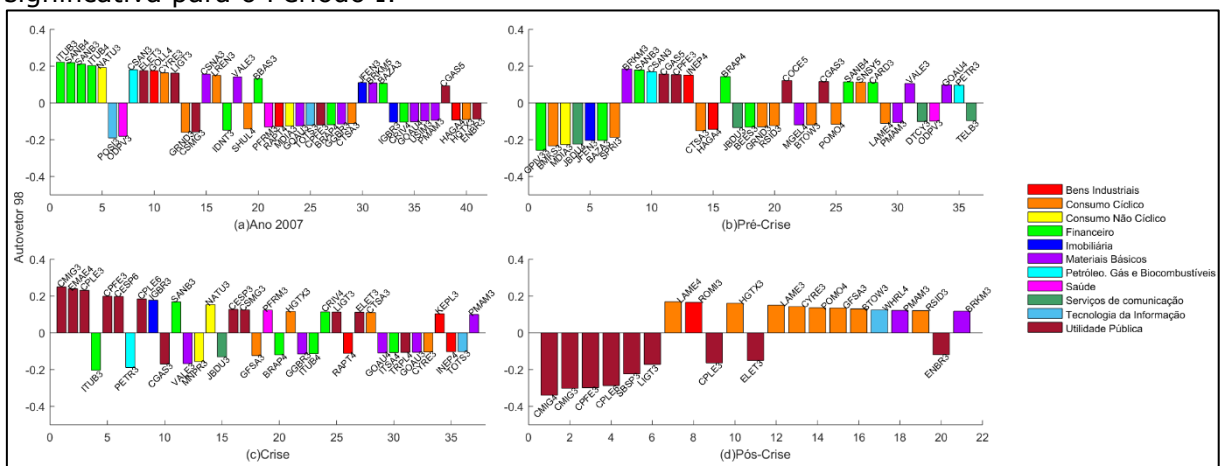
Fonte: Autor.

Percebe-se que o v_{99} do Período I, apresentado na Figura 43, é prioritariamente composto por ações do setor de Utilidade Pública durante grande parte deste período, contrastando com ações do setor de Materiais Básicos. Para o período de pós-crise, no entanto, o setor de Utilidade Pública perde força.

O v_{60} do Período II, visto na Figura 44, apresenta uma dependência temporal quanto sua composição. No entanto, o setor de Materiais Básicos ainda assim tem bastante influência durante todo o período.

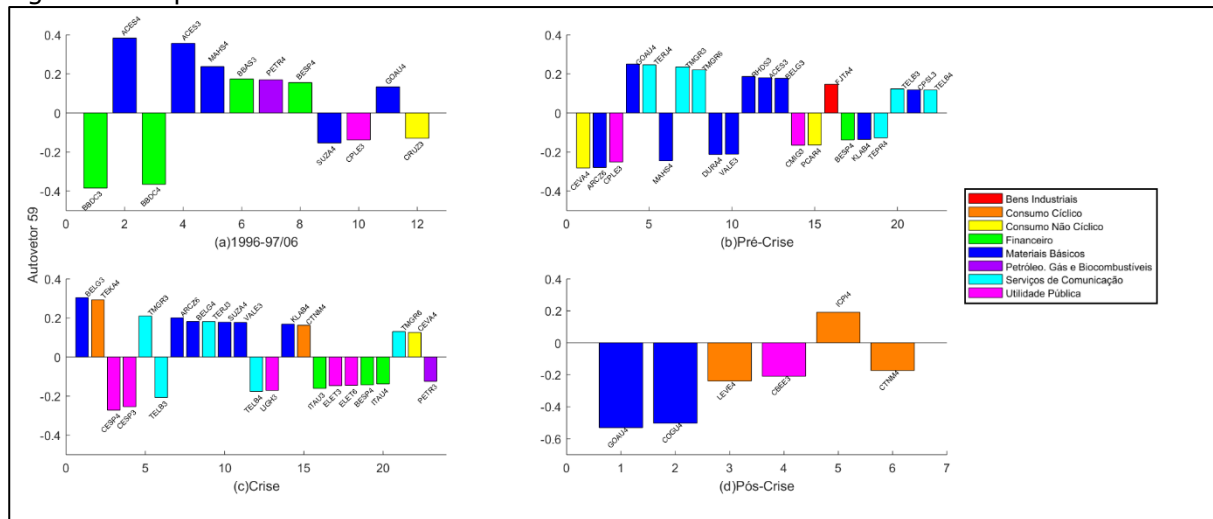
As Figuras 45 e 46 apresentam a composição de ações com participações significativas para os autovetores divergentes v_{98} do Período I e v_{59} do Período II respectivamente.

Figura 45 - Composição dos autovetores v_{98} considerando as ações com participação significativa para o Período I.



Fonte: Autor.

Figura 46 - Composição dos autovetores v_{59} considerando as ações com participação significativa para o Período II.



Fonte: Autor.

O vetor v_{98} do Período I apresentado na Figura 45, contém representantes da maioria dos setores econômicos estudados. Porém os subperíodos de crise e pós-crise são compostos prioritariamente por ações do setor de Utilidade Pública. Durante o período de 2007 o setor Financeiro é o mais influente.

Para o vetor v_{59} do Período II, apresentado na Figura 46, ainda há forte influência do setor de Materiais Básicos durante todo o período. Para o subperíodo de 1996-1997/06 o v_{59} apresenta uma oposição entre os setores de Materiais Básicos e Financeiros.

O autovetor v_{97} do Período I, apesar de ser divergente para os períodos de 2007 e pós-crise, apresenta uma composição de setores generalizada, pois era um autovetor relacionado a um autovalor muito próximo aos limites teóricos $[\lambda_{min}, \lambda_{max}]$.

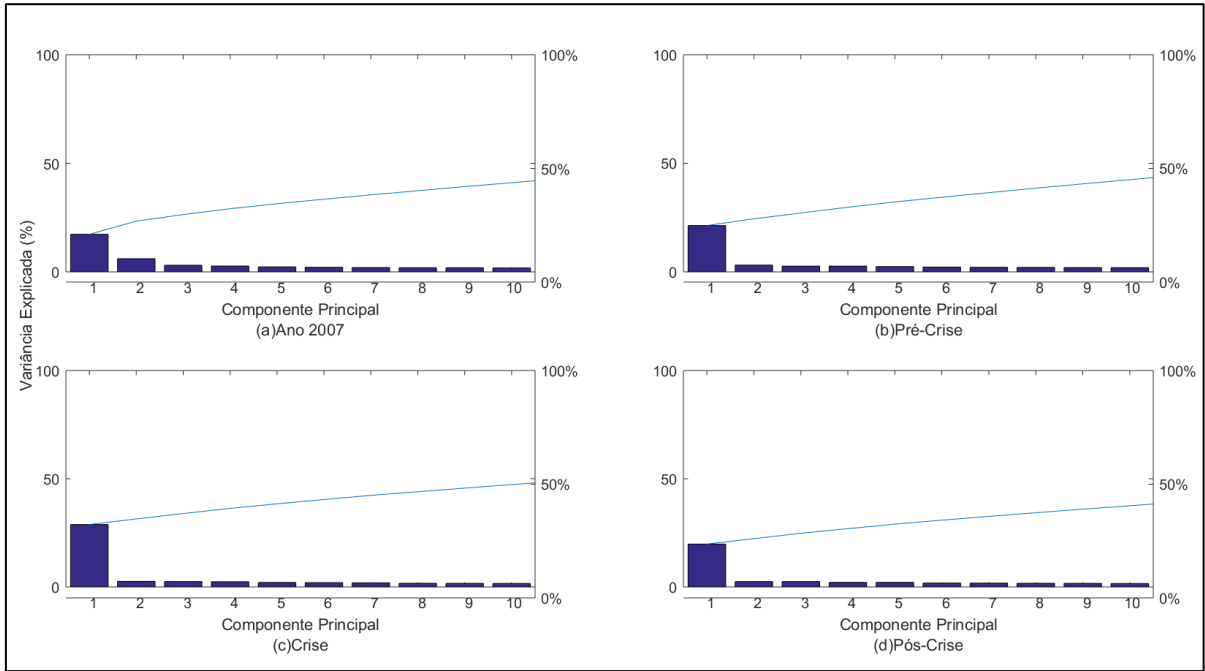
5.5 ACP para Correlação Empírica de Retornos

Para compreender de forma mais sucinta o comportamento das matrizes de correlação geradas durante a análise pode-se usar a Análise de Componentes Principais (ACP) apresentada na Seção 2.4.3.

As Figuras 47 e 48 apresentam um diagrama de Pareto que mostra a porcentagem da variação explicada pelos dez primeiros componentes principais para os dados empíricos formadas a partir das séries temporais

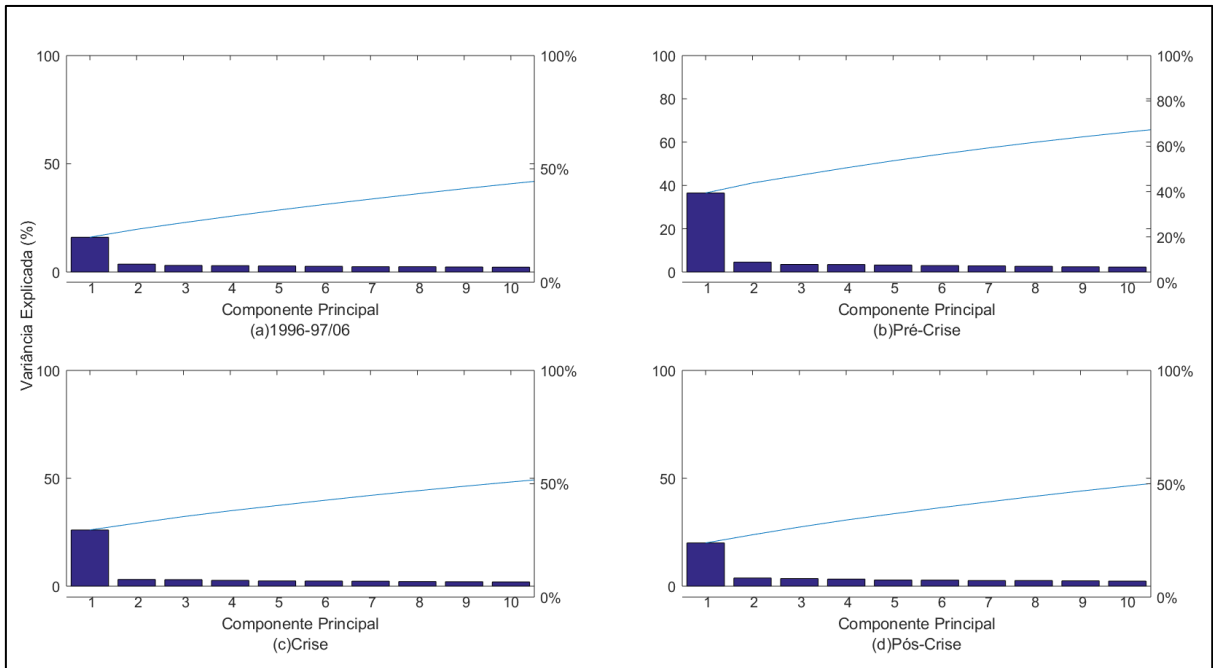
normalizadas de retornos diários das ações $r'_i(t)$ da Equação (2), os Períodos I e II respectivamente. A linha azul indica a porcentagem de variação acumulada.

Figura 47 - Diagrama de Pareto para os dez primeiros componentes principais do Período I.



Fonte: Autor.

Figura 48 - Diagrama de Pareto para os dez primeiros componentes principais do Período II.



Fonte: Autor.

A correlação explicada pelo primeiro componente principal de cada subperíodo é de uma ordem de grandeza muito maior comparada aos demais componentes principais, correspondendo a aproximadamente, para o Período I:

- 2007: 17,3%;
- pré-crise: 21,3%;
- crise: 28,8%;
- pós-crise: 19,8%.

Para o período II tem-se aproximadamente:

- 1996-1997/06: 16%;
- pré-crise: 36,6%;
- crise: 26%;
- pós-crise: 20%.

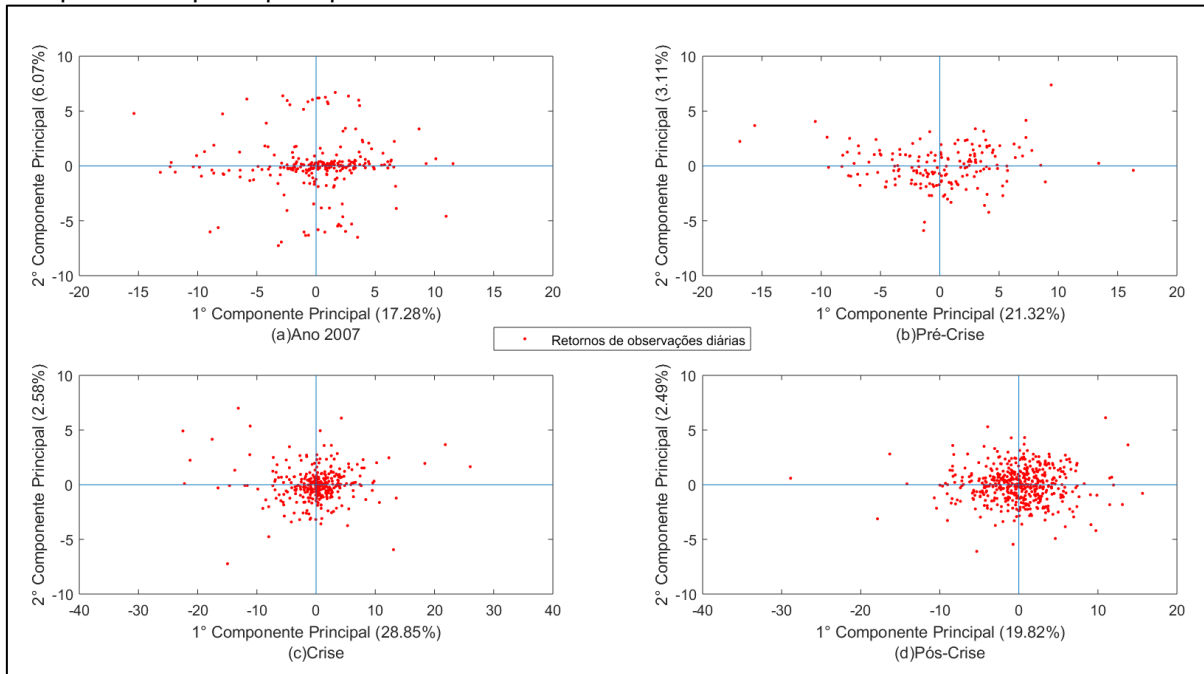
Desta forma, de acordo com os critérios de Johnson e Wichern (2002), pode-se dizer que o primeiro componente principal já seria suficientemente representativo em termos de ACP, uma vez que, segundo estes autores, para determinar o número apropriado de componentes principais, procura-se por um ponto no gráfico a partir do qual os autovalores restantes são relativamente pequenos e todos aproximadamente do mesmo tamanho. Além disso, o primeiro componente principal está relacionado com o autovetor que explica a maior parte da variância (o autovetor que mais diverge), desta forma, pode-se dizer que o primeiro componente principal explica o comportamento geral do mercado (COELHO *et al.* 2008; JOHNSON; WICHERN, 2002; PLEROU *et al.*, 2002; SENSOY; YUKSEL; ERTURK, 2013).

Nesta análise percebe-se mais uma vez o "efeito mola deformada", com uma alteração já previamente sentida pelo período de pré-crise e uma tendência do período de pós crise de retomada do comportamento habitual do mercado. Os primeiros componentes principais referente ao subperíodo de crise do Período I e pré-crise e crise do Período II, são os de maior magnitude, o que demonstra novamente como em períodos de instabilidade

financeira o mercado tende a se comportar de forma homogênea, com a crise no mercado afetando a maioria das empresas de forma semelhante.

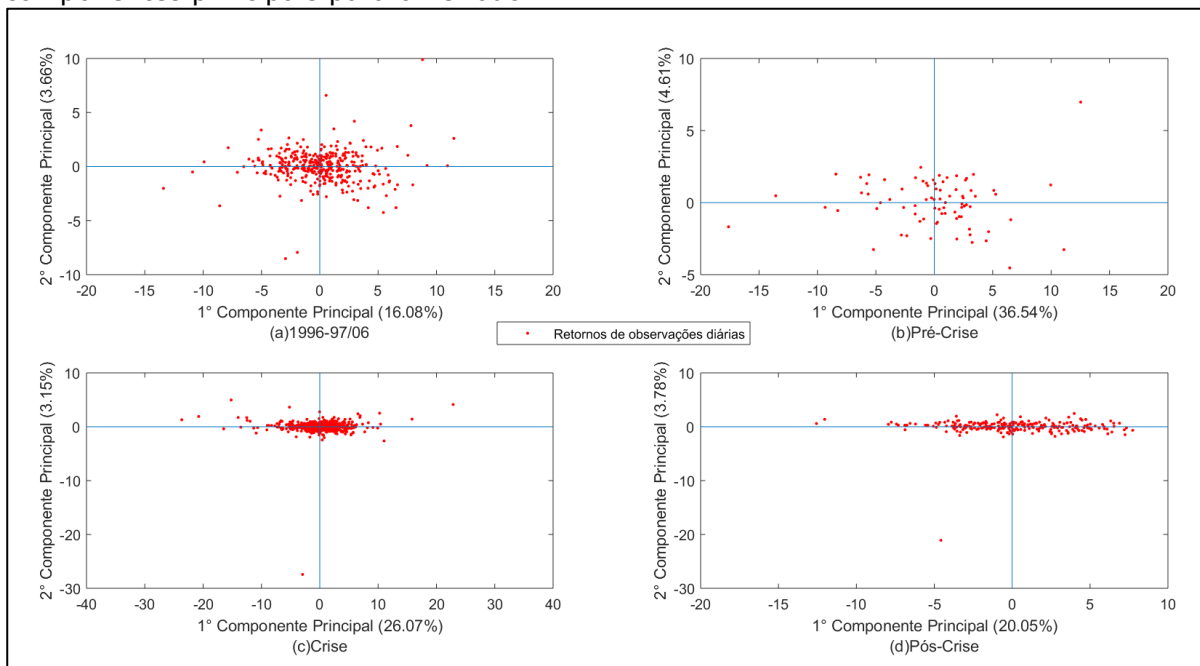
Esse colapso do mercado durante o período da crise também pode ser observado nas Figuras 49 e 50, onde as observações dos retornos diários (escores) de cada período foram reescalados para o novo sistema de base ortogonal com os dois primeiros componentes principais como eixos coordenados, sendo que cada ponto vermelho representa um dia de movimentação da bolsa considerando as ações selecionadas neste estudo.

Figura 49 - Escores de cada dia para ACP dos retornos diários sobre os dois primeiros componentes principais para o Período I.



Fonte: Autor.

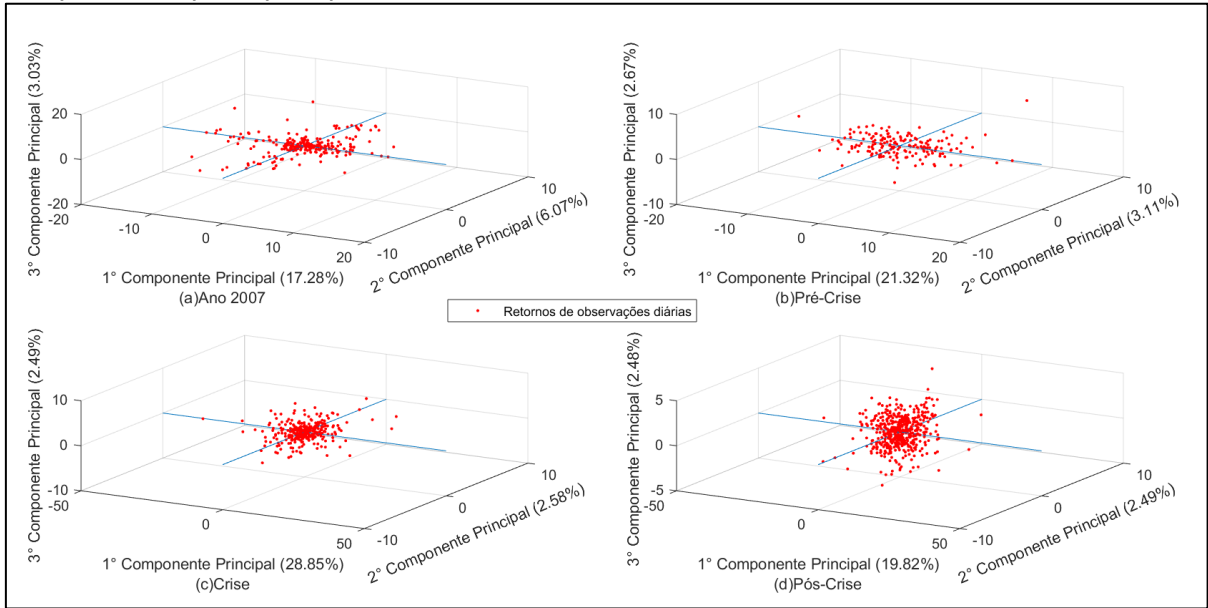
Figura 50 - Escores de cada dia para ACP dos retornos diários sobre os dois primeiros componentes principais para o Período II.



Fonte: Autor.

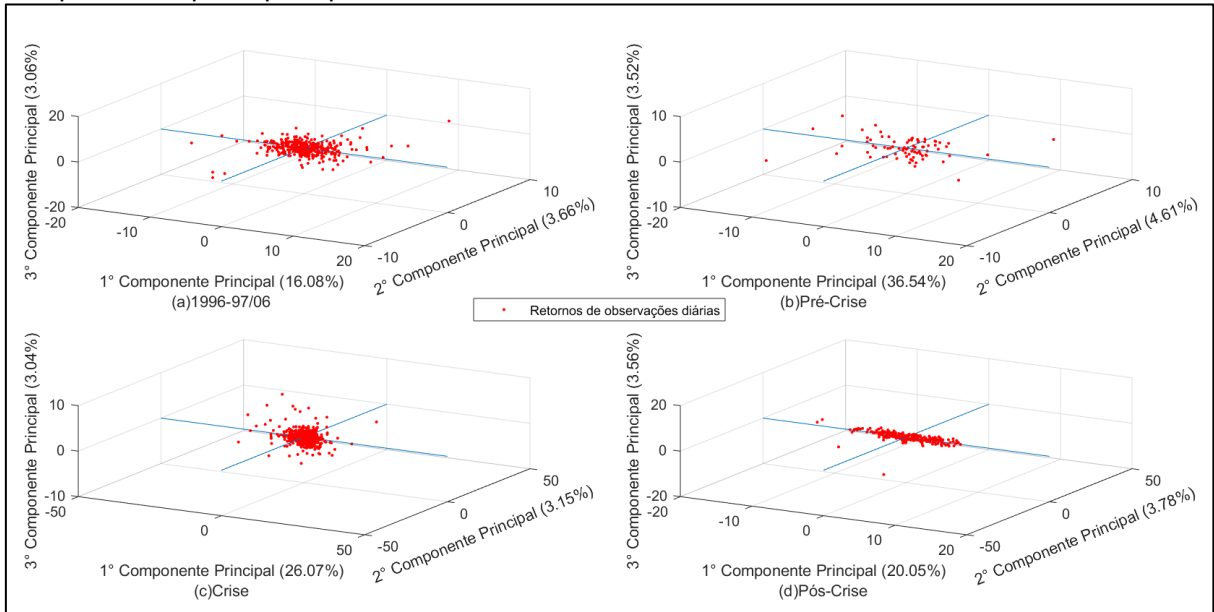
Observa-se que durante o período de crise o comportamento do mercado é genérico, apresentando um comportamento aglomerado. Durante o período pós-crise, passa a apresentar um espalhamento do comportamento das ações, tendendo a ter um comportamento semelhante ao apresentado anteriormente na crise ("efeito mola deformada"). As Figuras 51 e 52 apresentam as mesmas análises já apresentadas para os retornos diários, mas em uma base ortogonal tridimensional com os três primeiros componentes principais como eixos coordenados para uma melhor compreensão do comportamento das amostras dos retornos diários ao longo do tempo para ambos os períodos analisados.

Figura 51 - Escores de cada dia para ACP dos retornos diários sobre os três primeiros componentes principais para o Período I.



Fonte: Autor.

Figura 52 - Escores de cada dia para ACP dos retornos diários sobre os três primeiros componentes principais para o Período II.



Fonte: Autor.

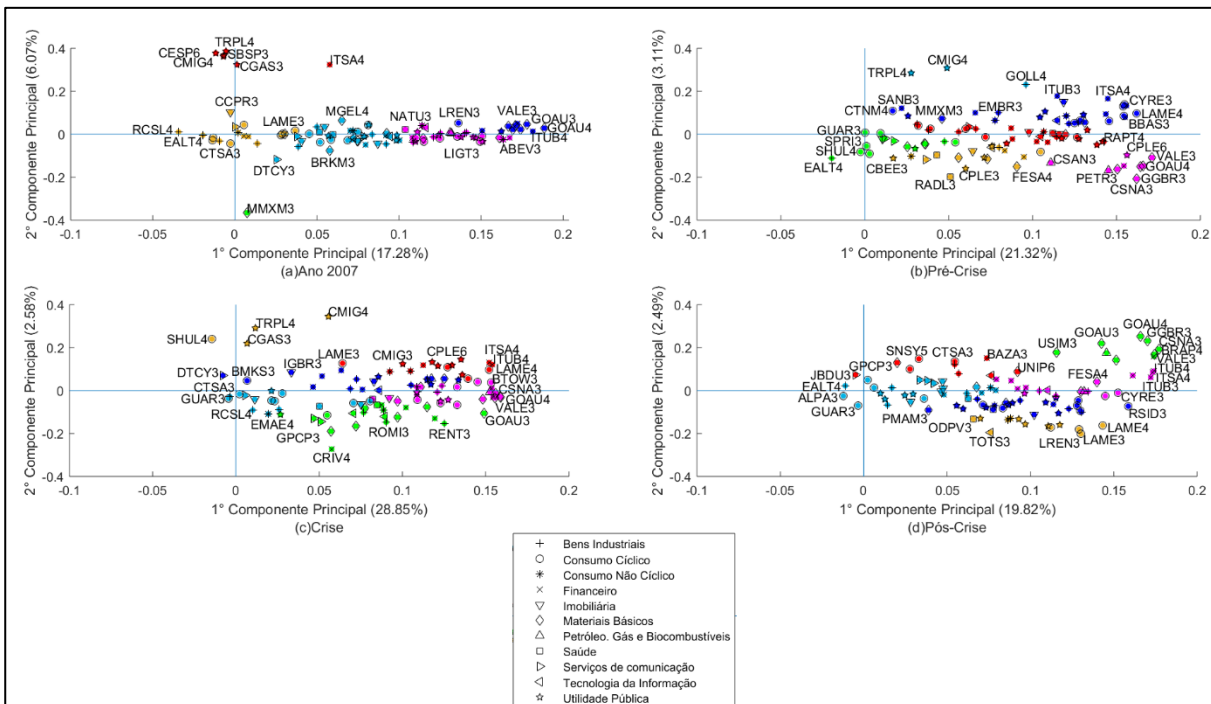
Para melhor entendimento do comportamento dos papéis e da correlação entre eles, as Figuras 53 e 54 apresentam cada ação reescalada para a nova base ortogonal formada pelos dois primeiros componentes principais. Os diferentes ícones representam os diferentes setores aos quais cada empresa pertence (veja a lista detalhada apresentada previamente nos Quadros 1 e 2, Seção 4.1). As cores representam seis grupos definidos

pelo método de agrupamento k-Means, o qual realiza o agrupamento a partir de centroides (MATHWORKS, 2020). Algumas ações foram nomeadas para auxiliar na compreensão dos agrupamentos.

Perceba que uma cor não representa o mesmo grupo de um subperíodo para outro, sendo que o importante nesta análise é a diferenciação do comportamento dos diferentes grupos para cada subperíodo. Perceba ainda que dentro de um grupo não há papéis fixos ao longo do tempo, sendo que, em cada período, os grupos são formados a partir de ações que são mais próximas de um dos seis centroides previamente definidos.

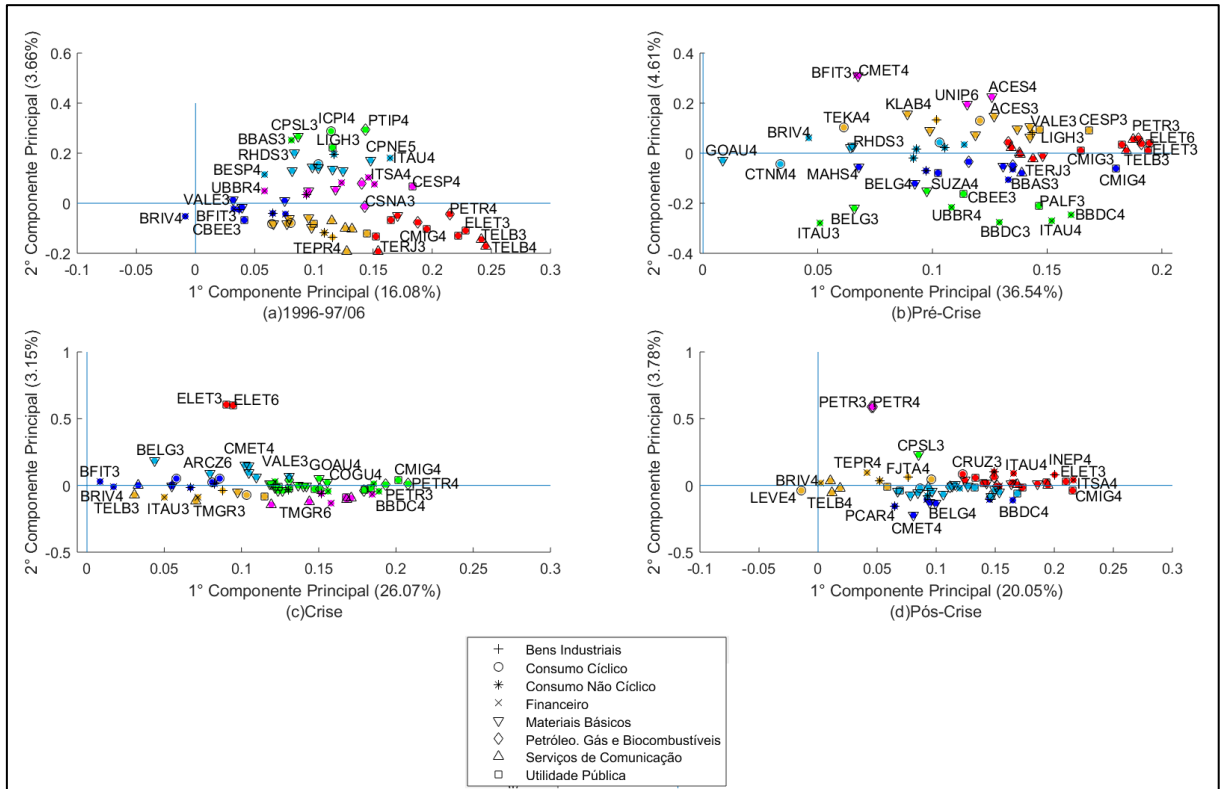
Assim, o agrupamento por k-Means está funcionando aqui apenas como um artifício visual para facilitar a compreensão do comportamento das ações que estão próximas ou distantes umas das outras.

Figura 53 - Coeficientes de cada ação para ACP com agrupamento por setores de mercado e por grupos k- Means sobre os dois primeiros componentes principais para p Período I.



Fonte: Autor.

Figura 54 - Coeficientes de cada ação para ACP com agrupamento por setores de mercado e por grupos k- Means sobre os dois primeiros componentes principais para p Período II.



Fonte: Autor.

Para este trabalho, escolheu-se a métrica de distância euclidiana quadrada como método de agrupamento k-Means, onde cada centroide é a média dos pontos naquele grupo, definido da seguinte maneira (MATHWORKS, 2020)

$$d(x, c) = (x - c)(x - c)'. \tag{18}$$

O agrupamento ajuda a visualizar os grupos de ações que apresentam comportamentos semelhantes ou opostos e, assim, possibilita montar uma carteira de investimentos mais equilibrada e com menos riscos. Analisando as Figuras 53 e 54, a maioria das ações tem coeficientes positivos para o primeiro componente principal, mostrando que esse componente principal aponta para a direção do comportamento geral do mercado.

Além disso, percebe-se que existem grupos (representados pelas diferentes cores) que apresentam comportamentos opostos, ou seja, mais distantes uns dos outros, por exemplo no subperíodo de 2007 do Período I, tem-se os grupos representados pelas cores amarelo e azul escuro como os

mais afastados entre si em relação ao primeiro componente principal e em relação ao segundo componente principal tem-se os grupos mais distantes representados pelas cores vermelho e verde. Uma análise semelhante também pode ser feita para todos os outros subperíodos. Pode-se perceber que não há um comportamento generalizado dentro do mesmo setor, porém existem subgrupos de empresas que se comportam de forma constante ao longo do tempo.

Ao analisar o Período I, verifica-se que o subgrupo que se encontra no extremo direito do eixo do primeiro componente principal, durante os quatro subperíodos, é um subgrupo composto por principalmente empresas do setor Materiais Básicos e algumas do setor Financeiro, significando que estes são os setores de maior influência no mercado durante este período. No entanto, durante o período de pré-crise e crise, o setor de Consumo Cíclico passa a ter uma influência comparativamente maior ao que tinha durante o subperíodo de 2007, porém tende a voltar a um comportamento semelhante ao apresentado antes da crise no subperíodo de pós-crise. Este mesmo fenômeno também pode ser observado para algumas empresas do setor de Petróleo, Gás e Biocombustível.

Além disso, existe um subgrupo composto por algumas empresas do setor de Bens Industriais, Consumo Cíclico e Serviços de Comunicação que apresenta um comportamento antagônico àquele anteriormente citado, ou seja, sempre aparecem na extrema esquerda do eixo do componente principal para o Período I, sendo assim o grupo de empresas com menor força de influência do mercado.

Ainda analisando-se o Período I, em relação ao segundo componente principal, percebe-se que existe um subgrupo do setor de Utilidade Pública que apresentou forte influência durante quase todo o período. No entanto, ele perde esta influência no período pós-crise. Não existe um subgrupo claramente anticorrelacionado a este ao longo do tempo. Para o período de 2007, porém, houve um forte contraste entre este subgrupo e a ação MMXM3, contraste este já evidenciado em análises anteriores na Seção 5.3.

Ao analisar-se o Período II, percebe-se que existe um subgrupo composto por empresas dos setores de Utilidade Pública, Petróleo, Gás e Biocombustível e Serviços de Comunicação, os quais sempre aparecem no extremo direito do eixo com primeiro componente principal, sendo este o subgrupo de maior influência no mercado durante o Período II. No entanto, percebe-se que estes setores revezam a liderança na influência do mercado entre si ao decorrer do tempo. Percebe-se, contudo, uma ascensão da influência de algumas empresas dos setores Financeiro a partir do subperíodo de crise, que se mantém durante o subperíodo de pós-crise, juntamente com uma representante do setor de Bens Industriais.

Em relação as ações de menor influência no mercado para o Período II, sempre há representantes dos setores de Financeiro e Consumo Cíclico. Para os subperíodos de 1996-1997/06 e pré-crise, no entanto, via-se um subgrupo do setor de Materiais Básicos com pouca influência no mercado, porém, este comportamento característico deste subgrupo foi substituído por um subgrupo do setor Serviços de Comunicações a partir do subperíodo de crise. Veja por exemplo o que aconteceu com os papéis TELB3 e TELB4, os quais eram as ações de maior influência no mercado no subperíodo de 1996-1997/06 e passaram a ser uma das de menor influência a partir da crise e permanecendo nesta condição durante o subperíodo de pós-crise.

Analisando-se o segundo componente principal do Período II, vê-se que um subgrupo de ações dos setores Consumo Cíclico, Financeiro e Materiais Básicos teve forte influência durante os subperíodos 1996-1997/06 e pré-crise. Porém durante o subperíodo de crise, o segundo componente principal apresenta forte influência de representantes do setor de Utilidade Pública e, durante o subperíodo de pós-crise, o segundo componente principal é influenciado principalmente por representantes do setor de Petróleo, Gás e Biocombustível.

Pode-se dizer, então, que a utilização de ACP, juntamente com técnicas de agrupamento, confirmou a análise anteriormente realizada para os autovetores mais divergentes na Seção 5.3 e nos ajudou a entender melhor a interação entre as ações.

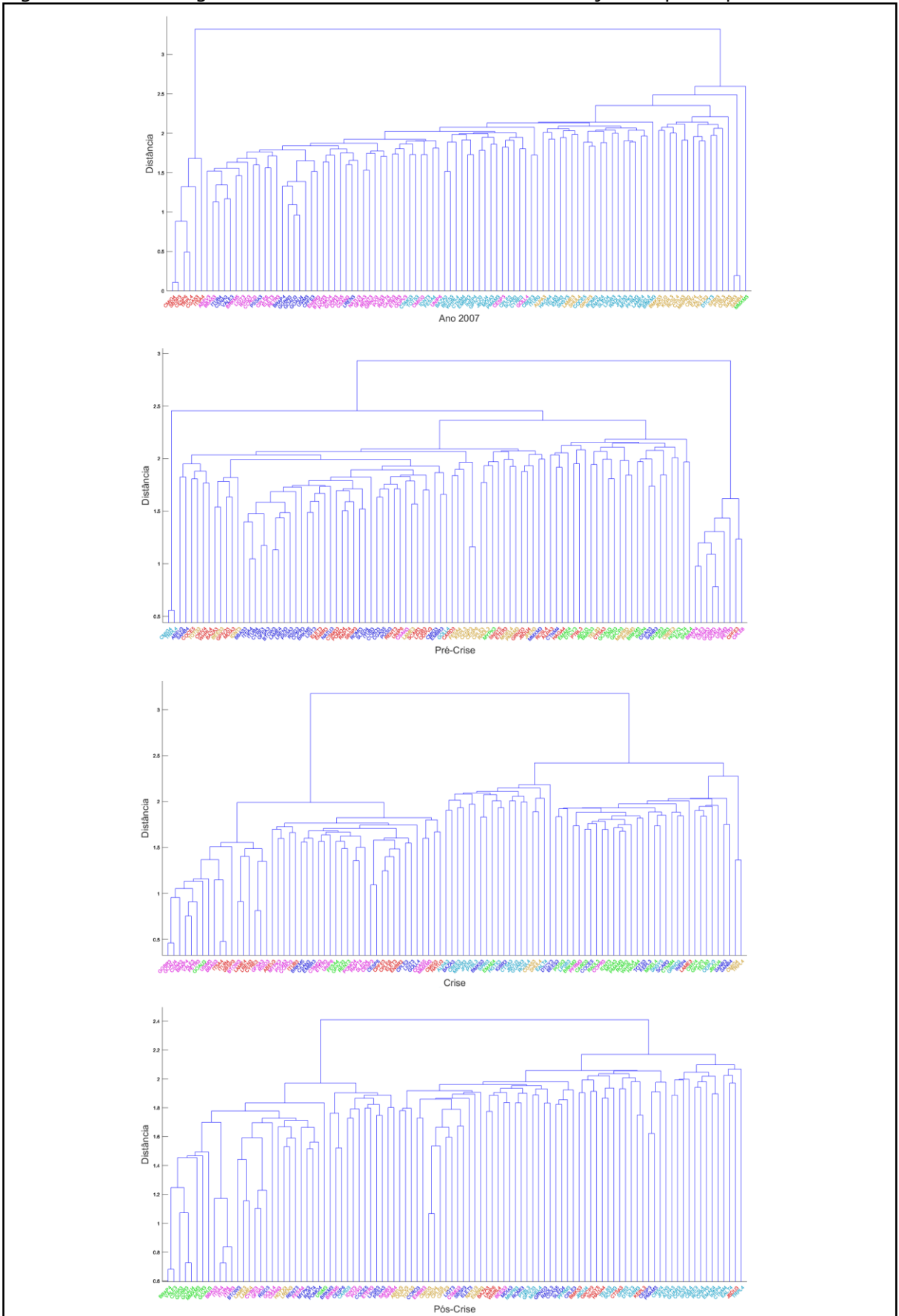
As Figuras 55 e 56 apresenta os dendrogramas das Árvores de Agrupamento Hierárquico Aglomerativas (do inglês *Agglomerative Hierarchical Cluster Tree* - AHCT) (MathWorks 2020) gerados a partir das matrizes de correlação empírica de retornos diários das ações $r'_i(t)$ da Equação (2) para ambos os períodos analisados, a qual traz uma visão mais detalhada e hierárquica da estrutura de conexão dos grupos já apresentados nas Figuras 53 e 54, considerando que foram utilizados as mesmas cores para representar cada grupo em ambos os agrupamentos. Desta forma, pode-se acompanhar detalhadamente a distribuições de cada grupo.

Para tanto, foi necessário criar uma rede de distâncias com base nestas correlações. Assim sendo, a seguinte métrica foi usada (MANTEGNA, 1999; SENSOY *et al.*, 2013)

$$d_{ij} = \sqrt{2(1 - c_{ij})}, \quad (19)$$

onde d_{ij} são os elementos de uma matriz de distância D de dimensões $N \times N$ e pode variar de 0 a 2 com pequenas distâncias correspondendo a altas correlações e vice-versa, a qual é uma métrica euclidiana válida (SENSOY *et al.*, 2013). Esta matriz de distância D foi usada para construir as AHCTs.

Figura 55 – Dendrograma da AHCT das Matrizes de Correlação Empírica para o Período I.



Fonte: Autor.

Uma vez que se deve conectar os grupos da AHCT (“folhas” e “galhos”) usando um método escolhido (MATHWORKS, 2020), neste trabalho foi escolhido o método de ligação de média ponderada que usa uma distância recursiva entre dois grupos, definido como (MATHWORKS, 2020)

$$d(r, s) = \frac{d(p, s) + d(q, s)}{2}, \quad (20)$$

onde r , s , p e q são grupos com pelo menos um nó (“folha” ou “galho”). Se o cluster r foi criado a partir da combinação dos clusters p e q , a distância entre r e outro grupo s é definida como a média das distâncias entre p e s e entre q e s (MATHWORKS, 2020). Este método de ligação foi escolhido porque mostrou maior similaridade com as formações de agrupamentos gerados pelo método de agrupamento k-Means apresentado anteriormente nas Figuras 53 e 54.

A construção AHCT começa com o par de nós com a distância mais curta de acordo com a Equação (20), conectando-os, então a segunda menor distância é identificada e adicionada ao AHCT e assim por diante. A partir do AHCT, pode-se traçar o gráfico de dendrograma associado a ele, o qual é apresentado nas Figuras 55 e 56.

Ao analisar as Figuras 55 e 56, pode-se concluir que elas confirmam o agrupamento apresentado anteriormente nas Figuras 53 e 54 uma vez que os elementos dos grupos coloridos daquelas figuras ainda estão próximos uns dos outros nas Figuras 55 e 56. Além disso, os grupos de comportamento opostos nas Figuras 53 e 54 permanecem separados nas Figuras 55 e 56, por exemplo pode-se verificar o caso da ação MMXM3 do subperíodo de 2007 do Período I, que é um *outlier* e tem comportamento oposto ao cluster vermelho (composto pelos papéis SBSP3, CESP6 e TRPL4 por exemplo) aparecem distantes em ambas as análises. Outro exemplo é o grupo amarelo (composto pelos estoques RCSL4, EALT4 e SPRI3 por exemplo) e azul escuro (composto pelos estoques VALE3, GOAU4 e ITUB4 por exemplo) do subperíodo de 2007 do Período I na Figura 53, os quais

também permanecem separados na Figura 55. Uma análise semelhante também pode ser feita para os demais subperíodos, onde as análises de agrupamento por k-Means das Figuras 53 e 54 são conformadas pelas análises de agrupamento por AHCT.

Uma vez que é sabido que uma carteira de ações equilibrada deve conter ações de empresas com comportamento oposto (anticorrelacionado) para reduzir os riscos (BOUCHAUD; POTTERS, 2009; OH, 2011; SENSOY; YUKSEL; ERTURK, 2013), deseja-se agora encontrar quais subgrupos de empresas são anticorrelacionados.

Assim, pode-se ter uma carteira de ações equilibrada que contenha um ou mais elementos pertencentes a conjuntos verdadeiramente anticorrelacionados entre si durante todo o período, mesmo considerando a existência de alta volatilidade em algum momento.

Para o Período I, por exemplo, o par de as ações RCSL4 e EALT4 e o grupo de ações GOAU4, VALE3 e ITUB4 demonstraram ter esta característica e poderiam formar uma carteira equilibrada visando a redução do risco. Além disso, para todos os subperíodos sempre existe um subgrupo dos setores de Materiais Básicos e Financeiro que contrasta com o comportamento de um outro subgrupo formado pelos setores de Bens Industriais, Consumo Cíclico e Serviços de Comunicação, mesmo que as ações que formem estes subgrupos possam variar com o tempo, o contraste entre representantes destes setores é sempre presente.

Para o Período II, pode-se destacar o par de as ações BRIV4 e BFIT3 que apresentam comportamento semelhante entre si e contrastam com o subgrupo formado pelas ações ELET3 e ELET6 ao longo de todo o período. O par de as ações BRIV4 e BFIT3 ainda contrasta durante todo o período com o subgrupo de ações BBDC4 e BBDC3. Desta forma pode-se escolher membros entre estes grupos contrastantes para reduzir o risco de um portfólio. Além disso, para todos os períodos existe um subgrupo formado pelos setores de Utilidade Pública, Petróleo, Gás e Biocombustível e Serviços de Comunicação que contrasta com o comportamento de um outro subgrupo formado pelos setores Financeiro e Consumo Cíclico, mesmo que

haja intercâmbios entre as ações que formem estes subgrupos com o tempo, o contraste entre representantes destes setores é sempre visível.

Desta forma, observando estas recomendações, pode-se compor uma carteira equilibrada visando redução de risco durante um certo período.

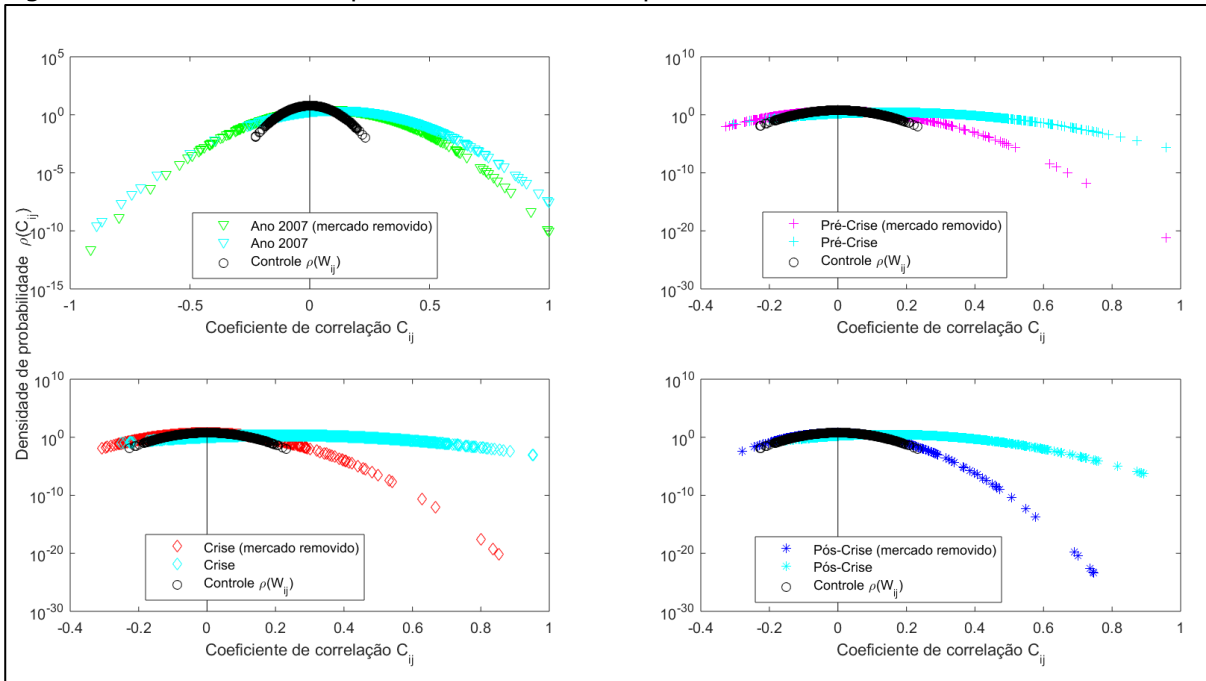
5.6 Remoção da Influência do Mercado

Conforme resultados apresentados na Seção 5.3, o autovetor que mais diverge é um índice de mercado. Necessita-se, no entanto, averiguar o significado dos demais autovetores que também divergem do comportamento previsto pela TMA, sobre os quais, busca-se saber, se poderiam ser índices de setores de mercado (PLEROU *et al.*, 2002).

No entanto, ao realizar análises de autovetores, a volatilidade devido ao movimento do mercado pode dificultar a tarefa de interpretação, uma vez que se obtém uma estrutura de dependência entre os ativos e o mercado, e não apenas a dependência entre os ativos (SANDOVAL JUNIOR; BORTOLUZZO; VENEZUELA, 2014). Uma forma bastante utilizada para remoção desta influência do mercado, é a utilização do Modelo de Índice Único (do inglês *Single Index Model* - SIM) apresentado na Equação (17) onde o resíduo, ou seja, a diferença entre a reta de regressão linear e os reais valores de retorno são os valores de retorno considerados sem a interferência do mercado (SANDOVAL JUNIOR; BORTOLUZZO; VENEZUELA, 2014; PLEROU *et al.*, 2002). Dessa forma, a fim de remover o efeito do mercado sobre os retornos de cada ação $r_i(t)$, foi realizado a regressão da Equação (17), e calculou-se os resíduos ϵ_i . Para tanto, utilizou-se ambas as series temporais (projeção $r_{100}(t)$ da Equação (16) e volatilidade diária do Ibovespa $r_{Ibovespa}(t)$ gerada conforme Equação (1)) como uma aproximação para $M(t)$ da Equação (17). Em seguida, foram calculadas as matrizes de correlação C^{100} e $C^{Ibovespa}$, conforme Equações (1) e (2), a partir dos resíduos ϵ_i^{100} (gerado a partir da remoção da influência de $r_{100}(t)$) e dos resíduos $\epsilon_i^{Ibovespa}$ (gerado a partir da remoção da influência de $r_{Ibovespa}(t)$) respectivamente, para o Período I. As Figuras 57 e 58 a seguir apresenta a

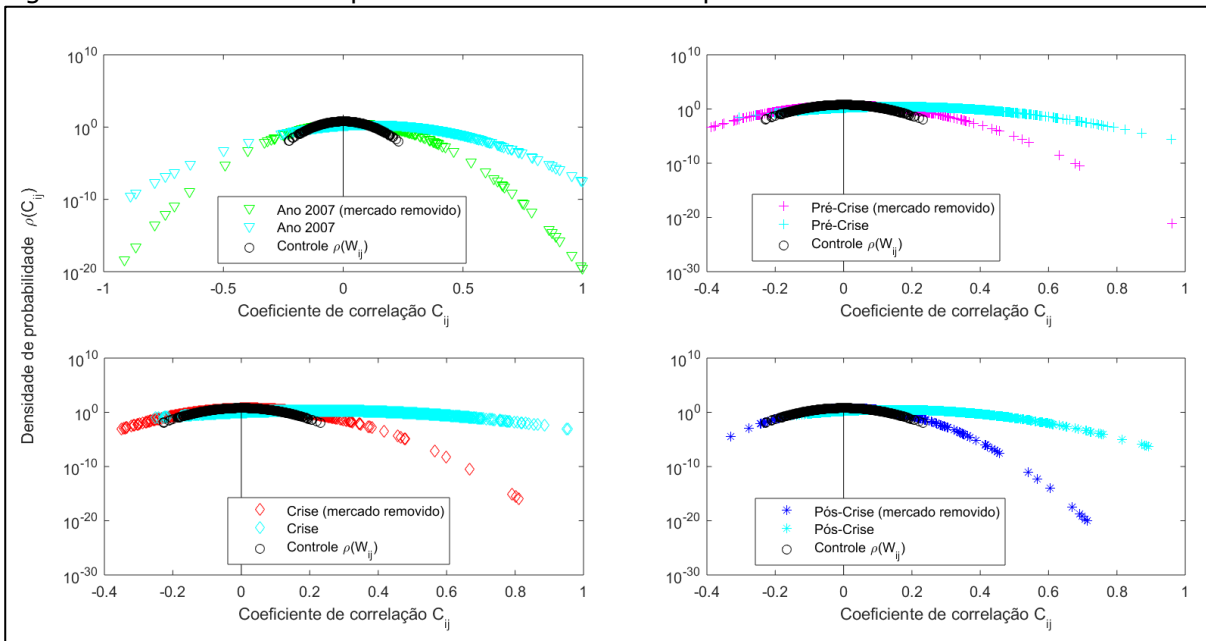
densidade de probabilidade das matrizes de correlação resultantes deste processo para o Período I.

Figura 57 - Densidade de probabilidade de C^{100} para o Período I.



Fonte: Autor.

Figura 58 - Densidade de probabilidade de $C^{Ibovespa}$ para o Período I.



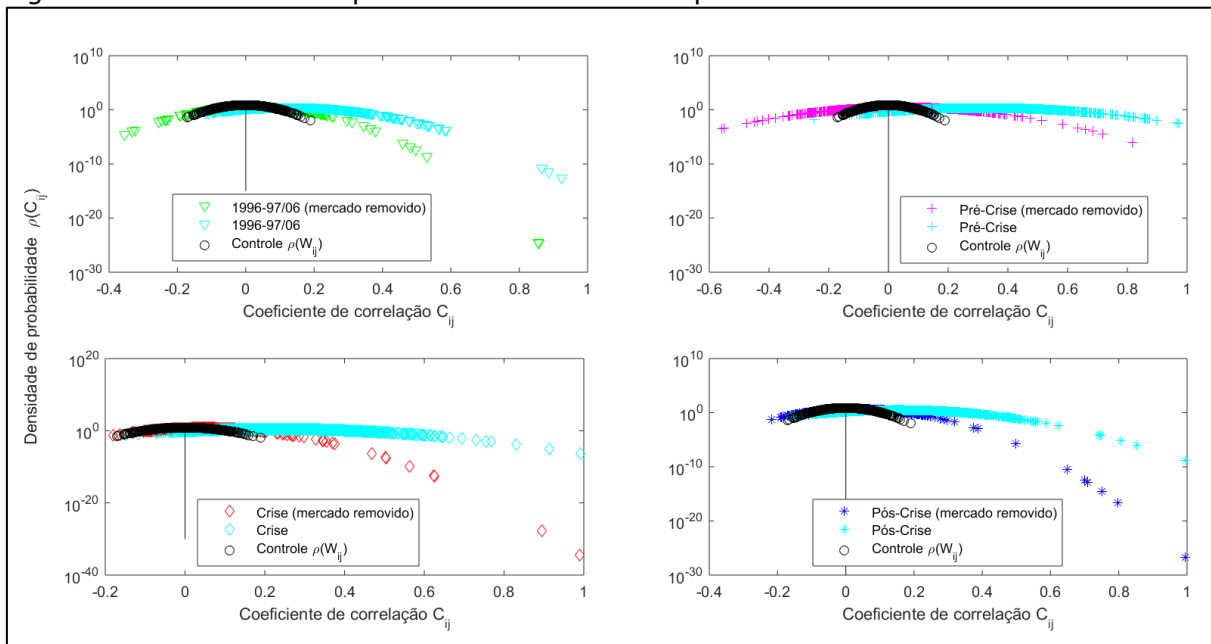
Fonte: Autor.

Percebe-se que ambas as abordagens de remoção da influência do mercado (remoção da influência de $r_{100}(t)$ ou da influência de $r_{Ibovespa}(t)$)

levam a correlações com semelhantes distribuições, sendo que o subperíodo que apresenta uma diferença visivelmente aparente é o subperíodo de 2007, com redução da média e da variância da matriz de correlação formada a partir de $\epsilon_i^{Ibovespa}$ quando comparada a matriz de correlação formada a partir de ϵ_i^{100} . Isto pode ser explicado uma vez que o coeficiente de correlação entre $r_{Ibovespa}(t)$ e $r_{100}(t)$ não é tão forte para o subperíodo de 2007 quando comparada aos demais períodos, conforme apresentado na Seção 5.3.

A Figura 59 apresenta a densidade de probabilidade da matriz de correlação para o Período II resultante do processo de remoção da influência do mercado representado pela volatilidade diária do Ibovespa $r_{Ibovespa}(t)$ gerada conforme Equação (1) como uma aproximação para $M(t)$ da Equação (17). Esta estratégia foi utilizada uma vez que, para todos os subperíodos, o coeficiente de correlação entre $r_{Ibovespa}(t)$ e a projeção $r_{61}(t)$ calculado conforme Equação (16) é alto, não resultando, portanto, em uma diferença muito grande nos resultados ao utilizar-se $r_{61}(t)$ como uma aproximação para $M(t)$ da Equação (17).

Figura 59 - Densidade de probabilidade de $C^{Ibovespa}$ para o Período II.



Fonte: Autor.

Ao analisar os resultados apresentados na Figuras 58 e 59, percebe-se que no geral, ao remover a influência do mercado, as médias das

correlações são levadas mais próximas de zero e as variâncias foram reduzidas. Aproximadamente temos, para o Período I:

- 2007: variâncias reduzidas pela metade;
- pré-crise: variâncias reduzidas a um terço;
- crise: variâncias reduzidas para menos de um quarto;
- pós-crise: variâncias reduzidas em mais de um terço.

Para o Período II, tem-se aproximadamente:

- 1996-1997/06: variância reduzida em menos da metade;
- pré-crise: variância reduzida em menos da metade;
- crise: variâncias reduzidas em mais de um terço;
- pós-crise: variâncias reduzidas em mais da metade.

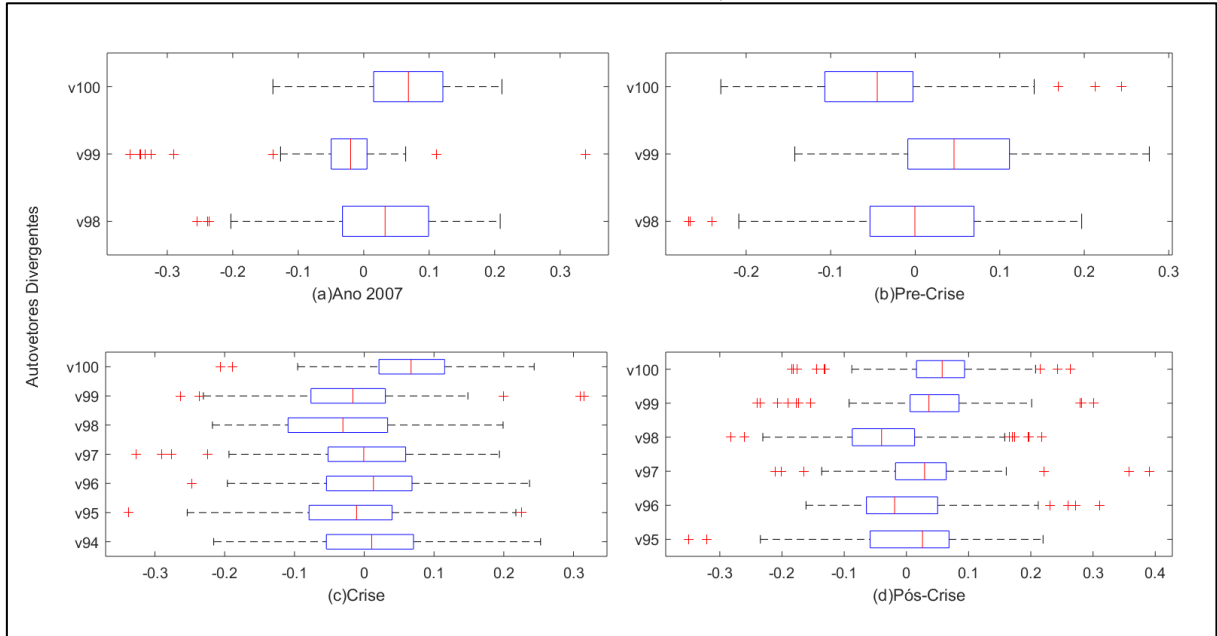
Desta forma, novamente ve-se o padrão do “efeito mola deformada”, com uma antecipação do comportamento semelhante a crise já apresentado no período de pré-crise e com uma tendência do período pós-crise de se aproximar ao comportamento anterior à crise, porém ainda apresentando resquícios de uma crise. Percebe-se também que a redução da variância para os subperíodos de crise é proporcionalmente maior que para os demais subperíodos, o que era esperado uma vez que a influência do mercado é muito maior e homogênea durante períodos de crise (PLEROU *et al.*, 2002; KULKARNI; DEO, 2007; DALKIR, 2009).

Calculou-se, então, os autovalores e autovetores relacionados. Para diferenciação de notação, os autovetores resultantes deste processo serão chamados de $v_i^{Ibovespa}$, para os autovetores gerados a partir dos resíduos $\epsilon_i^{Ibovespa}$. Em seguida foi realizada a análise das ações com participação significativa (*ITP e TP* da Equação (15)) nestes autovetores resultantes.

As Figuras 60 e 61 trazem os resultados de $v_i^{Ibovespa}$ para os Períodos I e II respectivamente, através do diagrama de caixa para os autovetores divergentes relativos aos autovalores que ultrapassaram o limite teórico superior λ_{max} da Equação (6). Conforme discutido na Seção 5.3, um autovetor com poucos *outliers* apresenta uma distribuição homogênea entre

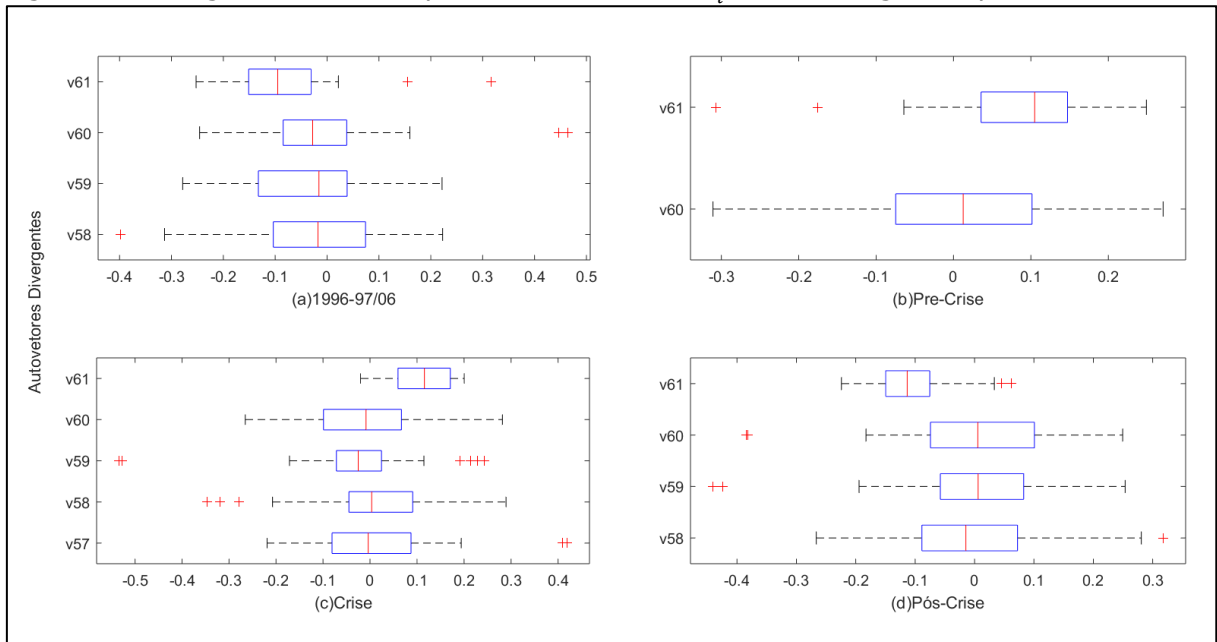
as ações, significando que ainda há uma influência do mercado sobre tal autovetor.

Figura 60 - Diagrama de Caixa para os autovetores $v_i^{Ibovespa}$ divergentes para o Período I.



Fonte: Autor.

Figura 61 - Diagrama de Caixa para os autovetores $v_i^{Ibovespa}$ divergentes para o Período II.



Fonte: Autor.

Percebe-se que após a remoção da influência do mercado sobre os retornos das ações, a quantidade de autovetores divergentes mudou (vide Seção 5.3), sendo que a quantidade de autovetores divergentes do Período I em cada subperíodo:

- 2007: passou de um total de quatro autovetores divergentes para três;
- pré-crise: passou de dois para três autovetores divergentes;
- crise: a quantidade de autovetores divergentes mais que dobrou, passando de três para sete;
- pós-crise: eram quatro autovetores divergentes e passou a ser seis.

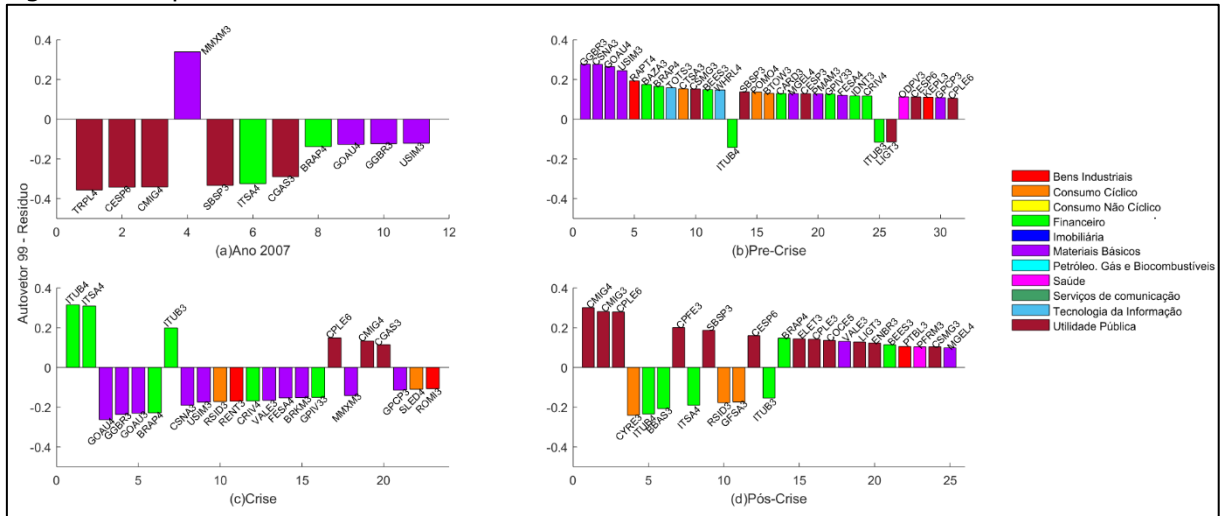
A quantidade de autovetores divergentes do Período II em cada subperíodo:

- 1996-1997/06: passou de dois autovetores divergentes para quatro;
- pré-crise: passou de um para dois autovetores divergentes;
- crise: a quantidade de autovetores divergentes passando de três para cinco;
- pós-crise: eram dois autovetores divergentes e passou a ser quatro.

Percebe-se novamente um padrão de comportamento relativo ao “efeito mola deformada”, onde após a crise o mercado tente a voltar a seu comportamento anterior ao período de instabilidade, porém ainda com resquícios da crise.

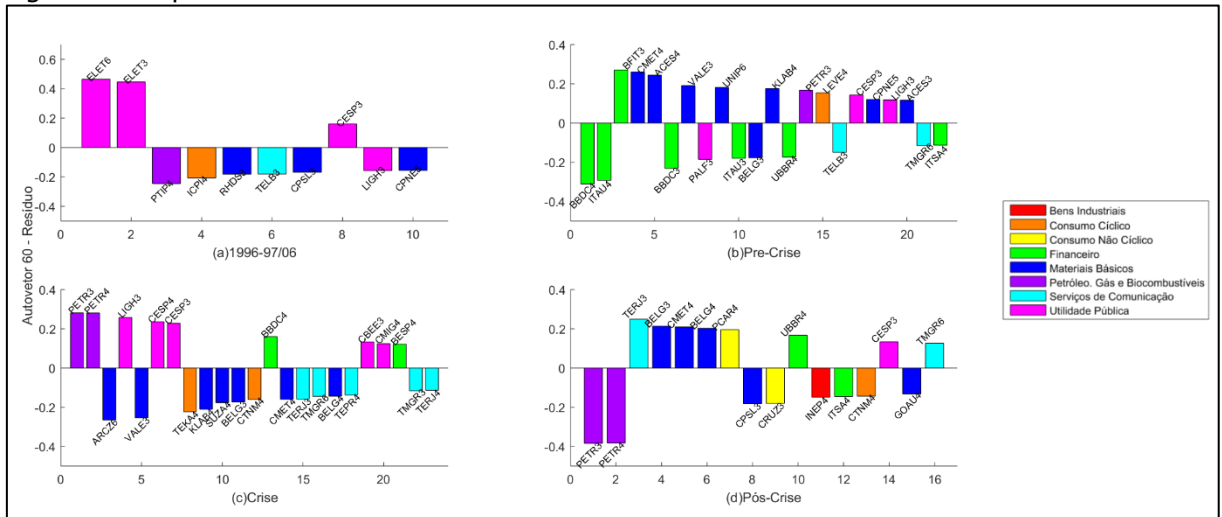
As Figuras de 62 a 63 mostram a composição dos autovetores divergentes $v_{99}^{Ibovespa}$ e $v_{60}^{Ibovespa}$ para os Períodos I e II respectivamente, referentes às empresas com participação significativa, os quais pode-se comparar com suas versões originalmente sem filtragem já apresentadas anteriormente na Seção 5.3.

Figura 62 - Composição dos autovetores $v_{99}^{Ibovespa}$ considerando as ações com participação significativa para o Período I.



Fonte: Autor.

Figura 63 - Composição dos autovetores $v_{60}^{Ibovespa}$ considerando as ações com participação significativa para o Período II.



Fonte: Autor.

Para o Período I apresentado na Figura 62, percebe-se que em comparação a análise realizada na Seção 5.3, o autovetor $v_{99}^{Ibovespa}$ para o subperíodo de 2007 continua praticamente idêntico, sendo que este vetor é composto majoritariamente pelo setor de Utilidades Públicas com influência do setor Financeiro, trazendo um forte contraste com a ação MMXM3 do setor de Materiais Básicos. Porém com relação aos demais subperíodos, houve grandes mudanças de composição ao retirar a influência do mercado. O $v_{99}^{Ibovespa}$ para o subperíodo de pré-crise é composto majoritariamente por empresas do setor de Materiais Básicos. Este setor

ainda apresenta força durante o subperíodo de crise, sendo contrastado com o setor Financeiro. Durante o subperíodo de pós-crise, no entanto, o setor de Materiais Básicos, o que vinha apresentando uma importante influência durante todo o período até então, praticamente desaparece, cedendo lugar ao setor de Utilidade Pública, o qual volta a ter importância da mesma forma em que tinha no subperíodo de 2007.

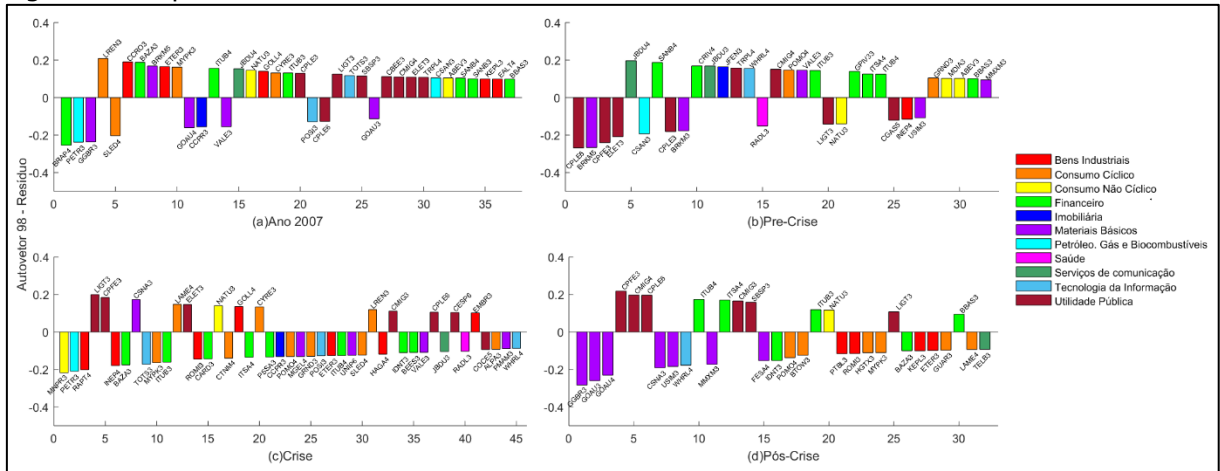
Apesar dos setores de maior influência neste autovetor $v_{99}^{Ibovespa}$ não serem sempre os mesmos, no entanto é possível observar um revezamento de influência entre três setores específicos. Pode-se então nomear este autovetor de "Contraste entre os setores de Utilidade Pública, Materiais Básicos e Financeiro".

Para o Período II, as ações que compuseram o autovetor v_{60} são praticamente as mesmas que compõem o autovetor $v_{60}^{Ibovespa}$, com algumas diferenças na proporção. O setor de Utilidade Pública é o mais influente durante o subperíodo de 1996-1997/06. Porém durante o subperíodo de pré-crise, o setor que aparece com força é o setor Financeiro. Durante o subperíodo de crise, o setor de Petróleo, Gás e Biocombustível é o mais forte e o setor de Utilidade Pública novamente retorna a ter participação significativa. Durante o pós-crise, o setor de Petróleo, Gás e Biocombustível ainda apresenta força, porém com componentes negativos, talvez influenciado pela crise da Rússia. O setor de Materiais Básicos, no entanto, é sempre presente como uma influência representativa durante todo o período. Desta forma pode-se nomear este autovetor de "Contraste entre os setores de Utilidade Pública, Materiais Básicos, Financeiro e Petróleo, Gás e Biocombustível".

Percebe-se que, mesmo para períodos distintos, o segundo autovetor que mais diverge em cada período representa praticamente o mesmo conjunto de setores empresariais, podendo, assim, indicar uma característica constante do mercado brasileiro presente durante um longo período.

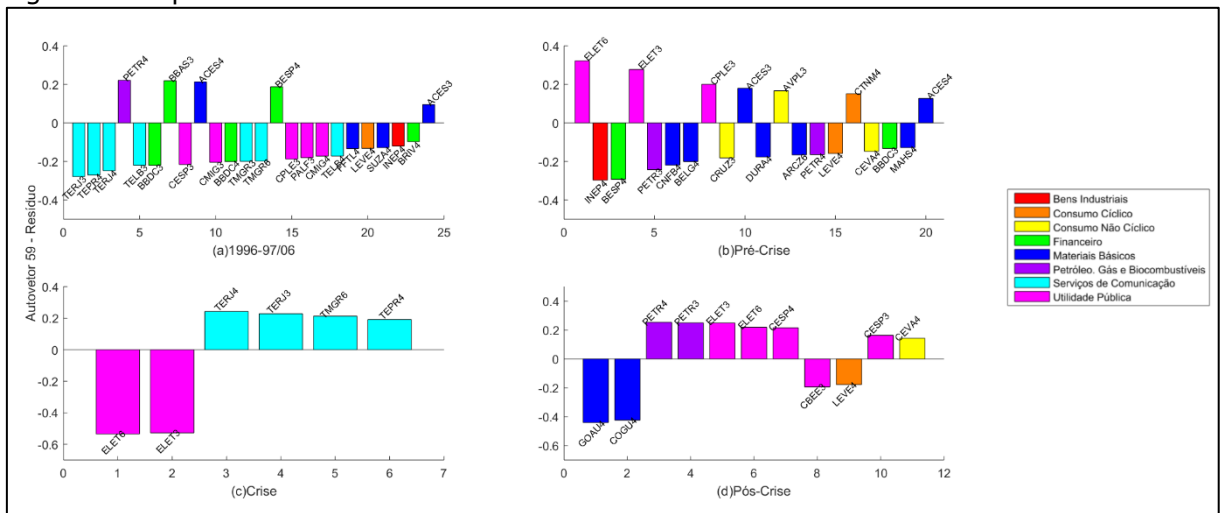
As Figuras 64 e 65, apresentam os autovetores $v_{98}^{Ibovespa}$ e $v_{59}^{Ibovespa}$ para os Períodos I e II respectivamente.

Figura 64 - Composição dos autovetores $v_{98}^{Ibovespa}$ considerando as ações com participação significativa para o Período I.



Fonte: Autor.

Figura 65 - Composição dos autovetores $v_{59}^{Ibovespa}$ considerando as ações com participação significativa para o Período II.



Fonte: Autor.

Para o Período I, os autovetores $v_{98}^{Ibovespa}$ para os subperíodos de 2007, pré-crise e crise apresentam composição genérica. Já para o período de pós-crise, apresenta uma composição mais definida, com influência do setor de Materiais Básicos, contrastando com o setor de Utilidade Pública. Todavia, o período de pré-crise já havia demonstrado indícios de quais setores iriam ser influentes após a crise. Apesar de ser bastante genérico, o nome que melhor descreve o $v_{98}^{Ibovespa}$ durante todo o período seria “Contraste entre os setores Utilidade Pública e Materiais Básicos”

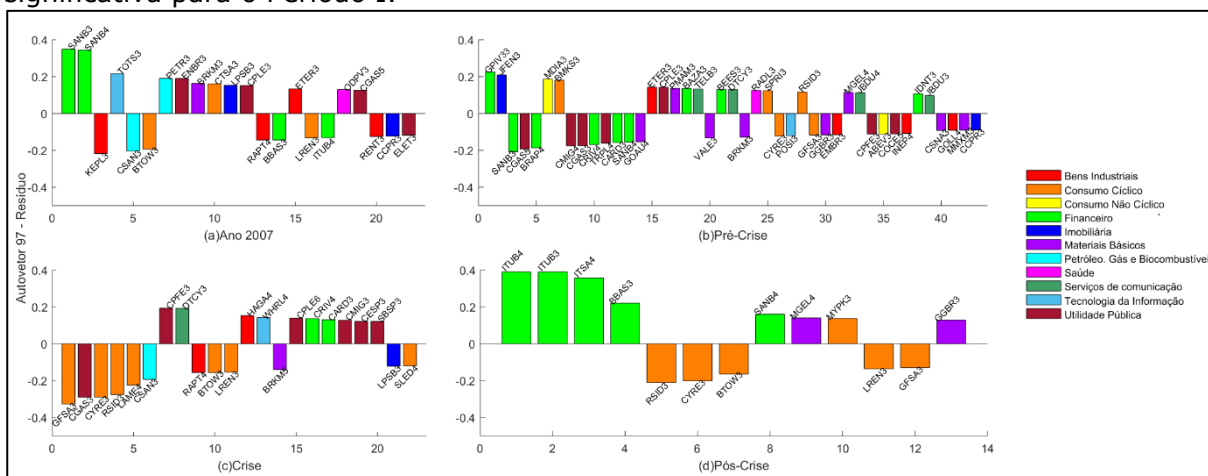
Para o Período II, o autovetor $v_{59}^{Ibovespa}$ para o subperíodo 1996-1997/06 apresenta uma forte influência do setor de serviços de

Comunicação. Durante o subperíodo de pré-crise, porém, esta influência praticamente desapareceu, sendo que o setor de Utilidade Pública é o setor de maior influência durante o subperíodo de pré-crise. Observa-se ainda que o $v_{59}^{Ibovespa}$ para o subperíodo de pré-crise apresenta uma composição genérica, o que pode ser explicado uma vez que tal autovetor para este subperíodo não é divergente, conforme apresentado previamente na Figura 61. Tal influência do setor de Utilidade Pública permanece até o fim do período e durante o subperíodo de crise é contrastado com o setor de Serviços de Comunicação. Durante o período de pós-crise, no entanto, o setor de Materiais Básicos se torna o mais influente. Observa-se ainda que as ações PETR4 ou PETR3 tiveram importante influência durante todo o período, exceto durante o subperíodo de crise, podendo ser um efeito colateral da crise russa.

Olhando o Período II como um todo, pode-se nomear o autovetor $v_{59}^{Ibovespa}$ de “Contraste entre os setores de Serviço de Comunicação, Utilidade Pública e Materiais Básicos”

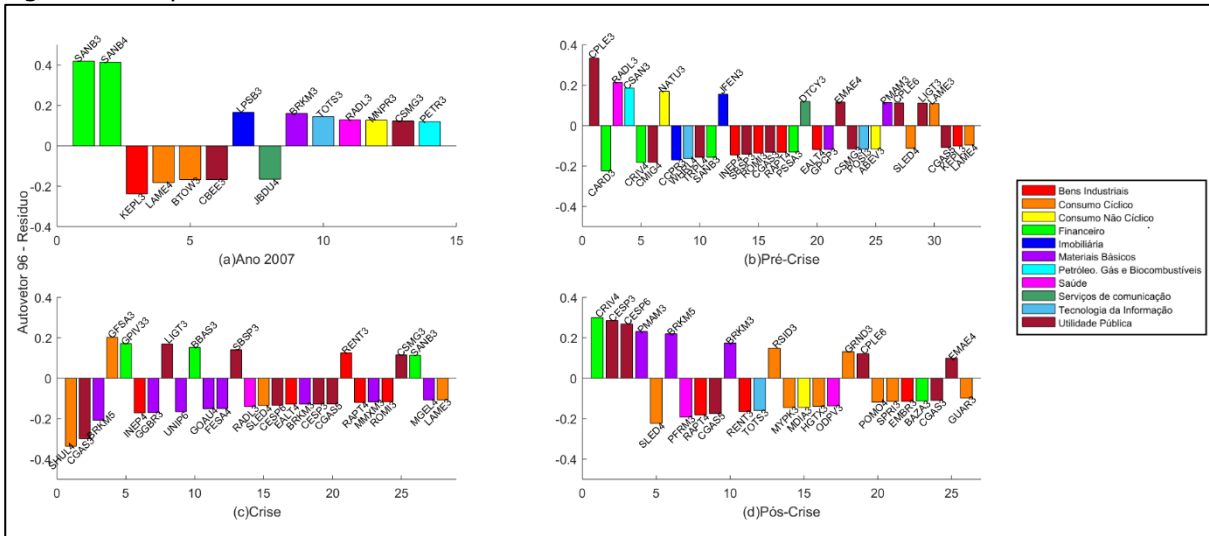
A seguir será apresentada a composição dos autovetores $v_{97}^{Ibovespa}$ e $v_{58}^{Ibovespa}$ dos Período I e II nas Figuras 66 e 67 respectivamente.

Figura 66 - Composição dos autovetores $v_{97}^{Ibovespa}$ considerando as ações com participação significativa para o Período I.



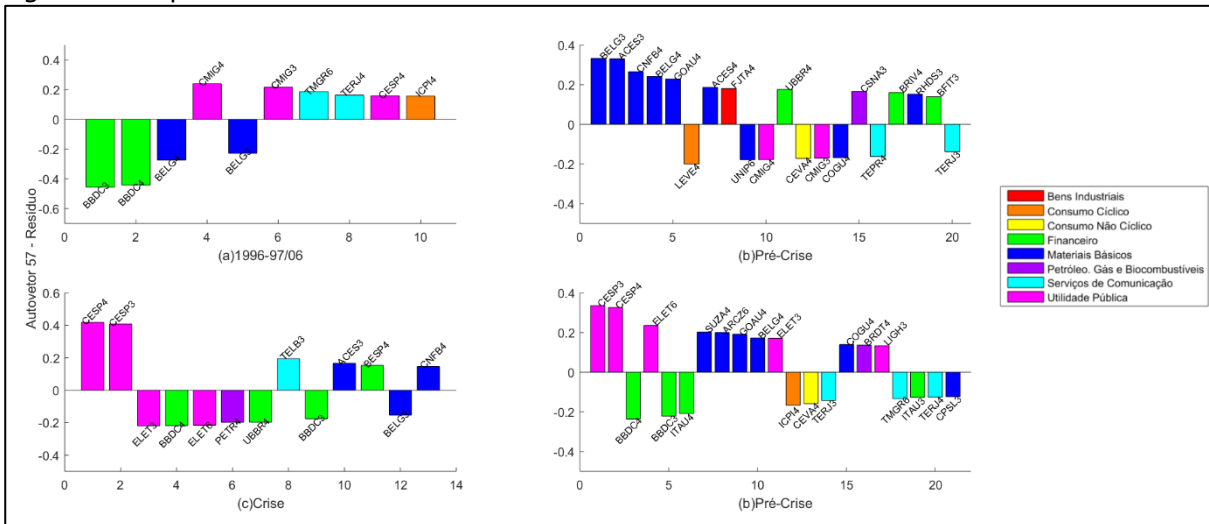
Fonte: Autor.

Figura 68 - Composição dos autovetores $v_{96}^{Ibovespa}$ considerando as ações com participação significativa para o Período I.



Fonte: Autor.

Figura 69 - Composição dos autovetores $v_{57}^{Ibovespa}$ considerando as ações com participação significativa para o Período II.



Fonte: Autor.

O autovetor $v_{96}^{Ibovespa}$ para o Período I durante o subperíodo de 2007 tem forte influência do setor Financeiro. Este setor continua influente durante todo o período. Para o período pré-crise o setor mais influente é o setor de Utilidade Pública, o qual também continua influente durante todo o período. Para o subperíodo de crise, o setor mais influente é o setor de Consumo Cíclico, seguido pelo setor de Utilidade Pública e Materiais Básicos. Por fim, para o subperíodo de pós-crise, os setores mais influentes são o Financeiro e Utilidade Pública. Pode-se nomear este autovalor como

“Contraste entre os setores Financeiro, Utilidade Pública, Consumo Cíclico e Materiais Básicos”, sendo este autovetor é mais genérico.

Analisando o autovetor $v_{57}^{Ibovespa}$ para o Período II, tem-se durante o subperíodo 1996-1997/06 uma forte influência do setor Financeiro. Já durante o período pré-crise, o autovetor é predominantemente composto pelo setor de Materiais Básicos, o qual foi influente durante todo o período. Os setores Utilidade Pública e Financeiro se demonstraram contrastantes durante os subperíodos de crise e pós-crise. Pode-se denominar este autovetor de “Contraste entre os setores Financeiro, Materiais Básicos e Utilidade Pública”.

Os demais autovetores divergentes apresentaram composições mais genéricas e de difícil análise, uma vez que estes estavam relacionados a autovalores muito próximos dos limites teóricos, o que pode ter influenciado seu comportamento.

Durante esta análise, pode-se perceber que os setores econômicos Materiais Básicos e Utilidade Pública constantemente estiveram presentes e com grande peso ajudam a compor os autovetores. Isto poderia ser uma característica do mercado brasileiro, uma vez que estes setores são uns dos mais influentes do mercado (vide Figuras 39 e 40). No entanto, percebeu-se que existe uma dependência temporal quanto a composição dos autovetores de forma geral, não podendo-se fixar de forma categórica, para os dados analisados neste trabalho, a qual grupo de empresas cada autovetor representa durante um longo período, uma vez que, de forma geral, existe uma alternância entre os setores que compõem significativamente um autovetor ao passar do tempo.

6. Conclusões

Uma das contribuições da presente tese ao estado da arte foram aplicação da Teoria das Matrizes Aleatórias (TMA) aos valores de retorno das ações da bolsa de valores brasileira em dois períodos: crise do *subprime* (anos 2000) e crise asiática (anos 1990), uma vez que não encontrou-se na literatura estudos que aplicassem a TMA sobre o mercado financeiro brasileiro no período da crise asiática, e foi possível encontrar apenas um trabalho que aplicou a TMA no mercado brasileiro durante o período da crise de 2008, porém com outro foco e abordagem (SANDOVAL JUNIOR; BORTOLUZZO; VENEZUELA, 2014). Pode-se observar que o método da TMA se demonstrou resiliente para análise das ações no mercado brasileiro, uma vez que: aplicou-se este método para períodos distintos no tempo os quais continham crises financeiras com diferentes características (a crise *subprime* de 2008 de proporções mundiais e a crise asiática-russa a qual atingiu com maior força os mercados emergentes); colocou-se obstáculos durante a análise, como a cotação do dólar, a qual atuou como um papel anticorrelacionado ao mercado como um todo; utilizou-se diferentes tamanhos de janelas para os subperíodos. Mesmo assim, a aplicação da TMA foi realizada com êxito durante toda a análise, podendo seus resultados serem ainda confirmados por algumas técnicas de análise multivariada.

Contribuiu-se também com a abordagem de análise de dados, ao subdividir cada período a ser analisado em quatro janelas temporais, o que possibilitou uma análise comparativa entre diferentes momentos característicos do comportamento do mercado financeiro brasileiro, ou seja, período sem volatilidade, período que antecede uma crise, período de crise e período posterior a uma crise. Esta divisão permitiu acompanhar com maior atenção os desdobramentos do mercado ao longo do tempo e como estes refletem sobre os dados analisados e respondem com diferentes aderências a previsão teórica dada pela TMA. Em especial, o período denominado de pré-crise teve uma grande importância na análise, uma vez que é um período intermediário que apresenta uma volatilidade um pouco maior que um período anterior estável, porém comparativamente com

volatilidade menor que o período de crise propriamente dita. Isto nos permite supor que seria possível em trabalhos futuros utilizar tal janela temporal para prever uma crise futura. Considera-se que esta subdivisão dos períodos em quatro janelas temporais considerando um subperíodo de pré-crise e os resultados decorrentes como uma contribuição ao estado da arte na área de análise de séries temporais financeiras.

Outra contribuição foi a observação e apontamento do “efeito mola deformada” nas duas crises, o qual seria uma marca das crises nas correlações das séries temporais que é observável pela aplicação da TMA. Tal efeito se trata de uma alteração substancial e consistente do comportamento das matrizes de correlação, dos autovalores e dos autovetores no período que antecede uma crise, apresentando sintomas que poderiam sinalizar a crise futura. Tais sintomas se intensificam durante a crise, onde o espectro de correlações de ações e a distribuição de autovalores e autovetor empíricos tendem a se desviar do comportamento previsto pela TMA. Após a crise, o comportamento do mercado tende a retornar ao seu comportamento original como um movimento de regressão ou um “efeito mola”. No entanto, esse movimento de retrocesso não retoma totalmente o comportamento que antecedeu a crise, em acordo com a literatura (SENSOY; YUKSEL; ERTURK, 2013; DALKIR, 2009), o que pode significar que o mercado ainda apresentava resquícios do período de crise ou mesmo significando uma mudança permanente no comportamento do mercado. O termo “efeito mola deformada” foi proposto nesta tese, uma vez que uma mola pode ser esticada, mas, uma vez que é reformada, não retorna totalmente a seu formato original.

Percebeu-se ainda que em cada subperíodo existe ao menos uma distribuição de componentes de autovetor significativamente diferente daquela prevista pela TMA, referente ao autovetor que mais divergente. Esse caso está relacionado com o maior autovalor que ultrapassara o limite teórico. Além disso, a análise da taxa de participação significativa dos componentes dos autovetores mostrou que esses casos apresentam uma taxa muito maior de participação significativa do que o número padrão

esperado pela TMA, o que, juntamente com o fato de as projeções destes autovetores sobre as séries temporais empíricas de retornos diários mostrar-se extremamente correlacionados com o Ibovespa, revela que estes autovetores são um índice do comportamento geral do mercado.

Contribui-se ainda com uma Análise de Componentes Principais do comportamento do mercado e da mudança de estrutura de dispersão dos retornos das ações da bolsa de valores brasileira, o que, juntamente com as técnicas de agrupamento, proporcionou uma análise rápida para identificar grupos de ações anticorrelacionadas as quais podem compor um portfólio equilibrado durante um longo período, reduzindo-se riscos mesmo em períodos de crise.

E, por fim, considera-se como uma contribuição ao estado da arte, a descrição da metodologia detalhada, apresentando todos os passos e cálculos utilizados e resultados decorrentes, uma vez que não foi possível encontrar uma bibliografia que detalhasse tal processo de aplicação, principalmente quanto ao cálculo do método *unfolding* dos autovalores da matriz de correlação dos dados empíricos, para o qual foi proposta neste trabalho uma nova maneira de abordar tal processo. Desta forma, buscou-se elucidar os passos de aplicação da TMA para posteriores aplicações em trabalhos futuros nas diferentes áreas de interesse que possam existir.

A aplicação do Modelo de Índice Único (do inglês *Single Index Model* - SIM) permitiu a filtragem das séries temporais de retornos diários o que possibilitou a análise dos demais autovetores que divergem. Foi possível perceber que os setores predominantes na composição de um autovetor não permanecem sempre os mesmos ao longo de um extenso período, desta forma estes autovetores não poderiam ser índices fixos de um determinado setor de mercado, porém poderiam ser índices da contradição entre alguns setores da economia, apesar dos setores Materiais Básicos e Utilidade Pública estarem constantemente presentes na composição da maioria dos autovetores para ambos os períodos. Desta forma, este seria um desafio para análises futuras, uma vez que esta investigação deve ser aprofundada.

Para trabalhos futuros, indica-se também a comparação entre resultados aqui obtidos pelo método da TMA com outras técnicas de análise de séries temporais. Indica-se, ainda, a continuação dos estudos sobre o período pré-crise com relação ao tamanho de uma janela ótima e uma possível utilização deste subperíodo para a previsão de uma crise futura. Outra possível verificação que pode ser realizada em trabalhos futuros seria observar se a cotação do dólar pode ser realmente utilizada como um fator anticorrelacionado com o mercado em outros períodos ou até mesmo para outros mercados utilizando a cotação do dólar local. Vê-se ainda, a necessidade de testar a eficiência em termos de risco das carteiras de ações construídas a partir da ACP na seção 5.5 utilizando, como por exemplo o índice Sharpe, e compará-la com outros métodos como Teoria de Markowitz. Pode-se ainda avançar na análise dos autovetores, verificando-se, por exemplo, qual a correlação dos demais autovetores divergentes com o índice Bovespa. Pode-se explorar, ainda, o comportamento da aplicação da metodologia de TMA em períodos de crise de origens adversas com o intuito de verificar se os resultados obtidos são suficientemente expressivos para caracterizar crises de tipos ou origens diferentes.

Referências Bibliográficas

- ABEPRO, Comissão de Graduação da. **Áreas e Sub-áreas de Engenharia de Produção**. 2008. Disponível em: < <https://www.abepro.org.br>>. Acesso em: jan. 2017.
- B3, Brasil Bolsa Balcão. **Relações com Investidores**. 2017. Disponível em: <<http://ri.bmfbovespa.com.br/>>. Acesso em: out. 2017.
- B3, Brasil Bolsa Balcão. **Séries Históricas**. 2020. Disponível em: <http://www.b3.com.br/pt_br/market-data-e-indices/servicos-de-dados/market-data/historico/mercado-a-vista/series-historicas/>. Acesso em: out. 2020.
- BM&FBOVESPA. **Sobre a BM&FBovespa**. 2016. Disponível em: <<http://www.bmfbovespa.com.br>>. Acesso em: out. 2017.
- BORÇA JUNIOR, G. R.; TORRES FILHO, E. T. Analisando a crise do subprime. **Revista do BNDES**, Rio de Janeiro, v. 15, n. 30, p. 129-159, 2008.
- BOUCHAUD, J. P.; POTTERS, M. Financial Applications of Random Matrix Theory: a Short Review, **ArXiv e-prints**, 2009.
- BRODY, T. A. *et al.* Random-matrix physics: spectrum and strength fluctuations. **Reviews of Modern Physics**, v. 53, n. 3, p. 385, 1981.
- CAMBRIDGE DICTIONARY. short selling. Cambridge University Press, 2020. Disponível em <<https://dictionary.cambridge.org/pt/dicionario/ingles/short-selling>> Acesso em: mar. 2020.
- CHEN, J. Short Sale. Investopedia, 2020. Disponível em <<https://www.investopedia.com/terms/s/shortsale.asp>> Acesso em: mar. 2020.
- COELHO, R.; RICHMOND, P.; HUTZLER, S. A random-matrix-theory-based analysis of stocks of markets from different countries. **Advances in Complex Systems**, v. 11, n. 5, p. 655-668, 2008.
- CONLON, T.; RUSKIN, H. J.; CRANE, M. Cross-correlation dynamics in financial time series. **Physica A: Statistical Mechanics and its Applications**, v. 388, n. 5, p. 705-714, 2009.
- DALKIR, M. Revisiting stock market index correlations. **Finance Research Letters**, v. 6, n. 1, p. 23-33, 2009.
- DA SILVA, E. L.; MENEZES, E. M. **Metodologia da pesquisa e elaboração de dissertação**. Florianópolis, UFSC, 4ed., 2005.
- DENG, W. *et al.* On the Application of the Cross-Correlations in the Chinese Fund Market: Descriptive Properties and Scaling Behaviors. **Advances in Complex Systems**, v. 14, n.

01, p. 97-109, 2011.

DYSON, F. J. Statistical theory of the Petróleo. Gás e Biocombustíveis levels of complex systems. **I. Journal of Mathematical Physics**, v. 3, n. 1, p. 140-156, 1962.

FABBENDER, H. *et al.* Hamiltonian square roots of skew-Hamiltonian matrices. **Linear algebra and its applications**, v. 287, n. 1-3, p. 125-159, 1999.

EL ALAOUI, M.; BOURI, E.; AZOURY, N. The Determinants of the US Consumer Sentiment: Linear and Nonlinear Models. **International Journal of Financial Studies**, v. 8, n. 3, p. 38, 2020.

FENN, D. J. *et al.* Temporal evolution of financial-market correlations. **Physical review E**, v. 84, n. 2, p. 026109, 2011.

FOLHA DE SÃO PAULO. Crash global abala mercados financeiros. **Agência Folha De São Paulo: Retrospectiva 97**. São Paulo, 1997. Disponível em: <https://www1.folha.uol.com.br/fof/retro97/retro97_economia1.htm>. Acesso em: jan. 2021.

FOLHA DE SÃO PAULO. Em 1997, crise na Ásia atingiu emergentes. **Folha de São Paulo - Mercado: Memória**. São Paulo, 2006. Disponível em: <<http://www.folha.uol.com.br/>>. Acesso em: jan. 2021.

G1. Entenda o que é recessão técnica. **G1 Economia**. São Paulo, 2014. Disponível em <<http://g1.globo.com/economia/noticia/2014/08/entenda-o-que-e-recessao-tecnica.html>>. Acesso em: set. 2020.

GARCÍA, A. Global financial indices and twitter sentiment: A random matrix theory approach. **Physica A: Statistical Mechanics and its Applications**, v. 461, p. 509-522, 2016.

GIL, A. C. **Como elaborar projetos de pesquisa**. 4 ed. São Paulo: Atlas, 2007.

GOLDFAJN, I. The swings in capital flows and the Brazilian crisis. **SSRN**, 224274, 2000.

GUHR, T.; MÜLLER-GROELING, A.; WEIDENMÜLLER, H. A. Random-matrix theories in quantum physics: common concepts. **Physics Reports**, v. 299, n. 4-6, p. 189-425, 1998.

GUIMARÃES, L. Entenda o que é recessão técnica e relembre o histórico de crises no Brasil. **CNN Brasil Business**. São Paulo, 2020. Disponível em <<https://www.cnnbrasil.com.br/business/2020/09/01/o-que-e-recessao-tecnica-e-quando-isso-ja-aconteceu-no-brasil>>. Acesso em: set. 2020.

HAIR, J. F. *et al.* **Análise multivariada de dados**. 6 Ed. Porto Alegre: Bookman editora,

2009.

IBGE, Instituto Brasileiro de Geografia e Estatística. **Sistema de Contas Nacionais Trimestrais – SCNT**. 2020. Disponível em: <<https://www.ibge.gov.br/estatisticas/economicas/comercio/9300-contas-nacionais-trimestrais.html?=&t=downloads>>. Acesso em: set. 2020.

IEDI, Instituto de Estudos para o Desenvolvimento Industrial. Entendendo a Crise Subprime e suas Repercussões na Economia Mundial – 1ª Parte. **Carta IEDI**, 300 ed., 2008. Disponível em: <<https://www.iedi.org.br/cartas/>>. Acesso em: jun. 2020.

IEDI, Instituto de Estudos para o Desenvolvimento Industrial. O Brasil no ranking mundial de IED. **Carta IEDI**, 1020 ed., 2020. Disponível em: <<https://www.iedi.org.br/cartas/>>. Acesso em: jan. 2021.

IKRAMOV, K. D. Hamiltonian square roots of skew-Hamiltonian matrices revisited. **Linear Algebra and its Applications**, v. 325, n. 1-3, p. 101-107, 2001.

INVESTING.COM. **Usd/brl - dólar americano real brasileiro**, 2020. Disponível em: <<https://br.investing.com/currencies/usd-brl-historical-data>>. Acesso em: jun. 2020.

ISI, **International Statistical Institute**, DEVELOPING COUNTRIES. 2018. Disponível em <<https://www.isi-web.org/index.php/resources/developing-countries>>. Acesso em: mar. 2018.

JOHNSON, R. A.; WICHERN, D. W. **Applied multivariate statistical analysis**. Vol. 5. No. 8. Upper Saddle River, NJ: Prentice hall, 2002.

KENTON, W.; SCOTT, G. Financial Crisis. **Investopedia**, 2020

Disponível em <<https://www.investopedia.com/terms/f/financial-crisis.asp>> Acesso em: mar. 2020.

KRUGMAN, P. How did economists get it so wrong? **The New York Times**, 2009. Disponível em <<https://www.nytimes.com/2009/09/06/magazine/06Economic-t.html>>. Acesso em: mar. 2018.

KULKARNI, V.; DEO, N. Correlation and volatility in an Indian stock market: A random matrix approach. **The European Physical Journal B**, v. 60, n. 1, p. 101-109, 2007.

KUMAR, S.; DEO, N. Analysing correlations after the financial crisis of 2008 and multifractality in global financial time series. **Pramana**, v. 84, n. 2, p. 317-325, 2015.

KUMAR, S.; DEO, N. Correlation and network analysis of global financial indices. **Physical Review E**, v. 86, n. 2, p. 026101, 2012.

LAEVEN, L.; VALENCIA, F. **Systemic banking crises: A new database**. International Monetary Fund - Working Paper, 2008.

LEITE, H. P.; SANVICENTE, A. Z. **Índice Ibovespa: um padrão para os investimentos brasileiros**. São Paulo: Atlas, 1994.

LIM, G. *et al.* Structure of a financial cross-correlation matrix under attack. **Physica A: Statistical Mechanics and its Applications**, v. 388, n. 18, p. 3851-3858, 2009.

LIM, K. *et al.* Statistical properties of the stock and credit market: RMT and network topology. **Physica A: Statistical Mechanics and its Applications**, v. 407, p. 66-75, 2014.

LIMA, T. D.; DEUS, L. N. A crise de 2008 e seus efeitos na economia brasileira. **Revista Cadernos de Economia**, v. 17, n. 32, p. 52-65, 2013.

LIOU, H. I.; CAMARDA, H. S.; RAHN, F. Application of Statistical Tests for Single-Level Populations to Neutron-Resonance-Spectroscopy Data. **Physical Review C**, v. 5, n. 3, p. 1002, 1972.

LIVAN, G.; REBECCHI, L. Asymmetric correlation matrices: an analysis of financial data. **The European Physical Journal B**, v. 85, n. 6, p. 213, 2012.

MANTEGNA, Rosario N. Hierarchical structure in financial markets. **The European Physical Journal B-Condensed Matter and Complex Systems**, v. 11, n. 1, p. 193-197, 1999.

MARKOWITZ, H. Portfolio selection. **The journal of finance**, v. 7, n. 1, p. 77-91, 1952.

MARTINS, M. A. DOS S.; METTE, F.; MACEDO, G. R. DE. A Utilização de Redes Neurais Artificiais para a Estimção dos Preços da Petrobrás PN na Bovespa. **ConTexto**, v. 8, n. 14, p. 1-16, 2008.

MATHWORKS, Inc., **Statistics and Machine Learning Toolbox**, 2020. Disponível em <<https://www.mathworks.com/help/stats/>> Acesso em: mar. 2020.

MATUCHESKI, S.; CLEMENTE, A.; SANDRINI, J. C. Governança corporativa e volatilidade das ações negociadas na Bovespa na crise financeira de 2008. **Revista Brasileira de Estratégia - REBRAE**, v. 2, n. 2, p. 171-183, ago. 2009.

MEIYING, Zhu; BIN, Zhang. Analysis of Financial Correlation Matrix Using Random Matrix Theory. In: **Business Intelligence and Financial Engineering (BIFE), 2011 Fourth International Conference on**. IEEE, 2011. p. 311-314.

MEYER, H.; D'AURIAC, J. C. A.; BRUUS, H. Spectral properties of statistical mechanics models. **Journal of Physics A: Mathematical and General**, v. 29, n. 18, p. L483, 1996.

MEYER, K.; HALL, G.; OFFIN, D. **Introduction to Hamiltonian Dynamical Systems and the N-Body Problem**. Springer, New York, NY, 2009.

MINITAB, Suporte ao. **Interpretar todas as estatísticas e gráficos para Análise de componentes principais**, 2020. Disponível em <<https://support.minitab.com/pt-br/minitab/18/>>. Acesso em: mar. 2020.

NGUYEN, H.; TAO, T.; VU, V. Random matrices: tail bounds for gaps between eigenvalues. **Probability Theory and Related Fields**, p. 1-40, 2015.

OH, G. *et al.* Statistical properties of cross-correlation in the Korean stock market. **The European Physical Journal B**, v. 79, n. 1, p. 55-60, 2011.

PARTOVI, M. H.; Caputo, M. Principal portfolios: Recasting the efficient frontier. **Economics Bulletin**, v. 7, n. 3, p. 1-10, 2004.

PENG, Y. *et al.* Between nonlinearities, complexity, and noises: an application on portfolio selection using kernel principal component analysis. **Entropy**, v. 21, n. 4, p. 376, 2019.

PINTO, T. dos S. Neoliberalismo e crise russa de 1998. **Mundo Educação**. 2021. Disponível em <<https://mundoeducacao.uol.com.br/historiageral/crise-russa.htm>>. Acesso em: jan. 2021.

PLEROU, V. *et al.* A random matrix theory approach to financial cross-correlations. **Physica A: Statistical Mechanics and its Applications**, v. 3, n. 287, 2000.

PLEROU, V. *et al.* Collective behavior of stock price movements—a random matrix theory approach. **Physica A: Statistical Mechanics and its Applications**, v. 299, n. 1, p. 175-180, 2001.

PLEROU, V. *et al.* Random matrix approach to cross correlations in financial data. **Physical Review E**, v. 65, n. 6, 2002.

PLEROU, V. *et al.* Universal and nonuniversal properties of cross correlations in financial time series. **Physical Review Letters**, v. 83, n. 7, p. 1471, 1999.

REIS, T. Crise asiática de 1997: entenda como aconteceu a crise dos Tigres Asiáticos. **Suno Research**, 2019. Disponível em <<https://www.suno.com.br/artigos/crise-asiatica/>>. Acesso em: jun. 2020.

REMDE, M. Z. A crise brasileira de 1998-1999: análise sob a ótica do modelo de Krugman. Universidade Federal de Pelotas (**Dissertação**). Pelotas, 2013. Disponível em <<https://wp.ufpel.edu.br/ppgom/files/2017/06/monicaremde.pdf>> Acesso em: jun. 2020.

SANDOVAL JUNIOR, L. To lag or not to lag? How to compare indices of stock markets that operate on different times. **Physica A: Statistical Mechanics and its Applications**, v. 403, p. 227-243, 2014.

SANDOVAL JUNIOR, L.; BORTOLUZZO, A. B.; VENEZUELA, M. K. Not all that glitters is RMT in the forecasting of risk of portfolios in the Brazilian stock market. **Physica A: Statistical Mechanics and its Applications**, v. 410, p. 94-109, 2014.

SANDOVAL JUNIOR, L.; FRANCA, I. De P. Correlation of financial markets in times of crisis. **Physica A: Statistical Mechanics and its Applications**, v. 391, n. 1, p. 187-208, 2012.

SANDRONI, P. **Novíssimo dicionário de economia**. Editora Best Seller, 1999.

SENSOY, A.; YUKSEL, S.; ERTURK, M. Analysis of cross-correlations between financial markets after the 2008 crisis. **Physica A: Statistical Mechanics and its Applications**, v. 392, n. 20, p. 5027-5045, 2013.

SILVA, I. N. DA; SAPATTI, D. H.; FLAUZINO, R. A. **Redes Neurais Artificiais: Para Engenharia e Ciências Aplicadas**. São Paulo: Artliber, 2010.

SUPARNO, E. N. *et al.* Group identification in Indonesian stock market. In: **Journal of Physics: Conference Series**. IOP Publishing, 2016. p. 012037.

UTSUGI, A.; INO, K.; OSHIKAWA, M. Random matrix theory analysis of cross correlations in financial markets. **Physical Review E**, v. 70, n. 2, p. 026110, 2004.

UOL. Entenda o que causou a crise financeira de 2008. **UOL Economia**. 2016. Disponível em <<https://economia.uol.com.br/noticias/redacao/2016/02/27/entenda-o-que-causou-a-crise-financeira-de-2008.htm>>. Acesso em: jun. 2020.

VIVO, P. UnfoldingPositive: version 1.0. **File Exchange**, MathWorks, MATLAB Central, 2009. Disponível em <<https://www.mathworks.com/matlabcentral/fileexchange/24122-unfoldingpositive>>. Acesso em: mar. 2018.

WIGNER, E. P. Characteristic Vectors of Bordered Matrices With Infinite Dimensions, **Annal of Mathematics**, vol. 62, no. 3, p. 548-564, 1955.

YAHOO FINANÇAS. **Dados Históricos**, 2020. Disponível em <<https://br.financas.yahoo.com>>. Acesso em: jun. 2020.

ZAHAR, J. Crise no Sudeste Asiático. **Memoria Globo**. Rio de Janeiro, 2005. Disponível em <<https://memoriaglobo.globo.com/jornalismo/coberturas/crise-no-sudeste-asiatico/>>. Acesso em: jan. 2021.