



impa Instituto Nacional de Matemática Pura e Aplicada

---

# Aplicação do Método de Galerkin no Problema de Alocação Ótima de Ativos

Autora: Louise Albuquerque Bassan

Orientadora: Claudia Alejandra Sagastizábal

Rio de Janeiro  
Julho de 2015



Para os meus pais.



## Agradecimentos

Aos meus queridos pais, Catia e Márcio, pelo amor e carinho que sempre tiveram por mim e por estarem sempre ao meu lado neste meu percurso acadêmico e de vida. Com vocês as vitórias são mais especiais.

Aos meus amigos da Accenture, que sempre me apoiaram e incentivaram ao longo do curso. Em especial, ao Gustavo Veronese, por me liberar para assistir às aulas e por me incentivar a sempre continuar estudando.

À minha orientadora, Claudia Alejandra Sagastizábal, pela paciência e confiança em que depositou em mim e por todo o aprendizado que me proporcionou. Não poderia ter escolhido orientadora mais atenciosa e que além de ser uma excelente professora é uma pessoa que tenho profunda admiração.

Ao professor e amigo, Juan Pablo Luna, por ter me ajudado tão de perto com este trabalho de dissertação. Sou muito grata por ter tido a oportunidade de aprender com este professor exemplar e ter feito um grande amigo.

Ao professor e coordenador do curso, Jorge Passamani Zubelli, pelos dois anos de ensinamento que moldaram o meu raciocínio e conhecimento em finanças. Tenho muita admiração pelo seu entusiasmo em ensinar.

A todos os amigos da turma de mestrado, pelos momentos de descontração e confraternização que tivemos durante todo o curso e pelas diversas vezes que me ajudaram a solucionar problemas. Em especial, agradeço ao Douglas Vieira, por todas as vezes que estudamos juntos e por ser um amigo tão querido.

Aos professores, alunos, funcionários e a todos que contribuem para que o Instituto Nacional de Matemática Pura e Aplicada seja um lugar fantástico para adquirir conhecimento.



## Resumo

A escolha adequada de uma carteira de investimentos traz consigo um grande desafio às instituições financeiras. Em geral, as empresas buscam investimentos que possuam retornos aceitáveis e ao mesmo tempo minimizem os riscos associados às oscilações do mercado. Administrar tal escolha de maneira tratável e eficiente torna muitas vezes este problema não trivial, principalmente ao considerar a natureza incerta dos ativos de forma dinâmica. Neste contexto, o presente trabalho tem como objetivo analisar um novo método para a resolução do problema de programação estocástica associado. O método de Galerkin permite que sejam incorporadas visões dinâmicas de mercado ao problema de carteira ótima de tal modo que o conjunto de soluções viáveis seja menor e, desta forma, a dimensão do problema de otimização seja reduzida quando comparada com uma formulação multi-estágios. Ainda, ao considerar tais visões de mercado, é possível encontrar soluções mais realistas para o problema. O método proposto é comparado com sua versão determinística equivalente através das funções de desvio padrão e AVaR. Ao final do estudo são apresentadas as considerações a respeito do método e são listados possíveis trabalhos futuros.

**Palavras-chave:** Otimização de Carteiras, Método de Galerkin, Programação Estocástica, Funções de Desvio





## **Abstract**

For financial institutions, to appropriately choose an investment portfolio is a challenging problem. In general, companies seek investments which offer acceptable returns while minimizing the risks associated with market fluctuations. Handling that choice in a tractable and efficient way often makes this problem nontrivial, especially when considering dynamically the uncertain nature of the assets. This work proposes a Galerkin approach to approximately solve the stochastic optimization portfolio problem. With this strategy exogenous information, available in the form of visions of the market experts, can be incorporated. When compared with the multistage formulation, the feasible set is smaller, but still realistic. Additionally, the proposed method is compared with its equivalent deterministic version by standard and AVaR deviation. The study finishes with considerations regarding the method and possible future work.

**Key words:** Portfolio Optimization, Galerkin Method, Stochastic Programming, Deviation Risk Functionals



# Sumário

<b>Índice</b>	<b>ix</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Motivação e Objetivo . . . . .	1
1.2 Problema Proposto . . . . .	2
1.3 Contribuições . . . . .	3
1.4 Estrutura do Trabalho . . . . .	3
<b>2 Revisão Bibliográfica</b>	<b>5</b>
<b>3 Metodologia</b>	<b>9</b>
3.1 Descrição dos Dados . . . . .	9
3.2 Geração de Cenários . . . . .	11
3.2.1 Detecção de Outliers . . . . .	11
3.2.2 Elipsóide de Volume Mínimo . . . . .	12
3.2.3 Estimadores Robustos . . . . .	16
3.2.4 Método de Monte Carlo . . . . .	17
3.3 Medidas de Risco . . . . .	18
3.3.1 Definição . . . . .	18
3.3.2 Medidas Coerentes de Risco . . . . .	19
3.3.3 Funções de Desvio . . . . .	20
3.4 Fronteira Eficiente e Reta de Alocação de Capital . . . . .	22
3.5 Modelo Determinístico Equivalente . . . . .	24
3.5.1 Desvio Padrão . . . . .	24
3.5.2 Desvio <i>Average Value at Risk</i> . . . . .	24
<b>4 Método de Galerkin Para Otimização Estocástica</b>	<b>27</b>
4.1 Conceitos Iniciais . . . . .	27
4.2 Aproximações de Galerkin . . . . .	28
4.3 Especialistas . . . . .	31
4.3.1 Determinístico . . . . .	31
4.4 Problema $\alpha$ . . . . .	32
4.4.1 Desvio Padrão . . . . .	34
4.4.2 Desvio <i>Average Value at Risk</i> . . . . .	35
<b>5 Experimentos Numéricos</b>	<b>37</b>
5.1 Parâmetros de Entrada . . . . .	37
5.2 Análise de Consistência do Modelo Estático . . . . .	39
5.3 Análise de Consistência do Modelo Proposto . . . . .	47

5.4 Desempenho das Carteiras . . . . .	49
<b>6 Considerações Finais</b>	<b>55</b>
<b>Referências Bibliográficas</b>	<b>56</b>
<b>A Representações Matriciais</b>	<b>59</b>
A.1 Modelo Determinístico Equivalente com Desvio Padrão . . . . .	59
A.2 Modelo Determinístico Equivalente com Desvio <i>Average Value at Risk</i> . . . . .	61
A.3 Problema $\alpha$ com Desvio Padrão . . . . .	63
A.4 Problema $\alpha$ com Desvio <i>Average Value at Risk</i> . . . . .	64
<b>B Códigos Utilizados</b>	<b>65</b>

# Lista de Figuras

3.1	Gráficos históricos semanais das 15 ações negociadas na BM&FBovespa. . . . .	10
3.2	Estimadores amostrais: Falta de robustez. . . . .	12
3.3	Elipsóide de Volume Mínimo. . . . .	13
3.4	Detecção de Outliers. . . . .	14
3.5	Elipsóide de Volume Mínimo. . . . .	15
3.6	Volume do elipsóide como função do número de observações. . . . .	16
3.7	Simulação de Monte Carlo. . . . .	17
3.8	VaR e AVaR de uma carteira. . . . .	19
3.9	Fronteira Eficiente. . . . .	23
3.10	Fronteira Eficiente com Reta de Alocação de Capital. . . . .	23
5.1	Grupos de Ativos. . . . .	38
5.2	Fronteira eficiente, reta de alocação de capital e composição da carteira em $t = 1$ . . . . .	40
5.3	Distribuição dos log-retornos do portfólio eficiente em $t = 1$ . . . . .	41
5.4	Composição da carteira através das abordagens <i>Risk Averse</i> e <i>Risk Neutral</i> em $t = 4$ . . . . .	42
5.5	Distribuição dos log-retornos dos ativos GOAU4, EMBR3 e ITUB4 em $t = 4$ . . . . .	43
5.6	Distribuição dos log-retornos da carteira através das abordagens <i>Risk Averse</i> e <i>Risk Neutral</i> em $t = 4$ . . . . .	43
5.7	Fronteira eficiente, reta de alocação de capital e composição da carteira em $t = 1$ e $t = 4$ . . . . .	44
5.8	Distribuição dos log-retornos dos portfólios eficientes em $t = 1$ e $t = 4$ . . . . .	46
5.9	Distribuição dos log-retornos dos portfólios obtidos pelo especialista em $t = 1$ . . . . .	47
5.10	Distribuição dos log-retornos dos portfólios de Galerkin $\times$ Estático em $t = 1$ . . . . .	48
5.11	Proporção $\alpha$ para cada nível de retorno, função de desvio e percentual mínimo. . . . .	49
5.12	Fronteira eficiente, reta de alocação de capital e composição da carteira em $t = 1$ para o modelo estático (esquerda) e proposto (direita). . . . .	50
5.13	Distribuição dos log-retonos das carteiras eficientes do modelo estático $\times$ proposto com 10.000 simulações. . . . .	51
5.14	Preços e log-retornos das carteiras estática e de Galerkin no período <i>out-of-sample</i> de 37 semanas. . . . .	52
5.15	Preços e log-retornos das carteiras estática e de Galerkin no período <i>out-of-sample</i> de 9 semanas. . . . .	53



# Lista de Tabelas

2.1	Programação matemática aplicada ao problema de otimização de carteira. . . . .	8
3.1	Empresa, Grupo Econômico e Participação no Ibovespa obtido através do site da BM&FBovespa [21]. . . . .	11
4.1	Tempos de computação em segundos para o método de Galerkin com 20 estratégias de base e 5.000 cenários. . . . .	31
4.2	Estimativas da amostra dos valores objetivos obtidos com o método de Galerkin e discretizações baseadas em árvores bem como a melhor estratégia de base individual. A quarta coluna indica os limites inferiores estatísticos obtidos com o método de árvores aleatórias. . . . .	31
5.1	Risco, retorno esperado e alocação do portfólio eficiente (PE) em $t = 1$ . . . . .	40
5.2	Risco, retorno esperado e alocação do portfólio eficiente (PE) em $t = 1$ e $t = 4$ . . . . .	45
5.3	Estatísticas básicas das séries de log-retornos das carteiras com 10.000 simulações. . . . .	51
5.4	Estatísticas das séries de log-retornos das carteiras para o período de 37 semanas. . . . .	52
5.5	Estatísticas das séries de log-retornos das carteiras para o período de 9 semanas. . . . .	53
5.6	Tempos computacionais dos métodos estático e de Galerkin em segundos. . . . .	54





# Capítulo 1

## Introdução

O método de Galerkin é uma técnica clássica para a aproximação de problemas infinitodimensionais por problemas de dimensão finita e tem sido amplamente aplicado em diversos campos da matemática aplicada, com maior notoriedade no método de elementos finitos para equações diferenciais parciais. Em otimização estocástica, conforme descrito por Koivu e Penanen [1], a ideia por de trás do método de Galerkin é tomar um conjunto finito de funções base que são pontos viáveis no problema original, não necessariamente ótimos, e em seguida buscar os pontos ótimos entre suas combinações lineares ou simpliciais, dependendo do problema considerado. Este trabalho de dissertação considera o emprego das técnicas de Galerkin para lidar com a dinâmica das incertezas em problemas de carteira ótima.

### 1.1 Motivação e Objetivo

A escolha adequada de uma carteira de investimentos é um problema recorrente para as instituições financeiras. Em geral, as empresas buscam investimentos que possuam retornos aceitáveis e ao mesmo tempo minimizem os riscos associados às oscilações do mercado. Administrar tal escolha de maneira tratável e eficiente é muitas vezes um desafio, principalmente ao considerar a natureza incerta dos ativos de forma dinâmica.

Uma das dificuldades inerentes à otimização sob incerteza é lidar com uma representação discreta dessas incertezas que frequentemente fazem com que o problema de otimização a ser resolvido seja de grande porte. Neste contexto, o método de Galerkin aparece como uma alternativa aos métodos de resolução de problemas estocásticos usuais pois permite a combinação de visões de mercado específicas para cada cenário de forma a reduzir o espaço de soluções viáveis para o problema e, desta forma, reduzir a dimensão do problema de otimização quando comparada com uma formulação multi-estágios. Ainda, ao considerar tais visões de mercado, espera-se encontrar soluções mais realistas para o problema.

Com esta motivação, a presente pesquisa tem por objetivo resolver o problema de alocação ótima de ativos de forma a minimizar o risco de uma carteira de investimentos utilizando o método de Galerkin para a resolução do problema de otimização estocástica associado. Além disso, o trabalho avalia a qualidade das soluções obtidas por este método através de diversos testes numéricos e simulações realizadas com ativos do IBovespa. Para a comparação dos resultados será utilizado um modelo estático de otimização estocástica e funções de desvio para a avaliação do risco da carteira.

## 1.2 Problema Proposto

O presente trabalho de dissertação visa resolver o problema de alocação ótima de ativos a partir do método de Galerkin, considerando as ações do mercado brasileiro que possuem maior participação no IBovespa. O retorno de cada ativo com risco  $i$  no tempo  $t$  será dado pelo retorno unitário randômico  $r_i^t(\xi)$  considerando um período de investimento semanal. Será considerado também um ativo sem risco  $x_0$ , por exemplo, um título do Tesouro Direto com retorno semanal  $r_0 \geq 0$ . O retorno da carteira ao final de cada período de investimento  $t$  será dado por:

$$x_0^t r_0^t + H(x^t(\xi), r^t(\xi)),$$

Onde,

$$H(x^t(\xi), r^t(\xi)) := \sum_{i=1}^N x_i^t(\xi) r_i^t(\xi).$$

Logo, o retorno da carteira é linear e seu valor esperado também é linear. Desta forma, o problema de fronteira eficiente, similar ao problema (1.42) descrito por Shapiro, Dentcheva e Ruszczyński [2] na página 15, tem a seguinte forma:

$$\left\{ \begin{array}{ll} \min_{x \in \mathbb{R}^{N \times T}, x_0 \in \mathbb{R}^T} & \mathcal{D}(H(x(\xi), r(\xi))) & \text{(Para diferentes funções de desvio } \mathcal{D}) \\ \text{s.a.} & x_0^t r_0^t + \mathbb{E}[H(x^t(\xi), r^t(\xi))] \geq \tau^t & t = 1, \dots, T \text{ (Retorno médio mínimo)} \\ & x_0^t + \sum_i x_i^t(\xi) = 1 & t = 1, \dots, T \text{ (Restrição de capital)} \\ & \sum_{i \in P} x_i^t(\xi) \leq \pi & t = 1, \dots, T \text{ (Restrição de capital no setor } P) \\ & x(\xi) \geq 0 & \text{(Restrição de não } \textit{shortselling}) \\ & x_0 \geq 0 & \text{(Restrição de não } \textit{shortselling}), \end{array} \right. \quad (1.2.1)$$

onde,

-  $\xi$  representa o cenário de incerteza associado aos ativos com risco.

-  $x(\xi)$  representa o conjunto de alocações em ativos com risco. Existem  $N$  ativos propostos e  $T$  períodos de tempo. Cada variável  $x_i^t(\xi)$  representa a proporção de capital investido no  $i$ -ésimo ativo no  $t$ -ésimo período de tempo, onde  $0 \leq x_i^t(\xi) \leq 1$ .

-  $r(\xi)$  representa o conjunto de retornos randômicos das variáveis de investimento de risco. Cada variável  $r_i^t(\xi)$  representa o retorno associado ao  $i$ -ésimo ativo no  $t$ -ésimo período de tempo.

-  $\tau$  é o retorno mínimo requerido para a carteira. Cada variável  $\tau^t$  representa o retorno mínimo esperado para a carteira em cada período de tempo  $t$

-  $\pi$  é o percentual máximo permitido para a alocação em ativos de determinado grupo econômico  $P$ .

Desta forma, o problema proposto possui uma restrição de retorno mínimo para a carteira que pode ser configurado pelo investidor. Possui também uma restrição de capital, de forma que não seja possível ficar alavancado. Além disso, o problema contém restrição de capital em um ou mais grupos econômicos também especificados pelo investidor, e restrição de não *shortselling*, ou seja, o investidor não pode ficar vendido nos ativos.

A função objetivo  $\mathcal{D}(H(x(\xi), r(\xi)))$  a ser minimizada respresenta uma função de desvio. Após a fixação da função  $\mathcal{D}(H(x(\xi), r(\xi)))$ , o problema fica parametrizado pelo valor mínimo do retorno esperado  $\tau$  e pelo percentual máximo do capital  $\pi$  que pode ser investido em ativos de determinado grupo econômico  $P$ . As funções de desvio  $\mathcal{D}(H(x(\xi), r(\xi)))$  propostas serão apresentadas na seção (3.3.3).

## 1.3 Contribuições

A principal contribuição do presente trabalho é a apresentação de um método alternativo de programação estocástica que permite a combinação de diversas visões de mercado específicas para cada cenário, representadas pelas funções de base ou estratégias base, de forma a reduzir o espaço de soluções viáveis e conseqüentemente a dimensão do problema de otimização. Desta forma, o método proposto permite que o tempo de processamento para a resolução do problema seja menor comparado com métodos diretos de resolução que podem não ser solucionáveis em tempo computacional aceitável. Além disso, o método proposto traz benefícios em relação à qualidade das soluções obtidas através da incorporação do conhecimento de especialistas de mercado. Portanto, soluções mais realistas podem ser obtidas através deste método.

## 1.4 Estrutura do Trabalho

O presente trabalho está estruturado em seis capítulos seguidos de referências bibliográficas e apêndices. Abaixo segue um breve resumo sobre cada capítulo.

1. **Introdução:** Neste capítulo é apresentado o tema central da pesquisa juntamente com a motivação, a definição do problema proposto e as contribuições do trabalho.
2. **Revisão Bibliográfica:** O capítulo revisa o estado da arte na literatura relacionada com os assuntos abordados na pesquisa. Como o objetivo do trabalho é a implementação e a análise de um modelo de otimização estocástica aplicado ao problema de alocação ótima de ativos, o capítulo apresenta uma revisão dos modelos de programação matemática, métodos de solução e métricas de risco para a avaliação das carteiras.
3. **Metodologia:** Neste capítulo são definidos os ativos que serão utilizados nos experimentos numéricos, assim como o tipo de retorno e o intervalo de dados que serão considerados. Além disso, é descrita a metodologia de Elipsóide de Volume Mínimo para a detecção de *outliers* e o método de simulação de Monte Carlo para a obtenção dos cenários aleatórios. São apresentados também os conceitos de medidas de risco, medidas coerentes de risco e funções de desvio como métricas para a avaliação do risco das carteiras. Por fim, é apresentado o conceito de fronteira eficiente para a comparação das metodologias e são definidos os problemas determinísticos equivalentes que serão confrontados com o método proposto.
4. **Método de Galerkin para Otimização Estocástica:** Neste capítulo é definido o método de Galerkin para otimização estocástica. Para um melhor entendimento do método é apresentado um estudo de caso, que pode ser verificado no artigo base desta pesquisa [1], baseado em um problema de otimização de carteira. O capítulo também apresenta as definições para os especialistas de mercado que serão utilizados e o novo problema de otimização gerado através da aplicação do método proposto, chamado de problema  $\alpha$ .

5. **Experimentos Numéricos:** Neste capítulo são feitos testes de consistência nos modelos especificados com o objetivo de verificar se os métodos foram implementados corretamente. Além disso, o método de Galerkin é avaliado para as funções de desvio propostas e comparado com o problema determinístico equivalente.
6. **Considerações Finais:** O último capítulo apresenta as considerações finais da pesquisa e possíveis extensões do trabalho.
7. **Referências Bibliográficas.**
8. **Apêndice A - Representações Matriciais:** Neste apêndice encontram-se as representações matriciais necessárias para a implementação dos modelos de otimização propostos.
9. **Apêndice B - Códigos Utilizados:** Neste apêndice são disponibilizados todos os códigos do trabalho em linguagem MATLAB. As funções de otimização utilizadas são obtidas através do otimizador Gurobi.

# Capítulo 2

## Revisão Bibliográfica

O objetivo do trabalho é a implementação e a análise de um modelo de otimização estocástica aplicado ao problema de alocação ótima de ativos. Desta forma, foram pesquisadas referências sobre modelos de programação matemática assim como métricas de risco para a avaliação de carteiras.

O problema de alocação ótima de ativos tem sido um dos mais importantes tópicos de pesquisa na área de Finanças. Este problema está relacionado com a atribuição de capital sobre um conjunto de ativos disponíveis. O principal objetivo da seleção de carteiras é selecionar a melhor combinação de ativos que gera os maiores retornos esperados e, ao mesmo tempo, garante um nível aceitável de risco.

O primeiro modelo de seleção de carteiras foi proposto por Markowitz [3] baseado em programação matemática. A programação matemática é uma das áreas de pesquisa que procuram maximizar ou minimizar uma função de muitas variáveis sujeitas a um conjunto de restrições impostas pela natureza do problema a ser estudado e restrições de integralidade em algumas ou em todas as variáveis. O modelo proposto, conhecido como média-variância, assume que o retorno total da carteira pode ser descrito usando a média dos retornos dos ativos e a variância dos retornos entre esses ativos. Para um nível de risco dado, pode-se obter o melhor retorno esperado através de um problema de maximização ou para um retorno específico dado, pode-se obter o menor risco minimizando a variância da carteira.

Conforme visto na revisão literária feita por Mokhtar [4], após a introdução do modelo média-variância, técnicas de programação matemática tem se tornado ferramentas essenciais no suporte às decisões financeiras das empresas e vem sendo usadas cada vez mais. Nos últimos anos, o progresso na área de computação deu origem a novas técnicas de programação matemática para modelar o problema de alocação ótima de ativos. Muitos modelos baseados em programação matemática foram desenvolvidos para resolver os problemas de seleção de carteiras atuais que envolvem um complexo, mas realista, conjunto de restrições.

Para maiores detalhes sobre os modelos resumidos a seguir, Markowitz [3], Konno e Yamazaki [5], Young [6], Papahristodoulou e Dotzauer [7], Kamil e Ibrahim [8], Konno e Yamamoto [9], Chiodi et al. [10], Benati e Rizzi [11], Rockafellar e Uryasev [12], Angelelli et al. [13], Soleimani et al. [14], Golmakani e Fazel [15] e Ibrahim et al. [16].

A primeira aplicação da programação matemática em problemas de otimização de carteira foi proposta por Markowitz [3], na década de 50. O autor sugeriu um modelo de programação quadrática para a seleção de uma carteira diversificada de ações que requer o uso de algoritmos numéricos não-lineares para resolver o problema. A fim de simplificar o modelo de Markowitz, vários autores têm proposto modelos de programação linear através de uma definição diferente de função de desvio. Na década de 90, os autores Konno e Yamazaki [5] propuseram o desvio padrão absoluto em relação à média como uma função de desvio. Tal função de desvio, assim como as demais que serão apresentadas neste capítulo, serão definidas com mais detalhes na Seção 3.3.3. O modelo proposto é contudo equivalente ao modelo de Markowitz quando os retornos possuem uma distribuição normal multivariada. Por outro lado, Young [6] introduziu um modelo que maximiza o rendimento mínimo, conhecido como modelo Maximin, ou minimiza a perda máxima, conhecido como modelo Minimax. De acordo com o autor, a formulação Minimax pode ser um método mais apropriado em relação à formulação da média-variância quando os retornos são log-normalmente distribuídos.

A partir do ano de 2000, vários autores contribuíram com novos modelos para o problema de otimização de carteira. Os autores Papahristodoulou e Dotzauer [7] formularam dois diferentes modelos de programação linear com base em formulações de minimização do desvio médio absoluto e Maximin. Estes modelos foram então comparados com a formulação de programação quadrática clássica para testar até que ponto todas estas formulações fornecem carteiras semelhantes. Os resultados deste estudo demonstraram que a formulação Maximin produz o maior retorno e risco, enquanto a formulação quadrática oferece o menor risco. Além disso, todas as três formulações performaram melhor do que as melhores carteiras de fundos de ações na Suécia.

Outro modelo de programação linear para o problema de alocação ótima de ativos foi proposto por Kamil e Ibrahim [8]. Neste estudo, o problema foi modelado considerando a minimização do risco *downside* que é o risco definido apenas pelas perdas. O efeito da utilização do risco *downside* é analisado neste trabalho através do semi-desvio padrão inferior. A fim de avaliar o desempenho do modelo proposto, os autores compararam os resultados do modelo proposto com os resultados dos modelos média-variância e desvio padrão absoluto. De acordo com os seus resultados, o modelo proposto oferece melhores retornos do que os obtidos através dos modelos média-variância e desvio padrão absoluto.

Konno e Yamamoto [9] formularam um problema de otimização de carteira com custo não-convexo de transação, restrições de unidade mínima de transação e restrições de cardinalidade como um problema de programação não-linear inteira. O objetivo deste estudo é mostrar que esta classe de problemas pode ser resolvida com sucesso por abordagens de *state-of-the-art* para programação inteira se o desvio absoluto é usado como uma função de desvio. Os experimentos computacionais para o problema de tamanho médio utilizando o otimizador CPLEX resultaram em boas soluções com um período de tempo razoável.

Chiodi et al. [10] apresentaram um modelo para o problema da seleção de carteiras de fundos mútuos quando comissões são levadas em consideração. Esse problema foi formulado como um modelo de programação linear inteiro-mista usando o semi desvio médio absoluto. Os autores também sugeriram algumas abordagens heurísticas para resolver o problema de carteira ótima. Os resultados das experiências computacionais mostraram que o problema pode ser resolvido usando heurísticas de forma eficaz e eficiente.

Os autores Benati e Rizzi [11] propuseram um novo problema de otimização de carteira baseado em uma extensão do modelo de Markowitz com a medida de risco *Value at Risk* (VaR) substituindo a variância na função objetivo. Esse problema foi formulado como um modelo de programação linear inteiro-mista. Os autores mostram que o modelo proposto pode ser resolvido usando o otimizador CPLEX em uma quantidade razoável de tempo, se o número de observações passadas ou o número de ativos envolvidos no estudo é pequeno. Embora o VaR seja uma medida de risco muito popular, possui características matemáticas indesejáveis, tais como falta de sub-aditividade e convexidade, ver Artzener [17].

Neste contexto, Rockafellar e Uryasev [12] demonstraram que técnicas de programação linear podem ser utilizadas para problemas de otimização de carteira com respeito a minimização da perda esperada a cima do VaR. Esta medida de risco é conhecida pelos nomes de *Conditional Value at Risk* (CVaR) ou *Average Value at Risk* (AVaR) e é considerada uma medida de risco mais consistente do que o VaR pelo fato de ser uma medida coerente de risco. Este fato pode ser verificado na Seção 3.3.2. Um estudo de caso considerando uma carteira de opções utilizando a técnica de minimização do AVaR pode ser verificado na pesquisa destes autores. Os experimentos numéricos com os onze ativos propostos mostraram que o AVaR pode ser minimizado de maneira eficiente através de técnicas de programação linear.

Dois modelos diferentes de programação linear inteiro-mista foram propostos por Angelelli et al. [13] para resolver o problema de seleção de carteiras que leva em conta os lotes de transação mínimos, custos de transação e restrições de cardinalidade. O primeiro modelo é baseado na minimização do AVaR. O segundo modelo baseia-se na minimização do desvio médio absoluto. Apesar dos experimentos numéricos mostrarem que as carteiras obtidas através da medida AVaR possuem retornos mais estáveis comparadas com o segundo modelo, este problema de otimização requer um enorme tempo de processamento para ser resolvido.

O problema de otimização de carteiras com características reais do mercado financeiro também foi estudado por Soleimani et al. [14]. Os autores estenderam o modelo de média-variância para incluir os lotes mínimos de transação, restrição de cardinalidade e restrição de capitalização por grupo econômico. Ao considerar tais restrições, a modelagem do problema de seleção de carteira requer o uso de técnicas de programação inteiro-mista e desta forma, o modelo é classificado como um modelo de programação não-linear inteiro-mista. Para resolver o modelo, foi utilizado algoritmo genético (AG). Com base nos resultados computacionais, verificou-se que o modelo proposto e a abordagem de solução são aplicáveis e de confiança nos mercados reais com grande número de ações.

Outro estudo realizado por Golmakani e Fazel [15] foi semelhante ao de Soleimani et al. [14], mas eles consideraram limites nas restrições do modelo. Pelo fato do modelo em questão ser classificado como programação quadrática inteira-mista, os autores propuseram uma heurística baseada no método de Otimização por Enxame de Partículas (OEP) para lidar com a complexidade do modelo estendido. Os autores também compararam a sua abordagem com o AG e os resultados computacionais mostraram claramente que o OEP proposto supera efetivamente o AG, especialmente em problemas de grande escala.

Conforme observado por Shapiro [18], um progresso considerável foi feito na área de programação estocástica na última década. Um avanço notável foi o Método de Monte Carlo para a modelagem das incertezas. Neste contexto, as abordagens para a modelagem dos problemas de

programação estocástica tem sido baseadas em métodos de geração de cenários. Desta forma, um finito número, computacionalmente administrável, de cenários é gerado e, conseqüentemente, o problema de otimização construído é resolvido por métodos de tipo de decomposição. Um fato relevante é que ao considerar vários cenários para o problema é certamente melhor do que resolver o problema para apenas um cenário, que seria uma abordagem de otimização determinística. Ibrahim et al. [16] propuseram modelos de programação estocástica de um e dois estágios com o objetivo de minimizar o desvio *downside* máximo do retorno esperado. O objetivo deste estudo é comparar a carteira ótima dos dois modelos. Os resultados mostraram que o modelo com dois estágios supera o modelo de estágio único em ambas as análises, fora e dentro da amostra.

A Tabela 2.1 apresenta o resumo dos estudos que utilizam técnicas de programação matemática no modelo de otimização de carteira.<sup>1</sup>

Autores	Restrições Reais de Mercado	Programação Matemática				
		LP	IP	MIP	MIQP	SP
Golmakani e Fazel [15]	Sim				✓	
Soleimani et al. [14]	Sim				✓	
Ibrahim et al. [16]	Não					✓
Angelelli et al. [13]	Sim			✓		
Kamil e Ibrahim [8]	Não	✓				
Benati e Rizzi [11]	Não			✓		
Konno e Yamamoto [9]	Sim		✓			
Papahristodoulou e Dotzauer [7]	Não	✓				
Chiodi et al. [10]	Sim			✓		
Rockafellar e Uryasev [12]	Não	✓				

Tabela 2.1: Programação matemática aplicada ao problema de otimização de carteira.

A partir desta revisão da literatura dos modelos de programação matemática, podemos entender o contexto do presente trabalho. A proposta deste trabalho de dissertação é resolver o problema de otimização de carteira considerando o desvio padrão como uma abordagem clássica, e o desvio AVaR como uma abordagem mais usual. O problema proposto (1.2.1) não possui restrições de cardinalidade ou lotes de transação mínimos. Portanto, os problemas a serem resolvidos são caracterizados como problemas de otimização estocástica contínua. Desta forma, o problema de otimização de carteira considerando o desvio padrão será tratado como um problema de programação quadrática. Já o problema com o desvio AVaR será tratado como um problema de programação linear. O método de Galerkin, que será utilizado para a resolução do problema proposto, será apresentado no Capítulo 4.

<sup>1</sup>Nota: LP-Programação Linear, IP-Programação Inteira, MIP-Programação Inteira Mista, MIQP-Programação Quadrática Inteira Mista, SP-Programação Estocástica.



# Capítulo 3

## Metodologia

Neste capítulo serão apresentadas as definições e metodologias necessárias para o desenvolvimento dos próximos capítulos. Na Seção 3.1 serão apresentados os dados utilizados para as análises. A Seção 3.2 introduz a metodologia de Elipsóide de Volume Mínimo para a detecção de *outliers* e a metodologia de Simulação de Monte Carlo para a geração dos cenários aleatórios. Na Seção 3.3 serão apresentadas as definições de medida de risco, medida coerente de risco e funções de desvio. As funções de desvio, que serão utilizadas para a avaliação do risco das carteiras, também se encontram nesta seção. A Seção 3.4 apresenta o conceito de fronteira eficiente que será utilizado para a comparação dos métodos de solução. Por fim, a Seção 3.5 define os problemas estáticos que serão utilizados para a comparação com o método proposto.

### 3.1 Descrição dos Dados

Os preços das ações utilizados neste trabalho foram obtidos através do provedor de dados Bloomberg [19]. Foram escolhidas 15 ações negociadas na Bolsa de Valores de São Paulo, BM&FBovespa, que compõem o índice Bovespa e possuem grande volume de negociação diária. Foram examinados os desempenhos históricos dessas ações durante as 768 semanas correspondentes ao período de 03 de Janeiro de 2000 a 17 de Outubro de 2014. O período foi escolhido levando-se em consideração a disponibilidade dos dados de todos os ativos. Este período contempla a crise proveniente do *subprime* em 2008 e a crise na Zona do Euro no final de 2009. Além destes fatos, contempla também períodos de instabilidade política como, por exemplo, a troca de governo em 2002.

Para a análise do desempenho histórico semanal será utilizado o log-retorno, conforme definido a seguir:

$$R_t = \ln\left(\frac{P_t}{P_{t-1}}\right),$$

onde,  $P_t$  é o preço do ativo no tempo  $t$ .

A escolha do log-retorno para a análise do desempenho semanal é devida à sua propriedade de ser invariante de mercado. A definição de Meucci [20] para invariantes de mercado é apresentada a seguir.

**Definição 1.** *As variáveis aleatórias  $X_t$  são **invariantes de mercado**, considerando um tempo de partida  $t$  e um intervalo de tempo  $\tau$ , se elas são independentes e identicamente distribuídas e se uma realização  $x_t$  de  $X_t$  é disponibilizada no tempo  $t$ .*

Conforme verificado por Meucci [20], a representação mais conveniente para os invariantes de um mercado de ações é o log-retorno dos preços por duas razões. A primeira delas é devido ao fato da distribuição dos log-retornos poder ser facilmente projetada para qualquer horizonte de tempo e depois trazidas de volta para a distribuição dos preços de mercado no horizonte de tempo especificado. A segunda razão é que a distribuição dos retornos lineares ou retornos totais não são simétricas<sup>1</sup>. Já os log-retornos têm uma distribuição aproximadamente simétrica. Isto faz com que seja mais fácil modelar a sua distribuição de probabilidade.

O intervalo de tempo semanal foi escolhido pois reflete melhor a periodicidade com que uma empresa planeja a sua estratégia de investimento. Foi utilizado o preço de fechamento da semana de referência e o preço de fechamento da semana anterior para o cálculo do log-retorno semanal.

Na Figura 3.1 são mostrados os preços e os log-retornos semanais dos ativos escolhidos. Na Tabela 3.1 podem ser verificados o grupo econômico, participação no Ibovespa e principais estatísticas desses ativos.

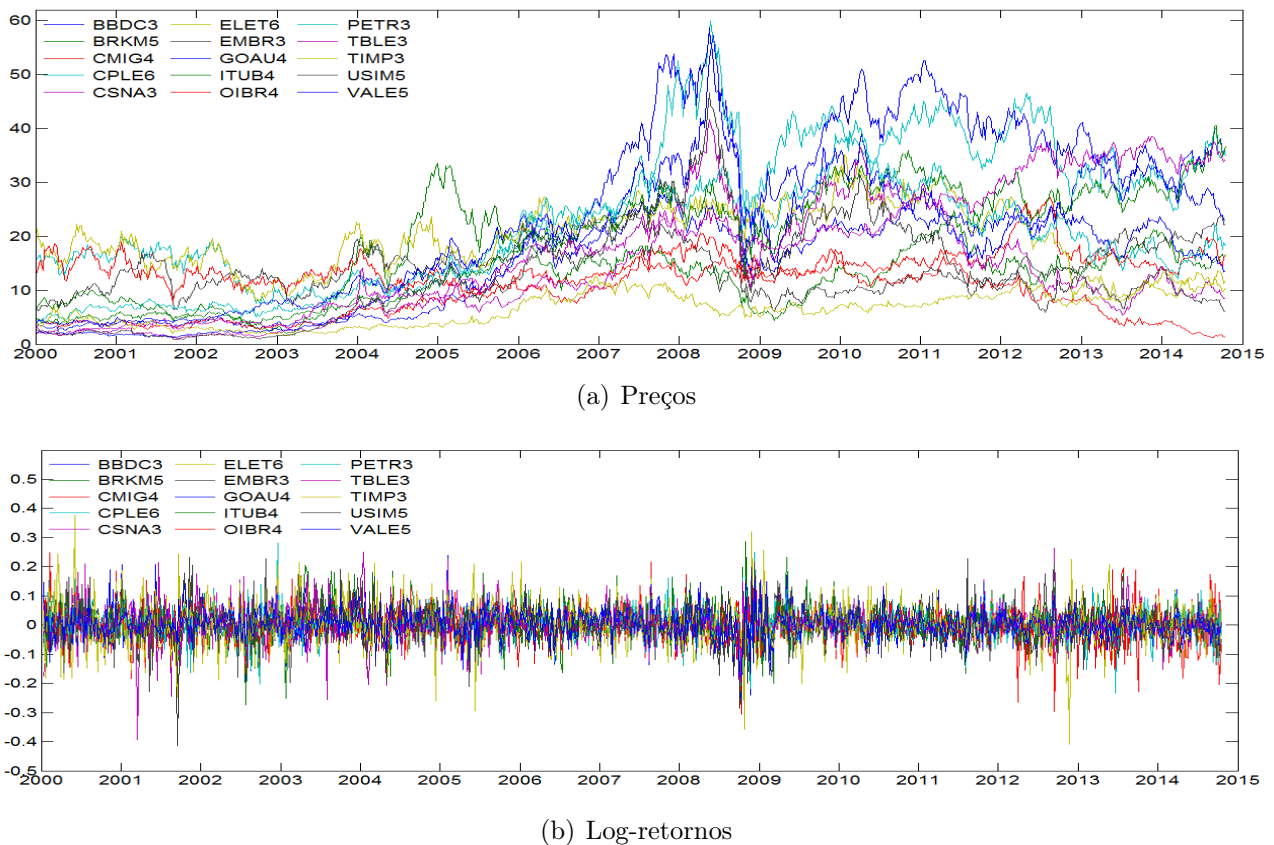


Figura 3.1: Gráficos históricos semanais das 15 ações negociadas na BM&FBovespa.

<sup>1</sup>Os retornos lineares e os retornos totais são definidos, respectivamente, por  $L_t = \frac{P_t}{P_{t-1}} - 1$  e  $H_t = \frac{P_t}{P_{t-1}}$ .

Ação	Empresa	Grupo Econômico	Part. (%)	Média	Desv. Pad.	Assimetria	Curtose	Mínimo	Máximo
BBDC3	BCO BRADESCO S.A.	Bancos	1.756	0.00319	0.05013	0.06451	7.094	-23.03%	26.98%
BRKM5	BRASKEM S.A.	Petroquímico	0.536	0.001183	0.06361	0.08878	4.431	-26.31%	23.31%
CMIG4	CEMIG	Energia Elétrica	1.146	0.001621	0.0511	-0.4421	5.229	-29.62%	16.97%
CPLE6	COPEL	Energia Elétrica	0.353	0.001035	0.05274	-0.2523	5.59	-27.22%	28.13%
CSNA3	CIA SIDERURGICA NACIONAL	Siderurgia	0.377	0.001201	0.06616	-0.3092	5.394	-39.40%	26.41%
ELET6	ELETRONBRAS	Energia Elétrica	0.209	-0.00093	0.063	-0.3731	6.049	-40.98%	22.57%
EMBR3	EMBRAER S.A.	Material de Transporte	1.933	0.001593	0.05512	-0.744	9.081	-41.27%	17.91%
GOAU4	METALURGICA GERDAU S.A.	Siderurgia	0.397	0.002433	0.0573	-0.1251	4.926	-28.59%	26.06%
ITUB4	ITAU UNIBANCO HOLDING S.A.	Bancos	10.385	0.002675	0.05207	-0.2262	7.166	-27.59%	28.63%
OIBR4	OI S.A.	Telecomunicações	0.421	-0.00308	0.06555	0.003443	4.573	-30.63%	24.68%
PETR3	PETROBRAS	Petróleo	3.999	0.001767	0.05333	-0.1452	4.694	-23.83%	24.98%
TBLE3	TRACTEBEL ENERGIA S.A.	Energia Elétrica	0.743	0.003738	0.05363	-0.04908	7.261	-25.99%	25.13%
TIMP3	TIM PARTICIPACOES S.A.	Telecomunicações	1.151	0.001514	0.06917	0.2126	6.575	-35.67%	37.63%
USIM5	USIMINAS	Siderurgia	0.286	0.001088	0.06853	-0.02289	4.19	-28.63%	23.16%
VALE5	VALE S.A.	Mineração	4.165	0.002304	0.04523	-0.1002	4.087	-17.54%	16.65%

Tabela 3.1: Empresa, Grupo Econômico e Participação no Ibovespa obtido através do site da BM&FBovespa [21].

## 3.2 Geração de Cenários

Nesta seção será apresentado o conceito de *outlier* e a importância da sua detecção para a estimação robusta dos parâmetros de localização e dispersão de uma distribuição. Será introduzida a metodologia de Elipsóide de Volume Mínimo para a detecção dos *outliers* e a metodologia de Simulação de Monte Carlo para a geração dos cenários aleatórios que utilizarão os estimadores robustos encontrados.

### 3.2.1 Detecção de Outliers

Um *outlier* pode ser entendido como uma observação inconsistente com os dados. Conforme descrito por Meucci [20], um único *outlier* pode distorcer a estimação dos parâmetros de uma distribuição. Neste contexto, o autor fornece uma intuição por de trás de uma estimação robusta utilizando o exemplo a seguir.

**Exemplo 1.** Considere a média amostral e a covariância amostral definidas, respectivamente, por (3.2.1) e (3.2.2):

$$\hat{\mu} = \frac{1}{T} \sum_{t=1}^T x_t. \quad (3.2.1)$$

$$\hat{\Sigma} = \frac{1}{T} \sum_{t=1}^T (x_t - \hat{\mu})(x_t - \hat{\mu})'. \quad (3.2.2)$$

Seja o elipsóide de localização-dispersão definido pela média amostral e covariância amostral de um conjunto de observações de invariantes de mercado:

$$\varepsilon_{\hat{\mu}, \hat{\Sigma}} = \{x \in \mathbb{R} \text{ tal que } (x - \hat{\mu})' \hat{\Sigma}^{-1} (x - \hat{\mu}) \leq 1\}. \quad (3.2.3)$$

Em seguida, adiciona-se uma observação discrepante, um *outlier*, e repete-se a estimação considerando a amostra ampliada. O novo elipsóide, que representa a nova média amostral e covariância amostral, é completamente diferente. Este fato pode ser verificado na Figura 3.2.

□

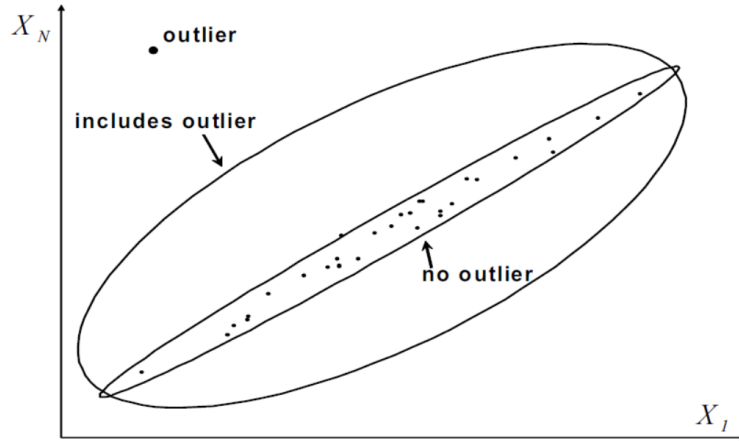


Figura 3.2: Estimadores amostrais: Falta de robustez.

Desta forma, o conceito de robustez está relacionado com a capacidade de um estimador de proporcionar estimativas razoáveis quando observações atípicas são encontradas na amostra. No experimento acima, podemos observar que os estimadores amostrais não são robustos. Porém, sabemos que a observação adicional é espúria. Em muitos casos, não sabemos a verdadeira distribuição subjacente e, mais importante, não temos nenhuma razão para acreditar que algumas observações podem ser espúrias. Neste sentido, uma questão interessante é: Qual a quantidade máxima de *outliers* que um certo estimador pode sustentar antes de quebrar? Ou seja, se houver um total de  $T = T_B + T_O$  observações, onde  $T_B$  são dados bons e  $T_O$  valores discrepantes, qual é a maior razão  $T_O/T$  que o estimador pode sustentar?

**Definição 2.** O *ponto de ruptura* de um estimador é o limite da razão  $T_O/T$ , quando o número de observações tende ao infinito.

Por definição, o ponto de ruptura é um número positivo que não pode ser superior a 0,5.

**Definição 3.** O estimador cujo o ponto de ruptura é próximo de 0,5 é chamado de *estimador de alta ruptura*.

Os estimadores de alta ruptura são úteis em aplicações financeiras porque eles nos permitem detectar os *outliers*. Um dos métodos para a construção de estimadores de alta ruptura de localização e dispersão é o Elipsóide de Volume Mínimo (EVM).

### 3.2.2 Elipsóide de Volume Mínimo

Suponha que sabemos que de um total de  $T$  observações,  $T_B$  são observações boas e  $T_O$  são observações discrepantes. Conforme descrito por Meucci [20], o menor elipsóide que contém os  $T_B$  dados bons é o menor entre os elipsóides que contém qualquer conjunto das  $T_O$  observações. A Figura 3.3 ilustra esta afirmação.

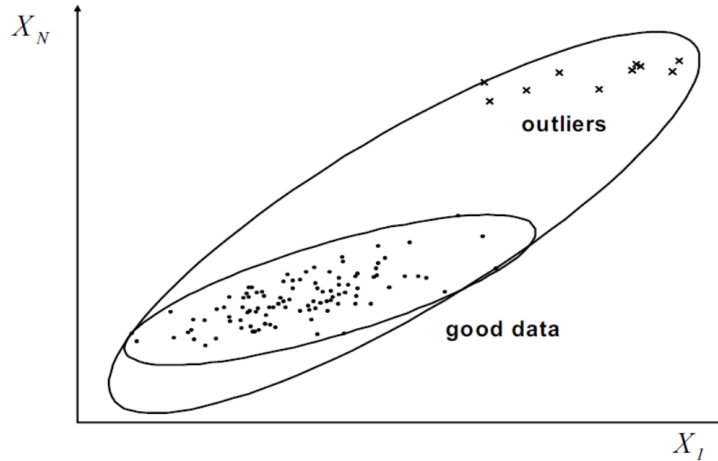


Figura 3.3: Elipsóide de Volume Mínimo.

Seja  $\mu$  um parâmetro de localização genérica, ou seja, um vetor  $N$ -dimensional, e uma matriz  $\Sigma$  de dispersão genérica, ou seja, uma matriz  $N \times N$  simétrica e positiva. Os parâmetros  $(\mu, \Sigma)$  definem um elipsóide como em (3.2.3). Podemos expandir este elipsóide da seguinte forma:

$$\varepsilon_{\mu, \Sigma}^q = \{x \in \mathbb{R}^N \text{ tal que } (x - \mu)' \Sigma^{-1} (x - \mu) \leq q^2\}. \quad (3.2.4)$$

Este lugar geométrico representa uma versão redimensionada do elipsóide original, onde todos os eixos são multiplicados por um fator  $q$ . O volume do elipsóide definido em (3.2.4) é dado por:

$$\text{Vol}\{\varepsilon_{\mu, \Sigma}^q\} = \gamma_N q^N \sqrt{|\Sigma|}, \quad (3.2.5)$$

onde,  $\gamma_N$  é o volume da esfera unitária definida por:

$$\gamma_N = \frac{\pi^{\frac{N}{2}}}{\Gamma(\frac{N}{2} + 1)}.$$

Para generalizar o valor absoluto em um ambiente multivariado, o autor introduz o conceito de Distância de Mahalanobis:

**Definição 4.** A *distância de Mahalanobis* de um ponto  $x$  ao ponto  $\mu$  através da métrica  $\Sigma$  é definido por  $Ma(x, \mu, \Sigma) \equiv \sqrt{(x - \mu)' \Sigma^{-1} (x - \mu)}$ , onde a métrica  $\Sigma$  é uma matriz simétrica e positiva.

Os pontos  $x$  que compartilham a mesma distância de Mahalanobis de  $\mu$  pertencem à superfície de um elipsóide centrado em  $\mu$ . Quanto maior os autovalores de  $\Sigma$ , menor é a distância de Mahalanobis de um ponto genérico  $x$  ao centro  $\mu$ . Portanto, a matriz  $\Sigma$  de fato fornece uma métrica para medir distâncias.

Considere o conjunto de distâncias de Mahalanobis de cada observação ao parâmetro de localização  $\mu$  através da métrica  $\Sigma$ :

$$\text{Ma}_t^{\mu, \Sigma} \equiv \text{Ma}(x_t, \mu, \Sigma) \equiv \sqrt{(x_t - \mu)' \Sigma^{-1} (x_t - \mu)}.$$

Podemos ordenar essas distâncias de forma crescente e considerar a  $T_B$ -ésima distância:

$$q_{T_B} \equiv \text{Ma}_{T_B:T}^{\mu, \Sigma}.$$

Por construção, o elipsóide  $\varepsilon_{\mu, \Sigma}^{q_{T_B}}$  contém apenas os  $T_B$  pontos e a partir de (3.2.5), o seu volume é dado por:

$$\text{Vol}\{\varepsilon_{\mu, \Sigma}^{q_{T_B}}\} = \gamma_N(\text{Ma}_{T_B:T}^{\mu, \Sigma})^N \sqrt{|\Sigma|}. \quad (3.2.6)$$

Observe que o produto do lado direito desta expressão não depende do determinante de  $\Sigma$ . Portanto, podemos impor a restrição de que o determinante de  $\Sigma$  é igual a 1.

Consequentemente, os parâmetros que dão origem ao menor elipsóide que contém as  $T_B$  observações são obtidos através da seguinte equação:

$$(\hat{\mu}_{T_B}, \hat{\Sigma}_{T_B}) = \underset{\mu, \Sigma \succeq 0, |\Sigma|=1}{\text{argmin}} \{ \text{Ma}_{T_B:T}^{\mu, \Sigma} \}, \quad (3.2.7)$$

onde a notação  $\Sigma \succ 0$  implica que  $\Sigma$  é simétrica e positiva. Uma vez que tenhamos calculado os parâmetros (3.2.7), podemos considerar como *outliers* todas as observações que não estão contidas no elipsóide (3.2.4) que será determinado por (3.2.7).

Na realidade não sabemos, a priori, o número verdadeiro  $T_B$  de dados bons. No entanto, se  $T_B$  é o maior conjunto de dados bons, o elipsóide de volume mínimo que contém  $T_B + 1$  observações tem um volume muito maior do que o elipsóide de volume mínimo que contém as  $T_B$  observações. Por esta razão, considera-se o volume do elipsóide de volume mínimo uma função do número de observações contidas no elipsóide:

$$T_B \rightarrow \gamma_N(\text{Ma}_{T_B:T}^{\hat{\mu}_{T_B}, \hat{\Sigma}_{T_B}})^N. \quad (3.2.8)$$

O verdadeiro número de dados bons é o valor  $T_B$  onde essa função exibe um salto abrupto, como pode ser verificado na Figura 3.4. Esses dados bons são utilizados para os cálculos das estimativas finais dos parâmetros de localização e dispersão, respectivamente, a média e a matriz de covariância.

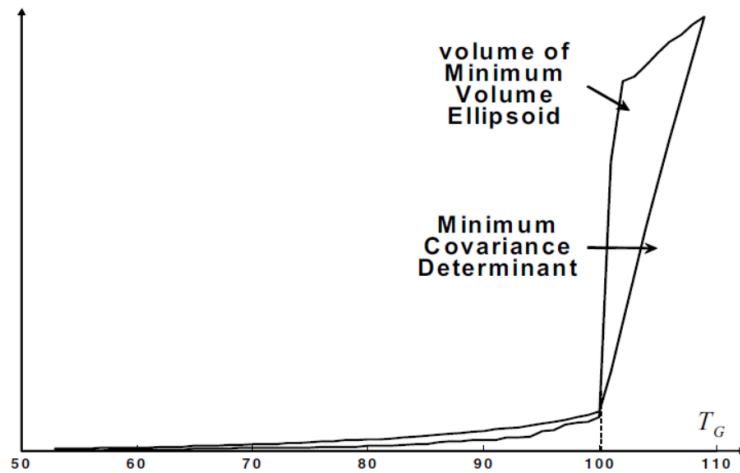


Figura 3.4: Detecção de Outliers.

Para a resolução computacional do problema, foram utilizados os códigos disponibilizados pelo autor através do site Advanced Risk and Portfolio Management, seção Multivariate Estimation (Non-Parametric, MLE, Shrinkage, Robust, ...) [22].

Na Figura 3.5 a seguir é mostrado o gráfico do elipsóide de volume mínimo considerando alguns pares de ativos e a retirada de 10 observações espúrias.

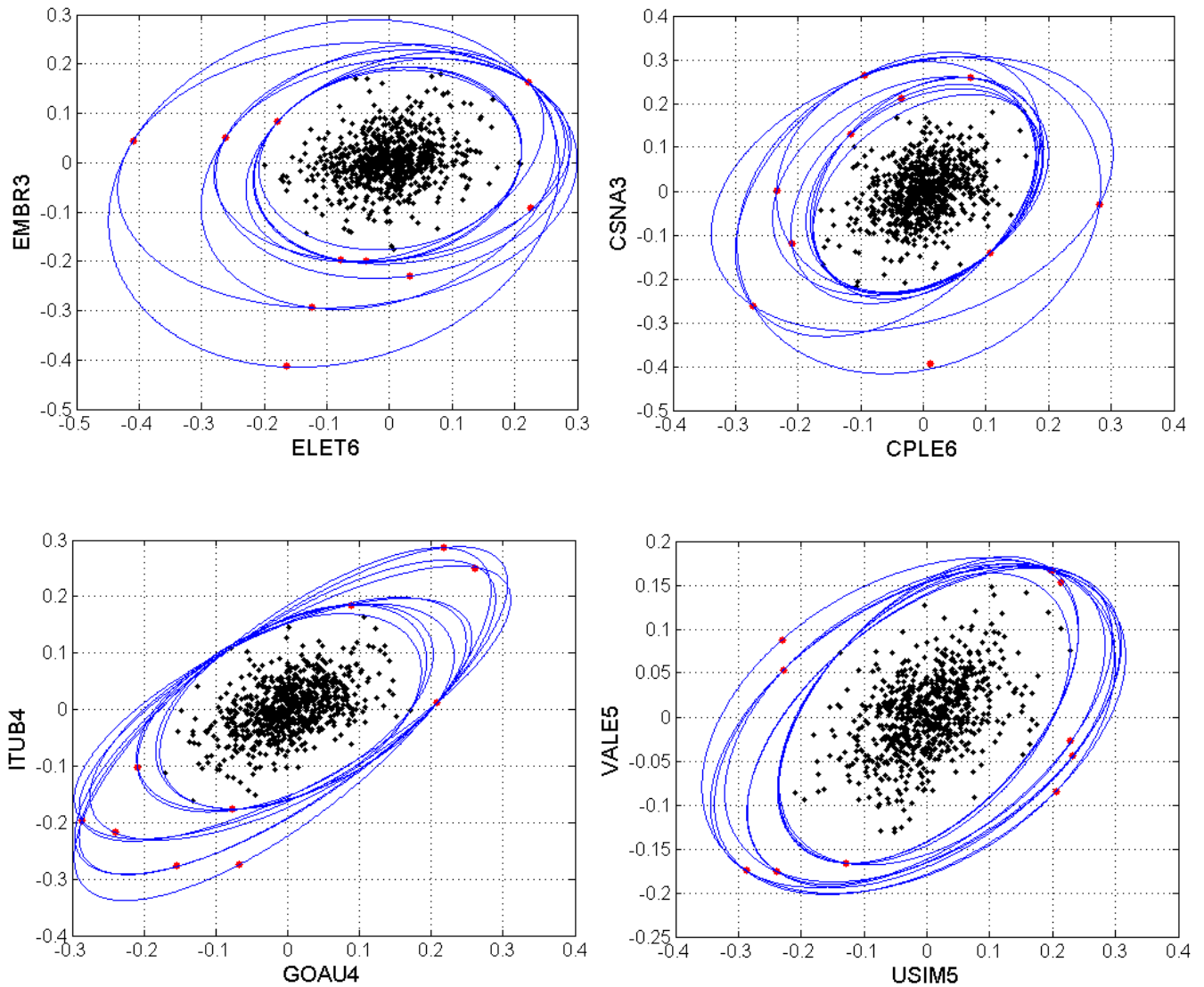


Figura 3.5: Elipsóide de Volume Mínimo.

Nos gráficos acima pode-se observar a variação dos elipsóides de volume mínimo considerando as observações extremas. O gráfico do volume do elipsóide formado pelos 15 ativos é apresentado na Figura 3.6 abaixo.

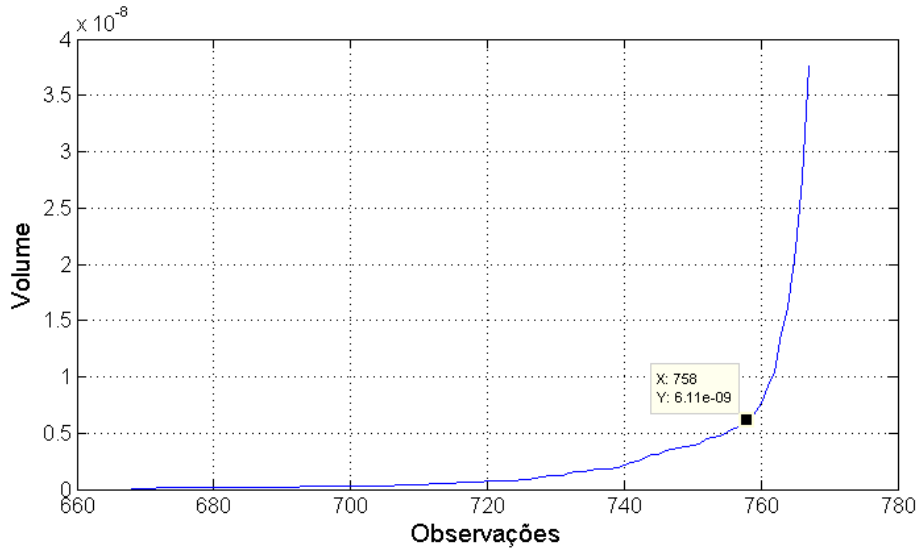


Figura 3.6: Volume do elipsóide como função do número de observações.

A partir do ponto assinalado pode ser observada uma variação abrupta no volume do elipsóide. Desta forma, foi escolhida a remoção das 10 observações responsáveis por este aumento abrupto no volume do elipsóide considerando um total de 768 observações. Este número parece razoável pois representa apenas 1,3% das observações fazendo com que a estimativa dos parâmetros de média e covariância não seja afetada devido à falta de informação.

### 3.2.3 Estimadores Robustos

Conforme verificado através do Exemplo 1, a média e a matriz de covariância amostrais, definidas respectivamente por (3.2.1) e (3.2.2), não são estimadores robustos para esses parâmetros. Porém, utilizando as observações sem *outliers*, a média e a matriz de covariância amostrais são pouco afetadas pela inclusão de novos dados. Sendo, portanto, medidas robustas desses parâmetros. Desta forma, foram escolhidos os estimadores amostrais para a média e a variância considerando a amostra de 758 observações sem *outliers*. Os valores encontrados para essas estimativas são mostrados a seguir:

$$\hat{\Sigma} = 10^{-3} \begin{bmatrix} 2.2 & 1.2 & 1.2 & 1.2 & 1.4 & 1.3 & 0.8 & 1.3 & 1.8 & 1.2 & 1.2 & 0.5 & 1.2 & 1.4 & 0.8 \\ 1.2 & 3.9 & 1.2 & 1.3 & 1.3 & 1.5 & 0.9 & 1.4 & 1.3 & 1.4 & 1.0 & 0.8 & 1.2 & 1.8 & 0.7 \\ 1.2 & 1.2 & 2.4 & 1.6 & 0.9 & 1.7 & 0.5 & 0.9 & 1.2 & 1.2 & 0.9 & 0.7 & 1.0 & 1.2 & 0.5 \\ 1.2 & 1.3 & 1.6 & 2.6 & 1.0 & 1.9 & 0.5 & 1.0 & 1.2 & 1.3 & 1.0 & 0.7 & 1.2 & 1.3 & 0.4 \\ 1.4 & 1.3 & 0.9 & 1.0 & 3.8 & 1.3 & 0.9 & 2.1 & 1.5 & 1.2 & 1.6 & 0.5 & 1.2 & 2.8 & 1.6 \\ 1.3 & 1.5 & 1.7 & 1.9 & 1.3 & 3.7 & 0.7 & 1.2 & 1.4 & 1.4 & 1.1 & 0.8 & 1.4 & 1.7 & 0.6 \\ 0.8 & 0.9 & 0.5 & 0.5 & 0.9 & 0.7 & 2.6 & 0.8 & 0.8 & 0.7 & 0.7 & 0.3 & 0.8 & 0.8 & 0.6 \\ 1.3 & 1.4 & 0.9 & 1.0 & 2.1 & 1.2 & 0.8 & 3.1 & 1.4 & 1.2 & 1.3 & 0.6 & 1.3 & 2.3 & 1.3 \\ 1.8 & 1.3 & 1.2 & 1.2 & 1.5 & 1.4 & 0.8 & 1.4 & 2.5 & 1.2 & 1.4 & 0.6 & 1.3 & 1.5 & 0.9 \\ 1.2 & 1.4 & 1.2 & 1.3 & 1.2 & 1.4 & 0.7 & 1.2 & 1.2 & 4.1 & 0.9 & 0.8 & 1.7 & 1.4 & 0.6 \\ 1.2 & 1.0 & 0.9 & 1.0 & 1.6 & 1.1 & 0.7 & 1.3 & 1.4 & 0.9 & 2.7 & 0.4 & 1.1 & 1.5 & 1.1 \\ 0.5 & 0.8 & 0.7 & 0.7 & 0.5 & 0.8 & 0.3 & 0.6 & 1.8 & 0.8 & 0.4 & 2.6 & 0.7 & 0.8 & 0.3 \\ 1.2 & 1.2 & 1.0 & 1.2 & 1.2 & 1.4 & 0.8 & 1.3 & 1.8 & 1.7 & 1.1 & 0.7 & 4.5 & 1.4 & 0.7 \\ 1.4 & 1.8 & 1.2 & 1.3 & 2.8 & 1.7 & 0.8 & 2.3 & 1.8 & 1.4 & 1.5 & 0.8 & 1.4 & 4.4 & 1.3 \\ 0.8 & 0.7 & 0.5 & 0.4 & 1.6 & 0.6 & 0.6 & 1.3 & 1.2 & 0.6 & 1.1 & 0.3 & 0.7 & 1.3 & 1.9 \end{bmatrix}$$



$$\hat{\mu} = 10^{-3} [2.8 \ 1.7 \ 2.5 \ 1.6 \ 1.9 \ 0.0 \ 3.0 \ 2.5 \ 2.4 \ -3.0 \ 1.9 \ 3.9 \ 1.1 \ 1.5 \ 2.1]^T$$

### 3.2.4 Método de Monte Carlo

O Método de Monte Carlo é utilizado em finanças e matemática financeira para avaliar e analisar instrumentos, carteiras e investimentos através da simulação das várias fontes de incerteza que afetam o seu valor e, em seguida, determinar o seu valor médio em toda a gama de valores resultantes. Isso geralmente é feito com a ajuda de modelos estocásticos de ativos. A vantagem do método de Monte Carlo sobre outras técnicas aumenta à medida que as dimensões do problema, fontes de incerteza, e o período especificado aumentam.

Conforme descrito por Pachamanova e Fabozzi [23], análises de risco são baseadas na modelagem de fontes de incerteza, e incerteza pode ser representada matematicamente por distribuições de probabilidade. As distribuições de probabilidade das fontes de incerteza garantem que esses ativos podem assumir diferentes valores com diferentes probabilidades. Essas distribuições de probabilidade são os blocos de construção para modelos de simulação, como o Método de Monte Carlo. Como os ativos utilizados neste trabalho são ações, as fontes de incerteza são os preços ou retornos dessas ações. Desta forma, foi assumido que os retornos das ações seguem uma distribuição Normal com os parâmetros estimados na Seção 3.2.3.

Na figura 3.7 são mostrados os gráficos das simulações para as ações TBLE3 e VALE5 em um horizonte de 4 semanas. Foram consideradas 1.000 simulações para cada ativo. O código em MATLAB utilizado para a geração das simulações pode ser consultado no Apêndice B deste trabalho.

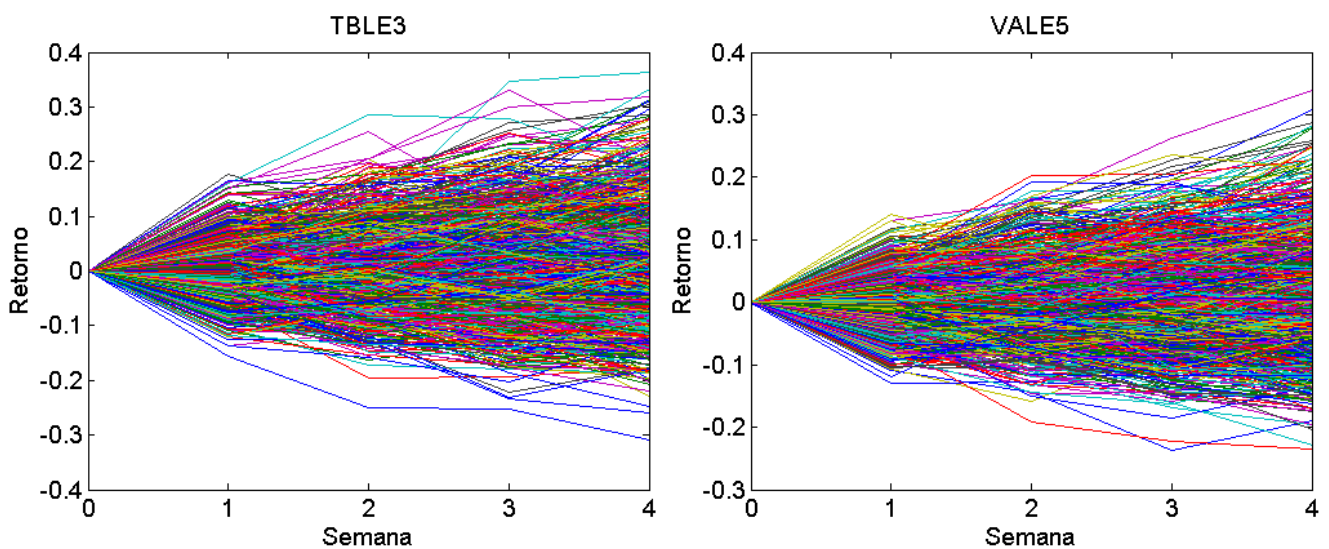


Figura 3.7: Simulação de Monte Carlo.

## 3.3 Medidas de Risco

### 3.3.1 Definição

Conforme descrito por McNeil, Frey e Paul [24], o conceito de risco financeiro pode ser entendido como qualquer evento ou ação que possa afetar adversamente a capacidade de uma organização de alcançar seus objetivos e executar suas estratégias. No contexto de risco de mercado, esses eventos estão associados às variações no câmbio, taxa de juros, preços das ações e preços de *commodities*. Uma variação muito grande ou inesperada de um desses fatores pode levar a organização a grandes perdas, ou até mesmo à falência. A definição matemática para risco pode ser encontrada em Kerkhof e Melenberg [25].

**Definição 5.** *Seja um espaço de probabilidade  $(\Omega, \mathcal{F}, \mathbb{P})$ , onde  $\Omega$  é o conjunto dos possíveis estados da economia,  $\mathcal{F}$  é uma  $\sigma$ -álgebra e  $\mathbb{P}$  é uma medida de probabilidade. Um **risco**  $X(\omega)$ ,  $\omega \in \Omega$  é uma variável aleatória real definida em  $\Omega$ .*

A partir da definição (5), podemos entender os log-retornos de ativos ou carteiras em determinado tempo  $t$  como riscos pois são variáveis aleatórias que dependem do estado da economia. Neste contexto, Kerkhof e Melenberg [25] também definem medida de risco.

**Definição 6.** *Seja um espaço de probabilidade  $m = (\Omega, \mathcal{F}, \mathbb{P})$  e seja  $\mathcal{M}(m)$  o conjunto dos possíveis riscos em  $\Omega$ . Uma **medida de risco** é uma função  $\rho : \mathcal{M}(m) \rightarrow \mathbb{R} \cup \{\infty\}$ .*

Desta forma, uma medida de risco é uma função que associa conjuntos de riscos a valores reais.

O *Value at Risk* (VaR) é uma medida de risco que resume, em um simples e facilmente compreensível número, o risco de uma instituição proveniente da flutuação dos preços no mercado financeiro. Essa é a razão para a sua rápida ascensão como ferramenta essencial para o gerenciamento de risco, levando os Comitês de Basileia I e II a autorizarem a sua utilização nos modelos internos das instituições financeiras.

Outra medida de risco bastante utilizada pelas instituições financeiras é o *Average Value at Risk* (AVaR), que também é conhecido pelos nomes *Conditional Value at Risk* (CVaR), *Tail Value at Risk* (TVaR) e *Expected Shortfall*. Essa medida de risco é definida em função do VaR, sendo obtida a partir da média das observações menores que o correspondente VaR estimado.

Abaixo segue a definição dos autores Rockafellar e Uryasev [26] para o VaR e o AVaR.

**Definição 7.** *O  $\mathbf{VaR}_\alpha$  de uma carteira com respeito a um nível de probabilidade especificado  $\alpha$  é o menor valor  $a^*$  tal que, com probabilidade  $\alpha$ , a perda não excederá  $a^*$ , enquanto o  $\mathbf{AVaR}_\alpha$  é a esperança condicional das perdas acima deste valor  $a^*$ .<sup>2</sup>*

Em termos práticos, o VaR de uma carteira é o quantil de nível  $(1 - \alpha)$  da função de distribuição dos retornos da carteira. A definição (7) garante que o VaR nunca é maior do que o AVaR fazendo com que o AVaR seja uma medida de risco mais conservadora. O VaR e o AVaR de uma carteira podem ser verificados na Figura 3.8.

<sup>2</sup>Para valores de  $\alpha$  tipicamente no intervalo de  $[0.9, 0.99]$ .

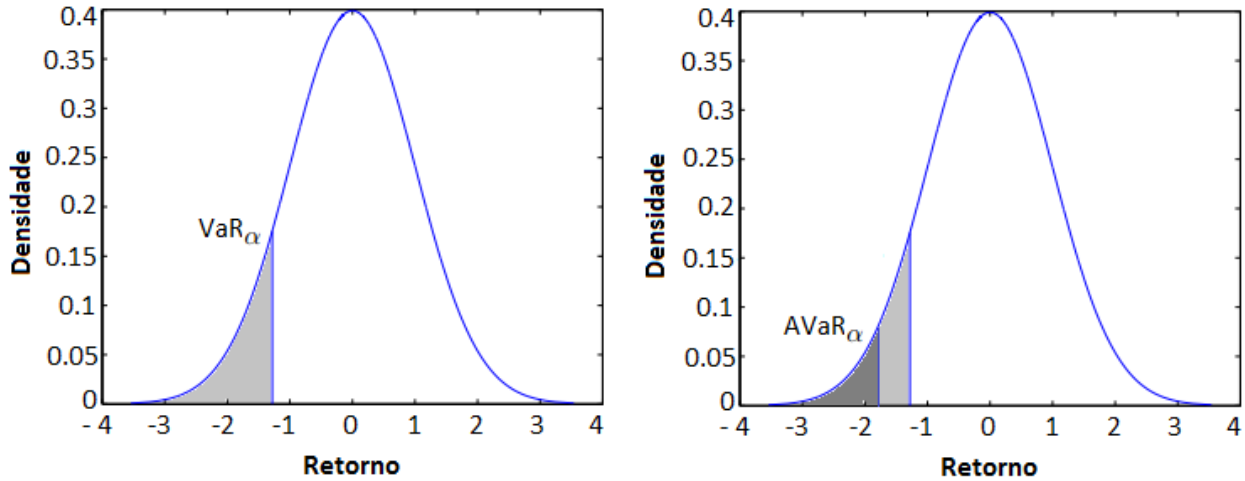


Figura 3.8: VaR e AVaR de uma carteira.

**Exemplo 2.** Considerando que um ativo  $X$  possui distribuição Normal com parâmetros  $\mu$  e  $\sigma$  para os retornos, podemos calcular o  $VaR_\alpha$  e o  $AVaR_\alpha$  do ativo para 1 período da seguinte forma:

$$VaR_\alpha(X) = -\mu - \sigma\Phi^{-1}(\beta)$$

$$AVaR_\alpha(X) = -\frac{1}{\beta} \int_{-\infty}^{\Phi^{-1}(\beta)} \frac{x}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = -\mu - \sigma \frac{\phi(\Phi^{-1}(\beta))}{\beta},$$

onde,  $\beta = 1 - \alpha$ . Além disso,  $\Phi(\cdot)$  é a função de distribuição de probabilidade e  $\phi(\cdot)$  é a distribuição de densidade de probabilidade da Normal Padrão.

**Exemplo 3.** Considerando que um ativo  $X$  possui distribuição T-Student com parâmetros  $\nu$ ,  $\mu$  e  $\sigma$  para os retornos, podemos calcular o  $VaR_\alpha$  e o  $AVaR_\alpha$  do ativo para 1 período da seguinte forma:

$$VaR_\alpha(X) = -\mu - \sigma t_\nu^{-1}(\beta)$$

$$AVaR_\alpha(X) = -\mu - \sigma \frac{g_\nu(t_\nu^{-1}(\beta))}{\beta} \left( \frac{\nu + (t_\nu^{-1}(\beta))^2}{\nu - 1} \right),$$

onde,  $\beta = 1 - \alpha$ . Além disso,  $t_\nu(\cdot)$  e  $g_\nu(\cdot)$  são, respectivamente, as funções de distribuição e densidade de probabilidade da T-Student com  $\nu$  graus de liberdade.

### 3.3.2 Medidas Coerentes de Risco

Pelo fato da definição (6) de medida de risco ser bastante abrangente, torna-se necessária a análise de algumas propriedades para a escolha adequada de uma medida de risco. Em fins da década de 90, Artzner [17] define o conceito de medidas coerentes de risco.

**Definição 8.** Uma *medida de risco* é dita *coerente* quando satisfaz as seguintes propriedades: *Invariância de Translação*, *Sub-aditividade*, *Homogeneidade Positiva* e *Monotonicidade*.

Essas propriedades juntamente com suas interpretações são apresentadas a seguir.

1. **Invariância de Translação:**  $X \in \mathcal{M}(m), k \in \mathbb{R} \Rightarrow \rho(X + k) = \rho(X) - k$   
Adicionar ativos livres de risco, cujo valor é  $k$ , à carteira  $X$  é equivalente à translação da distribuição das perdas em  $-k$  e portanto o risco será reduzido exatamente neste montante.
2. **Sub-aditividade:**  $X, Y, X + Y \in \mathcal{M}(m) \Rightarrow \rho(X + Y) \leq \rho(X) + \rho(Y)$   
A sub-aditividade reflete o efeito da diversificação das carteiras, ou seja, o risco de duas carteiras quando avaliadas conjuntamente deve ser no máximo igual à soma do risco individual de cada carteira, podendo ser menor. Desta forma, a diversificação tem o poder de reduzir o risco de uma carteira.
3. **Homogeneidade Positiva:**  $X \in \mathcal{M}(m), \lambda \in \mathbb{R} \Rightarrow \rho(\lambda X) = \lambda \rho(X)$   
O risco de uma carteira multiplicado por um escalar, um valor monetário, deve ser igual à multiplicação do escalar pelo risco individual da carteira.
4. **Monotonicidade:**  $X, Y \in \mathcal{M}(m), X \leq Y \Rightarrow \rho(X) \geq \rho(Y)$   
Se uma carteira  $X$  tem sempre retornos menores do que  $Y$  então o risco de  $X$  deve ser maior que o risco de  $Y$ .

Conforme demonstrado por Artzner [17], o AVaR é uma medida coerente de risco. Já o VaR é monotônico, positivamente homogêneo e invariante à translação mas não é sub-aditivo e, portanto, não é uma medida coerente de risco. Em termos práticos, este fato implica que a medida em questão não favorece a diversificação do portfólio.

### 3.3.3 Funções de Desvio

Em Rockafellar, Uryasev e Stanislav [12] um novo conceito de medida é definido. Esta nova medida é chamada de medida de desvio e referenciada em Pflug e Römisch [27] por funções de desvio através da seguinte definição:

**Definição 9.** *Uma função  $\mathcal{D}$  definida num espaço linear  $\mathcal{Y}$  de variáveis aleatórias em  $(\Omega, \mathcal{F}, \mathbb{P})$  é chamada de **função de desvio** se para todo  $Y \in \mathcal{Y}$  as propriedades de Invariância de Translação, Convexidade e Monotonicidade são satisfeitas.*

As propriedades de Invariância de Translação e Monotonicidade foram apresentadas na Definição 8. Já a propriedade de Convexidade pode ser entendida como a combinação das propriedades de Sub-aditividade e Homogeneidade Positiva, como pode ser verificado abaixo:

5. **Convexidade:**  $Y_1, Y_2 \in \mathcal{Y}, 0 \leq \lambda \leq 1 \Rightarrow \mathcal{D}(\lambda Y_1 + (1 - \lambda)Y_2) \leq \lambda \mathcal{D}(Y_1) + (1 - \lambda)\mathcal{D}(Y_2)$

**Definição 10.** *Uma função de desvio  $\mathcal{D}$  é chamada de **Positivamente Homogênea** se satisfaz:*

$$\mathcal{D}(\lambda Y) = \lambda \mathcal{D}(Y), \text{ para todo } \lambda \geq 0.$$

**Definição 11.** *Uma função de desvio  $\mathcal{D}$  é chamada de **Estrita** se satisfaz:*

$$\mathcal{D}(Y) \geq 0.$$

Uma função de desvio pode ser escrita a partir de uma medida de risco através da seguinte relação:  $\mathcal{D}(Y) = E(Y) - \rho(Y)$ . Desde que a medida de risco  $\rho$  seja coerente.

**Exemplo 4.** Exemplos de funções de desvio são o desvio padrão e o desvio *Average Value at Risk* que serão apresentadas a seguir. Já a função  $VaRD_\alpha(Y) = E(Y) - VaR_\alpha(Y)$  não é uma função de desvio no sentido da Definição 9 pelo fato da medida de risco  $VaR_\alpha(Y)$  não ter a propriedade de Sub-aditividade implicando a não convexidade desta medida de risco.

Outras funções de desvio usuais são:

- O Semi-Desvio Padrão Inferior:

$$\mathcal{D}(Y) := \sqrt{\mathbb{E}([Y - \mathbb{E}(Y)]_-)^2},$$

onde,  $[a]_- = -\min(a, 0)$ .

- O Semi-Desvio Padrão Superior:

$$\mathcal{D}(Y) := \sqrt{\mathbb{E}([Y - \mathbb{E}(Y)]_+)^2},$$

onde,  $[a]_+ = \max(a, 0)$ .

- O Desvio Médio Absoluto:

$$\mathcal{D}(Y) := \sqrt{\mathbb{E}|Y - \mathbb{E}(Y)|}$$

- O Desvio de Perda Mínima Quadrática:

$$\mathcal{D}(Y) := \min_{a \in \mathfrak{R}} \{\mathbb{E}([Y - a]_+^2) + c\mathbb{E}([Y - a]_-^2)\},$$

para um coeficiente  $c \neq 1$ . Neste caso, os desvios quadráticos positivos e negativos são ponderados de maneira diferente.

As funções de desvio  $\mathcal{D}$  que serão consideradas no problema proposto (1.2.1) são apresentadas com mais detalhes a seguir.

### 3.3.3.1 Desvio Padrão

Uma função de risco muito importante e amplamente utilizada é o Desvio Padrão definido pela raiz quadrada da variância:

$$\mathcal{D}(Y) := \sqrt{\mathbb{E}(Y - \mathbb{E}(Y))^2}$$

Para  $T$  períodos de tempo, consideraremos um desvio da forma:

$$\mathcal{D}(Y) = \sum_{t=1}^T \lambda_t \sqrt{\mathbb{E}(Y^t - \mathbb{E}(Y^t))^2}, \quad (3.3.1)$$

onde,  $\lambda_t \in [0, 1]$  é uma variável de controle de risco do tempo  $t$ . Por exemplo,  $\lambda_1 = 1$  e  $\lambda_{2, \dots, T} = 0$ , controlariam o risco do primeiro passo de tempo. Já  $\lambda_T = 1$  e  $\lambda_{1, \dots, T-1} = 0$  poderiam ser usados para controlar o risco no final do horizonte. Se o investidor necessita revisar a estratégia a cada passo de tempo, convém utilizar  $\lambda_t = 1$  para todo  $t$ . Em relação à dinâmica do risco, porém, empregar desvios ou medidas de risco simultaneamente para diferentes passos de tempo, pode não atender à condição chamada de *time consistency*.

### 3.3.3.2 Desvio *Average Value at Risk*

A variância penaliza desvios da média tanto por baixo quanto por acima. Porém uma carteira que varia acima da média, pode ser atrativa para se obter rendimentos maiores e não deveria ser penalizada. O objetivo do desvio *Average Value at Risk* é penalizar desvios unilaterais e para o problema proposto (1.2.1) é definido pela expressão:

$$\mathcal{D}(Y) := \mathbb{E}(Y) - AVaR_\alpha(Y) = \mathbb{E}(Y) - \max_{a \in \mathbb{R}} \left\{ a - \frac{1}{(1-\alpha)} \mathbb{E}([Y - a]^-) \right\}, \quad (3.3.2)$$

para valores do parâmetro  $\alpha$  tipicamente no intervalo de  $[0.9, 0.99]$ . A igualdade acima pode ser verificada em Pflüg e Römisch [27] através do teorema 2.34 da página 49.

Desta forma, temos um problema de otimização com variável de decisão  $a \in \mathbb{R}$ . Podemos escrevê-lo da seguinte forma:

$$\begin{aligned} AVaR_\alpha(Y) &= \max_{a \in \mathbb{R}} \left\{ a - \frac{1}{(1-\alpha)} \mathbb{E}([Y - a]^-) \right\} \\ &= - \min_{a \in \mathbb{R}} \left\{ -a + \frac{1}{(1-\alpha)} \mathbb{E}([Y - a]^-) \right\} \\ &= - \min_{a \in \mathbb{R}} \left\{ -a + \frac{1}{K(1-\alpha)} \sum_{k=1}^K \max(a - Y_k, 0) \right\} \end{aligned}$$

Desta forma, o desvio *Average Value at Risk* pode ser escrito como:

$$\mathcal{D}(Y) := \mathbb{E}(Y) - AVaR_\alpha(Y) = \mathbb{E}(Y) - \left[ - \min_{a \in \mathbb{R}} \left\{ -a + \frac{1}{K(1-\alpha)} \sum_{k=1}^K \max(a - Y_k, 0) \right\} \right]$$

Considerando  $T$  períodos de tempo, o desvio *Average Value at Risk* tem a seguinte forma:

$$\mathcal{D}(Y) = \sum_{t=1}^T \lambda_t \mathbb{E}(Y^t) + \sum_{t=1}^T \lambda_t \min_{a^t \in \mathbb{R}} \left\{ -a^t + \frac{1}{K(1-\alpha^t)} \sum_{k=1}^K \max(a^t - Y_k^t, 0) \right\}, \quad (3.3.3)$$

onde,  $\lambda_t \in [0, 1]$  é uma variável de controle de risco do tempo  $t$ , conforme explicado na Seção 3.3.3.1 anterior.

## 3.4 Fronteira Eficiente e Reta de Alocação de Capital

O modelo média-variância proposto por Markowitz [3] na década de 50 se tornou referência para a Teoria Moderna de Portfolios. A teoria proposta por Markowitz estabelece que decisões relacionadas à seleção de investimentos devam ser tomadas com base na relação risco  $\times$  retorno dos ativos. Tal teoria vem sendo aperfeiçoada considerando modelos de risco mais robustos. De modo a serem efetivos, tais modelos devem ser capazes de quantificar os níveis de risco e retorno dos investimentos.

A partir do modelo proposto por Markowitz, derivou-se o conceito de fronteira eficiente, que expressa diretamente a relação entre risco e retorno da carteira. A fronteira eficiente é obtida ao resolver diversas vezes o problema de carteira ótima, considerando níveis de retorno esperado distintos, conforme apresentado na Figura 3.9.

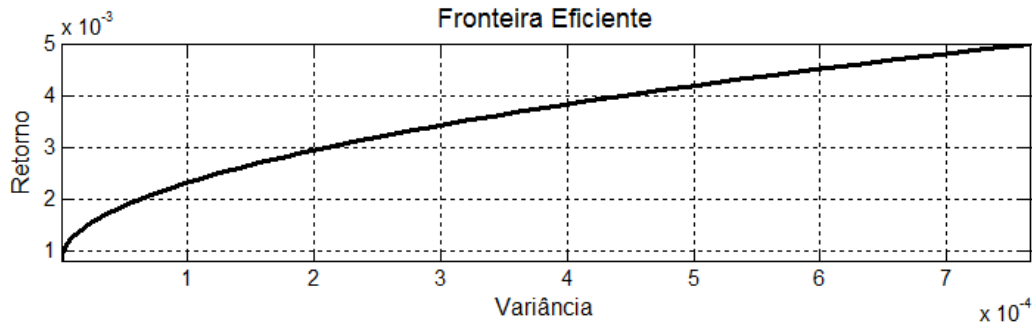


Figura 3.9: Fronteira Eficiente.

O conceito de fronteira eficiente é utilizado apenas para ativos com risco. Ao introduzirmos um ativo livre de risco, temos o conceito de reta de alocação de capital. Esta reta possui uma característica interessante, todas as carteiras contidas nela são uma combinação do ativo livre de risco e de uma carteira especial, chamada de portfólio eficiente, que se localiza no ponto de tangência entre a reta de alocação de capital e a fronteira eficiente, conforme pode ser verificado na Figura 3.9.

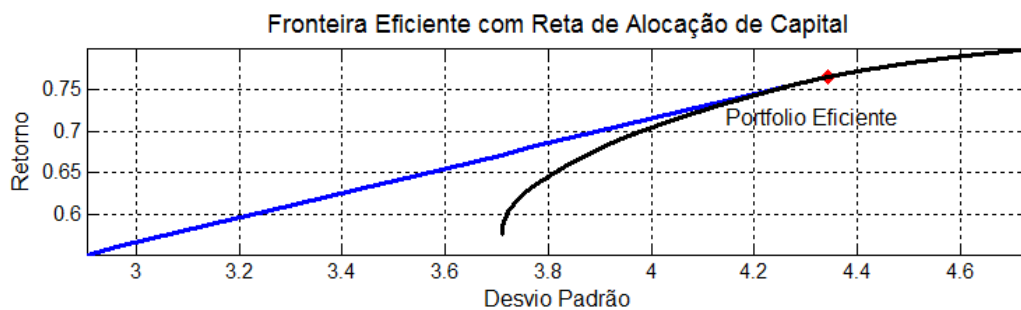


Figura 3.10: Fronteira Eficiente com Reta de Alocação de Capital.

O impacto no gráfico relacionado ao investidor poder emprestar ou tomar emprestado o capital associado ao ativo livre de risco pode ser verificado diretamente na reta de alocação de capital. Os pontos da reta anteriores ao portfólio eficiente são referentes a posições compradas no ativo livre de risco enquanto os pontos posteriores são referentes a posições vendidas. Caso o modelo não permita *shortselling* como o problema proposto (1.2.1), os pontos posteriores ao portfólio eficiente farão parte da fronteira eficiente, pelo fato de possuírem apenas ativos com risco.

Os conceitos de fronteira eficiente e reta de alocação de capital serão utilizados para a comparação do risco  $\times$  retorno obtidos pelos modelos analisados. Mais detalhes sobre o tema podem ser verificados em Elton e Gruber [28].

### 3.5 Modelo Determinístico Equivalente

Ao modelo estocástico gerado e resolvido como um modelo determinístico, sem o uso de técnicas especiais, dá-se o nome de *modelo determinístico equivalente*. Este modelo será utilizado para a comparação com o modelo proposto neste trabalho e será definido nesta seção.

Por razão de simplificação, será criado um conjunto com as restrições do problema proposto (1.2.1) da seguinte forma:

$$Z(\pi, \tau) := \left\{ \begin{array}{l} x \in \mathbb{R}^{N \times T} \text{ e } x_0 \in \mathbb{R}^T \mid \\ x_0^t r_0^t + \mathbb{E}[H(x^t(\xi), r^t(\xi))] \geq \tau^t, x_0^t + \sum_i x_i^t(\xi) = 1, \sum_{i \in P} x_i^t(\xi) \leq \pi, x(\xi) \geq 0 \text{ e } x_0 \geq 0, \\ \text{para } t = 1, \dots, T \end{array} \right\}.$$

Desta forma, o problema proposto (1.2.1) pode ser representado como:

$$\min \mathcal{D}(H(x(\xi), r(\xi))) \text{ sobre } x, x_0 \in Z(\pi, \tau). \quad (3.5.1)$$

As funções de desvio que serão utilizadas, definidas na Seção 3.3.3, são apresentadas novamente a seguir aplicadas ao problema proposto (3.5.1). Desta forma, com a definição da função de desvio  $\mathcal{D}$  e o percentual  $\pi$ , o problema será resolvido para diferentes valores de  $\tau$  de forma a obtermos uma fronteira eficiente conforme definida na seção anterior.

Será utilizada a seguinte notação para o valor esperado da carteira composta apenas pelos ativos com risco no tempo  $t$ :

$$\mu^t := \mathbb{E} \left[ H(x^t(\xi), r^t(\xi)) \right] = \frac{1}{K} \sum_{k=1}^K H(x^t(\xi_k), r^t(\xi_k)). \quad (3.5.2)$$

#### 3.5.1 Desvio Padrão

Substituindo a função de desvio em (3.5.1) pelo Desvio Padrão definido em (3.3.1), temos o seguinte problema:

$$\min \sum_{t=1}^T \lambda_t \sqrt{\mathbb{E} \left( H(x^t(\xi), r^t(\xi)) - \mu^t \right)^2} \text{ sobre } x, x_0 \in Z(\pi, \tau). \quad (3.5.3)$$

A representação matricial, para fins de implementação, do problema (3.5.3) pode ser verificada no Apêndice A.

#### 3.5.2 Desvio *Average Value at Risk*

Substituindo a função de desvio em (3.5.1) pelo Desvio *Average Value at Risk* definido em (3.3.3), temos o seguinte problema:



$$\left\{ \begin{array}{l} \min \quad \sum_{t=1}^T \lambda_t \mu^t + \sum_{t=1}^T \lambda_t \min_{a^t \in \mathbb{R}} \left\{ -a^t + \frac{1}{K(1-\alpha^t)} \sum_{k=1}^K \max(a^t - H(x^t(\xi_k), r^t(\xi_k)), 0) \right\} \\ \text{sobre } x, x_0 \in Z(\pi, \tau). \end{array} \right. \quad (3.5.4)$$

Em (3.5.4) temos um problema de minimização irrestrito dentro da função objetivo. Podemos juntar ambos os problemas de otimização para um problema agora de três variáveis de decisão:  $x, x_0 \in Z(\pi, \tau)$  e  $a \in \mathbb{R}^T$ . Além disso, podemos introduzir uma nova variável de decisão  $b \in \mathbb{R}^{K \times T}$  de forma a obter um problema de otimização linear. Desta forma, o problema (3.5.4) pode ser escrito como:

$$\left\{ \begin{array}{l} \min \quad \sum_{t=1}^T \lambda_t \mu^t + \sum_{t=1}^T \lambda_t \left( -a^t + \frac{1}{K(1-\alpha^t)} \sum_{k=1}^K b_k^t \right) \\ \text{sobre } x, x_0 \in Z(\pi, \tau), a \in \mathbb{R}^T \text{ e } b \in \mathbb{R}^{K \times T} \\ \text{s.a. } b_k^t \geq 0 \text{ para todo } k = 1, \dots, K \text{ e } t = 1, \dots, T \\ b_k^t \geq a^t - H(x^t(\xi_k), r^t(\xi_k)) \text{ para todo } k = 1, \dots, K \text{ e } t = 1, \dots, T. \end{array} \right. \quad (3.5.5)$$

A representação matricial, para fins de implementação, do problema (3.5.5) pode ser verificada no Apêndice A.



# Capítulo 4

## Método de Galerkin Para Otimização Estocástica

Neste capítulo é conceituado o método de Galerkin para otimização estocástica. A Seção 4.1 apresenta uma revisão de alguns conceitos de medida necessários para a definição do método proposto. A Seção 4.2 apresenta a definição do método de Galerkin aplicado ao problema proposto. A definição genérica do método pode ser encontrada no artigo de referência deste trabalho [1]. As Seções 4.3 e 4.4, apresentam a formulação matemática para o problema dos especialistas e para o problema  $\alpha$  que são gerados a partir da aplicação do método proposto.

### 4.1 Conceitos Iniciais

Para definirmos o método de Galerkin para otimização estocástica é necessário definirmos primeiramente alguns conceitos. Inicialmente, serão definidos os conceitos de  $\sigma$ -álgebra e espaço mensurável utilizando como referência o autor Bartle [29].

**Definição 12.** *Dado um conjunto  $\Omega$  não vazio, dizemos que uma família  $\mathcal{F}$  de subconjuntos de  $\Omega$  é uma  $\sigma$ -álgebra se:*

i  $\emptyset \in \mathcal{F}, \Omega \in \mathcal{F}$ .

ii Se  $A \in \mathcal{F}$  então seu complementar  $A^C \in \mathcal{F}$ .

iii Seja  $A_j$  uma sequência de conjuntos em  $\mathcal{F}$ , então qualquer união enumerável  $\cup_{j=1}^{\infty} A_j \in \mathcal{F}$ .

**Definição 13.** *Dado um conjunto  $\Omega$  não vazio e uma  $\sigma$ -álgebra, dizemos que o par  $(\Omega, \mathcal{F})$  é um **espaço mensurável**.*

A partir destes conceitos, podemos definir uma medida de probabilidade e conseqüentemente um espaço de probabilidade. As definições a seguir podem ser verificadas em Shreve [30].

**Definição 14.** *Seja  $(\Omega, \mathcal{F})$  um espaço mensurável. A **medida de probabilidade**  $\mathcal{P}$  é uma função que leva cada  $A \in \mathcal{F}$  a um número no intervalo  $[0, 1]$ , chamado de probabilidade de  $A$  e denotado por  $\mathcal{P}(A)$ . A medida de probabilidade possui as seguintes propriedades:*

i  $\mathcal{P}(\Omega) = 1$

ii Para qualquer sequência enumerável disjunta  $A_1, A_2, \dots \in \mathcal{F}$ ,

$$\mathcal{P}(\cup_{j=1}^{\infty} A_j) = \sum_{j=1}^{\infty} \mathcal{P}(A_j)$$

A tripla  $(\Omega, \mathcal{F}, \mathcal{P})$  é um *espaço de probabilidade*.

Finalmente, definimos uma filtração e um processo adaptado à filtração.

**Definição 15.** *Seja  $(\Omega, \mathcal{F})$  um espaço mensurável. Seja  $T > 0$  e assuma que para cada  $t \in [0, T]$  existe uma  $\sigma$ -álgebra  $\mathcal{F}_t$ . Quando para quaisquer  $s$  e  $t$ ,  $0 < t < s < T$  temos que  $\mathcal{F}_t \subseteq \mathcal{F}_s$ , então a família  $\{\mathcal{F}_t\}_{t \in I}$  de  $\sigma$ -álgebra é chamada de **filtração**. Por convenção,  $\mathcal{F}_0 = \{\emptyset, \Omega\}$ .*

Desta forma, podemos entender uma filtração como a ordenação da informação que temos disponível ao longo do tempo. Assim, com o passar do tempo, a  $\sigma$ -álgebra anterior é subconjunto da nova  $\sigma$ -álgebra, uma vez que teremos disponíveis não só as novas informações como também as informações que já tínhamos anteriormente. Por fim, a definição de processo adaptado à filtração é apresentada a seguir:

**Definição 16.** *Um processo  $\{X_t\}_{t \in I}$  é dito **adaptado à filtração**  $\{\mathcal{F}_t\}$  se para todo  $t$ ,  $X_t$  é  $\mathcal{F}_t$  mensurável.*

## 4.2 Aproximações de Galerkin

O método proposto neste trabalho para a resolução do problema (1.2.1) é o método de Galerkin. A seguir é apresentado o conceito de Aproximações de Galerkin aplicado ao problema proposto utilizando como referência a definição de Koivu e Pennanen [1].

Seja  $(\Omega, \mathcal{F}, \mathcal{P})$  um espaço de probabilidade e seja  $(\mathcal{F}_t)_{t=1}^T$  a filtração gerada pelo processo estocástico  $\xi = (\xi_t)_{t=1}^T$ . Considerando o problema proposto em (1.2.1), podemos escrevê-lo da seguinte forma:

$$\begin{cases} \min_{x \in \mathcal{N}} & \mathcal{D}(H(x(\xi)), r(\xi)) \\ \text{s.a.} & \mathbb{E}[\tau^t - x_0^t r_0^t - H(x^t(\xi), r^t(\xi))] \leq 0, \text{ para } t = 1, \dots, T \\ & x(\xi) \in \mathcal{X}(\xi) = \{x \mid \sum_i x_i^t(\xi) = 1, \sum_{i \in P} x_i^t(\xi) \leq \pi, x(\xi) \geq 0 \text{ e } x_0 \geq 0, t = 1, \dots, T\} \quad P\text{-q.c.} \end{cases} \quad (4.2.1)$$

Onde,  $\mathcal{N} = \{(x^t)_{t=1}^T \mid x^t \in L^2(\Omega, \mathcal{F}_t, \mathcal{P}; \mathbb{R}^{n_t})\}$  é o espaço das estratégias de decisão  $(\mathcal{F}_t)_{t=1}^T$ -adaptadas essencialmente limitadas com o tempo de decisão  $t$  sendo um vetor real  $n_t$ -dimensional.

Por questão de simplificação, podemos representar o problema de otimização (4.2.1) da seguinte forma:

$$\begin{cases} \min & \mathcal{D}(H(x(\xi)), r(\xi)) \\ \text{s.a.} & \mathbb{E}[\tau^t - x_0^t r_0^t - H(x^t(\xi), r^t(\xi))] \leq 0, \text{ para } t = 1, \dots, T \end{cases} \quad \text{sobre } x \in \mathcal{N} \cap \mathcal{X} \quad (4.2.2)$$

Onde,

$$\mathcal{X} = \{x \in L^2(\Omega, \mathcal{F}, \mathcal{P}; \mathbb{R}^{n_t}) \mid x(\xi) \in \mathcal{X}(\xi) \text{ } P\text{-q.c.}\} \quad (4.2.3)$$

Seja  $\{x_i \in \mathcal{N} \cap \mathcal{X}, i \in I\}$  um conjunto finito de *estratégias base*, conhecidas também como regras de decisão ou funções de base, e seja  $\Delta \in \mathbb{R}^I$  um conjunto convexo de *pesos* tal que:

$$\mathcal{N}_\Delta := \left\{ \sum_{i \in I} \alpha_i x_i \mid \alpha \in \Delta \right\} \subset \mathcal{N} \cap \mathcal{X} \quad (4.2.4)$$

Como  $\mathcal{N}$  é um espaço vetorial, temos que  $\mathcal{N}_\Delta \subset \mathcal{N} \cap \mathcal{X}$  tanto quanto  $\mathcal{N}_\Delta \subset \mathcal{X}$ . Desta forma, o conjunto  $\Delta$  deve ser escolhido em conformidade com as estratégias base  $x_i$  e o conjunto  $\mathcal{X}$ . Assumindo que as estratégias base são viáveis, temos que:

- $\Delta = \mathbb{R}^I$  quando  $\mathcal{X}$  é linear,
- $\Delta = \mathbb{R}_+^I$  quando  $\mathcal{X}$  é cônico,
- $\Delta = \{\alpha \in \mathbb{R}^I \mid \sum_{i \in I} \alpha_i = 1\}$  quando  $\mathcal{X}$  é afim,
- $\Delta = \{\alpha \in \mathbb{R}_+^I \mid \sum_{i \in I} \alpha_i = 1\}$  quando  $\mathcal{X}$  é convexo.

Conforme definido no problema proposto (4.2.2), o conjunto  $\mathcal{X}$  é convexo e desta forma temos que  $\Delta = \{\alpha \in \mathbb{R}_+^I \mid \sum_{i \in I} \alpha_i = 1\}$ .

Dado um conjunto  $\{x_i \mid i \in I\}$  de estratégias base viáveis e um conjunto convexo de pesos viáveis  $\Delta$ , a *Aproximação de Galerkin* para o problema proposto é o problema de otimização a seguir:

$$\begin{cases} \min & \mathcal{D}(H(x(\xi), r(\xi))) & \text{sobre } x \in \mathcal{N}_\Delta \\ \text{s.a} & \mathbb{E}[\tau^t - x_0^t r_0^t - H(x^t(\xi), r^t(\xi))] \leq 0, & \text{para } t = 1, \dots, T. \end{cases} \quad (4.2.5)$$

O mesmo pode ser escrito na forma finito-dimensional:

$$\begin{cases} \min & \mathcal{D}(\varphi_0(\alpha, \xi)) & \text{sobre } \alpha \in \Delta \\ \text{s.a} & \mathbb{E}[\varphi_1(\alpha, \xi)] \leq 0, \end{cases} \quad (4.2.6)$$

onde,

$$\varphi_0(\alpha, \xi) = H\left(\sum_{i \in I} \alpha_i x_i^t(\xi), r(\xi)\right) \text{ e } \varphi_1(\alpha, \xi) = \tau^t - x_0^t r_0^t - H\left(\sum_{i \in I} \alpha_i x_i^t(\xi), r^t(\xi)\right).$$

Desta forma, temos um problema de programação estocástica cuja função objetivo é convexa na variável de decisão  $\alpha$ . Isto sugere o uso de técnicas de programação estocástica que visam otimizar tais objetivos, avaliando o integrando e seus subgradientes com relação a  $\alpha$  ao longo de um número finito de cenários  $\xi$ . Uma avaliação de  $\varphi_j(\alpha, \xi)$  exige a avaliação de cada estratégia base  $x_i^t$  ao longo do cenário dado  $\xi$ . No entanto, uma vez que  $x_i^t(\xi)$  tenha sido avaliado para todo  $i \in I$ , podemos usá-los para avaliar  $\varphi(\cdot, \xi)$  para diferentes valores de  $\alpha$ .

Em resumo, para a utilização do método de Galerkin é necessário especificar os Especialistas que serão responsáveis por gerar as estratégias base  $x_i^t$  para cada cenário e o problema  $\alpha$  que, para o caso estudado, será dado pela equação (4.2.6). Nas Seções 4.3 e 4.4, são apresentadas as formulações matemáticas para o problema dos Especialistas e para o problema  $\alpha$  ao considerar as funções de desvio propostas.

Para ilustrar a utilização do Método de Galerkin será apresentado o exemplo 3.2 do artigo de referência deste trabalho [1].

**Exemplo 5.** O problema consiste na otimização de carteira a seguir:

$$\left\{ \begin{array}{ll} \min & \mathbb{E}(\exp(-\rho \sum_{j \in J} x_{T,j}(\xi))) \quad \text{sobre } x \in \mathcal{N} \\ \text{s.a.} & \sum_{j \in J} x_0, j \leq 1, \\ & \sum_{j \in J} x_{t,j}(\xi) \leq \sum_{j \in J} r_{t,j}(\xi) x_{t-1,j} \quad t = 0, \dots, T, \\ & x(\xi) \geq 0 \quad P\text{-q.c.}, \end{array} \right. \quad (4.2.7)$$

onde,

- $x_{t,j}$  é a quantidade de dinheiro investida no ativo  $j \in J$  no tempo  $t$ ,
- $r_{t,j}$  é o retorno randômico do ativo  $j \in J$  no período  $[t-1, t]$  e
- $\rho$  é o parâmetro de aversão ao risco.

O conjunto  $J$  contém 10 ativos e os retornos são dados por  $r_{t,j} = s_{t,j}/s_{t-1,j}$  onde o processo de preço  $s$  segue um movimento Browniano 10-dimensional.

O problema pode ser escrito como em (4.2.2) com  $n_t = 10$  para todo  $t = 0, \dots, T$ ,

$$\min \exp(-\rho \sum_{j \in J} x_{T,j}) \quad \text{sobre } x \in \mathcal{N} \cap \mathcal{X}, \quad (4.2.8)$$

que é independente de  $\xi$ , e

$$\mathcal{X}(\xi) = \{x \in \mathbb{R}_+^n \mid \sum_{j \in J} x_0, j \leq 1, \sum_{j \in J} x_{t,j}(\xi) \leq \sum_{j \in J} r_{t,j}(\xi) x_{t-1,j}, t = 0, \dots, T\}. \quad (4.2.9)$$

As *Aproximações de Galerkin* são construídas usando os especialistas  $i$  definidos a seguir:

$$x_t^i(\xi) = \begin{cases} \pi^i & \text{se } t = 0, \\ \pi^i \sum_{j \in J} r_{t,j}(\xi) x_{t-1,j}^i(\xi) & \text{caso contrário,} \end{cases} \quad (4.2.10)$$

onde,  $\pi^i \in \mathbb{R}^J$  é um vetor de pesos com as proporções investidas em cada ativo pelo especialista  $i$  de forma que as estratégias para cada tempo  $t$  e ativo  $j$  sejam definidas por esta proporção e o retorno nominal considerando a estratégia no tempo anterior e o retorno dos ativos no tempo  $t$ . Tais estratégias são chamadas de *fixed-mix*. Uma vez que  $x^i$  sejam estratégias viáveis, a escolha:

$$\Delta = \{\alpha \in \mathbb{R}_+^I \mid \sum_{i \in I} \alpha^i = 1\}, \quad (4.2.11)$$

irá garantir que todas as combinações lineares em  $\mathcal{N}_\Delta := \{\sum_{i \in I} \alpha^i x^i \mid \alpha \in \Delta\}$  são viáveis. Desta forma, com as estratégias  $x^i$  definidas e viáveis, podemos avaliar o problema em  $\alpha$ . Portanto, a Aproximação de Galerkin pode ser escrita como:

$$\min \mathbb{E}(\exp(-\rho \sum_{j \in J} \sum_{i \in I} \alpha^i x_{T,j}^i(\xi))) \quad \text{sobre } \alpha \in \Delta \quad (4.2.12)$$

O autor resolve o problema numericamente através de quadratura de integração. Os resultados são apresentados nas tabelas 4.1 e 4.2 abaixo. Conforme verificado pelo autor, a melhor estratégia de base individual supera a estratégia obtida com discretização baseada em árvore mostrando que a estratégia de Galerkin é um pouco melhor.

Simulação	Otimização
22	6

Tabela 4.1: Tempos de computação em segundos para o método de Galerkin com 20 estratégias de base e 5.000 cenários.

	Galerkin	Tree	Best basis	Lower bound
Média	0.2345	0.2348	0.2346	0.2304
99% Bound	0.2358	0.2378	0.2360	0.2303

Tabela 4.2: Estimativas da amostra dos valores objetivos obtidos com o método de Galerkin e discretizações baseadas em árvores bem como a melhor estratégia de base individual. A quarta coluna indica os limites inferiores estatísticos obtidos com o método de árvores aleatórias.

## 4.3 Especialistas

Conforme verificado por Koivu e Pennanen [1], a escolha de especialistas, ou funções de base, está relacionada com o bom conhecimento do problema subjacente. Essas funções podem conter informações valiosas que podem ser difíceis de serem construídas com técnicas de algoritmos genéricas. No Exemplo 5, os especialistas  $i$  são definidos na equação (4.2.10). Conforme visto, as estratégias de investimento para cada tempo  $t$  e ativo  $j$  são escolhidas por uma proporção fixa e o retorno nominal considerando a estratégia no tempo anterior e o retorno dos ativos no tempo  $t$ . Como pôde ser verificado neste exemplo, mesmo com escolhas bastante simples de estratégias para o problema, pode-se alcançar surpreendentemente boas soluções. A seguir será apresentada uma proposta para a construção dos especialistas que serão utilizados para a obtenção dos resultados numéricos no próximo capítulo.

### 4.3.1 Determinístico

O Especialista Determinístico considera que os tempos futuros irão se comportar da forma prevista pelo cenário base. Desta forma, o especialista busca maximizar o retorno da carteira considerando o cenário base e alguns critérios estipulados pelo investidor. Estes critérios são baseados em opiniões de mercado, tais como: perspectivas de aumento da energia elétrica,

aumento da taxa de juros ou outras informações que possam impactar no preço dos ativos analisados.

Será considerado que o espaço amostral  $\Xi$  é finito com  $K$  realizações  $(\xi_1, \dots, \xi_K)$ , onde  $\xi_k \in \mathbb{R}^T$ . A formulação matemática para o  $e$ -ésimo especialista determinístico aplicada ao problema proposto é apresentada a seguir:

$$\left\{ \begin{array}{l} \max_{x \in \mathbb{R}^n, x_0 \in \mathbb{R}} \sum_{t=1}^T \sum_{i=0}^N x_i^t(\xi_k, e) r_i^t(\xi_k) \\ s.a. \quad \sum_{i=0}^N x_i^t(\xi_k, e) = 1 \quad t = 1, \dots, T \\ \sum_{i \in P} x_i^t(\xi_k, e) \geq \pi \quad t = 1, \dots, T \\ x_i^t(\xi_k, e) \geq 0 \quad t = 1, \dots, T \text{ e } i = 0, \dots, N \\ x_0^t(\xi_k, e) = x_0^t, r_0^t(\xi_k) = r_0^t \quad t = 1, \dots, T \end{array} \right. \quad (4.3.1)$$

Conforme definido, cada especialista  $e$  possui preferência por determinado grupo econômico  $P$ . Desta forma, o especialista em questão decidirá investir pelo menos  $\pi\%$  em ações do grupo econômico de preferência. Com isto, o problema fica parametrizado pelo percentual mínimo  $\pi$  e é viável desde que  $\pi$  seja menor ou igual a 1.

Em termos de implementação, o modelo será representado como um modelo de minimização da perda do investidor. Essa alteração na prática só inverte o sinal do coeficiente da função objetivo, porém simplifica a representação do problema computacionalmente pois grande parte da literatura de otimização tem como convenção o estudo de problemas de minimização.

## 4.4 Problema $\alpha$

Para as definições seguintes serão utilizadas as notações a seguir:

- Índice do especialista, denotado por  $e$ , variando de 1 a  $E$ ;
- Índice do cenário, denotado por  $k$ , variando de 1 a  $K$ ;
- Índice do ativo, denotado por  $i$ , variando de 1 a  $N$ ;
- Índice do tempo, denotado por  $t$ , variando de 1 a  $T$ .

Dado um cenário  $\xi_k \in \mathbb{R}^T$ , cada especialista  $e$  retorna uma estratégia de investimento definida pela carteira:

$$x_i^t(\xi_k, e) \quad \text{para } i = 1, \dots, N \text{ e } t = 1, \dots, T.$$

O problema  $\alpha$  irá encontrar a melhor combinação convexa para as estratégias definidas pelos especialistas de acordo com a função de desvio escolhida. Essa combinação depende de um vetor  $\alpha \in \mathbb{R}^E$ , para  $\sum_{e=1}^E \alpha_e = 1$  e  $\alpha \geq 0$ , e é definido por:



$$\sum_{e=1}^E \alpha_e x_i^t(\xi_k, e).$$

Como  $\sum_{i=1}^N x_i^t(\xi_k, e) = 1$  e  $\sum_{e=1}^E \alpha_e = 1$ , pode-se verificar que:

$$\sum_{i=1}^N \sum_{e=1}^E \alpha_e x_i^t(\xi_k, e) = \sum_{e=1}^E \alpha_e \sum_{i=1}^N x_i^t(\xi_k, e) = \sum_{e=1}^E \alpha_e 1 = 1.$$

Para cada passo de tempo  $t$ , o retorno aleatório de cada ativo é dado por  $r_i^t(\xi_k)$  para  $k = 1, \dots, K$ . Fixado o tempo  $t$ , podemos construir uma matriz  $V^t$  com dimensão  $K \times E$  e com elementos:

$$v_{k,e}^t := \sum_{i=1}^N r_i^t(\xi_k) x_i^t(\xi_k, e),$$

retornando o ganho ou perda no tempo  $t$  para a carteira proposta pelo especialista  $e$  para o cenário  $k$ .

A combinação dos portfólios dos especialistas no tempo  $t$ , para cada cenário  $k$ , é o vetor  $K \times 1$  dado por  $V^t \alpha$ . Esta é a nossa variável aleatória no tempo  $t$ :

$$Y^t := V^t \alpha. \quad (4.4.1)$$

Note que a definição da matriz  $V^t$  depende dos cenários escolhidos, mas não do vetor  $\alpha$ .

Para uma variável aleatória com  $K$  realizações  $Y^t(\xi_k)$ , o vetor  $Y^t \in \mathbb{R}^K$  representa todas as realizações do espaço amostral e a notação  $\mathbb{E}[Y^t]$  representa o valor esperado:

$$\mathbb{E}[Y^t] = \frac{1}{K} \sum_{k=1}^K Y^t(\xi_k).$$

Em particular, o retorno médio no tempo  $t$  é dado por:

$$\mu^t(\alpha) := \mathbb{E}[Y^t] = \mathbb{E}[V^t(\xi) \alpha] = \mathbb{E}[V^t(\xi)] \alpha = \frac{1}{K} \sum_{k=1}^K (v_{k,\cdot}^t) \alpha = \frac{1}{K} \sum_{e=1}^E \sum_{k=1}^K v_{k,e}^t \alpha_e,$$

ou seja, na matriz  $V^t$  toma-se a média nas linhas para cada coluna  $e$ . O resultado é um vetor com  $E$  componentes, que são multiplicados por cada componente do vetor  $\alpha$  para a obtenção do número  $\mu^t$ .

A melhor combinação para os especialistas é dada pela solução  $\bar{\alpha}$  para o problema:

$$\left\{ \begin{array}{l} \min_{\alpha \in \mathbb{R}^E} \mathcal{D}(Y(\alpha)) \\ \text{s.a.} \quad 0 \leq \alpha_e \leq \gamma_e \text{ para } e = 1, \dots, E \\ \sum_{e=1}^E \alpha_e = 1 \\ \mu^t(\alpha) \geq \mu_{\min}^t \text{ para } t = 1, \dots, T, \end{array} \right. \quad (4.4.2)$$

Onde,  $\gamma_e$  representa o percentual máximo de proporção no especialista  $e$  e  $\mu_{\min}^t$  representa o retorno mínimo esperado no tempo  $t$ . Por questão de simplificação, podemos escrever o problema  $\alpha$  (4.4.2) da seguinte forma:

$$\min \mathcal{D}(Y(\alpha)) \quad \text{sobre } \alpha \in S(\mu_{\min}, \gamma), \quad (4.4.3)$$

onde,

$$S(\mu_{\min}, \gamma) := \left\{ \alpha \mid \sum_{e=1}^E \alpha_e = 1, 0 \leq \alpha_e \leq \gamma_e \text{ para } e = 1, \dots, E \text{ e } \mu^t(\alpha) \geq \mu_{\min}^t \text{ para } t = 1, \dots, T \right\}.$$

O problema  $\alpha$  será sempre viável desde que o retorno mínimo  $\mu_{\min}^t$  seja pelo menos o retorno obtido com o ativo livre de risco e no máximo o retorno obtido com o especialista com maior retorno no tempo  $t$ . Além disso, é necessário que  $0 \leq \gamma_e \leq 1$  e  $\sum_{e=1}^E \gamma_e \geq 1$ .

A seguir serão formulados os problemas  $\alpha$ , conforme definido em (4.4.3), de acordo com as funções de desvio escolhidas.

#### 4.4.1 Desvio Padrão

Conforme definido em (3.3.1), para diferentes períodos de tempo, o Desvio Padrão é dado por:

$$\mathcal{D}(Y(\alpha)) = \sum_{t=1}^T \lambda_t \sqrt{\mathbb{E} \left[ Y^t(\alpha) - \mathbb{E}[Y^t(\alpha)] \right]^2} = \sum_{t=1}^T \lambda_t \sqrt{\mathbb{E} \left[ V^t \alpha - \mathbb{E}[V^t \alpha] \right]^2}.$$

Desta forma, o problema  $\alpha$  (4.4.3) pode ser escrito como:

$$\min \sum_{t=1}^T \lambda_t \sqrt{\mathbb{E} \left[ V^t \alpha - \mathbb{E}[V^t \alpha] \right]^2} \quad \text{sobre } \alpha \in S(\mu_{\min}, \gamma),$$

De forma matricial, temos que:

$$\begin{aligned} \mathbb{E} \left[ V^t \alpha - \mathbb{E}[V^t \alpha] \right]^2 &= \mathbb{E} \left[ (V^t \alpha - \mathbb{E}[V^t \alpha])^T (V^t \alpha - \mathbb{E}[V^t \alpha]) \right] \\ &= \mathbb{E} \left[ (V^t \alpha - \mathbb{E}[V^t \alpha])^T (V^t \alpha - \mathbb{E}[V^t \alpha]) \right] \\ &= \mathbb{E} \left[ ((V^t - \mathbb{E}[V^t]) \alpha)^T ((V^t - \mathbb{E}[V^t]) \alpha) \right] \\ &= \alpha^T \mathbb{E} \left[ (V^t - \mathbb{E}[V^t]) (V^t - \mathbb{E}[V^t])^T \right] \alpha \\ &= \alpha^T \Sigma^t \alpha, \end{aligned}$$

onde,  $\Sigma^t$  é a covariância da matriz  $V^t$ . O vetor  $\alpha$  pode ser colocado para fora da esperança pois é determinístico.

Por questão de conveniência, a função objetivo será tratada sem a raiz quadrada. Na prática, isto afeta apenas o resultado obtido para a função objetivo e basta tomar a raiz quadrada deste resultado para que seja avaliado o desvio padrão. Esta prática é usada devido às funções de

otimização dos softwares trabalharemos com funções objetivo lineares ou quadráticas. Desta forma, temos um problema de programação quadrática dado por:

$$\min \sum_{t=1}^T \lambda_t \alpha^T \Sigma^t \alpha \quad \text{sobre } \alpha \in S(\mu_{\min}, \gamma). \quad (4.4.4)$$

A representação matricial do problema (4.4.4) para fins de implementação é apresentada no Apêndice A.

#### 4.4.2 Desvio *Average Value at Risk*

Conforme definido em (3.3.3), para diferentes períodos de tempo, o Desvio *Average Value at Risk* é dado por:

$$\mathcal{D}(Y) = \sum_{t=1}^T \lambda_t \mathbb{E}(Y^t) + \sum_{t=1}^T \lambda_t \min_{a^t \in \mathbb{R}} \left\{ -a^t + \frac{1}{K(1-\beta^t)} \sum_{k=1}^K \max(a^t - Y_k^t, 0) \right\},$$

onde, o nível de confiança adotado em cada período de tempo  $t$  é dado por  $\beta^t$  para que não haja duplicidade de notação em relação à variável de decisão  $\alpha$ . Desta forma, o problema  $\alpha$  (4.4.2) pode ser escrito como:

$$\min \sum_{t=1}^T \lambda_t \mu^t(\alpha) + \sum_{t=1}^T \lambda_t \left( -a^t + \frac{1}{K(1-\beta^t)} \sum_{k=1}^K \max(a^t - \sum_{e=1}^E v_{k,e}^t \alpha_e, 0) \right) \\ \text{sobre } \alpha \in S(\mu_{\min}, \gamma) \text{ e } a \in \mathbb{R}^T.$$

Podemos introduzir uma nova variável de decisão  $b \in \mathbb{R}^{K \times T}$ , de forma a obter o problema de otimização linear equivalente, conforme desejado:

$$\left\{ \begin{array}{l} \min \sum_{t=1}^T \lambda_t \mu^t(\alpha) + \sum_{t=1}^T \lambda_t \left( -a^t + \frac{1}{K(1-\beta^t)} \sum_{k=1}^K b_k^t \right) \\ \text{sobre } \alpha \in S(\mu_{\min}, \gamma), a \in \mathbb{R}^T \text{ e } b \in \mathbb{R}^{K \times T} \\ \text{s.a. } b_k^t \geq 0 \text{ para todo } k = 1, \dots, K \text{ e } t = 1, \dots, T \\ b_k^t \geq a^t - \sum_{e=1}^E v_{k,e}^t \alpha_e \text{ para todo } k = 1, \dots, K \text{ e } t = 1, \dots, T. \end{array} \right. \quad (4.4.5)$$

A representação matricial do problema (4.4.5) para fins de implementação é apresentada no Apêndice A.



# Capítulo 5

## Experimentos Numéricos

Este capítulo contém os resultados computacionais obtidos para os métodos estocásticos proposto e de comparação. Os resultados que serão apresentados neste capítulo foram obtidos através da ferramenta MATLAB R2013a juntamente com o otimizador Gurobi 5.6.3. O computador utilizado possui processador de 2.0 GHz e 8.0 GB de RAM.

Na Seção 5.1, são especificados os valores para os parâmetros de entrada dos modelos analisados. Nas Seções 5.2 e 5.3, são feitas análises de consistência para os modelos estático e proposto com o objetivo de validar a implementação dos mesmos. Por fim, na Seção 5.4, as carteiras obtidas através do modelo proposto são comparadas com as obtidas pelo modelo estático através de testes de performance.

### 5.1 Parâmetros de Entrada

O primeiro passo para a realização dos experimentos numéricos consiste na especificação dos parâmetros de entrada, que serão utilizados pelos modelos analisados. Observando o problema proposto descrito em (1.2.1), pode-se verificar que o problema fica bem definido através dos seguintes parâmetros: o horizonte de tempo  $T$ , o conjunto de ativos, a amostra de retornos  $r_i^t(\xi)$  para cada ativo  $i$  e tempo  $t$ , o retorno esperado mínimo  $\tau^t$  para a carteira para cada tempo  $t$ , a função de desvio  $\mathcal{D}$  para a avaliação do risco da carteira e o percentual máximo  $\pi$  que pode ser investido em ativos de determinado grupo econômico  $P$ .

O horizonte de tempo adotado para a escolha das carteiras foi de 4 semanas. A escolha da carteira ótima será avaliada semanalmente, considerando este horizonte. O conjunto de ativos utilizados é composto por 15 ações do IBovespa que possuem boa liquidez, apresentadas na Seção 3.1, e um ativo livre de risco hipotético. As ações foram escolhidas de modo que pudessem ser analisados diversos grupos econômicos e, com isto, as carteiras pudessem ser escolhidas considerando ativos diversificados. Desta forma, as ações escolhidas foram agrupadas de acordo com o grupo econômico associado e as análises foram consolidadas a partir desses grupos. As séries históricas de preço de fechamento semanal das ações foram obtidas para o intervalo de 03 de Janeiro de 2000 a 17 de Outubro de 2014, totalizando 768 observações para cada ativo.

A Figura 5.1 a seguir, mostra os grupos econômicos observados nas análises e as ações associadas a cada grupo. Também pode ser verificado que existe um grupo à parte para o ativo livre de risco. As abreviações mnemônicas para cada grupo de ativos será utilizada para os gráficos de fronteira eficiente.

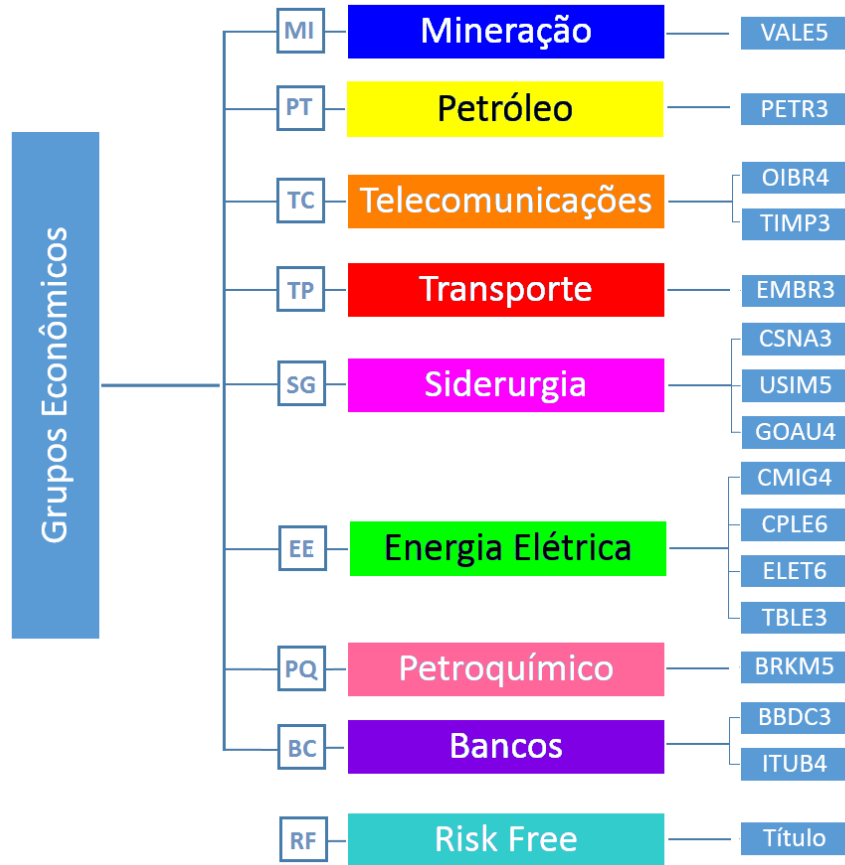


Figura 5.1: Grupos de Ativos.

Uma vez determinado o conjunto de ativos, a amostra dos retornos  $r_i^t(\xi)$  pôde ser construída através da simulação dos log-retornos, conforme descrito na Seção 3.2. Vale ressaltar que todas as séries utilizadas para a geração dos cenários estão ajustadas, ou seja, não consideram o efeito de eventos corporativos. Como exemplos de eventos corporativos, é possível citar o pagamento de dividendos, de juros sobre o capital próprio, o grupamento ou split das ações e as aquisições ou fusões com outras empresas. A importância desses eventos reside no efeito direto sobre os preços das ações. Pelo fato desta alteração nos preços ser fruto de decisões estratégicas e administrativas de cada empresa, estas geralmente não são incorporadas nos modelos de retornos.

O retorno mínimo esperado  $\tau^t$  para a carteira em cada tempo  $t$ , foi escolhido de forma que o problema proposto sempre fosse viável, ou seja, para o conjunto de cenários adotados analisou-se o menor e o maior retorno possível considerando as restrições do problema. A partir disto, foi criado um intervalo com 30 retornos esperados para cada tempo  $t$  considerando os limites de menor e maior retornos possíveis. Este intervalo de retornos esperados será utilizado para a construção da fronteira eficiente mencionada na Seção 3.4.

As funções de desvio  $\mathcal{D}$  que serão adotadas são: o desvio padrão, como uma abordagem clássica, e o desvio AVaR, como uma abordagem mais usual. Conforme visto na Seção 3.3.3, ao trabalhar com funções de desvio, é necessário especificar as variáveis de controle de risco  $\lambda_t$  para cada tempo  $t$ . As variáveis  $\lambda_t$  foram consideradas de forma que apenas o risco do tempo  $t$  observado fosse controlado, ou seja,  $\lambda_t = 1$  e  $\lambda_i = 0$ , onde  $i = \{1, \dots, \hat{t}, \dots, 4\}$ . Tal abordagem foi usada para que não fosse necessário contornar dificuldades relacionadas à não consistência

das funções de risco no tempo. Para o desvio AVaR é necessário a especificação de mais um parâmetro: o nível de confiança  $\alpha$ . A escolha de valores para o nível de confiança  $\alpha$  do AVaR não apresenta dificuldades. Nos testes foram adotados os valores típicos (95%, 97,5% e 99%).

Por fim, o percentual máximo  $\pi$  que pode ser investido em ativos de determinado grupo econômico  $P$  foi definido de acordo com o modelo analisado. Para o problema estático, foi definido que o percentual máximo que poderia ser alocado em ativos de cada um dos oito grupos econômicos observados era de 70%. Para o problema proposto, a dinâmica do percentual máximo  $\pi$  é tratada pelos especialistas apresentados na Seção 4.3. Desta forma, como os especialistas decidem como alocar os ativos para cada cenário, pode-se escolher o percentual máximo  $\pi$  de maneira mais interessante. Neste contexto, foram definidos oito especialistas, onde cada especialista  $e$  possui preferência por um determinado grupo econômico, conforme ilustrado na Figura 5.1. Com isto, cada especialista decide investir pelo menos  $\pi\%$  de seu capital em ativos do grupo econômico de interesse. Tal percentual será definido de acordo com cada teste. O percentual máximo que poderá ser escolhido da composição de cada especialista, também será configurado em cada teste.

Feitas estas parametrizações, os problemas estático e proposto podem ser resolvidos e analisados. Na próxima seção será apresentada uma análise de consistência para o modelo estático de comparação considerando as funções de risco propostas.

## 5.2 Análise de Consistência do Modelo Estático

No primeiro teste foi analisada a consistência das carteiras calculadas através do modelo estático em relação a três características: seu retorno esperado, sua composição e seu risco. Desta forma, foi utilizada a fronteira eficiente juntamente com um gráfico de alocações para ilustrar o comportamento do modelo. A formulação matemática para o modelo estático aqui analisado pode ser encontrada nas Seções 3.3.3.1 e 3.3.3.2.

O intervalo de retornos esperados necessário para a criação da fronteira eficiente foi definido através da resolução de um problema de minimização do risco da carteira. Neste problema são consideradas as restrições do problema estático sem a exigência de um retorno mínimo. Desta forma, obtém-se o retorno mínimo que pode ser alcançado com os cenários gerados através das alocações ótimas obtidas e o retorno esperado dos ativos. A partir disto, o retorno máximo da carteira, considerando os cenários gerados, também pode ser obtido através do ativo com maior retorno aplicado às restrições do modelo. Com os limites de retorno esperado mínimo e máximo bem definidos, particiona-se este intervalo de acordo com o número de carteiras desejadas para a análise. No caso do teste, foram escolhidas 30 carteiras.

Foram considerados 1.000 cenários para a geração dos retornos randômicos dos ativos e nível de confiança para o desvio AVaR de 95%. A seguir são mostradas as análises *Risk Neutral* e *Risk Averse* que consideram, respectivamente, a função objetivo igual ao valor esperado e a função objetivo igual ao desvio padrão ou desvio AVaR. Com o valor esperado, é de se esperar uma menor preocupação com a volatilidade, dado que o enfoque é a maximização do retorno da carteira. Já com a abordagem *Risk Averse* é de se esperar uma alocação maior no ativo livre de risco para que haja uma cobertura da volatilidade.

Na Figura 5.2 a seguir, pode-se verificar a fronteira eficiente com a reta de alocação de

capital e o gráfico da composição da carteira para o primeiro passo de tempo considerando a abordagem *Risk Averse*. A abscissa dos gráficos representa as funções de desvio e as ordenadas representam o retorno da carteira e sua composição consolidada por grupo econômico. A legenda no canto direito dos gráficos faz menção à notação definida na Figura 5.1, onde são relacionados os grupos econômicos com os ativos, suas abreviações e cores. Esta interpretação gráfica será utilizada para os demais gráficos de fronteira eficiente presentes neste trabalho.

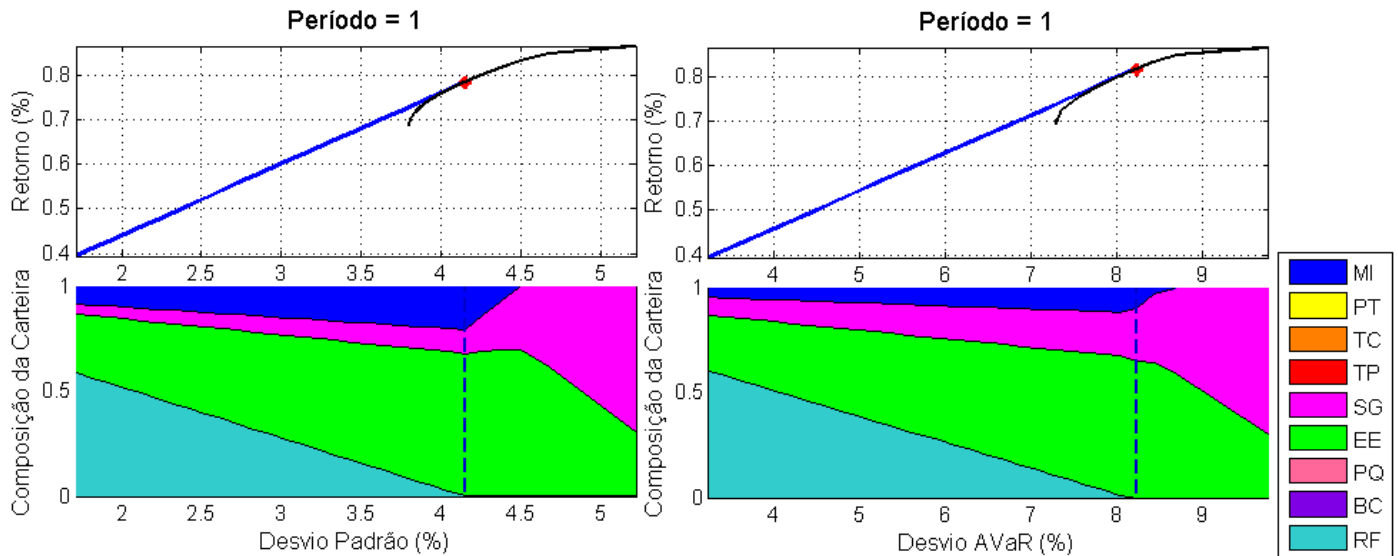


Figura 5.2: Fronteira eficiente, reta de alocação de capital e composição da carteira em  $t = 1$ .

Grupo Econômico	Ações	Desvio Padrão (%)	Desvio AVaR (%)	Log-Retorno Esperado (%)	Alocação PE DP (%)	Alocação PE DA (%)
[MI] Mineração	VALE5	4.59	8.82	0.64	20.58	9.64
[PT] Petróleo	PETR3	5.41	10.34	0.61	-	-
[TC] Telecomunicações	OIBR4	6.93	13.25	0.58	-	-
	TIMP3	6.56	12.11	-0.02	-	-
[TP] Transporte	EMBR3	4.99	10.02	0.53	-	-
[SG] Siderurgia	CSNA3	6.85	12.73	0.84	-	-
	USIM3	5.73	11.07	0.82	-	-
	GOAU4	6.26	11.94	0.87	11.45	25.60
[EE] Energia Elétrica	CMIG4	5.28	9.98	0.83	41.91	44.93
	CPLE6	6.27	12.30	0.63	-	-
	ELET6	5.24	10.11	0.77	13.16	12.47
	TBLE3	5.09	10.04	0.75	12.72	7.36
[PQ] Petroquímico	BRKM5	6.65	13.15	0.76	-	-
[BC] Bancos	BBDC3	5.22	10.48	0.69	-	-
	ITUB4	4.91	9.76	0.67	-	-
[RF] Risk Free	-	-	-	0.12	-	-

Tabela 5.1: Risco, retorno esperado e alocação do portfólio eficiente (PE) em  $t = 1$ .

Analisando a Figura 5.2 pode-se observar que as alocações propostas pelo modelo estático são bem próximas para as funções de desvio analisadas. Além disso, para os mesmos níveis de retorno, observa-se que o desvio AVaR apresenta sempre valores maiores em módulo em relação ao desvio padrão. Este fato é esperado, visto que o desvio AVaR considera apenas as perdas



enquanto o desvio padrão considera tanto perdas quanto ganhos. Feitas estas observações visuais, podemos analisar o risco e o retorno dos ativos de cada grupo econômico individualmente para entendermos se as alocações propostas são coerentes.

Através da Tabela 5.1, podemos analisar as alocações sugeridas pelo modelo. De fato, o setor de Energia Elétrica parece uma boa opção para reduzir o risco da carteira e ao mesmo tempo aumentar o retorno esperado da mesma, tendo em vista que as ações CMIG4, ELET6 e TBLE3 apresentam retornos esperados altos e riscos baixos comparado com os outros ativos. O grupo econômico com menor risco é o de Mineração. Comparado com os retornos dos ativos listados, a ação VALE5 apresenta um retorno razoável e estes fatos justificam a sua escolha. Pode ser observado também que os setores de Petróleo e Telecomunicações possuem retornos esperados menores que o de Mineração e com maiores riscos e, por este fato, nunca seriam escolhidos para minimizar o risco da carteira. Ainda existe uma pequena alocação no setor de Siderurgia que é verificada até o portfólio eficiente. O grupo em questão possui retorno e risco altos e por isto o modelo aloca uma pequena proporção do capital. Para alcançar o nível de retorno desejado, o modelo aloca o restante do capital no ativo *Risk Free* já que o mesmo não contribui para o risco da carteira. A partir do portfólio eficiente, pode-se verificar que as alocações convergem para 70% no grupo com maior retorno e 30% no segundo grupo com maior retorno. Este fato é esperado devido à restrição de alocação máxima em grupos econômicos imposta no problema.

A seguir são mostrados os gráficos da distribuição dos log-retornos do portfólio eficiente para o desvio padrão e o desvio AVaR considerando novas simulações. O gráfico foi obtido a partir da aplicação da alocação do portfólio eficiente aos retornos verificados para cada cenário. O objetivo da análise é verificar a eficiência das alocações para novos cenários.

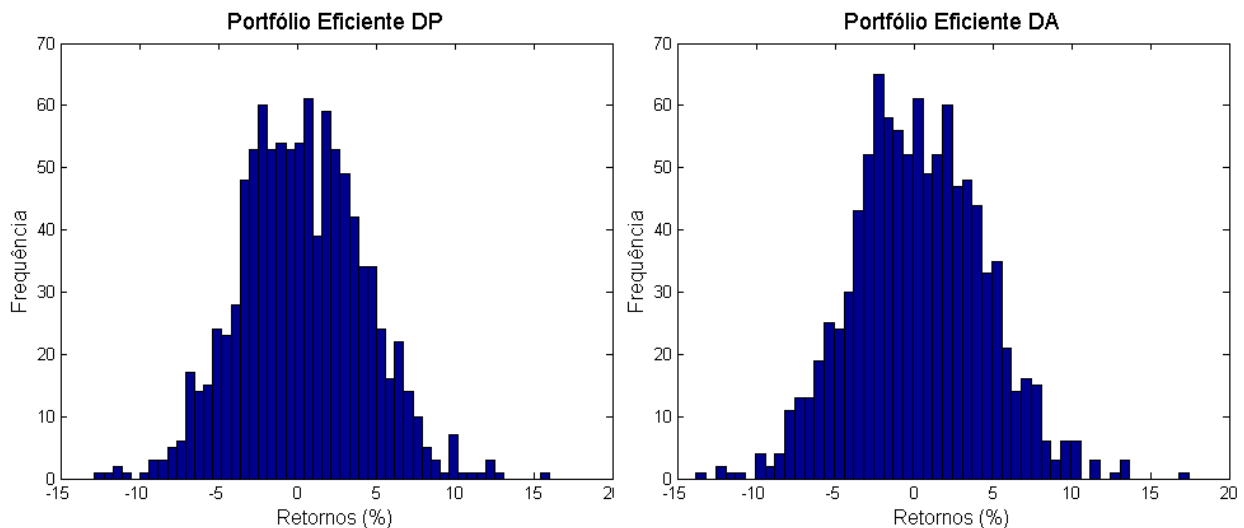


Figura 5.3: Distribuição dos log-retornos do portfólio eficiente em  $t = 1$ .

A partir da Figura 5.3, pode-se observar que as distribuições encontradas são parecidas. Utilizando o teste estatístico de Kolmogorov-Smirnov, verificou-se com um nível de 5% de significância que as distribuições são indistinguíveis. Este comportamento é esperado pelo fato das alocações propostas pelos modelos contemplarem os mesmos ativos com apenas uma ligeira mudança em seus percentuais. Ainda, o desvio padrão e o desvio AVaR para as distribuições

são próximos dos valores encontrados pelo modelo estático.

Desta forma, com base nas análises feitas e pelo fato do otimizador Gurobi indicar ter encontrado a solução ótima para o modelo estático configurado, é possível concluir que a carteira proposta por este modelo realmente apresenta risco mínimo e, portanto, é consistente.

O próximo teste foi realizado com novas simulações, considerando o último passo de tempo e um nível de confiança de 99% para o desvio AVaR. Para este teste, as variáveis de controle de risco foram configuradas da seguinte forma:  $\lambda_4 = 1$  e  $\lambda_{1,\dots,3} = 0$ . Além disso, foi considerada a abordagem *Risk Neutral*. Para este teste não será apresentado o gráfico de fronteira eficiente com a reta de alocação de capital pois as carteiras obtidas com a abordagem *Risk Neutral* não minimizam o risco da mesma e neste sentido não são consideradas eficientes. A análise será baseada na comparação das alocações propostas em relação à abordagem *Risk Averse*.

Na Figura 5.4, são mostradas as alocações sugeridas pelo modelo *Risk Averse*, gráficos superiores, e as alocações sugeridas pelo modelo *Risk Neutral*, gráficos inferiores.

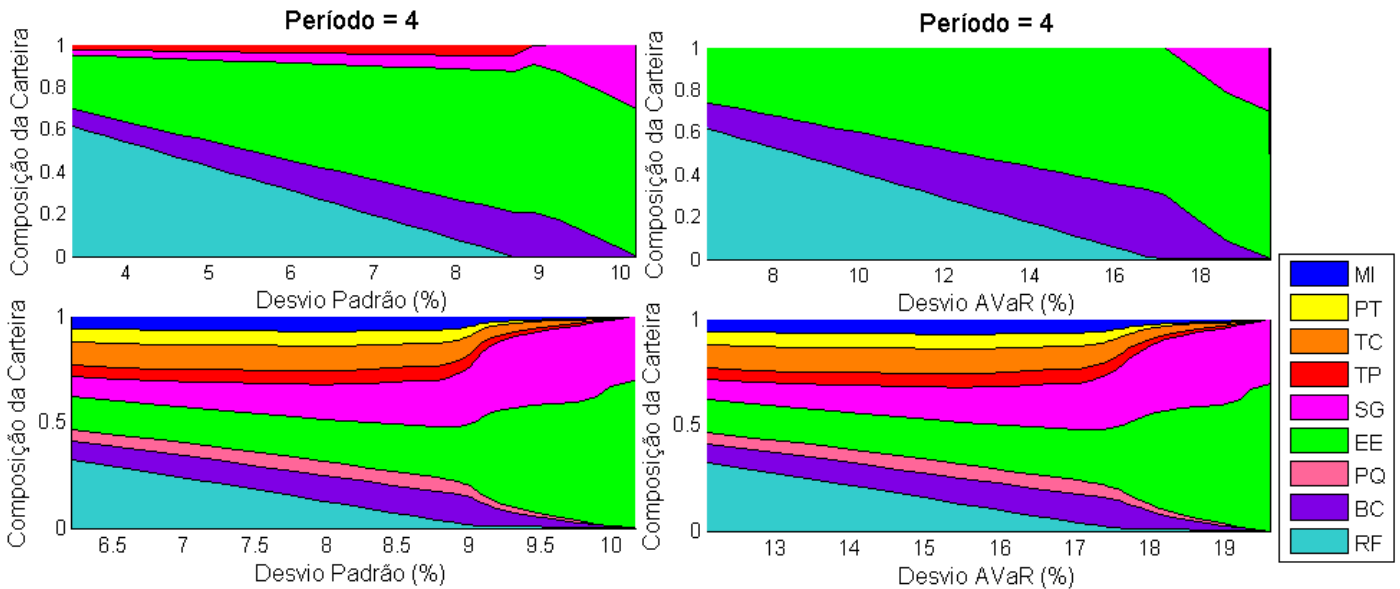


Figura 5.4: Composição da carteira através das abordagens *Risk Averse* e *Risk Neutral* em  $t = 4$ .

O modelo *Risk Neutral* consiste na resolução do problema (1.2.1) porém considerando a função objetivo igual ao retorno esperado da carteira. Além disso, a restrição de retorno mínimo para a carteira passa de uma restrição de desigualdade para uma restrição de igualdade. O intervalo de retornos utilizado para a resolução do problema foi o mesmo para as duas abordagens. Desta forma, é possível verificar que a abordagem *Risk Neutral* propõe alocações que possuem maior desvio padrão e desvio AVaR, considerando os mesmos retornos, quando comparada com a abordagem *Risk Averse*. Conforme esperado, pode-se observar que a abordagem *Risk Averse* faz mais alocações no ativo livre de risco para minimizar o risco da carteira enquanto a abordagem *Risk Neutral* não se preocupa em controlar a volatilidade da carteira. Para níveis de retorno altos, as abordagens convergem para a alocação em 70% no grupo com maior retorno e 30% no segundo grupo com maior retorno, conforme definido pela restrição de alocação máxima em grupos econômicos.

Uma observação que pode ser feita a respeito das alocações propostas através da abordagem *Risk Averse* é que, considerando o desvio padrão, o modelo propõe uma pequena alocação no grupo de Transportes e, considerando o desvio AVaR, o modelo não propõe esta alocação. O grupo em questão é composto pela ação EMBR3. Além da escolha da ação EMBR3, o modelo com desvio padrão também escolhe a ação GOAU4, que faz parte do grupo de Siderurgia, enquanto o modelo com desvio AVaR atribui essas alocações ao ativo ITUB4 do setor de Bancos. Para explicar essas alocações, foram analisadas as distribuições dos log-retornos desses ativos mostradas na Figura 5.5.

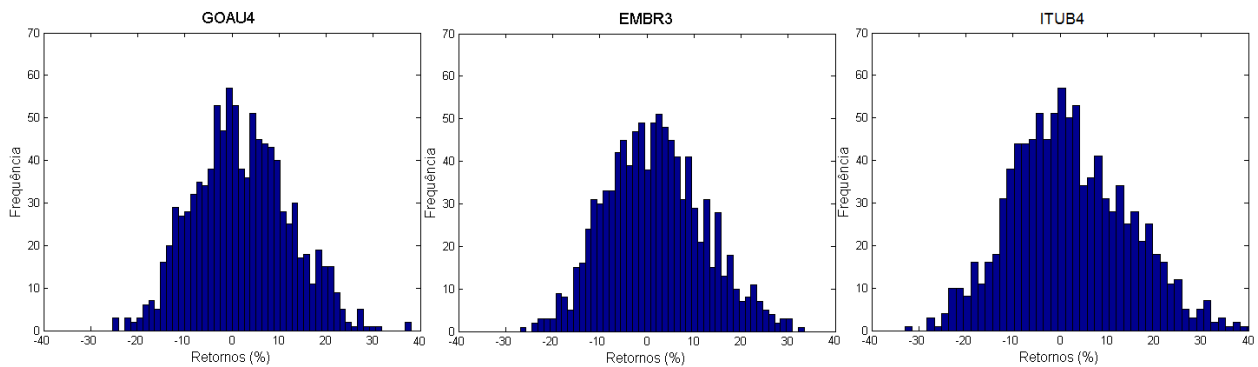


Figura 5.5: Distribuição dos log-retornos dos ativos GOAU4, EMBR3 e ITUB4 em  $t = 4$ .

Através das figuras, pode-se observar que as ações EMBR3 e GOAU4 possuem distribuições mais simétricas e desta forma são preferíveis pelo desvio padrão que procura manter a média das distribuições. Já a ação ITUB4, possui distribuição mais assimétrica com área maior para os retornos positivos. Portanto, este ativo é preferível ao considerar o desvio AVaR, visto que há mais chances de ganhos. Para os últimos níveis de retorno, ambos os modelos convergem em 70% para o ativo com maior retorno, CMIG4, e 30% para o segundo ativo com maior retorno, CSNA4, conforme esperado.

A Figura 5.6 apresenta a distribuição das carteiras propostas pelas duas abordagens considerando o primeiro nível de retorno e novos cenários. Para a carteira *Risk Averse* foram consideradas as alocações propostas para o desvio AVaR obtidas com as variáveis de controle de risco iguais a  $\lambda_4 = 1$  e  $\lambda_{1,\dots,3} = 0$ . Desta forma, é possível avaliar a performance das carteiras com novos dados *out-of-sample*.

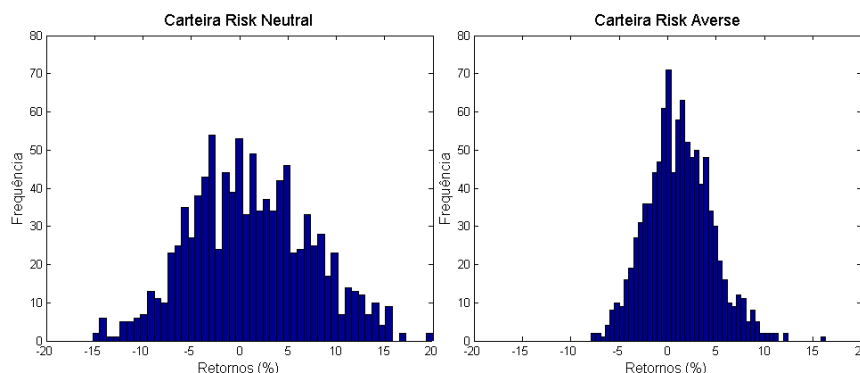


Figura 5.6: Distribuição dos log-retornos da carteira através das abordagens *Risk Averse* e *Risk Neutral* em  $t = 4$ .

A partir dos gráficos, podemos observar que a distribuição dos log-retornos considerando a alocação proposta pela abordagem *Risk Neutral* é mais dispersa comparada com a abordagem *Risk Averse* e por este fato apresenta maior risco.

A partir destas constatações e pelo fato do otimizador Gurobi indicar ter encontrado soluções ótimas para os modelos estáticos configurados, é possível concluir que as carteiras propostas para este problema são consistentes.

O último teste consiste na análise do modelo no primeiro e último tempo, com o objetivo de verificar o comportamento do risco, retorno e composição das carteiras ao longo do período através da abordagem *Risk Averse*. Para este teste foram considerados 2.500 cenários e um nível de confiança de 97,5% para o desvio AVaR. Além disso, para a análise do primeiro tempo foram consideradas as variáveis de controle de risco da seguinte forma:  $\lambda_1 = 1$  e  $\lambda_{2,\dots,4} = 0$ . Já para o último tempo, as variáveis foram configuradas com  $\lambda_4 = 1$  e  $\lambda_{1,\dots,3} = 0$ .

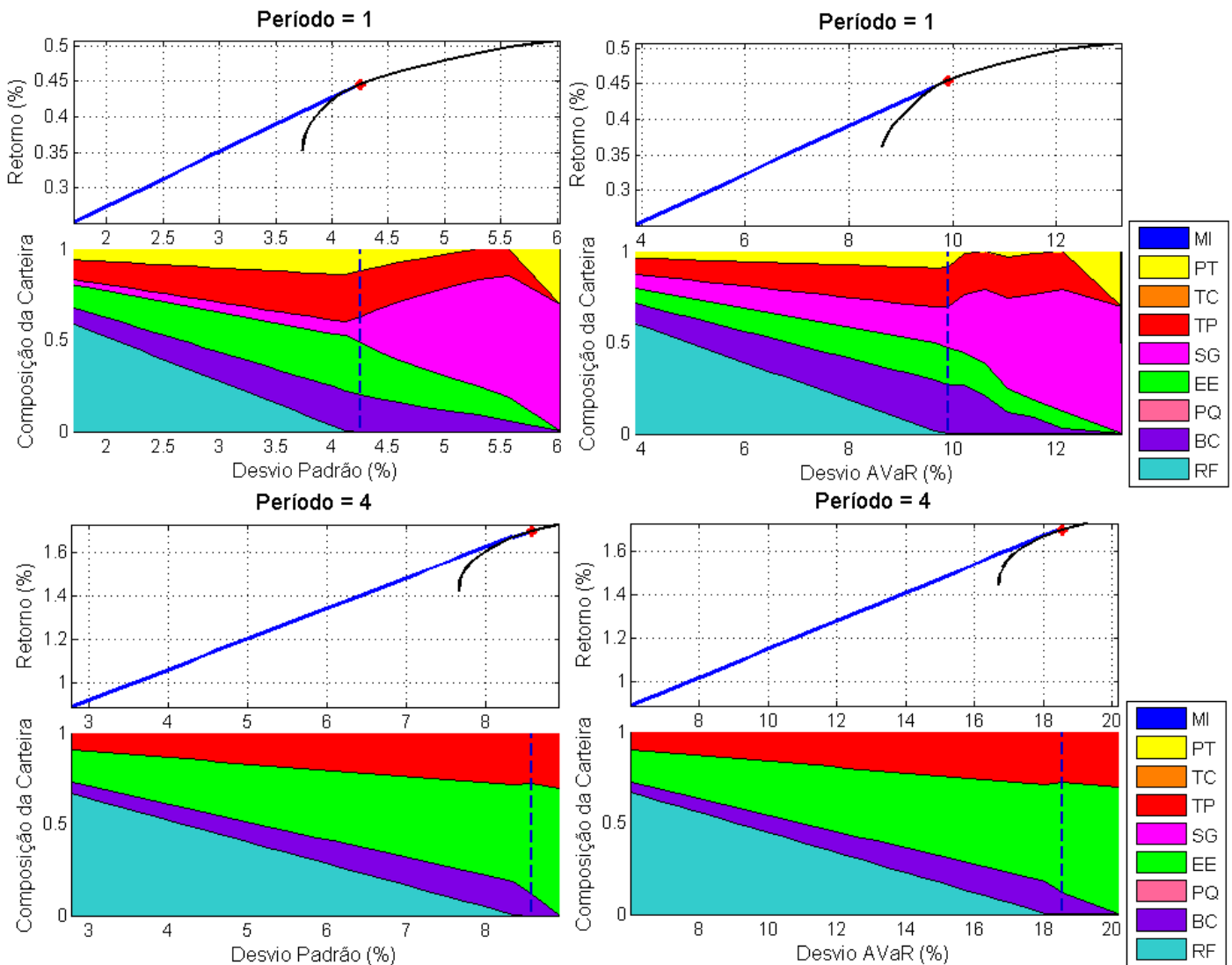


Figura 5.7: Fronteira eficiente, reta de alocação de capital e composição da carteira em  $t = 1$  e  $t = 4$ .

De maneira análoga à primeira análise, através da Figura 5.7 pode-se observar que as alocações propostas pelo modelo considerando o desvio padrão e o desvio AVaR são parecidas

no primeiro e último tempo. Além disso, verifica-se que para o último tempo são esperados retornos e riscos maiores.

Na Tabela 5.2 podem ser verificadas as informações do retorno, risco e composição do portfólio eficiente propostas pelo modelo no primeiro e último tempo.

Grupo Econômico $t = 1$	Ações	Desvio Padrão (%)	Desvio AVaR (%)	Log-Retorno Esperado (%)	Alocação PE DP (%)	Alocação PE DA (%)
[MI] Mineração	VALE5	4.49	9.81	0.28	-	-
[PT] Petróleo	PETR3	5.40	11.90	0.44	12.22	8.18
[TC] Telecomunicações	OIBR4	6.84	14.94	0.39	-	-
	TIMP3	6.49	13.91	-0.11	-	-
[TP] Transporte	EMBR3	5.14	11.72	0.42	24.66	22.17
[SG] Siderurgia	CSNA3	6.85	14.94	0.53	14.79	22.33
	USIM3	5.63	12.91	0.44	-	-
	GOAU4	6.33	13.73	0.42	-	-
[EE] Energia Elétrica	CMIG4	5.27	11.89	0.42	28.12	20.18
	CPLE6	6.20	13.63	0.30	-	-
	ELET6	5.26	11.44	-0.30	-	-
	TBLE3	5.02	11.28	0.37	-	-
[PQ] Petroquímico	BRKM5	6.31	13.63	0.44	-	-
[BC] Bancos	BBDC3	5.26	11.61	0.33	-	-
	ITUB4	4.91	11.14	0.44	20.21	27.14
[RF] Risk Free	-	-	-	0.12	-	-
Grupo Econômico $t = 4$	Ações	Desvio Padrão (%)	Desvio AVaR (%)	Log-Retorno Esperado (%)	Alocação PE DP (%)	Alocação PE DA (%)
[MI] Mineração	VALE5	9.27	19.95	1.14	-	-
[PT] Petróleo	PETR3	10.80	22.22	1.13	-	-
[TC] Telecomunicações	OIBR4	13.62	27.39	0.84	-	-
	TIMP3	12.77	25.22	-0.65	-	-
[TP] Transporte	EMBR3	10.40	21.41	1.55	27.32	27.57
[SG] Siderurgia	CSNA3	13.49	26.61	1.26	-	-
	USIM3	11.46	23.60	1.28	-	-
	GOAU4	12.89	26.24	1.30	-	-
[EE] Energia Elétrica	CMIG4	10.42	22.18	1.80	60.86	60.82
	CPLE6	12.50	25.14	0.78	-	-
	ELET6	10.56	22.64	-0.91	-	-
	TBLE3	10.12	21.97	1.42	-	-
[PQ] Petroquímico	BRKM5	12.75	25.63	1.40	-	-
[BC] Bancos	BBDC3	10.56	22.00	1.33	-	-
	ITUB4	9.80	20.61	1.50	11.82	11.61
[RF] Risk Free	-	-	-	0.50	-	-

Tabela 5.2: Risco, retorno esperado e alocação do portfólio eficiente (PE) em  $t = 1$  e  $t = 4$ .

As alocações no tempo  $t = 1$  indicam que ao investir o capital hoje, espera-se um risco e um retorno conforme proposto pelo método para a primeira semana. De maneira análoga, as alocações propostas no tempo  $t = 4$  indicam que ao investir o capital hoje, espera-se verificar um retorno e risco conforme proposto pelo método para a quarta semana. A escolha entre os tempos depende do perfil do investidor. Desta forma, caso o investidor deseje acompanhar o investimento semanalmente, deve-se escolher as alocações propostas no tempo  $t = 1$ , de acordo

com o nível de risco desejado, sendo necessário rodar novamente o programa para as alocações da semana seguinte. Caso o investidor queira se preocupar com o risco apenas uma vez por mês, deve-se escolher as alocações propostas pelo tempo  $t = 4$ .

Outro ponto que deve ser analisado para a escolha dos tempos, é que o investidor que deseja mexer periodicamente na carteira, deve levar em consideração os custos de transação. Quanto mais frequentemente se muda a composição da carteira, mais custos terá. Entretanto, melhor seria a probabilidade de acerto do modelo, pois utilizaria mais informações para a previsão.

A seguir são mostrados os gráficos da distribuição dos log-retornos do portfólio eficiente para o desvio padrão e o desvio AVaR.

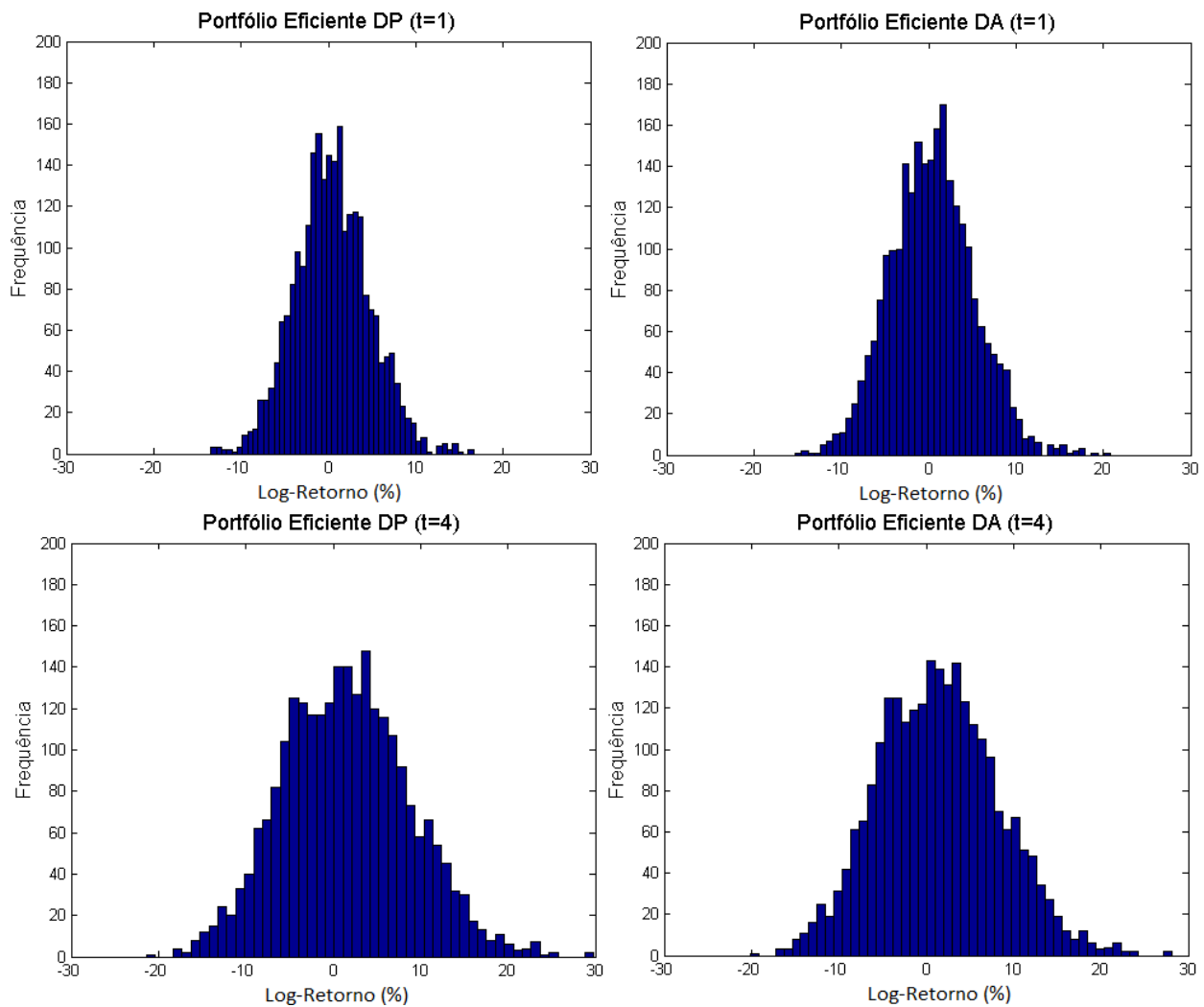


Figura 5.8: Distribuição dos log-retornos dos portfólios eficientes em  $t = 1$  e  $t = 4$ .

Analisando a Figura 5.8, pode-se verificar que as distribuições no tempo  $t = 4$  são mais dispersas, ou seja, possuem maior variância. Este fato também pode ser verificado através do Gráfico 5.7 e da Tabela 5.2. Este comportamento é esperado pois no tempo  $t = 4$  existe mais oscilação do mercado comparado a um horizonte de tempo menor. Desta forma, o modelo parece coerente ao considerar as variações dos cenários a cada tempo.

A partir das análises feitas para o modelo estático, pode-se concluir que o modelo é consistente considerando os horizontes de tempo, cenários e as abordagens *Risk Neutral* e *Risk Averse*.

### 5.3 Análise de Consistência do Modelo Proposto

Nesta seção será analisada a consistência do modelo proposto conforme feito para o modelo estático. A formulação matemática para o método de Galerkin pode ser encontrada nas Seções 4.4.1 e 4.4.2.

Para o modelo proposto, é necessário definir o comportamento dos especialistas de mercado conforme visto na Seção 4.3. Para o primeiro teste foi definido apenas um especialista com o objetivo de analisar o peso  $\alpha$  resultante que neste caso deve ser igual a 1. O especialista em questão possui o mesmo comportamento do modelo estático visto anteriormente, ou seja, deseja alocar no máximo 70% do capital em cada um dos grupos econômicos disponíveis. Desta forma, para este especialista em específico, a restrição de grupos econômicos do problema (4.3.1) teve sua desigualdade alterada de  $\geq$  para  $\leq$ . Foi considerado o primeiro passo de tempo, 1.000 simulações e um nível de confiança de 95% para o desvio AVaR.

Tanto para o problema  $\alpha$  considerando o desvio padrão na função objetivo quanto para o desvio AVaR foram obtidos valores de  $\alpha$  iguais a 1, para os diferentes níveis de retorno configurados, conforme esperado. A seguir é mostrado o gráfico da distribuição dos log-retornos encontrados pelo especialista para cada cenário.

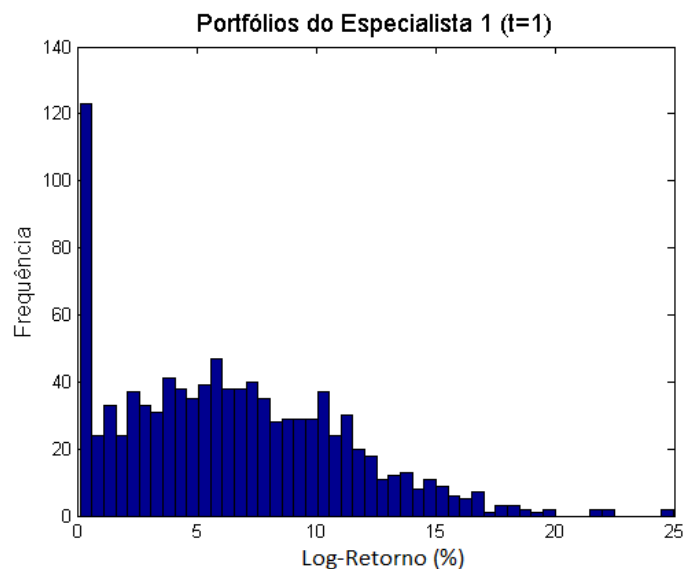


Figura 5.9: Distribuição dos log-retornos dos portfólios obtidos pelo especialista em  $t = 1$ .

Conforme definido, o especialista em questão busca maximizar o retorno da carteira em cada cenário observado e, por este fato, pode-se observar na Figura 5.9 que os retornos obtidos em cada cenário são sempre positivos. Caso os retornos esperados dos ativos para determinado cenário sejam negativos, o especialista aloca o capital no ativo livre de risco.

Para a análise de consistência com o modelo estático, foi obtida a média das alocações sugeridas pelo especialista. Com isto, é possível analisar a distribuição da carteira com a alocação média de Galerkin e comparar com a distribuição da carteira proposta pelo modelo estático.

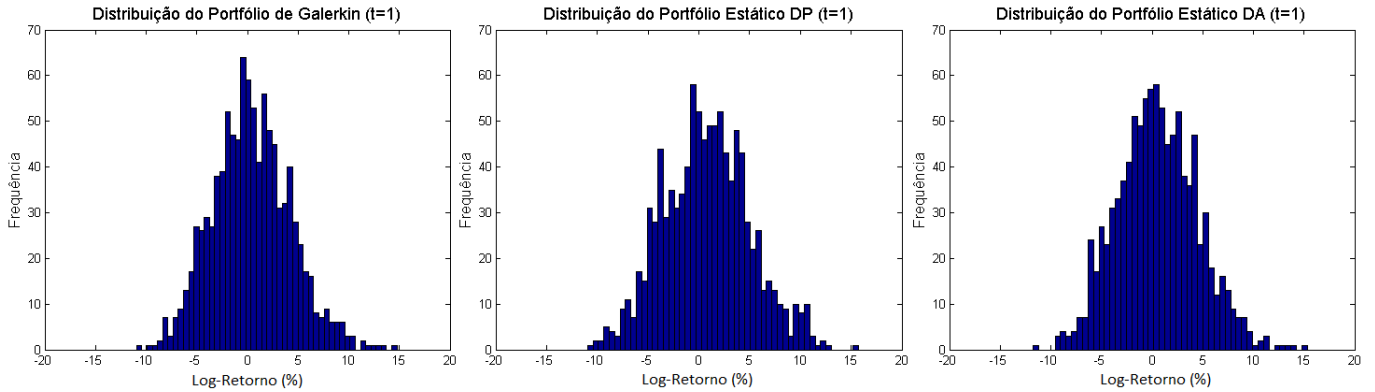


Figura 5.10: Distribuição dos log-retornos dos portfólios de Galerkin  $\times$  Estático em  $t = 1$ .

A Figura 5.10 mostra as distribuições das carteiras de Galerkin e Estáticas obtidas a partir da resolução do problema de minimização do desvio padrão e desvio AVaR. Como pode ser observado, as distribuições são muito parecidas. Utilizando o teste estatístico de Kolmogorov-Smirnov, verificou-se com um nível de 5% de significância que as distribuições são indistinguíveis. Este comportamento é esperado pelo fato do especialista ter sido configurado de acordo com as restrições do problema estático. Desta forma, o método parece ser consistente. Além disso, o otimizador Gurobi indicou ter encontrado soluções ótimas tanto para o problema do especialista quanto para o problema  $\alpha$ .

Para o próximo teste de consistência, foram definidos oito especialistas de mercado onde cada um dos especialistas possui interesse por um dos oito grupos econômicos disponíveis. O percentual mínimo de alocação no grupo de interesse de cada especialista, configurado no problema (4.3.1), será definido de três maneiras distintas (70%, 80% e 90%). O percentual máximo que poderá ser escolhido da composição de cada especialista foi configurado em 100%. O teste consiste em analisar o comportamento do vetor  $\alpha$  ao alterar o percentual mínimo dos especialistas. É esperado que para os mesmos cenários, o vetor  $\alpha$  possua valores próximos em cada coordenada, confirmando a preferência pelos mesmos especialistas.

A Figura 5.11 abaixo, mostra os valores do vetor  $\alpha$  obtidos ao resolver os problemas (4.4.4) e (4.4.5) para diferentes níveis de retorno e considerando os diferentes percentuais mínimos de alocação em grupo econômico. Pode-se verificar que para os diferentes percentuais mínimos, o modelo escolhe os mesmos especialistas com uma ligeira diferença em seus percentuais, tanto para o desvio padrão quanto para o desvio AVaR. Outro fato que pode ser observado é que considerando o mesmo percentual mínimo e retornos próximos, os  $\alpha$  encontrados pelo problema com desvio padrão e desvio AVaR são próximos, como verificado no modelo estático com as alocações nos grupos econômicos.



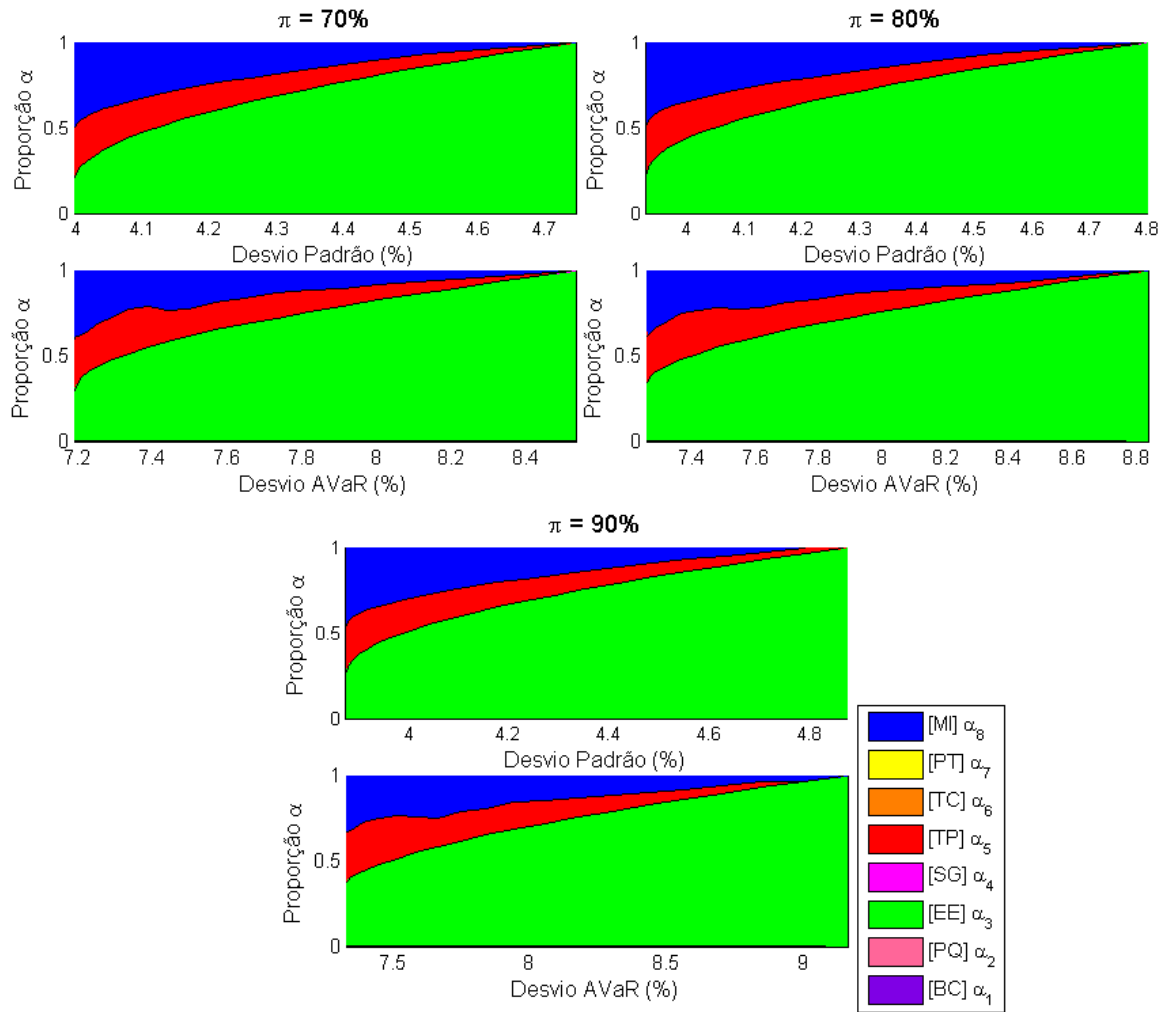


Figura 5.11: Proporção  $\alpha$  para cada nível de retorno, função de desvio e percentual mínimo.

Pelos fatos observados e também pelo otimizador Gurobi ter encontrado soluções ótimas tanto para o problema do especialista quanto para o problema  $\alpha$  conclui-se que o método é consistente ao considerar variações do percentual mínimo de alocação  $\pi$ .

## 5.4 Desempenho das Carteiras

Para a análise do desempenho das carteiras, foram definidos oito especialistas de mercado para o método de Galerkin, da mesma forma que foram definidos na análise anterior. Além dos especialistas que possuem preferência por cada um dos grupos econômicos, foi introduzida a possibilidade de escolha do ativo livre de risco diretamente no problema  $\alpha$ , (4.4.5). Desta forma, é possível construir a reta de alocação de capital juntamente com a fronteira eficiente para o problema em questão. A partir das análises de consistência do modelo estático e proposto, pôde-se verificar que as análises com o desvio AVaR são mais interessantes pelo fato do desvio padrão penalizar os retornos positivos. Com isto, as próximas análises serão feitas considerando apenas os problemas com desvio AVaR com o objetivo de diminuir a quantidade de informações e facilitar a interpretação dos resultados.

Para o primeiro teste de desempenho, o percentual mínimo de alocação no grupo de interesse, configurado no problema dos especialistas (4.3.1), foi definido em 70%. Além disso, foram considerados 1.000 cenários, nível de confiança de 97,5% e o primeiro passo de tempo com variáveis de controle de risco configuradas da seguinte forma:  $\lambda_1 = 1$  e  $\lambda_{2,\dots,4} = 0$ . O percentual máximo em cada especialista, configurado no problema  $\alpha$  (4.4.5), foi definido em 40%.

A Figura 5.12 mostra a fronteira eficiente com a reta de alocação de capital, tanto para o modelo estático quanto para o modelo proposto.

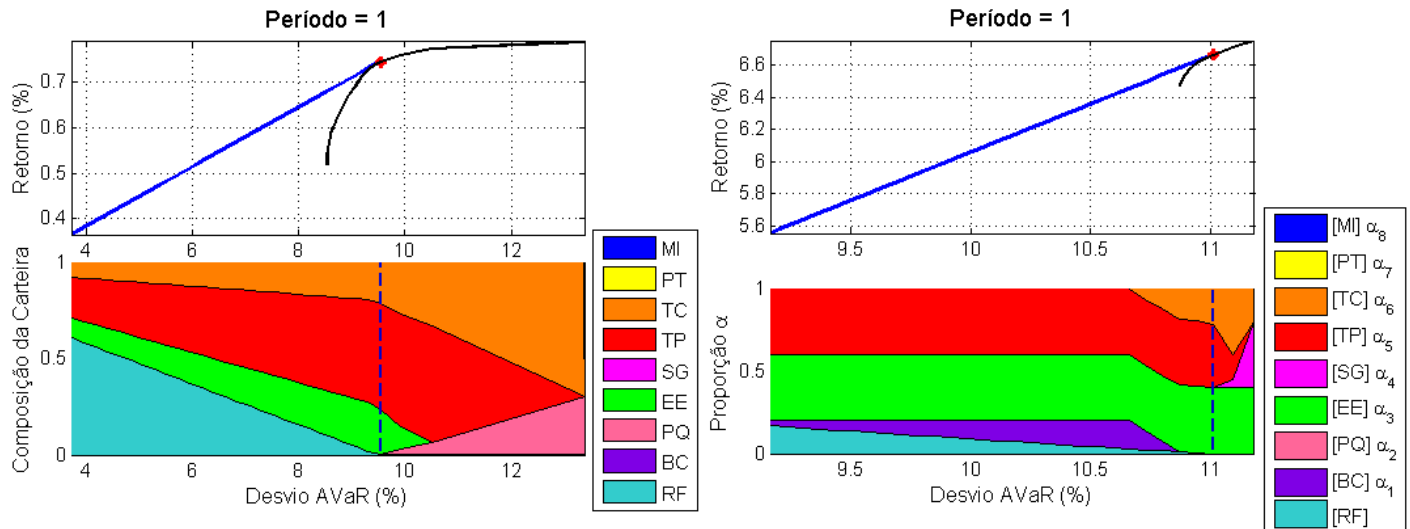


Figura 5.12: Fronteira eficiente, reta de alocação de capital e composição da carteira em  $t = 1$  para o modelo estático (esquerda) e proposto (direita).

O gráfico de composição do modelo proposto representa a proporção com que se deve investir em cada especialista de mercado definido. Na legenda do gráfico, pode-se verificar a preferência do grupo econômico de cada especialista. Conforme explicado, o ativo livre de risco é incorporado ao problema de forma que não seja necessária a consulta de especialistas. A partir da escolha do  $\alpha$  adequado, o gestor da carteira pode escolher uma abordagem para incorporá-lo à sua decisão. Uma dessas abordagens é assumir que cada especialista representa a intenção de investir em ativos do grupo econômico de interesse de maneira aproximada. Desta forma, a resposta  $\alpha_e$  de cada especialista  $e$  vai definir o percentual da carteira no grupo em questão. Esta seria uma decisão gerencial, porém clara e sempre a mesma. Tal abordagem será utilizada para a interpretação dos resultados.

A primeira observação que pode ser feita é a respeito dos retornos esperados por cada método. Pode-se verificar que os retornos obtidos pelo método proposto são muito superiores aos obtidos pelo método estático. Isto se deve ao fato do método de Galerkin construir o retorno esperado da proporção  $\alpha_e$  a partir da média dos retornos obtidos por cada especialista  $e$ . Como cada especialista busca maximizar o retorno em cada cenário, então ao tirar a média dos retornos nesses cenários, espera-se retornos altos através deste método. Já o método estático analisa apenas a média dos retornos dos ativos nos cenários.

Para a comparação do desempenho, foram utilizadas as composições dos portfólios eficientes considerando novas simulações para a análise das distribuições dos log-retornos. Desta forma, é possível analisar a performance das carteiras com novos dados *out-of-sample*, similar a um *backtesting* por simulação. Para a construção das carteiras de Galerkin, foi assumido que cada grupo econômico será representado pelo ativo com maior retorno. Então, o  $\alpha_e$  de cada especialista  $e$  será aplicado no ativo com maior retorno do grupo econômico de interesse. Para a análise das distribuições foram considerados 10.000 simulações. Com mais simulações espera-se ter distribuições mais contínuas. Os gráficos com as distribuições dos log-retornos e a tabela com as estatísticas básicas das séries são apresentados a seguir.

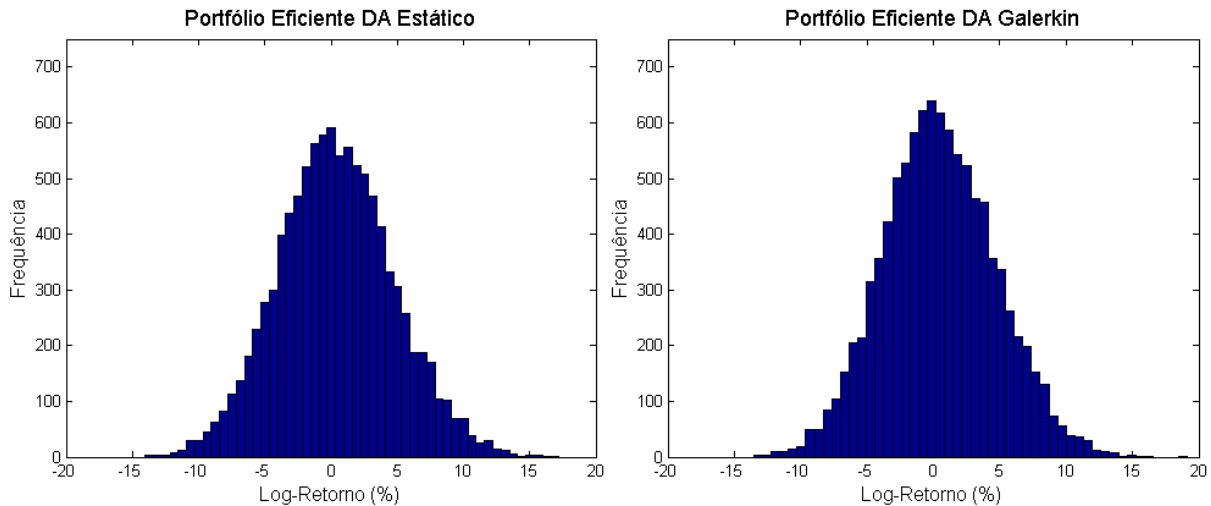


Figura 5.13: Distribuição dos log-retornos das carteiras eficientes do modelo estático  $\times$  proposto com 10.000 simulações.

	Média	Mínimo	Máximo	Desvio Padrão	Desvio AVaR	Assimetria	Curtose
Modelo Estático DA	0.38	-14.03	17.21	4.36	9.46	0.12	3.02
Modelo Galerkin DA	0.46	-13.53	19.16	4.22	9.03	0.12	2.98

Tabela 5.3: Estatísticas básicas das séries de log-retornos das carteiras com 10.000 simulações.

A partir da análise da Figura 5.13 e da Tabela 5.3, pode-se verificar que o modelo proposto é um pouco melhor pois possui maior log-retorno esperado, maior log-retorno mínimo, maior log-retorno máximo e desvios menores quando comparado ao modelo estático. Porém, as distribuições são muito parecidas e, ao aplicar o teste de Kolmogorov-Smirnov, verificou-se com um nível de 5% de significância que as distribuições são indistinguíveis. Tal fenômeno pode decorrer do fato dos especialistas serem determinísticos e dos modelos terem sido avaliados apenas com uma rodada de cenários. Desta forma, a análise não é muito conclusiva em relação à escolha de uma melhor formulação.

O próximo teste consiste em analisar o comportamento das carteiras através dos dados reais considerando as mesmas composições da análise anterior. Para a estimação dos parâmetros da distribuição dos log-retornos dos ativos e geração dos cenários, foi utilizado o período *in-sample* de 03 de Janeiro de 2000 a 17 de Outubro de 2014. Para a análise da performance das carteiras, será utilizado o período *out-of-sample* de 20 de Outubro de 2014 a 3 de Julho de 2015, totalizando 37 semanas.

A Figura 5.14 mostra os gráficos de preço e log-retorno das carteiras estática e de Galerkin no período *out-of-sample*.

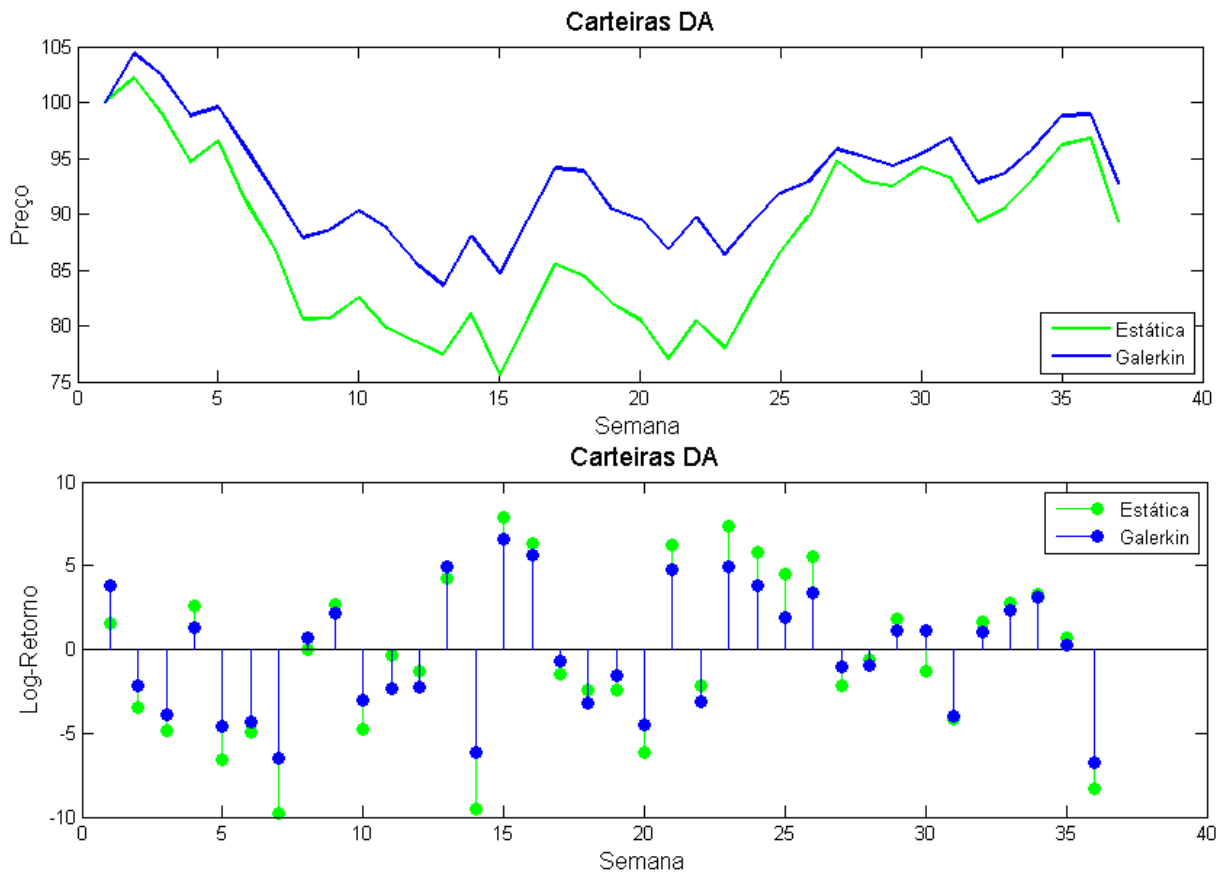


Figura 5.14: Preços e log-retornos das carteiras estática e de Galerkin no período *out-of-sample* de 37 semanas.

Analisando a Figura 5.14, verifica-se que a carteira de Galerkin possui melhores resultados comparado a carteira estática. O gráfico de preços mostra que apesar da carteira de Galerkin ter perdido valor durante o período, a mesma manteve-se a cima da carteira estática durante o período analisado. Através do gráfico dos log-retornos verifica-se que os ganhos e perdas são maiores com a carteira estática mostrando sua maior volatilidade. Na Tabela 5.4 são mostradas algumas informações sobre as séries dos log-retornos e preços das carteiras analisadas.

Estatísticas	Carteira Estática	Carteira Galerkin
Maior Log-Retorno	7.87%	6.55%
Menor Log-Retorno	-9.78%	-6.71%
# Log-Retornos Positivos	17	18
# Log-Retornos Negativos	19	18
Máximo Drawdown	-26.01%	-19.92%
—> # Log-Retornos	13	11
Log-Retorno Acumulado	-11.70%	-7.95%
Volatilidade Semanal Anualizada	33.45%	26.67%

Tabela 5.4: Estatísticas das séries de log-retornos das carteiras para o período de 37 semanas.

O Máximo Drawdown é a maior perda da carteira fixado um período de tempo, ou seja, é a perda medida a partir do seu valor máximo até o seu valor mínimo dentro do período. No

caso observado este período é de 37 semanas. A Volatilidade Semanal Anualizada foi calculada considerando 52 semanas. Analisando a Tabela 5.4 pode-se verificar que a carteira de Galerkin possui um melhor desempenho considerando tais estatísticas.

O próximo teste consiste na análise das carteiras DA considerando uma janela rolante de dados para os modelos. Para este teste foi considerado o período *in-sample* de 120 semanas de 02 de Janeiro de 2013 a 30 de Abril de 2015 e o período *out-of-sample* de 9 semanas de 08 de Maio de 2015 a 03 de Julho de 2015. A partir da primeira previsão, as observações subsequentes foram sucessivamente sendo incorporadas ao modelo, e as observações mais antigas foram sendo desconsideradas. Em cada rodada do programa, as alocações do portfólio eficiente dos modelos estático e proposto foram armazenadas para a análise da performance com os dados *out-of-sample*.

A Figura 5.15 mostra o gráfico de log-retornos das carteiras estática e de Galerkin no período *out-of-sample* e a Tabela 5.5 suas estatísticas. Como para este teste as alocações propostas mudam a cada semana, o gráfico de preços não será apresentado.

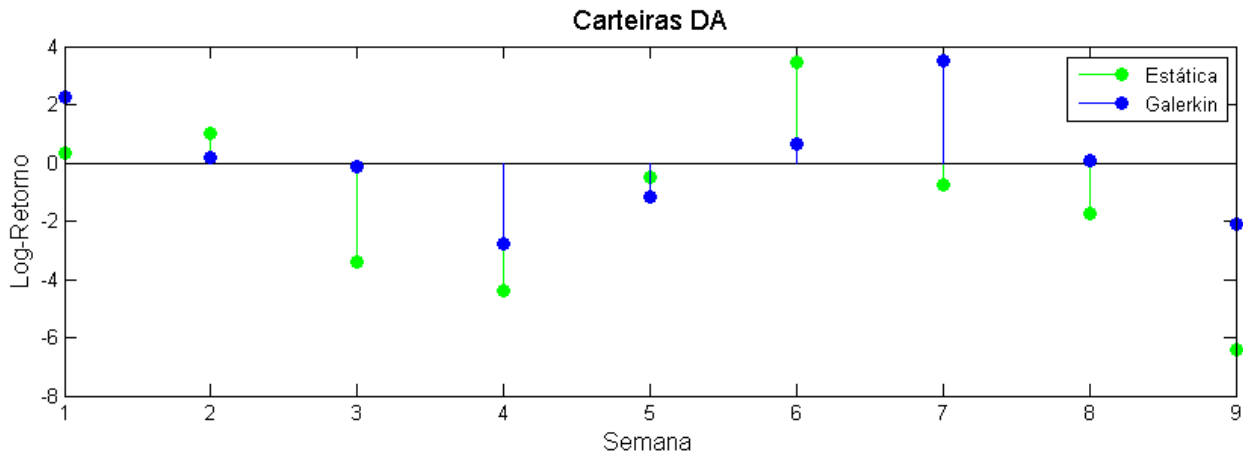


Figura 5.15: Preços e log-retornos das carteiras estática e de Galerkin no período *out-of-sample* de 9 semanas.

Estatísticas	Carteira Estática	Carteira Galerkin
Maior Log-Retorno	3.45%	3.48%
Menor Log-Retorno	-6.42%	-2.79%
# Log-Retornos Positivos	3	5
# Log-Retornos Negativos	6	4
Máximo Drawdown	-13.31%	-4.05%
—> # Log-Retornos	7	3
Log-Retorno Acumulado	-12.47%	0.48%
Volatilidade Semanal Anualizada	21.66%	14.25%

Tabela 5.5: Estatísticas das séries de log-retornos das carteiras para o período de 9 semanas.

A partir da análise gráfica e da tabela, pode-se concluir que as carteiras de Galerkin performam melhor do que as carteiras estáticas no período analisado, deixando visível a vantagem de se incorporar conhecimentos de mercado ao problema de otimização de carteiras.

Foram feitos testes de análise de desempenho das carteiras no tempo  $t = 4$ . Para a análise foram considerados nível de confiança para o desvio AVaR igual a 99%, 1.000 simulações e variáveis de controle de risco iguais a  $\lambda_4 = 1$  e  $\lambda_{1,\dots,3} = 0$ . A proporção máxima em cada especialista, configurada no problema (4.4.5), foi definida em 40%. O percentual mínimo de alocação no grupo de interesse de cada especialista, configurado no problema (4.3.1), foi definido em 80%. Ao utilizar as análises anteriores para a comparação das carteiras foram obtidos resultados semelhantes, confirmando a melhor performance do método de Galerkin quando comparado com uma abordagem estática.

Como última análise, foram observados os tempos computacionais médios para a resolução dos problemas a partir de cada método. Para a avaliação dos tempos, foram considerados o número de cenários e o número de especialistas, no caso do método de Galerkin. A Tabela 5.6 apresenta os resultados em segundos.

Modelos	Especialistas	Cenários		
		1.000	2.500	5.000
Estático DP	-	4.16	4.25	4.82
Estático DA	-	88.35	299.83	1235.77
Galerkin DP	1	11.08	26.62	52.90
Galerkin DA	1	13.61	45.20	934.56
Galerkin DP	4	37.11	90.63	200.02
Galerkin DA	4	37.85	121.91	1811.59
Galerkin DP	8	93.71	211.13	427.98
Galerkin DA	8	93.55	310.28	3080.59

Tabela 5.6: Tempos computacionais dos métodos estático e de Galerkin em segundos.

Como pode ser verificado, ambos os métodos possuem uma oscilação grande de tempo quando considerada a função DA e a quantidade de cenários. Isto se deve ao fato da variável  $v_{k,e}^t$ , que representa o retorno de cada especialista  $e$  no cenário  $k$  e tempo  $t$ , do problema DA (4.4.5) ser avaliada  $K \times T \times E$  vezes, onde  $K$  é o número de cenários,  $T$  é o número de tempos e  $E$  o número de especialistas. Para o método estático com desvio DA, este fato é similar dado que a variável  $H(x^t(\xi_k), r^t(\xi_k))$  do problema DA (3.5.5), que representa o retorno em cada cenário  $k$  e tempo  $t$ , é avaliada  $K \times T$  vezes. Considerando o problema DP, os tempos são mais estáveis em cada método.

Para o caso com oito especialistas, utilizado neste trabalho, observa-se que o problema DA com Galerkin possui tempos superiores quando comparado com sua versão estática. Apesar dos tempos obtidos com o enfoque de Galerkin serem maiores, ficam sempre inferiores à uma hora considerando uma plataforma sem compilação, como o MATLAB, e um *notebook*. É de se esperar melhoras significativas ao implementar tal método em linguagens mais avançadas, como o C++.

# Capítulo 6

## Considerações Finais

O objetivo deste trabalho de dissertação foi estudar o método de Galerkin, proposto por Koivu e Pennanen [1], como uma nova abordagem para a resolução do problema de carteira ótima. Além disso, o presente trabalho analisou aspectos teóricos e de implementação do modelo em questão.

A partir dos experimentos numéricos realizados, verificou-se que o modelo estocástico proposto foi capaz de utilizar as informações exógenas, na forma de visões de mercado, de forma eficiente, obtendo alocações com maior retorno quando comparado com sua versão determinística equivalente. O método proposto permite que tais visões de mercado sejam configuradas de acordo com as preferências do investidor. Para o caso de uso deste trabalho foram configurados especialistas de mercado a partir de problemas de otimização determinísticos. Estes especialistas podem ser configurados de maneira mais dinâmica, como por exemplo, através de cenários em pente. Desta forma, outra vantagem obtida pelo método é a flexibilidade de parametrização da dinâmica das incertezas inerentes ao problema.

Para a avaliação dos resultados, analisaram-se as carteiras propostas pelo método de Galerkin e o método estático através de testes de desempenho com dados simulados e dados reais. Para a definição da carteira do método de Galerkin, foi assumido que a proporção  $\alpha_e$  ótima representa a intenção de investir em ativos do grupo econômico de interesse do especialista  $e$  de tal forma que a proporção  $\alpha_e$  definisse o percentual da carteira no grupo em questão. Neste sentido, o método proposto também mostra-se flexível ao deixar a critério do investidor a abordagem para a definição da carteira. Outra possibilidade seria interpretar a proporção  $\alpha_e$  ótima como a proporção a ser investida na carteira obtida com as alocações médias do especialista  $e$ .

Em relação às funções objetivo avaliadas, a análise da Seção 5.2 confirma a superioridade do desvio AVaR quando comparado com o desvio padrão. Isto se deve ao fato do desvio padrão penalizar desvios da média tanto por baixo quanto por cima enquanto o desvio AVaR penaliza apenas desvios unilaterais.

Com relação ao tempo computacional, dependendo da quantidade de especialistas configurados para o problema, o método proposto pode apresentar tempo maior de processamento quando comparado com o método estático. Este resultado é esperado, dado que a versão estática não prevê qualquer dinâmica para o tratamento das incertezas e, por este fato, possui um tempo computacional menor. Conforme verificado por Koivu e Pennanen [1], comparado a métodos dinâmicos o método de Galerkin possui vantagens computacionais, dado que o mesmo reduz o conjunto de soluções viáveis do problema.

Como possibilidade de trabalhos futuros, o método de Galerkin pode ser aplicado ao problema de alocação ótima de ativos e comparado com métodos de otimização estocástica multi-estágios. Neste caso, teríamos uma árvore de cenários, que seria explorada à medida que a informação se revelasse e as decisões de portfólio fossem tomadas. A otimização desses modelos é complexa e necessita de técnicas especiais. Com este enfoque, poderiam ser implementados especialistas em pente com horizontes rolantes para o método de Galerkin. Desta forma, o modelo proposto incorpora uma dinâmica na tomada de decisões que fica mais próxima de uma formulação multi-estágio para comparação. Outra possibilidade seria a aplicação do método em outras áreas como, por exemplo, no setor de energia para o gerenciamento das fontes de energia e custos associados.



# Bibliografia

- [1] M Koivu e T Pennanen. “Galerkin methods in dynamic stochastic programming”. Em: *OPTIMIZATION* 59.3 (2010), pp. 339–354.
- [2] Alexander Shapiro, Darinka Dentcheva et al. *Lectures on stochastic programming: modeling and theory*. Vol. 9. SIAM, 2009.
- [3] Harry Markowitz. “Harry M. Markowitz”. Em: *Portfolio selection, Journal of Finance* 7.1 (1952), pp. 77–91.
- [4] M Mokhtar, A Shuib e D Mohamad. “Mathematical programing models for portfolio optimization problem: A Review”. Em: *Int. J. Social, Human Sci. Eng* 8 (2014), pp. 411–420.
- [5] Hiroshi Konno e Hiroaki Yamazaki. “Mean-absolute deviation portfolio optimization model and its applications to Tokyo stock market”. Em: *Management science* 37.5 (1991), pp. 519–531.
- [6] Martin R Young. “A minimax portfolio selection rule with linear programming solution”. Em: *Management science* 44.5 (1998), pp. 673–683.
- [7] Christos Papahristodoulou e Erik Dotzauer. “Optimal portfolios using linear programming models”. Em: *Journal of the Operational research Society* 55.11 (2004), pp. 1169–1177.
- [8] Anton Abdulbasah Kamil e Khalipah Ibrahim. “Mean-absolute deviation portfolio optimization problem”. Em: *Journal of Information and Optimization Sciences* 28.6 (2007), pp. 935–944.
- [9] Hiroshi Konno e Rei Yamamoto. “Integer programming approaches in mean-risk models”. Em: *Computational Management Science* 2.4 (2005), pp. 339–351.
- [10] Luca Chiodi, Renata Mansini e Maria Grazia Speranza. “Semi-absolute deviation rule for mutual funds portfolio selection”. Em: *Annals of Operations Research* 124.1-4 (2003), pp. 245–265.
- [11] Stefano Benati e Romeo Rizzi. “A mixed integer linear programming formulation of the optimal mean/value-at-risk portfolio problem”. Em: *European Journal of Operational Research* 176.1 (2007), pp. 423–434.
- [12] R Tyrrell Rockafellar, Stanislav P Uryasev e Michael Zabaranin. “Deviation measures in risk analysis and optimization”. Em: *University of Florida, Department of Industrial & Systems Engineering Working Paper 2002-7* (2002).
- [13] Enrico Angelelli, Renata Mansini e M Grazia Speranza. “A comparison of MAD and CVaR models with real features”. Em: *Journal of Banking & Finance* 32.7 (2008), pp. 1188–1197.

- [14] Hamed Soleimani, Hamid Reza Golmakani e Mohammad Hossein Salimi. “Markowitz-based portfolio selection with minimum transaction lots, cardinality constraints and regarding sector capitalization using genetic algorithm”. Em: *Expert Systems with Applications* 36.3 (2009), pp. 5058–5063.
- [15] Hamid Reza Golmakani e Mehrshad Fazel. “Constrained portfolio selection using particle swarm optimization”. Em: *Expert Systems with Applications* 38.7 (2011), pp. 8327–8335.
- [16] K Ibrahim, AA Kamil e A Mustafa. “Portfolio selection problem with maximum downside deviation measure: a stochastic programming approach”. Em: *Int. J. Math. Models Methods Appl. Sci* 1.2 (2008), pp. 123–129.
- [17] Philippe Artzner et al. “Coherent measures of risk”. Em: *Mathematical finance* 9.3 (1999), pp. 203–228.
- [18] Alexander Shapiro. “Topics in Stochastic Programming”. Em: *CORE Lecture Series, Université Catholique de Louvain* (2011).
- [19] *Bloomberg do Brasil*. URL: <http://www.bloomberg.com.br/> (acesso em 20/10/2014).
- [20] Attilio Meucci. *Risk and asset allocation*. Springer, 2009.
- [21] *BM&FBovespa*. URL: <http://www.bmfbovespa.com.br/indices/ResumoIndice.aspx?Indice=Ibovespa&Idioma=pt-br> (acesso em 20/10/2014).
- [22] *Symmys - Advanced Risk and Portfolio Management*. URL: <http://www.symmys.com/node/70> (acesso em 18/11/2014).
- [23] Dessislava Pachamanova e Frank J Fabozzi. *Simulation and optimization in finance: modeling with MATLAB, @ RISK, or VBA*. Vol. 173. John Wiley & Sons, 2010.
- [24] Alexander J McNeil, Rüdiger Frey e Paul Embrechts. *Quantitative risk management: concepts, techniques, and tools*. Princeton university press, 2010.
- [25] Jeroen Kerkhof e Bertrand Melenberg. “Backtesting for risk-based regulatory capital”. Em: *Journal of Banking & Finance* 28.8 (2004), pp. 1845–1865.
- [26] R Tyrrell Rockafellar e Stanislav Uryasev. “Optimization of conditional value-at-risk”. Em: *Journal of risk* 2 (2000), pp. 21–42.
- [27] Georg Ch Pflug e Werner Römisch. *Modeling, measuring and managing risk*. World Scientific, 2007.
- [28] Edwin J Elton et al. *Modern portfolio theory and investment analysis*. John Wiley & Sons, 2009.
- [29] Robert G Bartle. *The elements of integration and Lebesgue measure*. Vol. 92. John Wiley & Sons, 2011.
- [30] Steven E Shreve. *Stochastic calculus for finance II: Continuous-time models*. Vol. 11. Springer, 2004.

# Apêndice A

## Representações Matriciais

Neste apêndice encontram-se as representações matriciais necessárias para a implementação dos problemas propostos nos capítulos 3 e 4.

### A.1 Modelo Determinístico Equivalente com Desvio Padrão

Para a utilização da função *quadprog* do otimizador Gurobi é necessário escrever o problema  $\alpha$  apresentado em (3.5.3) no formato matricial a seguir:

$$\begin{cases} \min_{y \in Y} & y^T \Sigma y \\ & A_{eq} y = b_{eq} \\ & Ay \leq b \\ & lb \leq y \leq ub, \end{cases}$$

onde,

$$\Sigma = \begin{bmatrix} \lambda_1 \Sigma^1_{N \times N} & \mathbf{0}_{N \times N} & \dots & \mathbf{0}_{N \times N} \\ \mathbf{0}_{N \times N} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0}_{N \times N} \\ \mathbf{0}_{N \times N} & \dots & \mathbf{0}_{N \times N} & \lambda_T \Sigma^T_{N \times N} \end{bmatrix}_{(N \times T) \times (N \times T)}, \quad y = \begin{bmatrix} x_0^1 \\ \vdots \\ x_N^1 \\ \vdots \\ x_1^T \\ \vdots \\ x_N^T \end{bmatrix}_{(N \times T) \times 1},$$

Neste caso, as matrizes  $\Sigma^t$  representam as matrizes de covariância entre os retornos simulados no tempo  $t$ . Considerando o ativo livre de risco, a primeira linha e coluna da matriz  $\Sigma^t$  são definidas com zeros representando que o ativo livre de risco possui covariância zero com todos os demais ativos.

$$lb = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{(N \times T) \times 1}, \quad ub = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}_{(N \times T) \times 1}$$

$$A_{eq} = \begin{bmatrix} \mathbf{1}_{1 \times N} & \mathbf{0}_{1 \times N} & \cdots & \mathbf{0}_{1 \times N} \\ \mathbf{0}_{1 \times N} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0}_{1 \times N} \\ \mathbf{0}_{1 \times N} & \cdots & \mathbf{0}_{1 \times N} & \mathbf{1}_{1 \times N} \end{bmatrix}_{T \times (N \times T)}, \quad b_{eq} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}_{T \times 1}$$

$$A = \begin{bmatrix} -\mu^1 & \mathbf{0}_{1 \times N} & \cdots & \mathbf{0}_{1 \times N} \\ \mathbf{0}_{1 \times N} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0}_{1 \times N} \\ \mathbf{0}_{1 \times N} & \cdots & \mathbf{0}_{1 \times N} & -\mu^T \\ S & \mathbf{0}_{G \times N} & \cdots & \mathbf{0}_{G \times N} \\ \mathbf{0}_{G \times N} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0}_{G \times N} \\ \mathbf{0}_{G \times N} & \cdots & \mathbf{0}_{G \times N} & S \end{bmatrix}_{(T+T \times G) \times (N \times T)}, \quad b = \begin{bmatrix} -\tau^1 \\ \vdots \\ -\tau^T \\ \pi_1^1 \\ \vdots \\ \pi_1^T \\ \vdots \\ \pi_G^1 \\ \vdots \\ \pi_G^T \end{bmatrix}_{(T+T \times G) \times 1}.$$

Neste caso, a matriz  $\mu^t$  é definida como:

$$\mu^t = \begin{bmatrix} r_0^t & \frac{1}{K} \sum_{k=1}^K r_1^t(\xi_k) & \cdots & \frac{1}{K} \sum_{k=1}^K r_N^t(\xi_k) \end{bmatrix}_{1 \times N}$$

A matriz  $S$  representa a restrição associada aos grupos econômicos e depende da quantidade de ativos em cada um desses grupos. Desta forma, os elementos da matriz  $S$  podem ser definidos da seguinte forma:

$$s_{g,i} = \begin{cases} 1, & \text{se } i \in P_g \\ 0, & \text{caso contrário} \end{cases},$$

onde,  $i = 1, \dots, N$ ,  $g = 1, \dots, G$  e  $P_g$  é o conjunto dos ativos do grupo econômico  $g$ .

## A.2 Modelo Determinístico Equivalente com Desvio *Average Value at Risk*

Para a utilização da função *linprog* do otimizador Gurobi é necessário escrever o problema determinístico equivalente apresentado em (3.5.5) no formato matricial a seguir:

$$\begin{cases} \min_{y \in Y} f^T y \\ A_{eq} y = b_{eq} \\ Ay \leq b \\ lb \leq y \leq ub, \end{cases}$$

onde,

$$f = \begin{bmatrix} r_0^1 \\ \frac{1}{K} \sum_{k=1}^K r_1^1(\xi_k) \\ \vdots \\ \frac{1}{K} \sum_{k=1}^K r_N^1(\xi_k) \\ \vdots \\ \frac{1}{K} \sum_{k=1}^K r_N^T(\xi_k) \\ -\lambda_1 \\ \vdots \\ -\lambda_T \\ \frac{\lambda_1}{K(1-\alpha^1)} \mathbf{1}_{K \times 1} \\ \vdots \\ \frac{\lambda_T}{K(1-\alpha^T)} \mathbf{1}_{K \times 1} \end{bmatrix}_{(N \times T + T + K \times T) \times 1}, \quad y = \begin{bmatrix} x_0^1 \\ \vdots \\ x_N^1 \\ \vdots \\ x_1^T \\ \vdots \\ x_N^T \\ a^1 \\ \vdots \\ a^T \\ b_1^1 \\ \vdots \\ b_k^1 \\ \vdots \\ b_1^t \\ \vdots \\ b_k^t \end{bmatrix}_{(N \times T + T + K \times T) \times 1}$$

$$lb = \begin{bmatrix} \mathbf{0}_{(N \times T) \times 1} \\ -\infty_{T \times 1} \\ \mathbf{0}_{(K \times T) \times 1} \end{bmatrix}_{(N \times T + T + K \times T) \times 1}, \quad ub = \begin{bmatrix} \mathbf{1}_{(N \times T) \times 1} \\ \infty_{T \times 1} \\ \infty_{(K \times T) \times 1} \end{bmatrix}_{(N \times T + T + K \times T) \times 1}, \quad b_{eq} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}_{T \times 1}$$

$$A_{eq} = \begin{bmatrix} \mathbf{1}_{1 \times N} & \mathbf{0}_{1 \times N} & \cdots & \mathbf{0}_{1 \times N} & \mathbf{0}_{1 \times (T + K \times T)} \\ \mathbf{0}_{1 \times N} & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0}_{1 \times N} & \vdots \\ \mathbf{0}_{1 \times N} & \cdots & \mathbf{0}_{1 \times N} & \mathbf{1}_{1 \times N} & \mathbf{0}_{1 \times (T + K \times T)} \end{bmatrix}_{T \times (N \times T + T + K \times T)}, \quad b = \begin{bmatrix} -\tau^1 \\ \vdots \\ -\tau^T \\ \pi_1^1 \\ \vdots \\ \pi_G^T \\ \mathbf{0}_{(T \times K) \times 1} \end{bmatrix}_{(T + G \times T + K \times T) \times 1}$$

$$A = \begin{bmatrix} \begin{bmatrix} -\mu \end{bmatrix} & \mathbf{0}_{T \times T} & \mathbf{0}_{T \times (K \times T)} \\ \begin{bmatrix} G \end{bmatrix} & \mathbf{0}_{T \times T} & \mathbf{0}_{T \times (K \times T)} \\ \begin{bmatrix} -\mu^1(\xi) \end{bmatrix} & \begin{bmatrix} 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \dots & 0 \end{bmatrix} & \begin{bmatrix} -1 & 0 & \dots & \dots & \dots & \dots & 0 \\ \vdots & \ddots & \ddots & \dots & \dots & \dots & \vdots \\ 0 & \dots & -1 & 0 & \dots & \dots & 0 \end{bmatrix} \\ \vdots & \vdots & \vdots \\ \begin{bmatrix} -\mu^T(\xi) \end{bmatrix} & \begin{bmatrix} 0 & \dots & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & 1 \end{bmatrix} & \begin{bmatrix} 0 & \dots & \dots & \dots & -1 & \dots & 0 \\ \vdots & \dots & \dots & \dots & \dots & \ddots & \vdots \\ 0 & \dots & \dots & \dots & \dots & \dots & -1 \end{bmatrix} \end{bmatrix}_{(T+G \times T+K \times T) \times (N \times T+T+K \times T)}$$

Neste caso, as matrizes  $\mu$  e  $\mu^t(\xi)$  para  $t = 1, \dots, T$  são definidas da seguinte maneira:

$$\mu = \begin{bmatrix} \mu^1 & \mathbf{0}_{1 \times N} & \dots & \mathbf{0}_{1 \times N} \\ \mathbf{0}_{1 \times N} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0}_{1 \times N} \\ \mathbf{0}_{1 \times N} & \dots & \mathbf{0}_{1 \times N} & \mu^T \end{bmatrix}_{T \times (N \times T)}, \quad \mu^t = \begin{bmatrix} r_0^t & \frac{1}{K} \sum_{k=1}^K r_1^t(\xi_k) & \dots & \frac{1}{K} \sum_{k=1}^K r_N^t(\xi_k) \end{bmatrix}_{1 \times N}$$

$$\mu^t(\xi) = \begin{bmatrix} \mu^t(\xi_1) & \mathbf{0}_{K \times N} & \dots & \mathbf{0}_{K \times N} \\ \mathbf{0}_{K \times N} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0}_{K \times N} \\ \mathbf{0}_{K \times N} & \dots & \mathbf{0}_{K \times N} & \mu^t(\xi_K) \end{bmatrix}_{(K \times T) \times (N \times T)}, \quad \mu^t(\xi) = \begin{bmatrix} r_0^t & r_1^t(\xi_1) & \dots & r_N^t(\xi_1) \\ \vdots & \vdots & \dots & \vdots \\ r_0^t & r_1^t(\xi_K) & \dots & r_N^t(\xi_K) \end{bmatrix}_{K \times N}$$

A matriz  $G$  representa a restrição associada aos grupos econômicos e depende da quantidade de ativos em cada um desses grupos.

$$G = \begin{bmatrix} S & \mathbf{0}_{G \times N} & \dots & \mathbf{0}_{G \times N} \\ \mathbf{0}_{G \times N} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0}_{G \times N} \\ \mathbf{0}_{G \times N} & \dots & \mathbf{0}_{G \times N} & S \end{bmatrix}_{(G \times T) \times (N \times T)}$$

Desta forma, os elementos da matriz  $S$  podem ser definidos da seguinte forma:

$$s_{g,i} = \begin{cases} 1, & \text{se } i \in P_g \\ 0, & \text{caso contrário} \end{cases},$$

onde,  $i = 1, \dots, N$ ,  $g = 1, \dots, G$  e  $P_g$  é o conjunto dos ativos do grupo econômico  $g$ .

### A.3 Problema $\alpha$ com Desvio Padrão

Para a utilização da função *quadprog* do otimizador Gurobi é necessário escrever o problema  $\alpha$  apresentado em (4.4.4) no formato matricial a seguir:

$$\begin{cases} \min_{y \in Y} & y^T \Sigma y \\ & A_{eq} y = b_{eq} \\ & Ay \leq b \\ & lb \leq y \leq ub, \end{cases}$$

onde,

$$\Sigma = \begin{bmatrix} \sum_{t=1}^T \lambda_t cov(V_{k,1}^t, V_{k,1}^t) & \cdots & \sum_{t=1}^T \lambda_t cov(V_{k,1}^t, V_{k,E}^t) \\ \vdots & \ddots & \vdots \\ \sum_{t=1}^T \lambda_t cov(V_{k,E}^t, V_{k,1}^t) & \cdots & \sum_{t=1}^T \lambda_t cov(V_{k,E}^t, V_{k,E}^t) \end{bmatrix}_{E \times E}, \quad y = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_E \end{bmatrix}_{E \times 1},$$

$$lb = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{E \times 1}, \quad ub = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}_{E \times 1}$$

$$A_{eq} = [1 \ 1 \ \cdots \ 1]_{1 \times E}, \quad b_{eq} = [1]_{1 \times 1}$$

$$A = - \begin{bmatrix} \frac{1}{K} \sum_{k=1}^K \sum_{i=1}^N r_i^1(\xi_k) x_i^1(\xi_k, 1) & \cdots & \frac{1}{K} \sum_{k=1}^K \sum_{i=1}^N r_i^1(\xi_k) x_i^1(\xi_k, E) \\ \vdots & \ddots & \vdots \\ \frac{1}{K} \sum_{k=1}^K \sum_{i=1}^N r_i^T(\xi_k) x_i^T(\xi_k, 1) & \cdots & \frac{1}{K} \sum_{k=1}^K \sum_{i=1}^N r_i^T(\xi_k) x_i^T(\xi_k, E) \end{bmatrix}_{T \times E}, \quad b = - \begin{bmatrix} \mu_{\min}^1 \\ \vdots \\ \mu_{\min}^T \end{bmatrix}_{T \times 1}.$$

## A.4 Problema $\alpha$ com Desvio *Average Value at Risk*

Para a utilização da função *linprog* do otimizador Gurobi é necessário escrever o problema  $\alpha$  apresentado em (4.4.5) no formato matricial a seguir:

$$\begin{cases} \min_{y \in Y} f^T y \\ A_{eq} y = b_{eq} \\ Ay \leq b \\ lb \leq y \leq ub, \end{cases}$$

onde,

$$f = \begin{bmatrix} \frac{1}{K} \sum_{k=1}^K \sum_{t=1}^T \sum_{i=1}^N r_i^t(\xi_k) x_i^t(\xi_k, 1) \\ \vdots \\ \frac{1}{K} \sum_{k=1}^K \sum_{t=1}^T \sum_{i=1}^N r_i^t(\xi_k) x_i^t(\xi_k, E) \\ -\lambda_1 \\ \vdots \\ -\lambda_T \\ \frac{\lambda_1}{K(1-\alpha^1)} \mathbf{1}_{K \times 1} \\ \vdots \\ \frac{\lambda_T}{K(1-\alpha^T)} \mathbf{1}_{K \times 1} \end{bmatrix}_{(E+T+K \times T) \times 1}, \quad y = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_E \\ a^1 \\ \vdots \\ a^T \\ b_1^1 \\ b_2^1 \\ \vdots \\ b_1^t \\ \vdots \\ b_k^t \end{bmatrix}_{(E+T+K \times T) \times 1}$$

$$lb = \begin{bmatrix} \mathbf{0}_{E \times 1} \\ -\infty_{T \times 1} \\ \mathbf{0}_{(K \times T) \times 1} \end{bmatrix}_{(E+T+K \times T) \times 1}, \quad ub = \begin{bmatrix} \mathbf{1}_{E \times 1} \\ \infty_{T \times 1} \\ \infty_{(K \times T) \times 1} \end{bmatrix}_{(E+T+K \times T) \times 1}$$

$$A_{eq} = \begin{bmatrix} \mathbf{1}_{1 \times E} & \mathbf{0}_{1 \times (T+K \times T)} \end{bmatrix}_{1 \times (E+T+K \times T)}, \quad b_{eq} = \begin{bmatrix} 1 \end{bmatrix}_{1 \times 1}, \quad b = - \begin{bmatrix} \mu_{\min}^1 \\ \vdots \\ \mu_{\min}^T \\ \mathbf{0}_{(T \times K) \times 1} \end{bmatrix}_{(T+K \times T) \times 1}$$

$$A = \begin{bmatrix} -[\mu] & \mathbf{0}_{T \times T} & \mathbf{0}_{T \times (K \times T)} \\ -[V^1] & \begin{bmatrix} 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \dots & 0 \end{bmatrix} & \begin{bmatrix} -1 & 0 & \dots & \dots & \dots & \dots & 0 \\ \vdots & \ddots & \ddots & \dots & \dots & \dots & \vdots \\ 0 & \dots & -1 & 0 & \dots & \dots & 0 \end{bmatrix} \\ \vdots & \vdots & \vdots \\ -[V^T] & \begin{bmatrix} 0 & \dots & 0 & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 1 \end{bmatrix} & \begin{bmatrix} 0 & \dots & \dots & \dots & -1 & \dots & 0 \\ \vdots & \dots & \dots & \dots & \dots & \ddots & \vdots \\ 0 & \dots & \dots & \dots & \dots & \dots & -1 \end{bmatrix} \end{bmatrix}_{(T+K \times T) \times (E+T+K \times T)}$$



# Apêndice B

## Códigos Utilizados

Neste apêndice são apresentados os códigos utilizados neste trabalho em linguagem MATLAB.

```
1 % ----- Geracao dos Cenarios Aleatorios ----- %
2 % ----- IMPA 2014 - Mestrado em Metodos Matematicos em Financas ----- %
3 % --- Louise Albuquerque Bassan / Orientadora: Claudia Sagastizabal --- %
4
5 R = R_semanal;
6 P = P_semanal;
7
8 graf = 1; %Graficos On = 0, Off = 1
9
10 %% Determinando Outliers usando a tecnica da secao 4.6 de Meucci
11
12 if graf == 0
13     for i = 1:15
14         % Combinando os ativos dois a dois para a construcao dos graficos
15         if i ~= 15
16             Ativo = i; Nome1 = P_nomes(Ativo); Nome2 = P_nomes(Ativo+1);
17             CS = R(:, Ativo:Ativo+1);
18         else
19             Nome1 = P_nomes(15); Nome2 = P_nomes(1);
20             CS = [R(:,15) R(:,1)];
21         end
22
23         T = size(CS,1);
24         Index = 1:T;
25         N = 10;
26
27         % Funcao para detectar Outliers: Secao 4.6 de Meucci.
28         HighBreakdownMVE(CS,N, Index , Nome1 , Nome2)
29     end
30 end
31
32 if graf == 0
33     T = size(CS,1);
34     Index = 1:T;
35     N = 100;
36
37     % Funcao para detectar Outliers: Secao 4.6 de Meucci.
38     ElipsoideMinimo(R,N, Index)
39 end
```

```

41 %% Filtrando os Outliers e Calculando os Estimadores Robustos
43 R = R_semanal;
R = FilterOutliers(R,10);
45
norm_mu = mean(R);
47 norm_Sigma = cov(R);
49 %% Simulacoes de Monte Carlo
51 horizon = [1;2;4]; % horizonte de semanas
hzon = max(horizon); % maximo de semanas
53 npaths = 1000; % numero de trajetorias por periodo
55 samplelen = npaths*size(R,2);
sample = mvellrnd([hzon npaths], 'Gaussian', norm_mu, sqrt(norm_Sigma));
57 sample = reshape(sample, hzon, []);
sample = ret2price(sample(:,1:samplelen));
59 sample(:,1:samplelen) = sample(:,1:samplelen)-1;
61 S=size(P,2);
j=0; k=0;
63
% Mostra as primeiras 1000 simulacoes para cada ativo
65 if graf == 1
    for i=1:S
67
        if i <= k + 2
69            if i == k + 1
                figure;
71                end
                subplot(1,2,i-k)
73            elseif i == k + 3
                figure;
75                subplot(1,2,i-(k+2))
            else
77                subplot(1,2,i-(k+2))
                k = i;
79            end
            plot(repmat([0:4]', [1;1000]), sample(:, j+1:j+1000))
81            title(P_nomes(i)); set(gca, 'fontsize', 12);
83
            j = j + npaths;
85        end
    end
end

```

### Modelo Estático com Desvio Padrão

```

% ————— Problema Estatico ————— %
2 % ——— IMPA 2015 — Mestrado em Metodos Matematicos em Financas ——— %
% — Louise Albuquerque Bassan / Orientadora: Claudia Sagastizabal — %
4
% Construcao do Intervalo para a Fronteira Eficiente
6 Portfolio_Variancia_Min
8 % Incluindo o ativo livre de risco

```

```

f = zeros(4*(n+1),1);
10
RMean = [r0(1) RMean_T1 zeros(1,48);
12         zeros(1,16) r0(2) RMean_T2 zeros(1,32);
14         zeros(1,32) r0(3) RMean_T3 zeros(1,16);
16         zeros(1,48) r0(4) RMean_T4];

16 H = [cov1      cov_aux      cov_aux      cov_aux;
18      cov_aux      cov2      cov_aux      cov_aux;
20      cov_aux      cov_aux      cov3      cov_aux;
22      cov_aux      cov_aux      cov_aux      cov4];

lb = zeros(4*(n+1),1);
ub = ones(4*(n+1),1);

24 Aeq = [ones(1,16) zeros(1,48);
26         zeros(1,16) ones(1,16) zeros(1,32);
28         zeros(1,32) ones(1,16) zeros(1,16);
30         zeros(1,48) ones(1,16)];
beq = ones(4,1);

30 % Criacao da matriz que representa a restricao dos grupos economicos
A = zeros(8,n); a = 0; s = size(R_Final,2);
32 for j = 1:8
34     g = size(eval(char(Grupos(j))),2);
36     for l = 1:g
38         A(j,l+a) = 1;
40     end
42     a = a + g;

A_bloco = [zeros(8,1) A];
A_blank = zeros(8,n+1);
44 A = [-RMean;
46     A_bloco A_blank A_blank A_blank;
48     A_blank A_bloco A_blank A_blank;
50     A_blank A_blank A_bloco A_blank;
52     A_blank A_blank A_blank A_bloco];

54 % Resolucao do Problema de Otimizacao Estatico
options = optimoptions('quadprog','Algorithm','active-set');

56 for i = 1:length(Rdes)
58     b = [-Rdes(:,i);PercGE];
60     %[p,fval,exitflag] = quadprog(H,f,A,b,Aeq,beq,lb,ub,[],options);
62     [p,ef] = creategurobiqp(Aeq,beq,A,b,lb,ub,H,f);
64     carteira.pesos(i,:) = p';
66     carteira.retorno(i,:) = 100*(RMean*p);
    carteira.variancia_T1(i,:) = p(1:16)'*cov1*p(1:16);
    carteira.variancia_T2(i,:) = p(17:32)'*cov2*p(17:32);
    carteira.variancia_T3(i,:) = p(33:48)'*cov3*p(33:48);
    carteira.variancia_T4(i,:) = p(49:64)'*cov4*p(49:64);
end

carteira.desvio = 100*[sqrt(carreira.variancia_T1)...
    sqrt(carreira.variancia_T2)...
    sqrt(carreira.variancia_T3)...
    sqrt(carreira.variancia_T4)];

```

## Modelo Estático com Desvio AVaR

```

% ----- Problema Estatico - CVaR ----- %
% ----- IMPA 2015 - Mestrado em Metodos Matematicos em Financas ----- %
% ----- Louise Albuquerque Bassan / Orientadora: Claudia Sagastizabal ----- %
2
4
6 T = 4; NC = 0.99; graf = 1;
K = NS;
8 myInf = 1e2;
PercGE = 0.7*ones(8*4,1);
10
% Incluindo o ativo livre de risco
12 RMean = [r0(1) RMean_T1 zeros(1,48);
           zeros(1,16) r0(2) RMean_T2 zeros(1,32);
14           zeros(1,32) r0(3) RMean_T3 zeros(1,16);
           zeros(1,48) r0(4) RMean_T4];
16
% Limite inferior e superior das variaveis
18 lb = [zeros(T*(n+1),1); -myInf*ones(T,1); zeros(T*K,1)] ;
ub = [ones(T*(n+1),1); myInf*ones(T,1); myInf*ones(T*K,1)] ;
20
Aeq_Aux = [ones(1,16) zeros(1,48);
22           zeros(1,16) ones(1,16) zeros(1,32);
           zeros(1,32) ones(1,16) zeros(1,16);
24           zeros(1,48) ones(1,16)];
26
Aeq = [Aeq_Aux zeros(T,T+K*T)];
28
beq = ones(4,1);
30
% Criacao da matriz que representa a restricao dos grupos economicos
A = zeros(8,n); a = 0; s = size(R_Final,2);
32 for j = 1:8
    g = size(eval(char(Grupos(j))),2);
34     for l = 1:g
        A(j,l+a) = 1;
36     end
    a = a + g;
38 end
40
A_bloco = [zeros(8,1) A];
A_blank = zeros(8,n+1);
42
G = [A_bloco A_blank A_blank A_blank;
44     A_blank A_bloco A_blank A_blank;
     A_blank A_blank A_bloco A_blank;
46     A_blank A_blank A_blank A_bloco];
48
% Construcao da Matriz a
a1 = repmat([1 0 0 0],[K,1]);
50 a2 = repmat([0 1 0 0],[K,1]);
a3 = repmat([0 0 1 0],[K,1]);
52 a4 = repmat([0 0 0 1],[K,1]);
54
a = [a1; a2; a3; a4];

```

```

56 b1 = eye(T*K);
58 % Retornos em cada cenario
r01 = repmat(r0(1),K,1);
60 r02 = repmat(r0(2),K,1);
r03 = repmat(r0(3),K,1);
62 r04 = repmat(r0(4),K,1);
A_blank2 = zeros(K,n+1);
64
R_xi = [r01 R_Sim_T1 A_blank2 A_blank2 A_blank2;
66         A_blank2 r02 R_Sim_T2 A_blank2 A_blank2;
        A_blank2 A_blank2 r03 R_Sim_T3 A_blank2;
68         A_blank2 A_blank2 A_blank2 r04 R_Sim_T4];
70 % Criacao da matriz A
A = [-RMean zeros(T,T+K*T);
72       G      zeros(8*T,T+K*T);
      -R_xi  a  -b1];
74
% Funcao Objetivo
76 lbd1 = 1 ;lbd2 = 0; lbd3 = 0; lbd4 = 0;
Mi = [r0(1) RMean_T1 r0(2) RMean_T2 r0(3) RMean_T3 r0(4) RMean_T4];
78 at = [lbd1 lbd2 lbd3 lbd4];
bkt = ones(1,K)/(K*(1-NC));
80 bk = [lbd1*bkt lbd2*bkt lbd3*bkt lbd4*bkt];
82 f = [Mi -at bk];
84 % Resolucao do Problema de Otimizacao Estatico para o AVaR
for i = 1:length(Rdes)
86   b = [-Rdes(:,i); PercGE; zeros(T*K,1)];
      [p, ov, ef] = creategurobilp(Aeq, beq, A, b, lb, ub, f);
88   % [p, ef] = linprog(f,A,b,Aeq,beq,lb,ub);
      carteira.pCVaR(i,:) = p(1:T*(n+1))';
90   carteira.retornoCVaR(i,:) = 100*(RMean*p(1:T*(n+1)));
      carteira.CVAR_T1(i,:) = [lbd1*r0(1) lbd1*RMean_T1 lbd2*r0(2) lbd2*RMean_T2
92   lbd3*r0(3) lbd3*RMean_T3 lbd4*r0(4) lbd4*RMean_T4...
      -lbd1 0 0 0 lbd1*bkt 0*bkt 0*bkt 0*bkt]*p;
      carteira.CVAR_T2(i,:) = [lbd1*r0(1) lbd1*RMean_T1 lbd2*r0(2) lbd2*RMean_T2
94   lbd3*r0(3) lbd3*RMean_T3 lbd4*r0(4) lbd4*RMean_T4...
      0 -lbd2 0 0 0*bkt lbd2*bkt 0*bkt 0*bkt]*p;
      carteira.CVAR_T3(i,:) = [lbd1*r0(1) lbd1*RMean_T1 lbd2*r0(2) lbd2*RMean_T2
96   lbd3*r0(3) lbd3*RMean_T3 lbd4*r0(4) lbd4*RMean_T4...
      0 0 -lbd3 0 0*bkt 0*bkt lbd3*bkt 0*bkt]*p;
      carteira.CVAR_T4(i,:) = [lbd1*r0(1) lbd1*RMean_T1 lbd2*r0(2) lbd2*RMean_T2
98   lbd3*r0(3) lbd3*RMean_T3 lbd4*r0(4) lbd4*RMean_T4...
      0 0 0 -lbd4 0*bkt 0*bkt 0*bkt lbd4*bkt]*p;
100 end
102 carteira.desvioAVaR = 100*[carteira.CVAR_T1...
      carteira.CVAR_T2...
104   carteira.CVAR_T3...
      carteira.CVAR_T4];

```

## Funções Utilizadas

Estimadores de Alta Ruptura.

```

1 % this script shows how to use the minimum volume ellipsoid to detect outliers.
  See Sec. 4.6.1 of "Risk and Asset Allocation" – Springer (2005), by A. Meucci
  for the theory and the routine implemented below.
% Algumas alteracoes foram feitas para o plot dos graficos.
3
function HighBreakdownMVE(CorruptSample ,T, Index ,Nome1 ,Nome2)
5
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
7 % compute high-breakdown estimates
9 CorruptSample_aux = CorruptSample;
11 Sample_Length = [];
  Vol_MVE = [];
13
  for j = 1:T
15
16     % compute MVE
17     [MVE_Location , MVE_Dispersion] = ComputeMVE(CorruptSample);
19
20     % store results
21     Store(j).MVE_Location = MVE_Location;
22     Store(j).MVE_Dispersion = MVE_Dispersion;
23
24     Sample_Length = [Sample_Length size(CorruptSample , 1) ];
25     Vol_MVE = [Vol_MVE sqrt(det(MVE_Dispersion)) ];
26
27     Store(j).Index = Index;
28
29     % erase one outlier
30     Rejected = RejectOutlier(CorruptSample , Index);
31     CorruptSampleRejected(j , :) = CorruptSample(Rejected , :);
32
33     CorruptSample(Rejected , :) = [];
34     Index(Rejected) = [];
35
36     RejectedIndex(j) = Rejected;
37
38 end
39
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
40 % plot results
41
42 figure % data
43 h = plot(CorruptSample(:,1) , CorruptSample(:,2) , '.' );
44 xlabel(strcat(Nome1 , 'FontSize' , 14));
45 ylabel(strcat(Nome2) , 'FontSize' , 14);
46 set(gca , 'fontsize' , 12);
47 set(h , 'color' , 'k')
48 hold on
49 h = plot(CorruptSampleRejected(:,1) , CorruptSampleRejected(:,2) , '.' );
50 set(h , 'color' , 'r' , 'MarkerSize' , 16)
51
  for j = 1:T

```

```

53 hold on
E=TwoDimEllipsoid(Store(j).MVE_Location,Store(j).MVE_Dispersion,1,0,0);
55 set(E,'color','b','linewidth',1)
end

```

### Elipsóide de Volume Mínimo.

```

% this script shows how to use the minimum volume ellipsoid to detect outliers.
  See Sec. 4.6.1 of "Risk and Asset Allocation" – Springer (2005), by A. Meucci
  for the theory and the routine implemented below.
2 % Algumas alteracoes foram feitas para o plot dos graficos.

4 function Elipsoideminimo(CorruptSample,T,Index)
6 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
6 % compute high-breakdown estimates
8
8 CorruptSample_aux = CorruptSample;
10
10 Sample_Length=[];
12 Vol_MVE=[];
14
14 for j=1:T
16
16 % compute MVE
[MVE_Location,MVE_Dispersion]=ComputeMVE(CorruptSample);
18
18 % store results
20 Store(j).MVE_Location=MVE_Location;
Store(j).MVE_Dispersion=MVE_Dispersion;
22
22 Sample_Length=[Sample_Length size(CorruptSample,1)];
24 Vol_MVE=[Vol_MVE sqrt(det(MVE_Dispersion))];
26
26 Store(j).Index=Index;
28
28 % erase one outlier
Rejected=RejectOutlier(CorruptSample,Index);
30 CorruptSampleRejected(j,:)=CorruptSample(Rejected,:);
32
32 CorruptSample(Rejected,:)=[];
Index(Rejected)=[];
34
34 RejectedIndex(j)=Rejected;
36
36 end
38 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
40 % plot results
42 figure % volume of ellipsoid as function of sample length
plot(Sample_Length,Vol_MVE)
44 set(gca,'fontsize',12);
xlabel('Observacoes','FontSize',14);
46 ylabel('Volume','FontSize',14);
grid on

```

Problema  $\alpha$  com Desvio Padrão

```

1  % ----- Metodo de Galerkin – Problema Alpha DP ----- %
2  % ----- IMPA 2015 – Mestrado em Metodos Matematicos em Financas ----- %
3  % --- Louise Albuquerque Bassan / Orientadora: Claudia Sagastizebal --- %
4
5  function [p_alphas, desvP, retP] = Prob_Alpha(Xi, K, Rdes, Cenario)
6
7  myInf = 1e2;
8
9  n1 = size(Xi,1);
10 n2 = size(Xi,2);
11
12 % Numero de especialistas
13 E = n2/K;
14
15 % Limite inferior e superior das variaveis
16 lb = zeros(E,1);
17 ub = ones(E,1);
18
19 percE = 0.4*ones(E,1);
20
21 % Construcao da Matriz Vtke
22 v = 0; d = 0; C1 = ones(15,1); Vtke = [];
23 for t = 1:4
24     for e = 1:E
25         for k = 1:K
26             CenarioAux = Cenario(k,:)';
27             XiAux = Xi(:,k+d);
28
29             vtke(e,k) = (XiAux(1+v:v+15,:).*CenarioAux(1+v:v+15,:))'*C1;
30         end
31         mte(t,e) = vtke(e,:)*ones(K,1)/K;
32         d = d + K;
33     end
34     Vtke = [Vtke; vtke];
35     vtke = [];
36     v = v + 15;
37     d = 0;
38 end
39
40 % Matriz da restricao de desigualdade
41 A = [-mte; eye(E)];
42 %A = [];
43
44 % Construcao da Matriz Vtk
45 v = 0; Vtk = [];
46 for t = 1:4
47     for k = 1:K
48         vtk = Vtke(1+v:v+E,k);
49         Vtk = [Vtk; vtk'];
50     end
51     v = v + E;
52 end
53
54 V1 = Vtk(1:K,1:E);

```



```

55 V2 = Vtk(K+1:2*K,1:E);
V3 = Vtk(2*K+1:3*K,1:E);
57 V4 = Vtk(3*K+1:4*K,1:E);

59 % funcao objetivo
lbd1 = 1; lbd2 = 0; lbd3 = 0; lbd4 = 0;
61 H = lbd1*cov(V1) + lbd2*cov(V2) + lbd3*cov(V3) + lbd4*cov(V4);
f = zeros(E,1);

63 % Restricao de igualdade
65 Aeq = ones(1,E);
beq = ones(1,1);
67 %b = [];

69 options = optimoptions('quadprog','Algorithm','active-set');

71 for i = 1:length(Rdes)
    b = [-Rdes(:,i);percE];
73     %beq = [ones(1,1); Rdes(:,i)];
    %[alpha,fval,exitflag] = quadprog(H,f,A,b,Aeq,beq,lb,ub,[],options);
75     [alpha,ef] = creategurobiqp(Aeq,beq,A,b,lb,ub,H,f);
    p_alphas(:,i) = alpha;
77     desvP(i,:) = 100*sqrt(alpha'*H*alpha);
    retP(i,:) = 100*mte*alpha;
79 end
81 end

```

### Problema $\alpha$ com Desvio AVaR

```

1 % ----- Metodo de Galerkin - Problema Alpha ----- %
% ----- IMPA 2015 - Mestrado em Metodos Matematicos em Financas ----- %
3 % --- Louise Albuquerque Bassan / Orientadora: Claudia Sagastizabal --- %

5 function [p_alphas,desvAVaR,retAVaR] = Prob_Alpha_CVaR(Xi, K, Rdes, Cenario, NC)

7 myInf = 1e2;

9 n1 = size(Xi,1);
n2 = size(Xi,2);

11 % Numero de especialistas
13 E = n2/K;

15 % Limite inferior e superior das variaveis
lb = [zeros(E,1); -myInf*ones(4,1); zeros(4*K,1)];
17 ub = [ones(E,1); myInf*ones(4,1); myInf*ones(4*K,1)];

19 percE = 0.4*ones(E,1);

21 % Restricao de igualdade
Aeq = [ones(1,E) zeros(1,4) zeros(1,4*K)];
23 beq = 1;

25 % Construcao da Matriz Vtke
v = 0; d = 0; C1 = ones(15,1); Vtke = [];
27 for t = 1:4

```

```

29     for e = 1:E
30         for k = 1:K
31             CenarioAux = Cenario(k,:)';
32             XiAux = Xi(:,k+d);
33
34             vtke(e,k) = (XiAux(1+v:v+15,:).*CenarioAux(1+v:v+15,:))'*C1;
35         end
36         mte(t,e) = vtke(e,:)*ones(K,1)/K;
37         d = d + K;
38     end
39     Vtke = [Vtke; vtke];
40     vtke = [];
41     v = v + 15;
42     d = 0;
43 end
44
45 % Matriz da primeira restricao
46 A_blank = zeros(4,4*(K+1));
47 A1 = [-mte A_blank];
48
49 % Construcao da Matriz Vtk
50 v = 0; Vtk = [];
51 for t = 1:4
52     for k = 1:K
53         vtk = Vtke(1+v:v+E,k);
54         Vtk = [Vtk; vtk'];
55     end
56     v = v + E;
57 end
58
59 % Construcao da Matriz a
60 a1 = repmat([1 0 0 0],[K,1]);
61 a2 = repmat([0 1 0 0],[K,1]);
62 a3 = repmat([0 0 1 0],[K,1]);
63 a4 = repmat([0 0 0 1],[K,1]);
64
65 a = [a1; a2; a3; a4];
66
67 % Matriz b
68 b1 = eye(4*K);
69
70 % Matriz da segunda restricao
71 A2 = [-Vtk a -b1];
72
73 % Matriz com as restricoes de desigualdade
74 A = [A1; A2; eye(E) zeros(E,4) zeros(E,4*K)];
75
76 % Funcao Objetivo
77 lbd1 = 1; lbd2 = 0; lbd3 = 0; lbd4 = 0;
78 Mi = lbd1*mte(1,:)+lbd2*mte(2,:)+lbd3*mte(3,:)+lbd4*mte(4,:);
79 at = [lbd1 lbd2 lbd3 lbd4];
80 bkt = ones(1,K)/(K*(1-NC));
81 bk = [lbd1*bkt lbd2*bkt lbd3*bkt lbd4*bkt];
82
83 f = [Mi -at bk];
84
85 for i = 1:length(Rdes)
86     b = [-Rdes(:,i); zeros(4*K,1); percE];

```

```

[alpha , ov , ef] = creategurobilp(Aeq, beq, A, b, lb, ub, f);
87 % [alpha , ov] = linprog(f,A,b,Aeq,beq,lb,ub);
    p_alphas(:,i) = alpha;
89 desvAVaR(i,:) = 100*ov;
    retAVaR(i,:) = 100*mte*alpha(1:E);
91 end
93 end

```

### Exemplo de um dos especialistas determinísticos (Bancos)

```

1 % ----- Metodo de Galerkin – Expert 1 ----- %
2 % ----- IMPA 2015 – Mestrado em Metodos Matematicos em Financas ----- %
3 % — Louise Albuquerque Bassan / Orientadora: Claudia Sagastizabal — %
4
5 function Xi = Expert1_Bancos(r, P_nomes, pG)
6
7 % O Expert requer uma alocao de pelo menos pG% em bancos
8
9 n = 15;
10 PercGE = ones(8*4,1);
11 PercGE(1,1) = -pG; PercGE(9,1) = -pG;
12 PercGE(17,1) = -pG; PercGE(25,1) = -pG;
13
14 f = -r;
15
16 Aeq = [ones(1,15) zeros(1,45);
17        zeros(1,15) ones(1,15) zeros(1,30);
18        zeros(1,30) ones(1,15) zeros(1,15);
19        zeros(1,45) ones(1,15)];
20 beq = ones(4,1);
21
22 lb = zeros(4*(n),1);
23 ub = ones(4*(n),1);
24
25 % Criacao da matriz que representa a restricao dos grupos economicos
26 A = zeros(8,n); a = 0; s = size(r,2);
27 for j = 1:8
28     g = size(eval(char(Grupos(j))),2);
29     for l = 1:g
30         A(j,l+a) = 1;
31     end
32     a = a + g;
33 end
34
35 A(1,:) = -A(1,:);
36 A_bloco = A;
37 A_blank = zeros(8,n);
38 A = [A_bloco A_blank A_blank A_blank;
39      A_blank A_bloco A_blank A_blank;
40      A_blank A_blank A_bloco A_blank;
41      A_blank A_blank A_blank A_bloco];
42
43 b = PercGE;
44
45 %Xi = linprog(f,A,b,Aeq,beq,lb,ub);
[Xi, ov, ef] = creategurobilp(Aeq, beq, A, b, lb, ub, f);

```

```
47 | end
```