

Universidade Federal de Juiz de Fora, Instituto de Ciências Exatas
Programa de Pós-Graduação em Modelagem Computacional

João Daniel Madureira Yamim

Otimização de portfólios *Online* com métodos de projeção para controle
de risco

Juiz de Fora

2024

João Daniel Madureira Yamim

Otimização de portfólios *Online* com métodos de projeção para controle
de risco

Tese apresentada ao Programa de Pós-Graduação em Modelagem Computacional da Universidade Federal de Juiz de Fora, como requisito parcial para obtenção do título de Doutor em Modelagem Computacional.

Orientador: Carlos Cristiano Hasenclever Borges

Coorientador: Raul Fonseca Neto

Juiz de Fora

2024

João Daniel Madureira Yamim

Otimização de portfólios *Online* com métodos de projeção para controle de risco

Tese apresentada ao Programa de Pós-Graduação em Modelagem Computacional da Universidade Federal de Juiz de Fora, como requisito parcial para obtenção do título de Doutor em Modelagem Computacional.

Aprovada em 11 de setembro de 2024.

BANCA EXAMINADORA

Prof. Dr. Carlos Cristiano Hasenclever Borges -
Orientador
Universidade Federal de Juiz de Fora

Prof. Dr. Raul Fonseca Neto - Coorientador
Universidade Federal de Juiz de Fora

Prof. Dr. Vinicius Layter Xavier
Universidade do Estado do Rio de Janeiro

Prof. Dr. Marcos Garcia Todorov
Laboratório Nacional de Computação Científica

Profa. Dra. Fernanda Finotti Cordeiro
Universidade Federal de Juiz de Fora

Prof. Dr. Leonardo Goliatt da Fonseca
Universidade Federal de Juiz de Fora



UNIVERSIDADE FEDERAL DE JUIZ DE FORA
PRÓ-REITORIA DE PÓS-GRADUAÇÃO E PESQUISA



ATA DE DEFESA DE TRABALHO DE CONCLUSÃO DE PÓS-GRADUAÇÃO
STRICTO SENSU

PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM COMPUTACIONAL

Nº PPG: 66

Formato da Defesa: virtual

Ata da sessão pública referente à defesa da tese intitulada "Otimização de portfólios Online com métodos de projeção para controle de risco", para fins de obtenção do título de doutor em Modelagem Computacional, área de concentração Modelagem Computacional, pelo discente João Daniel Madureira Yamim (matrícula 103120067 - início do curso em 05/03/2018), sob orientação da Prof.ªDr.ª Carlos Cristiano Hasenclever Borges e coorientação da Prof.ªDr.ª Raul Fonseca Neto.

Ao 11 dia do mês de setembro do ano de 2024, às 08 horas e 30 minutos, de forma virtual conforme Resolução CSPP/UFJF nº 28/2023 e Portaria PROPP/UFJF nº 53/2024 da Universidade Federal de Juiz de Fora (UFJF), reuniu-se a Banca examinadora da tese em epígrafe, aprovada pelo Colegiado do Programa de Pós- Graduação, conforme a seguinte composição:

Titulação	Nome	Na qualidade de:
Prof. Dr.	Carlos Cristiano Hasenclever Borges	Orientador e Presidente da Banca
Prof. Dr.	Raul Fonseca Neto	Coorientador
Prof. Dr.	Vinicius Layter Xavier	Membro titular externo
Prof. Dr.	Marcos Garcia Todorov	Membro titular externo
Prof.ª Dr.ª	Fernanda Finotti Cordeiro	Membro titular interno
Prof. Dr.	Leonardo Goliatt da Fonseca	Membro titular interno
Prof. Dr.	Fabrizio Condé de Oliveira	Suplente externo
Prof. Dr.	Davi Riani Gotardelo	Suplente externo
Prof. Dr.	Heder Soares Bernardino	Suplente interno
Prof. Dr.	Saulo Moraes Villela	Suplente interno

*Obs: Conforme §2º do art. 54 do Regulamento Geral da Pós-graduação stricto sensu, aprovado pela Resolução CSPP/UFJF nº 28, de 7 de junho de 2023, "estando o(a) orientador(a) impedido(a) de compor a banca, a presidência deverá ser designada pelo Colegiado".

AVALIAÇÃO DA BANCA EXAMINADORA

Tendo o(a) senhor(a) Presidente declarado aberta a sessão, mediante o prévio exame do referido trabalho por parte de cada membro da Banca, o(a) discente procedeu à apresentação de seu Trabalho de Conclusão de Curso de Pós-graduação Stricto sensu e foi submetido(a) à arguição pela Banca Examinadora que, em seguida, deliberou sobre o seguinte resultado:

APROVADO

REPROVADO, conforme parecer circunstanciado, registrado no campo Observações desta Ata e/ou em documento anexo, elaborado pela Banca Examinadora

APROVADO CONDICIONALMENTE, sendo consenso da banca de que esta ata é provisória e condicional

Novo título da Dissertação/Tese (só preencher no caso de mudança de título):

Observações da Banca Examinadora caso haja necessidade de anotações gerais sobre a dissertação/tese e sobre a defesa, as quais a banca julgue pertinentes

Descrição de todas as condicionalidades para a aprovação (só preencher no caso de Aprovação Condicional):

Nada mais havendo a tratar, o(a) senhor(a) Presidente declarou encerrada a sessão de Defesa, sendo a presente Ata lavrada e assinada pelos(as) senhores(as) membros da Banca Examinadora e pelo(a) discente, atestando ciência do que nela consta.

INFORMAÇÕES

Para fazer jus ao título de mestre(a)/doutor(a), a versão final da dissertação/tese, considerada Aprovada, devidamente conferida pela Secretaria do Programa de Pós-graduação, deverá ser tramitada para a PROPP, em Processo de Homologação de Dissertação/Tese, dentro do prazo de de 60 ou 90 dias, para discentes aprovados condicionalmente, a partir da data da defesa. Após o envio dos exemplares definitivos, o processo deverá receber homologação e, então, ser encaminhado à CDARA.

Se as condições descritas nesta ata não forem atestadamente atendidas dentro do prazo de 90 dias, a aprovação condicional será convertida em reprovação.

Esta Ata de Defesa é um documento padronizado pela Pró-Reitoria de Pós-Graduação e Pesquisa. Observações excepcionais feitas pela Banca Examinadora poderão ser registradas no campo disponível acima ou em documento anexo, desde que assinadas pelo(a) Presidente(a).

Esta Ata de Defesa somente poderá ser utilizada como comprovante de titulação se apresentada junto à Certidão da Coordenadoria de Assuntos e Registros Acadêmicos da UFJF (CDARA) atestando que o processo de confecção e registro do diploma está em andamento.



Documento assinado eletronicamente por **Carlos Cristiano Hasenclever Borges, Professor(a)**, em 11/09/2024, às 11:49, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Fernanda Finotti Cordeiro Perobelli, Professor(a)**, em 11/09/2024, às 11:51, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Leonardo Goliatt da Fonseca, Professor(a)**, em 11/09/2024, às 11:54, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **MARCOS GARCIA TODOROV, Usuário Externo**, em 11/09/2024, às 12:02, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Raul Fonseca Neto, Professor(a)**, em 11/09/2024, às 13:08, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **João Daniel Madureira Yamim, Usuário Externo**, em 11/09/2024, às 16:38, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Vinicius Layter Xavier, Usuário Externo**, em 18/09/2024, às 16:02, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no Portal do SEI-Uffj (www2.ufff.br/SEI) através do ícone Conferência de Documentos, informando o código verificador **1969537** e o código CRC **C0A476B4**.

*A Deus em primeiro lugar. Aos meus pais, minhas irmãs, esposa, filho e amigos
pelo apoio incondicional.*

AGRADECIMENTOS

Agradeço ao Programa de Pós-graduação em Modelagem Computacional, à Universidade Federal de Juiz de Fora, aos professores, servidores e colegas, por este período de grande aprendizagem e compartilhamento de experiências. Em especial, agradeço aos meus orientadores e amigos Professor Carlos Cristiano Hasenclever Borges e Professor Raul Fonseca Neto pela paciência e ensinamentos que me auxiliaram a dar um grande passo na vida acadêmica.

Aos membros da banca, por aceitarem o convite e pelas valiosas contribuições que certamente enriquecerão este trabalho.

Aos meus pais, Cristina e David, que com muito carinho e apoio, não mediram esforços para que eu chegasse até esta etapa de minha vida.

Às minhas irmãs, Clara e Nívea, agradeço pelo zelo, incentivo, carinho, paciência e pela ajuda desde os primeiros anos de escola.

À minha esposa Rayanne, que percorreu comigo cada passo dessa trajetória, incentivando, animando e demonstrando seu carinho todos os dias.

Ao meu filho Francisco, que chegou em nossas vidas trazendo uma imensa alegria e um amor incondicional. Sua presença trouxe uma nova perspectiva e alegria, fazendo com que cada desafio enfrentado ao longo deste percurso valha ainda mais a pena. Prometo continuar lutando para ser um exemplo para você, assim como você, sem saber, é para mim. Obrigado por iluminar nossos dias com seu sorriso e por fazer parte desta jornada tão especial.

Acima de tudo, agradeço a Deus. O que seria de mim sem a fé que tenho Nele? Agradeço também à Virgem Maria e a São Pio de Pietrelcina, por sua intercessão e por serem fontes de inspiração e força ao longo de toda esta caminhada.

“A humildade é o primeiro degrau para a sabedoria”

São Tomás de Aquino

RESUMO

O objetivo deste trabalho é investigar o benefício da incorporação de métricas de risco econométricas na performance de algoritmos de aprendizado *online* para a solução do problema de alocação dinâmica de portfólio. O problema foi formulado como um processo de decisão *online* onde variantes do método *Online Gradient Descent* (OGD) são estudadas. Propõe-se três modificações no algoritmo OGD, cada uma utilizando uma forma distinta de projeção convexa. A primeira modificação inclui uma restrição baseada na estimação do beta variante no tempo durante a etapa de projeção, visando controlar o risco sistemático do portfólio. Reconhecendo que o beta por si só não captura todos os aspectos do risco, introduziu-se uma segunda modificação que incorpora o *Value at Risk* (VaR) como uma medida mais abrangente de risco. Essa abordagem permite um controle mais efetivo da exposição ao risco do portfólio, considerando tanto o risco sistemático quanto o não sistemático. A terceira modificação utiliza uma projeção convexa que minimiza o *Conditional Value at Risk* (CVaR). Esta abordagem apresentou resultados mais consistentes em termos de equilíbrio entre risco e retorno, proporcionando uma gestão de risco mais robusta e eficaz. Antes da implementação das modificações propostas, visando garantir a validade teórica e a eficiência computacional dos algoritmos desenvolvidos, demonstrou-se matematicamente a convexidade da projeção para cada uma dos métodos. O impacto da escolha da distribuição de probabilidade (normal e t de Student) na minimização do VaR durante a etapa de projeção e sua influência no desempenho e no perfil de risco das carteiras otimizadas também é investigado. As metodologias propostas são avaliadas empiricamente em diferentes conjuntos de dados brasileiros (IBOVESPA e *Small Caps*), considerando tanto ações de alta como de baixa capitalização em períodos de mercado distintos. Os resultados mostram que a incorporação de métricas de risco, especialmente o CVaR, no processo de otimização *online* permitiu um melhor controle do risco, ao mesmo tempo que mantém retornos competitivos em comparação com *benchmarks* de mercado e outras estratégias de otimização. A principal contribuição deste trabalho é o desenvolvimento e avaliação de metodologias que integram métricas de risco amplamente utilizadas na teoria de finanças com algoritmos de otimização *online*, proporcionando uma abordagem adaptativa para a gestão dinâmica de portfólios

em mercados voláteis. Futuros trabalhos utilizando dados de outros mercados e novos tipos de distribuições para os retornos poderão gerar mais robustez para as metodologias desenvolvidas.

Palavras-chave: Gestão de Risco, *Time varying CAPM*, *Value at Risk*, *Conditional Value at Risk*, Otimização de Portfólio *Online*, Algoritmo *Online Gradient Descent*.

ABSTRACT

The objective of this work is to investigate the benefits of incorporating econometric risk metrics into the performance of online learning algorithms for solving the dynamic portfolio allocation problem. The problem is formulated as an online decision process where variants of the Online Gradient Descent (OGD) method are studied. Three modifications to the OGD algorithm were proposed, each using a distinct form of convex projection. The first modification includes a time-varying beta estimation constraint during the projection step, aiming to control the systematic risk of the portfolio. Recognizing that beta alone does not capture all aspects of risk, a second modification introduces Value at Risk (VaR) as a more comprehensive risk measure. This approach allows for more effective control of portfolio risk exposure, considering both systematic and unsystematic risk. The third modification utilizes a convex projection that minimizes Conditional Value at Risk (CVaR). This approach demonstrated more consistent results in terms of balancing risk and return, providing a more robust and effective risk management strategy. Before implementing the proposed modifications, mathematical proof of the projections convexity for each method was provided to ensure theoretical validity and computational efficiency of the developed algorithms. The impact of the choice of probability distribution (normal and Student's t) on VaR minimization during the projection step and its influence on performance and risk profile of the optimized portfolios is also investigated. The proposed methodologies are empirically evaluated on different Brazilian datasets (IBOVESPA and Small Caps), considering both high and low capitalization stocks in different market periods. The results show that incorporating risk metrics, especially CVaR, into the online optimization process allowed for better risk control while maintaining competitive returns compared to market benchmarks and other optimization strategies. The main contribution of this work is the development and evaluation of methodologies that integrate widely used risk metrics in financial theory with online optimization algorithms, providing an adaptive approach for dynamic portfolio management in volatile markets. Future studies using data from other markets and new types of return distributions may further enhance the robustness of the developed methodologies.

Keywords: Risk Management, Time varying CAPM, Value at Risk, Conditional Value at Risk, Online Portfolio Optimization, Online Gradient Descent Algorithm.

LISTA DE GRÁFICOS

Gráfico 1 – Fronteira eficiente de Markowitz	47
Gráfico 2 – Distribuição de Retornos do Portfólio e VaR 1%	83
Gráfico 3 – Comparação de Distribuições de Retornos com Diferentes Vari- âncias e seus VaR 1%	84
Gráfico 4 – Simplificação do modelo de projeção de VaR para dois ativos	92
Gráfico 5 – Representação Visual das Restrições de Risco no Projected OGD com VaR	93
Gráfico 6 – Resultados do modelo OGD sem controlar o beta - comparação ao índice <i>Small Caps</i> e UCRP	103
Gráfico 7 – Resultados do modelo OGD com beta entre -0,3 e 0,1 - compa- ração ao índice <i>Small Caps</i>	105
Gráfico 8 – Evolução do portfólio OGD_beta - ativos pertencentes ao índice <i>Small Caps</i> e/ou livre de risco	106
Gráfico 9 – Resultados do modelo OGD com beta entre 1,5 e 3 - comparação ao índice <i>Small Caps</i>	108
Gráfico 10 – Resultados do modelo OGD com beta entre 1,9 e 3 - comparação ao índice <i>Small Caps</i> e UCRP	110
Gráfico 11 – Resultados do modelo OGD com beta entre -0,3 e 0,1 - compa- ração ao índice <i>Small Caps</i> e UCRP	112
Gráfico 12 – Resultados do modelo OGD sem controlar o beta - comparação ao índice Bovespa e UCRP	114
Gráfico 13 – Resultados do modelo OGD com beta entre 1,4 e 3 - comparação ao índice Bovespa e UCRP	116
Gráfico 14 – Resultados do modelo OGD com beta entre -0,3 e 0,1 - compa- ração ao índice Bovespa e UCRP	118
Gráfico 15 – Comparação entre o portfólio OGD, UCRP, I_SMALL e Mar- kowitz	122
Gráfico 16 – Análise Temporal da Alocação de Ativos na Carteira OGD: Concentração e Diversificação de Investimentos de 2011 a 2014 para dados <i>small caps</i>	124

Gráfico 17 – Comparação entre o portfólio OGD, OGD_80 Normal, OGD_40 Normal, OGD_20 Normal e OGD_10 Normal para dados <i>small caps</i>	126
Gráfico 18 – Desempenho das carteiras: TCAC (%) vs. VaR 99 (%) com Riqueza Terminal como tamanho do círculo para dados <i>small caps</i>	128
Gráfico 19 – Comparação dos pesos investidos nos portfólios OGD, OGD_80 Normal, OGD_40 Normal, OGD_20 Normal e OGD_10 Normal para dados <i>small caps</i>	129
Gráfico 20 – Retornos e VaR Normal para dados <i>small caps</i>	131
Gráfico 21 – <i>Box-Plots</i> dos retornos brutos das carteiras OGD, Markowitz, UCRP e I_SMALL	132
Gráfico 22 – <i>Box-Plots</i> dos retornos brutos das carteiras OGD, OGD_80 Normal, OGD_40 Normal, OGD_20 Normal e OGD_10 Normal para dados <i>small caps</i>	133
Gráfico 23 – Comparação entre os portfólios OGD, OGD_80 t Student, OGD_40 t Student, OGD_20 t Student e OGD_10 t Student para dados <i>small caps</i>	134
Gráfico 24 – Desempenho das carteiras: TCAC (%) vs. VaR 99 (%) com Riqueza Terminal como tamanho do círculo para dados <i>small caps</i>	138
Gráfico 25 – Comparação dos pesos investidos nos portfólios OGD_80 t Student, OGD_40 t Student, OGD_20 t Student e OGD_10 t Student para dados <i>small caps</i>	139
Gráfico 26 – Retornos e VaR t de Student para dados <i>small caps</i>	141
Gráfico 27 – <i>Box-Plots</i> comparativos dos retornos brutos das carteiras OGD_80, OGD_40, OGD_20 e OGD_10 considerando distribuições Normal e t de Student para dados <i>small caps</i>	141
Gráfico 28 – Comparação entre o portfólio OGD, IBOV, UCRP e Markowitz	145
Gráfico 29 – <i>Box-Plots</i> dos retornos brutos das carteiras OGD, OGD_80 Normal, OGD_40 Normal, OGD_20 Normal e OGD_10 Normal para dados <i>IBOV</i>	147

Gráfico 30 – Comparação de Desempenho e Risco: Análise das Carteiras s OGD_80 Normal, OGD_40 Normal, OGD_20 Normal e OGD_10 Normal para dados IBOV	148
Gráfico 31 – Comparação dos pesos investidos nos portfólios OGD, OGD_80 Normal, OGD_40 Normal, OGD_20 Normal e OGD_10 Normal para dados IBOV	150
Gráfico 32 – Retornos e VaR Normal para dados IBOV	151
Gráfico 33 – Comparação entre os portfólios OGD, OGD_80 t Student, OGD_40 t Student, OGD_20 t Student e OGD_10 t Student para dados IBOV	153
Gráfico 34 – Retornos e VaR t de Student para dados IBOV	154
Gráfico 35 – Desempenho das carteiras: TCAC (%) vs. VaR 99 (%) com Riqueza Terminal como tamanho do círculo para dados IBOV .	156
Gráfico 36 – Comparação dos pesos investidos nos portfólios OGD_80 t Student, OGD_40 t Student, OGD_20 t Student e OGD_10 t Student para dados IBOV	158
Gráfico 37 – Comparação entre os portfólios OGD, OGD_80 CVaR, OGD_40 CVaR, OGD_20 CVaR e OGD_10 CVaR para dados <i>small caps</i>	163
Gráfico 38 – Comparação dos pesos investidos nos portfólios OGD_80 CVaR, OGD_40 CVaR, OGD_20 CVaR e OGD_10 CVaR para dados <i>small caps</i>	167
Gráfico 39 – Desempenho das carteiras: TCAC (%) vs. VaR (%) com Riqueza Terminal como tamanho do círculo para dados <i>small caps</i>	168
Gráfico 40 – Análise Comparativa de Riqueza Terminal e CVaR entre a carteira OGD e estratégias de minimização de CVaR, e minimização de VaR com Distribuição t Student para dados <i>small caps</i>	170
Gráfico 41 – <i>Box-Plots</i> dos retornos brutos das carteiras OGD_80 CVaR, OGD_40 CVaR, OGD_20 CVaR e OGD_10 CVaR para dados <i>small caps</i>	171
Gráfico 42 – Comparação entre os portfólios OGD, OGD_80 CVaR, OGD_40 CVaR, OGD_20 CVaR e OGD_10 CVaR para dados IBOV . .	173

Gráfico 43 – Análise Comparativa de Riqueza Terminal e CVaR entre a carteira OGD e estratégias de minimização de CVaR, e minimização de VaR com Distribuição t Student para dados IBOV	176
Gráfico 44 – <i>Box-Plots</i> dos retornos brutos das carteiras OGD_80 CVaR, OGD_40 CVaR, OGD_20 CVaR e OGD_10 CVaR para dados IBOV	178
Gráfico 45 – Comparação dos pesos investidos nos portfólios OGD_80 CVaR, OGD_40 CVaR, OGD_20 CVaR e OGD_10 CVaR para dados IBOV	179

LISTA DE TABELAS

Tabela 1 – Comparação entre métricas de risco para o portfólio OGD sem controlar o beta, índice <i>Small Caps</i> e UCRP	104
Tabela 2 – Comparação entre métricas de risco para o OGD com beta entre -0,3 e 0,1.	107
Tabela 3 – Comparação entre métricas de risco para o OGD com beta entre 1,5 e 3.	109
Tabela 4 – Comparação entre as métricas de risco para o OGD com beta entre 1,9 e 3.	111
Tabela 5 – Comparação entre métricas de risco para o OGD com beta entre -0,3 e 0,1.	113
Tabela 6 – Comparação entre métricas de risco para o portfólio OGD sem controlar o beta índice Bovespa e UCRP	115
Tabela 7 – Comparação entre métricas de risco para o OGD com beta entre 1,4 e 3.	117
Tabela 8 – Comparação entre métricas de risco para o OGD com beta entre -0,3 e 0,1.	119
Tabela 9 – Comparação de Desempenho e Risco: Análise das Carteiras OGD, Markowitz, UCRP e I_SMALL	123
Tabela 10 – Comparação de Desempenho e Risco: Análise das Carteiras s OGD_80 Normal, OGD_40 Normal, OGD_20 Normal e OGD_10 Normal para dados <i>small caps</i>	127
Tabela 11 – Comparação de Desempenho e Risco: Análise das Carteiras OGD_80 t Student, OGD_40 t Student, OGD_20 t Student, OGD_10 t Student para dados <i>small caps</i>	135
Tabela 12 – Comparação do número de excedências para os modelos VaR Normal e VaR t de Student para dados <i>small caps</i>	142
Tabela 13 – Comparação de Desempenho e Risco: Análise das Carteiras OGD, Markowitz, UCRP e IBOV	145
Tabela 14 – Comparação de Desempenho e Risco: Carteiras OGD com suposição de distribuição Normal para dados IBOV	149

Tabela 15 – Comparação de Desempenho e Risco: Análise das Carteiras OGD_80 t Student, OGD_40 t Student, OGD_20 t Student, OGD_10 t Student para dados IBOV	155
Tabela 16 – Comparação de Desempenho e Risco: Análise das Carteiras OGD_80 CVaR, OGD_40 CVaR, OGD_20 CVaR, OGD_10 CVaR para dados <i>small caps</i>	164
Tabela 17 – Comparação de Desempenho e Risco: Análise das Carteiras OGD_80 CVaR, OGD_40 CVaR, OGD_20 CVaR, OGD_10 CVaR para os dados IBOV	174
Tabela 18 – Resumo de resultados de outros trabalhos focados em dados brasileiros e modelagem <i>machine learning</i>	187

LISTA DE ABREVIATURAS E SIGLAS

AA	<i>Aggregating-Type Algorithms</i>
AEM	Algoritmos Evolutivos Multiobjetivos
Anticor	<i>Anticorrelation</i>
AOLNPM	<i>Adaptive Online Net Profit Maximization</i>
BCRP	<i>Best Constant Rebalanced Portfolios</i>
BS	<i>Best Stock</i>
CAPM	<i>Capital Asset Pricing Model</i>
CDI	Certificado de Depósito Interbancário
CORN	<i>Correlation-Driven Nonparametric Learning Strategy</i>
CRP	<i>Constant Rebalanced Portfolio</i>
CVaR	<i>Conditional Value at Risk</i>
DJIA	<i>Dow Jones Industrial Average Index</i>
DD	<i>Drawdown</i>
EG	<i>Exponential Gradiente Algorithm</i>
EGARCH	<i>Exponential Generalized Autoregressive Conditional Heteroskedasticity</i>
ES	<i>Expected Shortfall</i>
EZTC3	ações da EZ TEC Empreendimentos e Participações
FLH	<i>Follow the Leading History</i>
FTL	<i>Follow-the-Loser</i>
FTW	<i>Follow-the-Winner</i>

FU	<i>Fast Universalization</i>
GARCH	<i>Generalized Autoregressive Conditional Heteroskedasticity</i>
HBOR3	Ações da Helbor Empreendimentos
HML	<i>High Minus Low</i>
IBOV	índice Bovespa
IBrX-100	Índice Brasil 100
IID	independentemente e identicamente distribuídos
I_SMALL	índice <i>Small Cap</i> SMLL B3
LCA	Letra de Crédito do Agronegócio
LCI	Letra de Crédito Imobiliário
LASSO	<i>Least Absolute Shrinkage and Selection Operator</i>
MDD	<i>Máximo Drawdown</i>
MLA	<i>Meta-Learning Algorithms</i>
MV	Modelo Média-Variância
NKB	<i>Nonparametric Kernel-based Log-Optimal Strategy</i>
NNN	<i>Nonparametric Nearest Neighbor Log-Optimal Strategy</i>
OCAPM	<i>Omega Capital Asset Pricing Model</i>
OGD	<i>Online Gradient Descent</i>
OLS	<i>Ordinary Least Squares</i>
ONS	<i>Online Newton Step</i>
PAMR	<i>Passive Aggressive Mean Reversion</i>
PMA	<i>Pattern-Matching Approaches</i>

RMR	<i>Robust Median Reversion</i>
SMB	<i>Small Minus Big</i>
SCR	<i>Successive Constant Rebalanced Algorithm</i>
SVM	<i>Support Method Machine</i>
SOCP	<i>Second-Order Cone Programming</i>
TCAC	Taxa de Crescimento Anual Composta
TGARCH	<i>Threshold Generalized Autoregressive Conditional Heteroskedasticity</i>
TPPT	<i>Trend Peak Price Tracing</i>
UCRP	<i>Uniform Constant Rebalanced Portfolio</i>
VaR	<i>Value at Risk</i>
VL	<i>Value Line</i>
WASA	<i>Weak Aggregating Specialist Algorithm</i>
WUP	<i>Weighted Universal Portfolio</i>

LISTA DE SÍMBOLOS

β	beta do mercado, risco sistemático ou sensibilidade de um ativo em relação ao período de estimativa
$p_{i,t}$	preço do ativo i no período t
$r_{i,t}$	retorno do ativo i no período t
$\boldsymbol{\mu}$	vetor de dimensão $n \times 1$ de retornos médios
$\sigma_{i,j}$	covariância entre os ativos ou retorno dos ativos i e j
$\boldsymbol{\Sigma}$	a matriz de covariância formada por elementos $\sigma_{i,j}$
σ_i^2	variância do ativo ou retorno do ativo i
\mathbf{b}	vetor coluna que denomina uma alocação em um portfólio
b_{it}	proporção de portfólio do ativo i no período de negociação t e elemento de \mathbf{b}_t
$\mathcal{N}_{\boldsymbol{\mu},\boldsymbol{\Sigma}}$	distribuição normal multivariada com parâmetros $\boldsymbol{\mu}$ e $\boldsymbol{\Sigma}$ representando média e covariância, respectivamente.
$r_{m,t}$	retorno médio do mercado no período t
$r_{f,t}$	retorno de um ativo livre de risco no período t
β_i	sensibilidade do ativo i em relação ao período de estimativa
$\varepsilon_{i,t}$	choque idiossincrático ligado ao ativo i no período t
$\beta_{i,t}$	sensibilidade do ativo i em t
φ_t	variável aleatória gaussiana em t
P_t	parâmetro no período t do modelo espaço estado
Q_t	parâmetro no período t do modelo espaço estado
K_t	ganho de Kalman no período t

$F(\cdot)$	função de distribuição cumulativa ou acumulada
γ	nível de confiança
$VaR_\gamma(L)$	<i>Value at Risk</i> ligado a perda L e nível de confiança γ
$F^{-1}(\cdot)$	função quantílica ou quantil ligado a distribuição F
$\mathcal{N}_{\mu, \sigma^2}$	distribuição normal univariada com parâmetros μ e σ^2 representando média e variância, respectivamente
Φ	função distribuição acumulada ligado a distribuição normal padrão
$\Phi^{-1}(\cdot)$	função quantílica ou quantil ligado a distribuição normal padrão
Z_γ	percentil correspondente ao nível de confiança γ em uma distribuição normal padrão
ν	graus de liberdade em uma distribuição t de Student
$\Gamma(\cdot)$	função Gamma
$t_{\gamma, \nu}$	percentil correspondente ao nível de confiança γ em uma distribuição t de Student com ν graus de liberdade
T	número de períodos / horizonte de investimento
T^{anu}	número de períodos / horizonte de investimento em termos anuais
W_t	riqueza no final do período de negociação t
$E(X A)$	esperança de X condicionada a A
$CVaR_\gamma(L)$	<i>Conditional Value at Risk</i> ligado a perda L e nível de confiança γ
$f_t(\cdot)$	função de perda ou custo avaliada em t
A_i	um ativo i
Y_i^{inic}	quantidade inicial de ações de cada ativo i no tempo $t = 0$

Y_{it}	quantidade de ações do ativo i até o tempo t , após o rebalanceamento
y_{it}	quantidade de ações do ativo i negociadas em t
$\nabla f_t(\cdot)$	gradiente da função de perda ou custo avaliada em t
η_t	taxa de aprendizado de uma ALG no tempo t
$\ \mathbf{x}\ $	norma euclidiana padrão aplicada a um vetor coluna \mathbf{x}
D	diâmetro do conjunto viável de um ALG
G	cota superior para a norma do gradiente das funções de perda ou custo
$O(n)$	Ordem de um algoritmo
\mathcal{D}	subconjunto convexo do \mathbb{R}^n
$\Pi_{\mathcal{D}}$	uma projeção de \mathbb{R}^n para um conjunto convexo $\mathcal{D} \subset \mathbb{R}^n$
\mathfrak{B}_n	simplex com n variáveis de decisão
\mathbb{N}	conjunto dos números naturais
\mathbb{N}^*	conjunto dos números naturais diferentes de zero
\mathbf{b}'	uma alocação de portfólio intermediária do simplex \mathfrak{B}_n
\mathbf{b}^*	alocação de portfólio ótimo do simplex \mathfrak{B}_n
Ω	número de diferentes alocações em um simplex discretizado
α	parâmetro de um modelo ou ALG
ξ	valor de um VaR em um ALG
\mathcal{S}_t	conjunto de alocações viáveis em um ALG em t usando informações até o momento $t - 1$

\mathbf{C}_t matriz formada pelos vetores característicos de uma matriz positiva semi-definida $\mathbf{\Sigma}_t$

$\mathbf{\Lambda}_t$ matriz formada pelos valores característicos de uma matriz positiva semi-definida $\mathbf{\Sigma}_t$

SUMÁRIO

1	INTRODUÇÃO	28
1.1	OBJETIVOS	30
1.1.1	OBJETIVOS GERAIS	30
1.1.2	OBJETIVOS ESPECÍFICOS	30
1.2	MOTIVAÇÃO	32
1.3	ESTRUTURA DO TRABALHO	34
2	REVISÃO DE LITERATURA	35
3	FUNDAMENTAÇÃO TEÓRICA	45
3.1	MODELO MÉDIA-VARIÂNCIA	45
3.2	MODELO CAPM: BETA DO MERCADO	49
3.2.1	Filtro de Kalman	52
3.3	RISCO	54
3.3.1	<i>Value at Risk</i> (VaR)	55
3.3.2	<i>Conditional Value at Risk</i> (CVaR)	58
3.3.3	Métricas de Análise de Risco	61
3.4	OTIMIZAÇÃO CONVEXA <i>ONLINE</i>	63
3.4.1	<i>Constant Rebalanced Portfolio</i>	64
3.4.2	Seleção de Portfólio <i>Online</i>	66
3.4.3	<i>Online Gradient Descent</i> (OGD)	68
4	TÉCNICAS DE PROJEÇÃO PARA CONTROLE DE RISCO	74
4.1	PROJEÇÃO EM UM SIMPLEX	75
4.2	PROJEÇÃO USANDO BETA VARIANTE NO TEMPO	79
4.3	<i>VALUE AT RISK</i> COMO RESTRIÇÕES CÔNICAS DE SEGUNDA ORDEM	81
4.3.1	VaR e sua relação com a distribuição dos retornos	82
4.3.2	Incorporação do VaR como restrição no modelo de otimização	85

4.4	<i>CONDITIONAL VALUE AT RISK</i> COMO RESTRIÇÃO	94
4.4.1	Incorporação do CVaR como restrição no modelo de otimização	95
5	DADOS E APLICAÇÃO EMPÍRICA	100
6	RESULTADOS	102
6.1	RESULTADOS DO OGD COM BETA VARIANTE NO TEMPO	102
6.1.1	OGD com beta variante no tempo para dados <i>Small Caps</i> - Resultados para períodos específicos de tempo	109
6.1.2	OGD com beta variante no tempo para dados IBOV - Resultados Gerais	113
6.2	RESULTADOS DO OGD COM RESTRIÇÃO <i>VALUE AT RISK</i>	119
6.2.1	OGD com restrição VaR para dados <i>Small Caps</i> - Resultados considerando retornos com distribuição normal	122
6.2.2	OGD com restrição VaR para dados <i>Small Caps</i> - Resultados considerando retornos com distribuição t de Student	134
6.2.3	OGD com restrição VaR para dados IBOV - Resultados conside- rando retornos com distribuição normal	144
6.2.4	OGD com restrição VaR para dados IBOV - Resultados conside- rando retornos com distribuição t de Student	152
6.3	RESULTADOS DO OGD COM RESTRIÇÃO <i>CONDITIONAL VALUE AT RISK</i>	161
6.3.1	OGD com restrição CVaR para dados <i>Small Caps</i>	162
6.3.2	OGD com restrição CVaR para dados IBOV	172
6.4	RESUMO DE RESULTADOS E COMPARAÇÕES COM OU- TROS TRABALHOS	181
7	CONCLUSÕES	193
7.1	TRABALHOS FUTUROS	196
	REFERÊNCIAS BIBLIOGRÁFICAS	199

1 INTRODUÇÃO

O problema de seleção de portfólios é um processo decisório no qual o investidor deve alocar uma quantidade de riqueza em um conjunto finito de ativos dentro de um horizonte de tempo. Para solucionar o problema, o investidor decide quanto de sua riqueza irá destinar para cada um dos ativos disponíveis no mercado. Cada ativo representa uma oportunidade de investimento distinta e a decisão realizada para uma alocação é um portfólio. A escolha do portfólio ótimo é um problema tão antigo quanto o próprio mercado de ações, e uma vez que o retorno dos ativos são variáveis aleatórias, sujeito a incertezas, torna o problema de alocação ainda menos claro. Foi com o trabalho seminal de Markowitz [1] que o problema de alocação de carteiras ganhou uma abordagem rigorosa, sendo posta como um problema matemático cujo objetivo é minimizar risco e maximizar retorno, ao invés do senso comum da época de olhar apenas para valores esperados. Após o trabalho de Markowitz, abordagens modernas para esse problema buscam obter uma boa relação entre o retorno esperado e o risco do portfólio.

Nas últimas décadas, abordagens baseadas em técnicas de aprendizado de máquina têm sido intensamente aplicadas, tornando-se uma área de pesquisa importante e ativa, por oferecer a possibilidade de definir estratégias *online* de investimento que permitem maximizar a riqueza sem a utilização de pressupostos estatísticos sobre o preço dos ativos. No entanto, em contraste com as abordagens econométricas modernas, os algoritmos de otimização *online* possuem o aspecto negativo de buscarem maximizar o retorno esperado do portfólio, ignorando completamente a noção de risco-retorno introduzida por Markowitz [2, 3, 4]. Existem, portanto, duas comunidades que abordam esse problema de maneiras distintas: a comunidade financeira e a comunidade de aprendizado de máquina. A comunidade financeira é aquela que se baseia no trabalho de Markowitz e outros como [5, 6, 7], focando em modelos de otimização de variáveis aleatórias para maximizar uma medida de desempenho considerando o risco e a riqueza. Já a comunidade de aprendizado de máquina utiliza dados de séries temporais para desenvolver algoritmos que sugerem alocações de portfólio baseadas em informações históricas de preços dos ativos. Estes algoritmos visam maximizar a riqueza ao final do

horizonte de investimento.

Nesta pesquisa, avança-se no problema de seleção de portfólios ao estabelecer uma conexão entre a teoria financeira e os métodos contemporâneos de aprendizado de máquina. São propostas três metodologias distintas que integram métricas de risco amplamente utilizadas na teoria de finanças com algoritmos de otimização *online*. Essas abordagens demonstram a capacidade de controlar eficazmente a qualidade no processo de construção dos portfólios.

A primeira metodologia leva em conta o comportamento relativo dos ativos em relação ao mercado, capturado pelo coeficiente beta, crucial para a composição de um portfólio robusto, de maneira dinâmica impondo limites para a combinação dos betas de ativos que compõem uma carteira. Nela, o beta variante no tempo foi incorporado e estimado por meio do Filtro de Kalman, sendo inserido no processo de escolha num modelo de otimização *online*. Desta forma, é um algoritmo que respeita a natureza dinâmica do mercado e permite que o investidor controle o perfil de risco do portfólio de forma mais precisa, ao levar em conta possíveis momentos de queda de mercado (betas negativos) e/ou de crescimento do mesmo (betas positivos).

Já a segunda metodologia integra um mecanismo de controle de risco que opera dentro do contexto do *Value at Risk* (VaR)[8]. Ao introduzir uma restrição de VaR no processo de otimização *online*, o algoritmo proposto não só acompanha a trajetória de alocação ótima em termos de retorno esperado, mas também atenua o potencial de perdas extremas, alinhando-se com um perfil de risco pré-definido.

Por fim, a terceira metodologia é uma evolução natural da segunda dado que controla o risco no contexto do *Conditional Value at Risk* (CVaR) [9, 10], uma medida mais robusta que o VaR por ser coerente no sentido de permitir, dentre outras coisas, diversificação ligada a menor risco em casos onde a distribuição têm caudas extremamente pesadas [11]. Logo, além das vantagens do VaR destacadas acima, tende a se adequar melhor na formação de carteiras em mercados com maior presença de ativos com retornos extremos, um fato estilizado em alguns países [12, 13] e em períodos de crise como por exemplo a *subprime* e o início da guerra Rússia-Ucrânia [14, 15, 15, 16, 17].

Vale ressaltar que as três metodologias propostas utilizam o rigor da teoria de portfólios e as vantagens adaptativas dos métodos de aprendizado de máquina, contribuindo, assim, para uma prática de investimento que é ao mesmo analiticamente rigorosa e flexível às condições de mercado em constante mudança. Ao incorporar aspectos de risco diretamente no processo de decisão, busca-se aproveitar o melhor de dois mundos: a capacidade adaptativa dos algoritmos de otimização *online* e a consideração de medidas de risco-retorno realistas, amplamente utilizadas na teoria financeira.

Nesse sentido, esse trabalho contribui para a literatura de otimização *online*, ao propor metodologias que integram métricas de risco no processo de decisão, visando obter estratégias de investimento robustas e alinhadas com as práticas de gestão de risco no mundo real.

1.1 OBJETIVOS

Esta seção visa enumerar os objetivos gerais e específicos desenvolvidos nessa tese.

1.1.1 OBJETIVOS GERAIS

- Desenvolver três metodologias que incorporaram métricas de risco econométricas - o coeficiente beta (β), o *Value at Risk* (VaR) e o *Conditional Value at Risk* (CVaR) - no processo de otimização de portfólios utilizando algoritmos de aprendizado *online*, visando aprimorar a gestão de risco e a adaptabilidade na alocação dinâmica de ativos.

1.1.2 OBJETIVOS ESPECÍFICOS

- Investigar a viabilidade da primeira metodologia, isto é, incorporar o coeficiente beta variante no tempo, estimado por meio do Filtro de Kalman, no processo de otimização de portfólios com algoritmos de aprendizado *online*, buscando capturar a dinâmica temporal da correlação entre os ativos e o mercado.

- Analisar a capacidade da primeira metodologia em adaptar a alocação de ativos de acordo com as condições do mercado, explorando a relação entre o beta do portfólio e os períodos de alta e baixa do mercado.
- Investigar a viabilidade da segunda metodologia, isto é, incorporar restrições ao risco por meio do VaR no processo de otimização de portfólios com algoritmos de aprendizado *online*, buscando maximizar os retornos ao limitar o risco das carteiras via uma métrica comumente utilizada no mercado financeiro.
- Investigar a viabilidade da terceira metodologia, isto é, restringir a escolha ao utilizar o CVaR como medido de risco chave no processo de otimização de portfólios com algoritmos de aprendizado *online*, objetivando maximizar retornos via medida mais adequada a mercados com distribuições com caudas pesadas.
- Comparar o desempenho e as características de risco das carteiras otimizadas com a incorporação do VaR, CVaR e do beta com *benchmarks* de mercado e carteiras obtidas por meio de outras estratégias de otimização, como a carteira uniformemente ponderada *Uniform Constant Rebalanced Portfolio* (UCRP) e a carteira de mínima variância (Markowitz).
- Analisar os impactos de modificações no algoritmo *Online Gradient Descent* (OGD) ao se incluir restrições de risco baseadas no VaR, CVaR e no coeficiente beta durante a etapa de projeção, permitindo um controle mais efetivo da exposição ao risco do portfólio e da correlação com o mercado.
- Avaliar a eficácia das metodologias propostas em diferentes conjuntos de dados, considerando ações de alta e baixa capitalização (IBOVESPA e *Small Caps*) e períodos de mercado distintos (alta e baixa), a fim de verificar sua robustez e aplicabilidade em diversos cenários.
- Investigar o impacto da escolha da distribuição de probabilidade (normal e t de Student) na estimação do VaR e sua influência no desempenho e no perfil de risco das carteiras otimizadas.

- Avaliar a aplicabilidade das metodologias propostas na gestão de investimentos, destacando suas vantagens, limitações e possíveis extensões para pesquisas futuras, considerando a avaliação de outros mercados, uso de outras distribuições e restrições adicionais, como limites de alavancagem e restrições de venda a descoberto (*short selling*).

1.2 MOTIVAÇÃO

Os mercados financeiros contemporâneos são bastante complexos e dinâmicos, alimentados por avanços tecnológicos, globalização e a interconexão crescente entre economias globais. Neste ambiente de alta volatilidade e incerteza, a gestão eficaz de portfólios de investimento tornou-se um desafio crítico, exigindo estratégias adaptativas e robustas que possam responder rapidamente às mudanças nas condições de mercado.

Dentro deste contexto, os algoritmos de seleção de portfólio *online* surgiram como uma solução promissora, recebendo crescente atenção tanto da academia quanto da indústria financeira. Baseados em princípios de aprendizado de máquina e otimização convexa *online*, esses algoritmos têm demonstrado bons desempenhos em estudos empíricos, superando *benchmarks* tradicionais e estratégias passivas¹, mesmo em cenários de alta volatilidade.

A principal vantagem dos algoritmos baseados na estratégia de otimização *online* reside na capacidade de absorver e se adaptar a novas informações de mercado de forma contínua, ajustando as alocações de portfólio em tempo real. Isso os torna particularmente adequados para ambientes não estacionários, onde os padrões e relações entre ativos estão em constante evolução. Contudo, uma limitação crítica nas abordagens atuais é a incorporação insuficiente de métricas de risco durante o processo de otimização. Embora o foco tradicional seja maximizar o retorno ou minimizar o arrependimento (*Regret*) em relação a *benchmarks*, essas estratégias

¹ Estratégias passivas de investimento, também conhecidas como indexação, são abordagens que visam replicar o desempenho de um índice de mercado específico, como o S&P 500 ou o IBOVESPA. Estas estratégias geralmente envolvem menor rotatividade de ativos e custos mais baixos em comparação com estratégias ativas, mas não buscam superar o desempenho do mercado. Exemplos comuns incluem fundos de índice e ETFs (*Exchange-Traded Funds*) que acompanham índices de mercado.

em alguns casos negligenciam a consideração dinâmica e explícita do risco, o que pode expor os portfólios a volatilidades excessivas ou a perdas significativas em cenários adversos.

Essa falha em tratar o risco de maneira robusta pode resultar em portfólios que, embora lucrativos em média, estão expostos a níveis inaceitáveis de risco ou que não atendem a critérios regulatórios e institucionais, limitando sua aplicabilidade prática. Por isso, é importante desenvolver abordagens que integrem o risco de forma dinâmica e eficiente, garantindo que os portfólios não apenas busquem maximizar retornos, mas também mantenham um perfil de risco controlado e compatível com as expectativas dos investidores.

Este trabalho visa somar a literatura por meio da incorporação de métricas de risco avançadas no *framework* de seleção de portfólio *online*. Especificamente, são exploradas três metodologias que respeitam a convexidade do problema de otimização:

1. A utilização de um beta variante no tempo, estimado via filtro de Kalman, para capturar dinamicamente o risco sistemático, adaptando o modelo às mudanças do mercado.
2. A incorporação do Value at Risk (VaR) como uma restrição cônica de segunda ordem no processo de otimização, permitindo um controle mais preciso do risco.
3. A implementação do Conditional Value at Risk (CVaR) para o controle de risco extremo.

Embora essas técnicas sejam bem estabelecidas na literatura de gestão de risco tradicional, elas ainda não foram plenamente integradas ao paradigma de seleção de portfólio *online*. A principal hipótese avaliada é que sua incorporação não apenas preservará a adaptabilidade e a performance dos métodos existentes, mas também resultará em estratégias com perfis de risco-retorno aprimorados e mais alinhados com as exigências dos investidores institucionais e reguladores.

Assim, ao abordar essa temática, esta pesquisa pretende contribuir para o desenvolvimento de futuros algoritmos de seleção de portfólio *online*, que sejam eficazes na geração de retornos e, ao mesmo tempo, robustos e confiáveis frente às incertezas e volatilidades dos mercados financeiros.

1.3 ESTRUTURA DO TRABALHO

O restante do trabalho é apresentado de forma que, primeiramente, no capítulo 2, é realizada uma revisão da literatura, onde trabalhos anteriores a esse são destacados. Depois, no capítulo 3, é efetuada uma fundamentação teórica para os objetivos propostos, com discussões acerca de modelos comumente vistos no mercado e na literatura financeira (Média-Variância, CAPM, VaR e CVaR) e sobre otimização convexa *online*. Já no capítulo 4 são apresentados os algoritmos propostos e utilizados. Depois, no capítulo 5, há uma descrição dos dados e parâmetros utilizados nos resultados apresentados no capítulo 6. Por fim, no capítulo 7, há conclusões e uma descrição de trabalhos futuros.

2 REVISÃO DE LITERATURA

A otimização de portfólios *online* é uma área bastante estudada, com contribuições significativas de pesquisadores que abordam o problema sob diferentes perspectivas. No entanto, uma limitação observada em grande parte dos trabalhos existentes é a ausência de uma consideração explícita do risco durante a otimização dos portfólios.

Os algoritmos de otimização frequentemente focam na maximização do retorno, mas negligenciam os aspectos de risco, o que pode levar a decisões sub-ótimas em cenários de alta volatilidade ou em mercados adversos. Em seu trabalho seminal, Cover [18] introduziu o conceito de Portfólios Universais. Neste trabalho, propõe um algoritmo para seleção de portfólios que, assintoticamente, pode superar o desempenho da melhor ação individual no mercado. A estratégia proposta é adaptativa e não requer conhecimento prévio dos retornos futuros; em vez disso, ajusta-se dinamicamente com base nos retornos passados.

O principal objetivo do algoritmo de Cover é alcançar uma performance que se aproxime à do melhor portfólio reequilibrado constante, definido como aquele que teria sido idealmente escolhido com o conhecimento dos retornos futuros. A estratégia é robusta e não depende de suposições estatísticas específicas sobre o comportamento do mercado, permitindo assim que o portfólio se adapte a diferentes sequências de retornos.

Cover demonstrou que a riqueza acumulada pela estratégia universal converge, em termos de taxa de crescimento exponencial, para a do portfólio ótimo, que maximiza o crescimento da riqueza ao longo do tempo. Ele mostrou que a diferença na taxa de crescimento entre a estratégia universal e o portfólio ótimo diminui para zero à medida que o número de períodos de negociação aumenta. Em outras palavras, demonstrou que existe uma estratégia de alocação constante que proporciona o maior crescimento da riqueza a longo prazo (*Best Constant Rebalanced Portfolio - BCRP*).

O Trabalho de Cover teve influências diretas e contribuições fundamentais das teorias da informação, probabilidade e finanças. O ponto de partida para essa

formulação foi o critério de Kelly [19]. Kelly mostrou que, em um contexto de apostas, a estratégia que maximiza o crescimento logarítmico do capital ao longo do tempo também maximiza a taxa de transmissão de informação em um canal de comunicação. Este critério, frequentemente referido como "critério de Kelly" ou "aposta proporcional", estabelece que a alocação ótima do capital é proporcional à vantagem percebida. Para especialistas em finanças, o critério de Kelly se traduz em uma estratégia de alocação de ativos que busca maximizar o crescimento composto do portfólio, ao invés de simplesmente maximizar o retorno esperado ou minimizar o risco em um único período.

A generalização desse conceito foi expandida por Breiman [20], que examinou sistemas de apostas ótimos em jogos favoráveis. Breiman introduziu a ideia de que, ao adotar uma abordagem de diversificação, onde o capital é distribuído entre várias apostas (ou ativos), é possível maximizar o crescimento esperado do capital. Sua análise se baseia em princípios da teoria da probabilidade como a lei dos grandes números, que é diretamente aplicável à teoria moderna de portfólios, onde a diversificação é um dos princípios fundamentais. Breiman demonstrou que, em um jogo favorável, a estratégia ótima é aquela que maximiza o valor esperado do logaritmo do capital, resultando em um crescimento exponencial do portfólio ao longo do tempo.

Em um estudo de caráter mais aplicado, Finkelstein e Whitley [21] contribuíram com uma perspectiva mais prática, focando na eficiência das estratégias de investimento em mercados financeiros reais. Eles analisaram como as estratégias baseadas no critério de Kelly poderiam ser aplicadas para maximizar o crescimento do capital em condições de mercado voláteis e incertas. A robustez dessas estratégias foi destacada, demonstrando que, mesmo em cenários de mercado adversos, uma abordagem baseada na maximização do crescimento logarítmico do capital poderia superar estratégias tradicionais que não levam em conta a variabilidade do mercado.

O trabalho de Cover influenciou significativamente o desenvolvimento de vários algoritmos importantes na literatura de seleção de portfólios *online*, em especial os algoritmos da categoria *Follow the Winner*, cujo princípio fundamental é alocar mais capital nos ativos ou estratégias que demonstraram melhor desempenho

recente, baseando-se na premissa de que esse bom desempenho tende a persistir no curto prazo. Esses algoritmos geralmente tentam acompanhar ou superar o desempenho de outros que geralmente tentam acompanhar ou superar o desempenho do Melhor Portfólio Constante Rebalanceado (*Best Constant Rebalanced Portfolio* - BCRP) em retrospecto. Dentre os diversos trabalhos dessa área, abaixo são destacados alguns pelo seus graus de importância.

Helmhold et al. [22] propuseram um algoritmo que estende o conceito de portfólios universais, introduzido inicialmente por Cover. Este algoritmo utiliza atualizações multiplicativas para selecionar portfólios em uma abordagem *online*, onde a performance passada de cada portfólio influencia diretamente a seleção futura. A inovação principal está na adaptação das distribuições de Dirichlet para calcular a média ponderada dos portfólios possíveis, levando em consideração o desempenho anterior e o conhecimento a priori.

Vovk e Watkins [23] apresentaram uma generalização do algoritmo de portfólios universais de Cover, aplicando o *Aggregating-Type Algorithms* (ATA) a um conjunto não enumerável de portfólios reequilibrados constantemente. Esta abordagem expande os limites estabelecidos por Cover, permitindo a aplicação em mercados mais complexos, como aqueles com posições curtas, e estendendo a robustez do algoritmo para taxas de aprendizagem não superiores a 1.

Hazan e outros [24] desenvolveram um algoritmo que combina as ideias de portfólios universais com técnicas de otimização convexa *online*, resultando em uma garantia de desempenho logarítmico. Este algoritmo é capaz de ajustar dinamicamente as alocações de ativos em resposta às variações do mercado, maximizando o crescimento do capital ao longo do tempo, mesmo sem pressupor distribuições estatísticas específicas para os retornos dos ativos.

O trabalho de Singer [25] introduz um algoritmo de portfólios comutáveis, que oferece uma abordagem dinâmica para a seleção de portfólios em mercados financeiros voláteis. Em contraste com os portfólios reequilibrados constantes, que mantêm a mesma distribuição de riqueza ao longo do tempo, os portfólios comutáveis permitem que o investidor alterne entre diferentes estratégias de investimento com base nas mudanças observadas no mercado. Este método é particularmente eficaz

em mercados onde os padrões de desempenho dos ativos mudam ao longo do tempo, como ilustrado pelos exemplos de ações altamente voláteis discutidos no artigo. O algoritmo é capaz de rastrear a estratégia de investimento ótima de um investidor presciente sem necessidade de conhecimento prévio sobre a forma exata da estratégia ótima.

Em [26], Levina e Shafer propõem um método de seleção de portfólios que utiliza a maximização da entropia relativa para construir uma carteira que maximiza o crescimento esperado da riqueza. Este método baseia-se na ideia de que a entropia relativa pode ser utilizada como uma medida de incerteza no contexto da teoria da informação, e pode ser aplicada para ajustar dinamicamente a distribuição de investimentos entre diferentes ativos, maximizando o retorno esperado ao mesmo tempo em que controla o risco associado às mudanças no mercado. A abordagem de Levina e Shafer se destaca por sua capacidade de lidar com uma ampla gama de condições de mercado, ajustando-se automaticamente às flutuações na volatilidade dos ativos.

Por outro lado, Kozat e Singer [27] explora a ideia de portfólios universais com comutação sob custos de transação, estendendo o conceito de portfólios reequilibrados constantes para cenários onde há custos associados às transações financeiras. Em outro trabalho, Kozat e Singer [28] desenvolvem um algoritmo sequencial que é capaz de alcançar a riqueza do melhor portfólio reequilibrado por partes, ajustado para as comissões sobre transações. Este algoritmo não requer conhecimento prévio da duração do investimento ou do número de transações necessárias, e pode competir com um número exponencial de estratégias de investimento comutáveis usando apenas a complexidade linear em relação ao comprimento dos dados.

Baseado em evidências empíricas de reversão à média nos mercados financeiros e inspirado pela estratégia de Portfólios Constantemente Rebalanceados (CRP), uma nova abordagem para seleção de portfólio online emergiu, contrastando com as estratégias *Follow the Winner* anteriores. Esta nova classe de algoritmos, conhecida como *Follow the Loser*, foi inaugurada com o trabalho de Borodin [29], que introduziu o algoritmo *Anticorrelation* (Anticor). Enquanto as estratégias *Follow the Winner* buscavam capitalizar sobre a persistência de tendências, alocando

mais capital aos ativos com melhor desempenho recente, o Anticor adotou uma abordagem contrária. Ele foi projetado para explorar a tendência de reversão à média, transferindo riqueza dos ativos com melhor desempenho para aqueles com pior desempenho, na expectativa de que esta tendência se inverteria no futuro próximo. Esta mudança de paradigma foi motivada por observações empíricas nos mercados financeiros, que sugeriam que os preços dos ativos frequentemente reverterem para uma média após movimentos extremos. O Anticor incorporou esta ideia, utilizando medidas de correlação entre janelas de mercado para identificar oportunidades de reversão. Embora o Anticor não oferecesse as mesmas garantias teóricas robustas de alguns algoritmos *Follow the Winner*, seu desempenho empírico superior em muitos conjuntos de dados reais chamou a atenção da comunidade de pesquisa. Isso levou a um novo foco no desenvolvimento de estratégias de reversão à média, marcando uma virada significativa na pesquisa de seleção de portfólio *online*.

Baseado no algoritmo Anticor, outros pesquisadores começaram a explorar e expandir a ideia de reversão à média em contextos de seleção de portfólios *online*. Um dos avanços mais notáveis foi a introdução do algoritmo *Passive Aggressive Mean Reversion* (PAMR) por Li et al. [30]. O PAMR adota uma abordagem diferente para a reversão à média, combinando a ideia de ajuste passivo e agressivo, onde a alocação de capital é ajustada não apenas para maximizar retornos, mas também para mitigar perdas em situações de mercado adversas. Este algoritmo equilibra o risco e o retorno ao corrigir de modo agressivo as alocações quando os desvios da média são significativos, ao mesmo tempo em que se mantém passivo em cenários mais estáveis.

Dando continuidade a essa linha de pesquisa, Li et al. [31] introduziram uma extensão do PAMR baseada em intervalos de confiança, criando um modelo que se adapta de forma mais dinâmica às incertezas do mercado. Este modelo ajusta as alocações de portfólio com base em intervalos de confiança, proporcionando uma estratégia de investimento mais robusta em mercados voláteis, onde a previsibilidade dos retornos pode ser baixa.

Outro avanço importante foi a estratégia *Robust Median Reversion* (RMR), proposta por Huang et al. [32]. A *RMR* aprimora a estimativa de reversão à média

utilizando o estimador robusto mediano, que é menos sensível a ruídos e *outliers* presente nos dados do mercado. Essa abordagem resultou em um algoritmo que não só oferece melhor desempenho em condições de mercado variadas, como também mantém uma complexidade computacional linear, tornando-o altamente aplicável a cenários de negociação em grande escala.

O Trabalho de Li et al. [33] sobre a seleção de portfólios *online* consolidou essas ideias em um *framework* abrangente que combina técnicas de aprendizado *online* com otimização convexa. Este *framework* visa maximizar o crescimento do capital ao longo do tempo, ajustando continuamente as alocações de ativos em resposta às condições do mercado, e representa uma evolução significativa no campo, integrando conceitos de reversão à média com metodologias modernas de otimização.

Influenciados por resultados anteriores de trabalhos como Breiman [20], Algoet e Cover [34], além de Cover [18], que exploraram a seleção de portfólios em mercados com processos estocásticos estacionários e de sistemas dinâmicos, emergiu uma nova e importante classe de algoritmos conhecida como *Pattern Matching-based*. Esta abordagem representou uma mudança na estratégia de seleção de portfólio *online*, em especial por levar em conta que os mercados financeiros frequentemente exibem padrões recorrentes. A ideia fundamental é identificar sequências históricas de mercado similares à sequência atual e usar essa informação para fazer previsões sobre o comportamento futuro do mercado.

O trabalho seminal nesta categoria foi introduzido por Györfi e Schäfer [35] em 2003, com o desenvolvimento da estratégia de seleção de amostra baseada em histograma não-paramétrico. Este estudo introduziu estratégias de investimento sequenciais baseadas em métodos não-paramétricos, especificamente utilizando técnicas de *kernel*, projetadas para garantir uma taxa de crescimento ótima do capital sob suposições mínimas sobre o comportamento do mercado.

Esta base teórica permitiu o desenvolvimento de estratégias tanto teoricamente robustas quanto empiricamente eficazes. Um aspecto crucial de trabalho inicial de Györfi e Schäfer foi a discussão sobre a universalidade destas estratégias, demonstrando que elas podem atingir a taxa de crescimento ótima assintótica sem

conhecimento prévio das propriedades estatísticas do processo subjacente. Esta universalidade foi provada através do uso de estratégias baseadas em histogramas e em *kernel*, analisadas tanto de maneira teórica quanto empírica.

Seguindo este trabalho pioneiro, uma série de estudos expandiu e refinou a abordagem de *Pattern Matching*. Györfi et al. Em [36] foi introduzido a *Nonparametric Kernel-based Log-Optimal Strategy* (NKB), que aprimorou a abordagem original usando técnicas de *kernel* mais sofisticadas. Posteriormente, Györfi et al. [37] propuseram a *Nonparametric Nearest Neighbor Log-Optimal Strategy* (NNN), que utilizava uma abordagem de vizinhos mais próximos para identificar padrões similares.

Outras estratégias importantes incluem a *Nonparametric Kernel-based Semi-Log-Optimal Strategy* de Györfi et al. [38], que buscava um equilíbrio entre a otimização log e a praticidade computacional, e a *Nonparametric Kernel-based Markowitz-Type Strategy* de Ottucsák e Vajda [39], que incorporou princípios da teoria moderna de portfólio de Markowitz na estrutura de *Pattern Matching*.

Um avanço significativo veio com o trabalho de Li et al. [40], que introduziu a *Correlation-Driven Nonparametric Learning Strategy* (CORN). Esta abordagem inovou ao usar correlações entre janelas de mercado para identificar padrões similares, oferecendo uma nova perspectiva na seleção de amostras relevantes. Györfi e Vajda [41] também contribuíram com a *Nonparametric Kernel-based GV-Type Strategy*, que considerava custos de transação em sua formulação, abordando uma limitação prática importante das estratégias anteriores.

Uma nova classe de algoritmos tem início com o trabalho de Vovk [42], que inaugura uma outra abordagem para a seleção de portfólio *online*. Essa classe, conhecida como *Meta-Learning*, caracteriza-se pela combinação de múltiplos algoritmos ou especialistas para formar uma estratégia final mais adaptável. Um dos diferenciais do *Meta-Learning* reside em sua capacidade de agregar o desempenho de diferentes algoritmos base, ajustando-se as condições de mercado, tentando mitigar fraquezas inerentes a cada algoritmo individual e proporcionando uma generalização superior em diferentes conjuntos de dados.

O trabalho seminal de Vovk introduziu o Aggregating Algorithm (AA), que

estabeleceu as bases teóricas para muitos dos avanços subsequentes na área. O AA propôs uma abordagem de ponderação exponencial para a combinação das previsões de múltiplos especialistas, garantindo teórica e empiricamente um desempenho robusto.

Seguindo esta linha de pesquisa, Akcoglu et al. [43] propuseram o algoritmo *Fast Universalization* (FU). O FU amplia a ideia de universalização de Cover para uma classe ainda mais ampla de estratégias de investimento, oferecendo uma implementação que é tanto computacionalmente eficiente quanto capaz de lidar com um grande número de estratégias base.

Das e Banerjee [44] deram uma contribuição significativa ao introduzir dois algoritmos: *Online Gradient Updates* e *Online Newton Updates*. Estes métodos aplicam técnicas de otimização online para combinar estratégias base de forma dinâmica e adaptativa. A importância destes algoritmos está na sua habilidade de ajustar rapidamente os pesos atribuídos aos especialistas, em resposta às mudanças nas condições de mercado.

Hazan e Seshadhri [45] propuseram o algoritmo *Follow the Leading History* (FLH), especificamente desenvolvido para enfrentar o desafio dos ambientes não estacionários em seleção de portfólio. O FLH é projetado para adaptar-se a mudanças nas distribuições subjacentes ao longo do tempo, uma característica essencial para aplicações práticas em mercados financeiros dinâmicos.

Trabalhos mais atuais seguem de maneira geral uma tendência de abordar a seleção de portfólio *online* com o objetivo de maximizar a riqueza acumulada enquanto controlam custos de transação. Salas [46] desenvolveu uma nova família de algoritmos de aprendizado bayesianos para seleção de portfólio *online*, superando limitações das técnicas tradicionais ao permitir a modelagem da incerteza e previsões probabilísticas. Khedmati e Azin [47] propuseram um algoritmo de seleção de portfólio *online* baseado em abordagens de *clustering*, considerando custos de transação e maximizando o retorno ajustado ao risco. Em seguida, Guo et al. [48] introduziram o algoritmo *Adaptive Online Net Profit Maximization* (AOLNPM), integrando um método de média móvel *online* adaptativa com um modelo de maximização de lucro líquido, otimizando a estratégia de investimento ao longo do

tempo.

Outros trabalhos recentes, como o de Zhang et al. [49], propuseram uma abordagem sensível a custos utilizando aprendizado por reforço profundo, que considera a complexidade das séries temporais não estacionárias e a correlação entre ativos, enquanto introduzem um mecanismo de recompensa para maximizar o retorno acumulado. Dai et al. [50] propuseram uma estratégia baseada em aprendizado por *ensemble* denominada *Trend Peak Price Tracing* (TPPT), combinando previsão de preços de três estados com retropropagação de erro rápido para enfrentar as dificuldades em mercados financeiros voláteis. Dai et al. [51] continuaram nessa linha com a introdução do algoritmo *DSF-RI-PPT*, que integra decomposição empírica, análise de componentes principais e redes LSTM para maximizar a acumulação de riqueza.

Mais recentemente, Zhang et al. [52] propuseram uma nova abordagem para o mercado *long-short*, utilizando o algoritmo de *Mirror Gradient Descent*, que reformula o problema clássico de seleção de portfólio *online* para permitir operações *long/short*, enfrentando desafios como alta complexidade computacional e alta variabilidade de gradientes. He et al. [53] propuseram o algoritmo *Weak Aggregating Specialist Algorithm* (WASA), que desenvolve uma estratégia de agregação seletiva de conselhos de especialistas ativos, equilibrando o retorno esperado com os custos de transação. Por fim, He e Li [54] propuseram o algoritmo *Anticor-DTW*, que utiliza *Dynamic Time Warping* para melhorar a identificação e exploração da reversão à média dos ativos, superando o algoritmo Anticor original em termos de riqueza acumulada e robustez em mercados com altos custos de transação.

Apesar dos avanços significativos na área de seleção de portfólios *online*, conforme evidenciado pela evolução das estratégias desde os trabalhos seminais de Cover até as abordagens mais recentes de *meta-learning* e *pattern matching*, pouco enfoque foi dado na incorporação de métricas de risco robustas no processo de otimização.

A maioria dos algoritmos discutidos nesta revisão, incluindo as estratégias *Follow the Winner*, *Follow the Loser*, *Pattern Matching* e *Meta-Learning*, focam primariamente na maximização do retorno ou na minimização do arrependimento

Regret em relação aos modelos *benchmarks*. Embora essas abordagens tenham demonstrado resultados promissores em termos de performance, elas frequentemente negligenciam o risco de maneira explícita e/ou sua dinâmica durante o processo de construção e ajuste do portfólio.

Esta omissão é particularmente problemática no contexto dos mercados financeiros contemporâneos, caracterizados por alta volatilidade, eventos extremos e rápidas mudanças nas condições de mercado. A falta de um tratamento robusto do risco pode resultar em portfólios que, embora potencialmente lucrativos em média, estão expostos a níveis altos de volatilidade ou a perdas significativas em cenários adversos.

Além disso, a falta de uma consideração adequada do risco limita a aplicabilidade prática dessas estratégias em ambientes institucionais, onde requisitos regulatórios e perfis de risco específicos dos investidores devem ser rigorosamente observados. Esta limitação representa uma barreira significativa para a adoção generalizada de algoritmos de seleção de portfólios *online* em contextos de investimento do mundo real.

Portanto, desenvolver abordagens que integrem de forma eficaz e dinâmica métricas de risco avançadas no *framework* de seleção de portfólios *online* mostra-se de grande valia. Tais abordagens devem não apenas visar a maximização dos retornos, mas também garantir que os portfólios mantenham um perfil de risco controlado e alinhado com as expectativas dos investidores e requisitos regulatórios.

Esta lacuna na literatura atual serve como principal motivação para a presente pesquisa, que busca investigar e propor novas metodologias para incorporar métricas de risco robustas no processo de otimização de portfólios *online*, mantendo a convexidade do problema e a eficiência computacional característica dessas abordagens. Ao abordar esta questão, espera-se contribuir para o desenvolvimento de estratégias de investimento que sejam não apenas eficazes na geração de retornos, mas também confiáveis e resilientes frente às incertezas e volatilidades inerentes aos mercados financeiros modernos.

3 FUNDAMENTAÇÃO TEÓRICA

3.1 MODELO MÉDIA-VARIÂNCIA

A formalização matemática da estratégia de diversificação de portfólio foi introduzida por Harry Max Markowitz, delineando o fundamento do modelo Média-Variância (MV) [1]. Este modelo enfatiza que os investidores, ao compor um portfólio, buscam não apenas maximizar os retornos, mas equilibrar retorno esperado e risco, sendo este último modelado pela variância dos retornos esperados.

Usando a formulação mais usual do MV [40], considere que haja n ativos com índice i possíveis variantes no tempo t com retorno linear r_{it} e o vetor de retornos médios $\boldsymbol{\mu}$ definidos por:

$$r_{it} = \frac{p_{it} - p_{i(t-1)}}{p_{i(t-1)}} \quad (3.1)$$

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{bmatrix} \quad (3.2)$$

Além disso, seja $\sigma_{i,j}$ a covariância entre os quaisquer retornos dos ativos i e j , $\boldsymbol{\Sigma}$ a matriz de covariância dos retornos e σ^2 a variância do portfólio considerando uma alocação qualquer $\mathbf{b} = (b_1, b_2, \dots, b_n)^\top$ onde $\sum_{i=1}^n b_i = 1$ e $\mathbf{r}_t = (r_{1t}, r_{2t}, \dots, r_{nt})^\top$ são independentemente e identicamente distribuídos (IID) com distribuição normal multivariada $\mathcal{N}_{\boldsymbol{\mu}, \boldsymbol{\Sigma}}$. Assim, dado a linearidade implícita nas equações acima:

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{1,1} & \sigma_{1,2} & \cdots & \sigma_{1,n} \\ \sigma_{2,1} & \sigma_{2,2} & \cdots & \sigma_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n,1} & \sigma_{n,2} & \cdots & \sigma_{n,n} \end{bmatrix}, \quad (3.3)$$

$$\sigma^2 = \mathbf{b}^\top \boldsymbol{\Sigma} \mathbf{b}. \quad (3.4)$$

Logo, considerando que se saiba a média e a variância dos retornos,² o problema é resolvido ao escolher a alocação ótima que maximize retorno do portfólio dado níveis de risco (σ^{2*}) no período único; isto é:

$$\mathbf{b}^* = \arg \max_{\mathbf{b}} \mathbf{b}^\top \boldsymbol{\mu} \text{ sujeito a } \mathbf{b}^\top \boldsymbol{\Sigma} \mathbf{b} = \sigma^{2*}, \sum_{i=1}^n b_i = 1 \text{ e } b_i \geq 0 \forall i \quad (3.5)$$

A representação visual do problema descrito acima, expondo todas combinações ótimas de ativos, onde cada ponto representa um portfólio que oferece o máximo retorno esperado para um dado nível de risco (fronteira eficiente de Markowitz), é exemplificada abaixo no Gráfico 1 no espaço bidimensional.

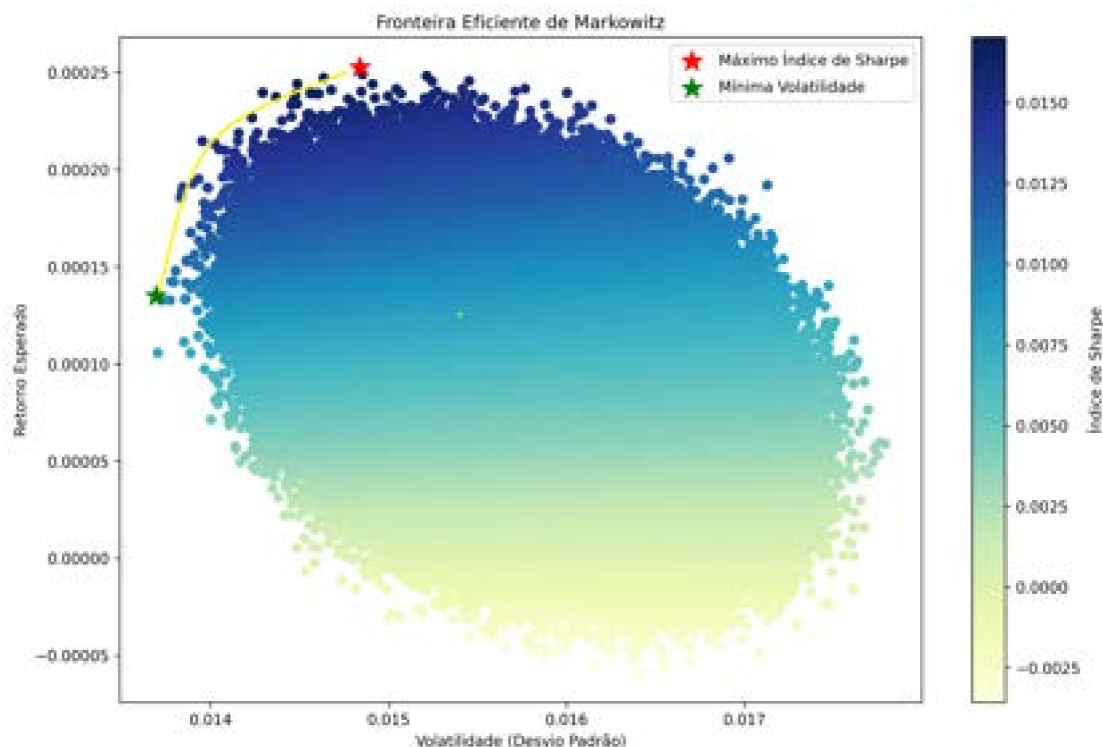
Para sua realização, foi efetuada uma simulação de Monte Carlo, gerando 10.000 portfólios hipotéticos com ponderações de ativos aleatórias. Cada portfólio foi avaliado em termos de seu retorno esperado, volatilidade (risco), e Índice de *Sharpe*, que mede o retorno ajustado ao risco. Esses portfólios são então plotados em um espaço bidimensional onde o eixo x representa a volatilidade e o eixo y, o retorno esperado. A coloração dos pontos reflete o Índice de *Sharpe*, proporcionando uma indicação visual do desempenho do portfólio em relação ao risco assumido.

A fronteira eficiente, exemplificada pela linha amarela no gráfico acima, é a que contorna a borda superior da região onde os portfólios são plotados, indicando a série de portfólios ótimos que os investidores podem escolher de acordo com sua tolerância ao risco. Esta análise permite identificar a composição dos portfólios que se situam no ponto de máxima eficiência, que é o ponto na fronteira onde se consegue o maior retorno para um dado nível de risco ou o menor risco para um dado nível de retorno.

Embora esta abordagem seja matematicamente elegante, ela pode não refletir adequadamente as complexidades e nuances das preferências dos investidores, que podem ser influenciadas por uma miríade de fatores comportamentais e contextuais. De fato, uma crítica fundamental ao modelo MV, conforme discutido por [55],

² Dado a suposição de retornos IID, seria possível estimar de maneira consistente tais parâmetros via métodos como a máxima verossimilhança usando uma série de preços históricos $\{p_{i,t}\}$ dos ativos por meio da aproximação $\log(p_{i,t}/p_{i,t-1}) = \log(1 + r_{i,t}) \approx r_{i,t}$

Gráfico 1 – Fronteira eficiente de Markowitz



Fonte: Elaborado pelo autor (2024).

reside na sua representação simplista da utilidade do investidor, sem usar outras características intrínsecas como o grau de aversão, limitando, assim, os objetivos de investimento à média e variância dos retornos.

Adicionalmente, a presunção de que os retornos dos ativos obedecem a uma distribuição normal tem sido objeto de críticas substanciais na literatura financeira. Estudos empíricos têm consistentemente mostrado que os retornos dos ativos podem, na realidade, apresentar caudas pesadas e, em certos casos, assimetria, o que destoia do pressuposto de normalidade [56, 57]. Essas características das distribuições de retorno indicam que a probabilidade de observar retornos extremos pode ser substancialmente maior do que a sugerida por uma distribuição normal. Desta forma, ao depender da suposição de normalidade, o modelo MV pode oferecer uma visão subestimada e potencialmente distorcida do risco envolvido nas estratégias de investimento. Além disso, a adoção da hipótese de distribuição de retornos

normais pode ser problemática, pois, na prática, os mercados frequentemente exibem movimentos que são inconsistentes com esta suposição, fenômeno esse que foi validado em múltiplas temporalidades e mercados por diferentes estudos [58].

A aplicabilidade do modelo MV também é questionada quando consideramos investidores com objetivos de longo prazo, como fundos de pensão e dotações. O modelo, em sua forma clássica conforme detalhado acima, é intrinsecamente um *framework* de único período, o que o torna potencialmente inadequado para tratar de questões pertinentes a investidores com horizontes temporais estendidos e obrigações futuras, como pagamentos de benefícios de pensão [59].

Além disso, a crítica à abordagem MV também se estende à sua utilidade prática na alocação de ativos, com alguns argumentando que a simulação de ativos e passivos é uma abordagem superior, proporcionando uma visão mais holística e integrada da gestão de portfólio, alinhando mais estreitamente os investimentos com as obrigações futuras [60].

No entanto, talvez a crítica mais penetrante ao modelo MV seja sua instabilidade e ambiguidade inerentes. Os otimizadores MV, ao funcionarem como um sistema de decisão de investimento caótico, muitas vezes produzem portfólios que são altamente sensíveis a pequenas alterações nas premissas de entrada [55]. Essa sensibilidade pode resultar em portfólios que, apesar de teoricamente ótimos, podem ter pouco valor de investimento confiável na prática [61].

Os desafios do modelo MV, tais como sua instabilidade e ambiguidade, como delineado anteriormente, propiciam um campo fértil para explorar alternativas que possam harmonizar teoria e prática no âmbito da gestão de portfólio. Inúmeros investidores e gestores financeiros têm migrado para métodos intuitivos ou adotado novas abordagens que, embora possam abdicar da matemática bela e rigorosa, se propõem a oferecer estratégias mais robustas e aplicáveis na realidade volátil dos mercados. Uma dessas alternativas que ganharam espaço na discussão acadêmica e prática de mercado é o Modelo de Precificação de Ativos de Capital - *Capital Asset Pricing Model* (CAPM).

A popularidade e adoção generalizada do CAPM, especialmente na avaliação de ativos e na definição de estratégias de investimento, sugerem que este modelo

atende a necessidades práticas particulares que talvez o modelo MV não seja capaz de suprir integralmente - sendo este um contraponto à elegância matemática, porém, em alguns contextos, à impraticabilidade do modelo MV [62]. De fato, o CAPM oferece uma medida de risco prontamente compreensível e aplicável, o beta, que tem atraído tanto acadêmicos quanto praticantes por sua facilidade de interpretação e implementação [63]. Vale ressaltar, porém, que a sugestão do CAPM de que o retorno esperado de um ativo ou portfólio é proporcional ao seu beta tem sido constantemente posta à prova, com vários estudos destacando anomalias e inconsistências quando aplicadas a dados históricos [64].

Com isso em mente, o CAPM, melhor discutido na próxima seção, apresenta uma ponte entre a teoria e prática, permitindo, assim, que gestores de portfólio alinhem expectativas de retorno com uma mensuração quantitativa do risco sistemático, mesmo que essa métrica por vezes seja vista como simplista ou insuficiente para capturar a totalidade da complexidade do mercado [62].

3.2 MODELO CAPM: BETA DO MERCADO

Com base no modelo clássico proposto em [1], os trabalhos de [5] e [65] deram início ao modelo de precificação de ativos chamado classicamente de *Capital Asset Pricing Model* (CAPM). Até meados da década de 1950, o consenso era que o desempenho deveria ser medido pelo retorno ao longo de um período sem ajuste de risco. O CAPM introduziu duas novas premissas ao modelo clássico de Markowitz: expectativas homogêneas e a taxa livre de risco. A premissa de expectativas homogêneas afirma que todos os investidores têm as mesmas perspectivas de retorno esperado, desvio padrão e covariância dos ativos (premissa do mercado eficiente). A taxa livre de risco pressupõe que existe um investimento no mercado cuja remuneração é assegurada exatamente como esperado, e fatores econômicos e cíclicos não podem afetar sua liquidez.

O modelo CAPM é descrito pela seguinte equação:

$$r_i - r_f = \beta_i(r_m - r_f) + \varepsilon_i, \quad (3.6)$$

onde:

- r_i é o retorno de um ativo i .
- r_f é o retorno de um ativo livre de risco.
- r_m é o retorno médio do mercado.
- $\sigma_{i,m}$ é a covariância entre retorno do ativo i e o retorno médio do mercado.
- σ_m é a volatilidade do mercado.
- β_i é definido como $\beta_i = \frac{\sigma_{i,m}}{\sigma_m^2}$, representando o risco sistêmico do ativo i .
- ε_i é um choque idiossincrático com $\mathbb{E}[\varepsilon_i] = 0$ e $\mathbb{E}[\varepsilon_i^2] = \sigma_i^2$.

A equação (3.6) assume que os choques idiossincráticos, $\varepsilon_{i,t}$, têm média zero, variância constante no tempo σ_i^2 e IID. Em [66] é vista uma revisão detalhada de tal modelo.

A implementação *clássica* do modelo CAPM é geralmente estimada por meio do *Ordinary Least Squares* (OLS), partindo do pressuposto de que o beta, ou risco sistemático, de um ativo ou carteira permanece constante ao longo do tempo. Contudo, é pertinente considerar uma hipótese alternativa que contempla a variação do beta de uma empresa ao longo do tempo. Essa variação pode ser influenciada tanto por fatores microeconômicos — como mudanças operacionais internas ou no ambiente de negócios específico da empresa — quanto por fatores macroeconômicos — incluindo a taxa de inflação, o estado geral dos negócios e as expectativas quanto a eventos futuros significativos. Uma análise aprofundada sobre estes tópicos já foram apontadas em trabalhos como [67], [68] e [69]. Evidências que corroboram a hipótese de variação do risco sistemático com o tempo também já foram amplamente estudadas em trabalhos como [70, 71, 72].

Em revisões abrangentes [73, 74, 75, 76] mostra-se que no CAPM, o beta tende a ser volátil ao longo do tempo. Conforme discutido em [77] e [78], as cargas de risco no retorno excessivo do mercado, assim como os fatores HML (*High Minus Low*, que captura a diferença de retornos entre ações de valor e ações de crescimento) e SMB (*Small Minus Big*, que mede o diferencial de retornos entre empresas de pequena e grande capitalização), apresentam forte variação temporal.

Os trabalhos [79, 80] afirmam que a principal razão para as limitações do CAPM incondicional é que o modelo não reproduz o processo de aprendizado dos investidores. Argumentam que a inobservabilidade do beta que varia ao longo do tempo induzirá o processo de aprendizado dos investidores. No entanto, a regressão de séries temporais via OLS não consegue reproduzir com sucesso o processo de aprendizado dos investidores, o que leva à diferença entre a verdadeira expectativa dos investidores sobre beta e o beta estimado.

A estimação do beta condicional no CAPM envolve técnicas que reconhecem e capturam a variabilidade ao longo do tempo, em resposta a mudanças nas condições do mercado e variáveis econômicas. Muitos métodos foram testados, como: regressão recursiva e janelas deslizantes como em [81]; uso de métodos de regressão baseado em instrumentos levando em conta aspectos do mercado como o ciclo real de negócios (algumas *proxies* comumente usadas na literatura são *default spread*, *term spread* e *dividend yield* [82, 83]; métodos bayesianos supondo que os betas se comportam como um processo estocástico de reversão na média [84]; dentre outros. Já os modelos GARCH (*Generalized Autoregressive Conditional Heteroskedasticity*) são frequentemente utilizados para modelar e prever a volatilidade dos retornos de um ativo, permitindo que o beta seja ajustado dinamicamente com base na volatilidade estimada [85, 86, 87, 88]. Variações desses modelos, como o EGARCH (*Exponential Generalized Autoregressive Conditional Heteroskedasticity*) ou TGARCH (*Threshold Generalized Autoregressive Conditional Heteroskedasticity*), podem capturar assimetrias e efeitos de alavancagem na volatilidade, que por sua vez afetam o beta condicional [89, 90].

Uma abordagem alternativa frequentemente adotada é a aplicação do Filtro de Kalman, um estimador recursivo que proporciona a atualização contínua do parâmetro β mediante a incorporação de novos dados. Essa técnica é especialmente vantajosa em contextos onde o sistema modelado está sujeito à influência de estados latentes e/ou quando os parâmetros do modelo são variáveis ao longo do tempo. Uma sucessão de pesquisas comparativas [81, 91, 92, 93] destacam a superioridade da estimação via Filtro de Kalman utilizando dados para diferentes economias, incluindo a brasileira [94].

Na próxima seção, apresenta-se detalhes da estimação do CAPM condicional

usando o Filtro de Kalman. Essa discussão fornecerá um entendimento abrangente de como essa técnica se mostra eficiente quando adaptada para funcionar na etapa de projeção de algoritmos de seleção de portfólios *online*.

3.2.1 Filtro de Kalman

Para a estimação de modelos estruturais, um método comum é a utilização de Filtros de Kalman. Sob a suposição de normalidade dos dados observados, é viável desenvolver uma estrutura para a dinâmica dos coeficientes beta dos ativos e realizar a extração usando algoritmos de otimização [95].

Desenvolvido no contexto da teoria de controle dos anos 1960, o filtro de Kalman estima parâmetros que variam no tempo em sistemas dinâmicos, apresentando benefícios para modelos como o CAPM, cujo coeficiente beta é suscetível a variações temporais. O cerne dessa metodologia é definida no modelo de espaço de estados, que se compõe de uma equação de observação, representando as variáveis mensuráveis, e uma equação de transição, descrevendo a evolução do estado sistêmico. Esta técnica se destaca na integração de variáveis não observáveis e na aplicação de estruturas temporais ao beta do CAPM, atualizando-se recursivamente e gerando séries de interceptos e coeficientes beta condicionais do modelo linear.

A flexibilidade na definição da equação de transição é um ponto crucial do filtro, permitindo formatos diversos como um processo auto-regressivo ou passeio aleatório, sendo o último frequentemente mais preciso na caracterização de beta em evolução. Além disso, sua capacidade de refinar estimativas com novas informações o faz especialmente valioso para o mercado financeiro, onde as condições são voláteis.

Incorporar essa ferramenta na otimização de portfólios possibilita ajustes refinados das características do portfólio, fundamentados em uma medida de risco sistemático atual e precisa. A adaptabilidade e exatidão na estimação de parâmetros que mudam com o tempo podem enriquecer significativamente a gestão de portfólio.

Para estimar o beta de cada ativo (β_i), emprega-se um modelo de espaço de estados no qual o β_i é tratado como uma variável latente ou um fator oculto, seguindo um processo de passeio aleatório com perturbações gaussianas, resultando em um modelo de nível estocástico. O problema é formulado da seguinte forma: o

retorno de um ativo específico i no período t é dado por

$$r_{i,t} = \beta_{i,t}r_{m,t} + r_{f,t} + \varepsilon_{i,t}, \quad (3.7)$$

onde: $r_{m,t}$ é o retorno do mercado no período t , $r_{f,t}$ é a taxa livre de risco no período t e $\varepsilon_{i,t}$ é um choque idiossincrático no período t no ativo i com média zero e variância σ_i^2 que difere entre os ativos. Além disso, $\beta_{i,t}$ é definido como um passeio aleatório, ou seja,

$$\beta_{i,t+1} = \beta_{i,t} + \varphi_{i,t}, \quad (3.8)$$

onde $\varphi_{i,t}$ é uma variável aleatória gaussiana com média zero e variância $\sigma_{\beta,i}^2$. Os parâmetros do modelo σ_i^2 , $\sigma_{\beta,i}^2$ e a condição inicial $\beta_{i,0}$ podem ser estimados por máxima verossimilhança, e o vetor de estado $\beta_{i,t}$ de cada ativo i pode ser extraído usando as equações recursivas do filtro de Kalman. Para o modelo proposto, as equações de atualização do filtro de Kalman serão dadas por:

$$\varepsilon_{i,t} = r_{i,t} - \hat{r}_{i,t}, \quad (3.9)$$

$$Q_t = P_t + \sigma_i^2, \quad (3.10)$$

$$\beta_{i,t|t-1} = \beta_{i,t} + K_t \varepsilon_{i,t}, \quad (3.11)$$

$$P_{t|t} = P_t - K_t P_t, \quad (3.12)$$

onde $K_t = P_t Q_t^{-1}$ é o ganho de Kalman, P_t é a variância de $\beta_{i,t}$ condicional ao conjunto $\{r_{i,1}, r_{i,2}, \dots, r_{i,t-1}\}$ e Q_t é a variância de $\varepsilon_{i,t}$. As equações de previsão são:

$$\beta_{i,t+1} = \beta_{i,t|t-1}, \quad (3.13)$$

$$P_{t+1} = P_{t|t-1} + \sigma_{\beta,i}^2, \quad (3.14)$$

$$\hat{r}_{i,t+1} = \beta_{i,t} + r_{f,t}, \quad (3.15)$$

O Filtro de Kalman, portanto, emerge como uma ferramenta importante na análise de séries temporais financeiras, permitindo a estimação precisa de parâmetros dinâmicos e a análise de risco em tempo real. De fato, sua implementação que possibilita a utilização de fatores tidos como os importantes na literatura (como o HML e SMB), variantes no tempo, além do possível uso de métodos baseados em LASSO (*Least Absolute Shrinkage and Selection Operator*) para mitigar o possível problema de má identificação ao penalizar fatores mais fracos como visto em [96], demonstram o poder de tal metodologia.

Na seção 4.2, mais detalhes sobre a utilização dessa técnica na etapa de projeção de algoritmos para seleção de portfólios *online* são destacados.

3.3 RISCO

O termo risco é frequentemente utilizado na linguagem cotidiana e carrega múltiplas conotações. No contexto do mercado de ações, risco pode ser definido como a probabilidade de uma perda, a chance de não receber o retorno esperado ou a possibilidade de que um investidor não alcance a taxa de retorno desejada em um investimento.

O mercado de ações é dinâmico e sujeito a inúmeras variáveis, tornando os retornos dos investimentos em ações voláteis. A variação nos retornos é muitas vezes imprevisível e, para muitos profissionais do mercado financeiro, o risco é definido como a possibilidade de os retornos reais diferirem dos retornos esperados [97].

Assim como existem várias maneiras de definir risco, há várias maneiras de categorizar os elementos do risco. Os profissionais do mercado frequentemente dividem o risco em riscos sistemáticos e riscos específicos. Riscos sistemáticos, como o beta do mercado discutido na última seção, afetam o mercado de ações como um todo, enquanto riscos específicos estão associados a empresas ou setores individuais.

As instituições financeiras, notadamente, também estão expostas a uma diversidade de riscos, a saber: de mercado, de crédito e operacionais. Entre eles, o risco de mercado, isto é, o potencial de uma instituição sofrer perdas devido

a variações indesejáveis nos preços de mercado, tem dominado as preocupações. Nesse cenário, o *Value at Risk* (VaR), tem ganhado destaque como uma métrica imprescindível na gestão de riscos de mercado.

3.3.1 *Value at Risk* (VaR)

O *Value at Risk* (VaR) emergiu como uma resposta à crescente necessidade de quantificar e gerenciar riscos financeiros de forma mais eficaz. Suas raízes remontam aos anos 1970 e 1980, quando instituições financeiras começaram a desenvolver modelos internos para avaliar sua exposição ao risco. A evolução do VaR foi impulsionada por uma série de fatores, incluindo a globalização dos mercados financeiros, o aumento da volatilidade e a ocorrência de grandes perdas financeiras que destacaram a importância do gerenciamento de riscos. Contribuições significativas vieram de várias fontes, incluindo o trabalho pioneiro do *Bankers Trust* com o RAROC (*Risk-Adjusted Return on Capital*) e os esforços de acadêmicos em estatística e finanças que desenvolveram as bases teóricas para a modelagem de riscos.

O conceito de VaR foi introduzido na década de 1990 por profissionais do J.P. Morgan, um dos maiores bancos de investimento do mundo, integrando seu projeto *RiskMetrics*. Esse instrumento rapidamente se estabeleceu como o padrão na mensuração do risco de mercado no âmbito do setor financeiro. Não se restringiu apenas aos bancos, mas foi igualmente adotado por fundos de *hedge*, empresas de seguros e órgãos reguladores financeiros.

A principal função do VaR é proporcionar uma quantificação do risco de mercado, ou seja, ele estima a perda máxima esperada para um portfólio em particular sob condições normais de mercado, durante um período de tempo definido e em um nível de confiança específico [8]. Para elucidar, se o VaR de um portfólio é de R\$1 milhão em um dia com um nível de confiança de 95%, isso significa que existe 5% de probabilidade de que as perdas ultrapassem R\$1 milhão em um dia [98].

O VaR inovou ao apresentar uma visão quantitativa do risco de um portfólio inteiro, expressando-o em termos monetários e oferecendo uma perspectiva sobre o

potencial de perdas em um dado período de tempo com um determinado nível de confiança. Este foi um avanço importante porque permitiu que os gerentes de risco, pela primeira vez, quantificassem o risco de forma padronizada e compreensível [99].

Através do *RiskMetrics* e do crescente reconhecimento de sua utilidade, o VaR rapidamente ganhou popularidade e foi adotado por muitas instituições financeiras. O sucesso desta iniciativa levou a Basileia, um comitê de supervisão bancária, a incluir o VaR como um componente chave em seus regulamentos prudenciais [100].

Formalmente, para um dado nível de confiança γ , se $F(\cdot)$ é a função de distribuição cumulativa da perda L , então o VaR é definido como:

$$VaR_\gamma(L) = \inf\{l \in \mathbb{R} : F(l) > 1 - \gamma\} \text{ onde } 0 \leq \gamma \leq 1 \quad (3.16)$$

Isso significa que a probabilidade de uma perda exceder o VaR é $1 - \gamma$. Já as suposições mais comuns vistas na literatura são enumeradas abaixo [99]:

- (i) **Estabilidade do portfólio:** O VaR presume que a composição do portfólio permanece inalterada durante o período de tempo considerado.
- (ii) **Liquidez do mercado:** O VaR pressupõe que todos os ativos no portfólio podem ser comprados ou vendidos sem impactar seus preços.
- (iii) **Normalidade dos retornos:** Os retornos dos ativos são normalmente distribuídos, isto é, tem distribuição $\mathcal{N}_{\mu, \Sigma}$ com distribuição cumulativa Φ .

Apesar das duas primeiras suposições poderem ser quebradas na prática, por exemplo, durante crises pode haver mudanças no perfil de risco da instituição e no preço dos ativos dentro do período avaliado, respectivamente, a normalidade dos retornos parece ser o maior alvo de críticas na literatura. De fato, apesar de na prática facilitar os cálculos, como pode ser visto abaixo ao considerar os retornos sendo normais IID; isto é, $r_t = \mu + \varepsilon_t$ com ε_t tendo distribuição $\mathcal{N}_{0, \sigma^2}$ e distribuição cumulativa Φ onde $q_\gamma(L) = F_L^{-1}(\gamma)$ é a função quantílica, além de Z_γ definido como

o percentil correspondente ao nível de confiança γ em uma distribuição normal padrão,

$$\begin{aligned} VaR_\gamma(L) &= \inf\{l \in \mathbb{R} : \Phi(l) > 1 - \gamma\} = \Phi_L^{-1}(1 - \gamma) = q_{1-\gamma}(L) = -q_\gamma(-L) \\ &= -q_c(-\mu - \varepsilon_t) = \mu + q_\gamma(\varepsilon_t) = \mu + \sigma Z_\gamma \end{aligned} \quad (3.17)$$

onde, muitos retornos de ativos apresentam *skewness* (assimetria) e *kurtosis* (curtose) excessiva, o que significa que eles têm caudas mais pesadas que a distribuição normal. Nesse contexto, outras distribuições podem ser pensadas, como a t de Student e a t de Student assimétrica. Por exemplo, considerando que ε_t seja IID com uma distribuição t de Student com parâmetro de localidade 0, parâmetro de dispersão σ^2 , $\nu > 0$ os graus de liberdade, $\Gamma(\cdot)$ sendo a função Gamma, a sua função de distribuição e o cálculo do VaR seriam, após aplicar procedimento similar ao visto acima, respectivamente:

$$\begin{aligned} f(\varepsilon, \nu, \sigma^2) &= \frac{\Gamma((\nu + 1)/2)}{\sqrt{\nu\pi\sigma^2} \Gamma(\nu/2)} (1 + \varepsilon^2/\nu\sigma^2)^{-(\nu+1)/2} e \\ VaR_\gamma(L) &= \mu + \sigma t_{\gamma,\nu} \text{ onde } t_{\gamma,\nu} \text{ é o percentil da distribuição aos } \gamma \% , \end{aligned} \quad (3.18)$$

onde fica claro que o parâmetro ν impacta diretamente a variância do processo e, conseqüentemente, o VaR.

Dentre os vários métodos para o cálculo do VaR [101], destacamos:

1. **Método da Variância-Covariância (Paramétrico):** pressupõe que os retornos são IID e seguem uma distribuição normal. Portanto, seguindo o procedimento visto na equação 3.17, o VaR é calculado como:

$$VaR_\gamma = \mu + \sigma Z_\gamma \quad (3.19)$$

Assim, por construção, tende a subestimar a perda L caso a distribuição seja anormal (problema mitigado usando outras distribuições); além de depender da modelagem dos retorno, em especial no que tange a estrutura da variância

condicional - problema podendo ser reduzido usando modelagem tipo GARCH e similares.

2. **Simulação Histórica:** usa a distribuição empírica dos retornos passados para calcular o VaR. Assumindo que temos n observações de retornos, o VaR é o $(1 - \gamma)$ -ésimo percentil dos retornos

$$VaR_\gamma = F^{-1}(1 - \gamma), \quad (3.20)$$

onde F^{-1} é a função quantil da distribuição empírica. Nesse contexto, a maior crítica é mitigada por construção, pois não se pressupõe uma distribuição. Porém, por natureza, é um método totalmente dependente dos dados. Desta forma, caso, por exemplo, a informação *prior* seja baseada num período de crise, haverá uma superestimação do risco caso o momento atual for de aquecimento - uso de janelas deslizantes pode ajudar a mitigar tal problema.

3. **Simulação de Monte Carlo:** modela os movimentos futuros dos preços dos ativos usando simulação computacional. O VaR é então calculado a partir da distribuição resultante dos retornos simulados. Assim como o método discutido acima, mas simulando diversos cenários (inclusive utilizando fatores que afetam o risco de portfólios de maneira não linear e no tempo), estima os quantis de interesse sem pré-definir uma distribuição. Vale ressaltar, porém, que dependendo do número de fatores levados em conta, o custo computacional pode ser alto, o que pode ser reduzido usando método de componentes principais - ver [102].

Na próxima subseção será discutido uma alternativa para o VaR, o *Conditional Value at Risk* (CVaR), tido na literatura como mais robusta.

3.3.2 *Conditional Value at Risk* (CVaR)

O *Conditional Value-at-Risk* (CVaR), frequentemente equiparado ao *Expected Shortfall* (ES), embora com sutis diferenças técnicas em certos contextos, surgiu como uma resposta às limitações identificadas no Value-at-Risk (VaR) no final da década de 1990 e início dos anos 2000. Enquanto o VaR havia se estabelecido

como uma medida padrão de risco no setor financeiro, pesquisadores e profissionais começaram a reconhecer suas deficiências, particularmente sua incapacidade de capturar adequadamente o risco de eventos extremos e sua falta de subaditividade, uma propriedade crucial para medidas de risco coerentes [103].

A necessidade de uma medida de risco mais robusta e informativa tornou-se ainda mais evidente após uma série de crises financeiras, incluindo o colapso do *Long-Term Capital Management* em 1998. Estes eventos destacaram a importância de compreender e quantificar melhor os riscos associados as caudas, que o VaR não abordava adequadamente [104].

O desenvolvimento do CVaR foi impulsionado por trabalhos teóricos sobre medidas de risco coerentes, notadamente o artigo seminal de [103]. Este trabalho estabeleceu as propriedades desejáveis para medidas de risco, fornecendo uma base teórica sólida para o desenvolvimento de alternativas ao VaR. Os trabalhos de Rockafellar e Uryasev (2000, 2002) desempenharam um papel fundamental na formalização e popularização do CVaR [9, 10]. Eles não apenas definiram rigorosamente o CVaR, mas também demonstraram suas propriedades matemáticas superiores e desenvolveram métodos eficientes para sua otimização, tornando-o uma ferramenta prática para gestão de riscos e otimização de portfólio.

O CVaR rapidamente ganhou aceitação tanto na academia quanto na indústria devido à sua capacidade de fornecer uma visão mais completa do risco de cauda, sua coerência matemática e sua tratabilidade computacional. Ao contrário do VaR, que apenas indica um limiar de perda, o CVaR quantifica a perda esperada além desse limiar, oferecendo uma perspectiva mais abrangente do risco potencial [104].

A adoção do CVaR foi ainda mais impulsionada por mudanças regulatórias. O Comitê de Basileia de Supervisão Bancária, reconhecendo as limitações do VaR, recomendou a transição para o Expected Shortfall (ES)³ em suas diretrizes de

³ O Conditional Value-at-Risk (CVaR) e o Expected Shortfall (ES) são frequentemente tratados como sinônimos na literatura financeira, e para distribuições contínuas, são de fato matematicamente equivalentes. Ambos medem a perda esperada nos piores casos de um portfólio. No entanto, existem sutis diferenças técnicas, especialmente para distribuições discretas ou com descontinuidades. O CVaR, introduzido por Rockafellar

Basileia III, solidificando o papel do CVaR como uma medida de risco padrão na indústria financeira [105].

Formalmente, o CVaR é definido como a expectativa condicional das perdas que excedem o VaR para um dado nível de confiança. Para um retorno excessivo $r_t - r_f$ e um nível de confiança γ , o CVaR é expresso como:

$$CVaR_\gamma = E [r_t - r_f | r_t - r_f \leq -VaR_\gamma(r_t - r_f)] \quad (3.21)$$

Esta definição implica que o CVaR fornece uma estimativa da perda média nos piores $1 - \gamma$ casos, oferecendo uma visão mais abrangente do risco de cauda. Uma propriedade crucial do CVaR é sua coerência como medida de risco, satisfazendo os axiomas propostos por Artzner et al. [103]. Em particular, o CVaR possui a propriedade de subaditividade, melhor discutida na seção 4.4 e ausente no VaR, o que o torna mais adequado para a modelagem de riscos em portfólios diversificados pois incorpora a ideia de que a diversificação está de fato ligada a uma redução de risco. Para distribuições contínuas, o CVaR pode ser calculado através da seguinte fórmula:

$$CVaR_\gamma = -\frac{1}{1-\gamma} \int_\gamma^1 VaR_x(r_t - r_f) dx \quad (3.22)$$

Esta representação integral demonstra como o CVaR incorpora informações de toda a cauda da distribuição além do VaR, não se limitando a um único ponto de corte. No caso de retornos excessivos normalmente distribuídos, com $r_t - r_f = \mu + \varepsilon_t$ e $\varepsilon_t \sim \mathcal{N}(0, \sigma^2)$, o CVaR tem uma forma analítica simples:

$$CVaR_\gamma = -\mu + \sigma \frac{\phi(Z_\gamma)}{1-\gamma} \quad (3.23)$$

onde $\phi(\cdot)$ é a função de densidade da distribuição normal padrão e Z_γ é o γ -ésimo percentil da distribuição normal padrão. Para distribuições com caudas mais pesadas, como a t de Student, o cálculo do CVaR se adapta para refletir a

e Uryasev [9], foi inicialmente formulado no contexto de problemas de otimização e possui propriedades matemáticas particularmente úteis nesse âmbito, como a convexidade. O ES, por outro lado, foi desenvolvido com um foco mais estatístico. Na prática regulatória, o termo ES é preferido, como evidenciado nas diretrizes do Comitê de Basileia. Para os propósitos desta tese, que envolve otimização algorítmica, o termo CVaR é utilizado devido às suas propriedades matemáticas específicas.

maior probabilidade de eventos extremos. Isso torna o CVaR uma medida mais conservadora e realista em muitos cenários práticos de gestão de riscos.

Com a discussão até aqui em mente, um dos objetivos do atual trabalho, utilizar um método de aprendizado de máquina no contexto de otimização *online* com e sem uso de janelas deslizantes em conjunto com métodos paramétricos (VaR e CVaR) baseados em fatores tidos como determinantes do risco; diminuindo, assim, os custos computacionais, mostra-se de grande valia.

3.3.3 Métricas de Análise de Risco

Na literatura há diversas métricas utilizadas para a análise de risco dentre diferentes métodos e modelos. Dessa forma, abaixo são delineadas as utilizadas ao longo dos resultados, sendo comumente vistas na literatura de finanças. Revisões sobre assunto podem ser vistas em [106, 107, 108].

Para tal, dado a sua generalidade, preferiu-se discutir as métricas num contexto onde o objetivo é calcular tais métrica em uma janela *out-of-sample* $t \in \{T^{in} + 1, \dots, T\}$ onde T^{in} o tamanho da janela *in-sample*.

1. **Volatilidade anualizada:** Considerando o desvio padrão ligada ao portfólio entre o momento $T^{in} + 1$ e $T > 0$ medido em dias (ou meses, trimestres, ...) como $\hat{\sigma}$ e T^{anu} sendo $T - T^{in}$ em termos anuais:

$$\text{Volatilidade anualizada} = \hat{\sigma} \sqrt{T^{anu}} \quad (3.24)$$

Como por construção são sempre positivos, o menor valor possível deste é o desejado.

2. **Semi-Desvio:** Considerando r_t^{out} sendo o retorno *out-of-sample* em t ligada ao portfólio, além de \bar{r} sendo o seu retorno médio no período:

$$\text{Semi-Desvio} = \sqrt{E [\min(r_t^{out} - \bar{r})^2 | T^{in} + 1 \leq t \leq T]} \quad (3.25)$$

a interpretação é similar a anterior, com valores menores sendo também melhores.

3. **Máximo Drawdown (MDD):** Considerando a perda *out-of-sample* de um pico observado, *drawdown* (DD), como sendo $DD_t = \frac{W_t - \max_{T^{in}+1 \leq \tau \leq t} W_\tau}{\max_{T^{in}+1 \leq \tau \leq t} W_\tau}$ com $W_t = W_{t-1} (1 + r_t^{out})$ onde r_t^{out} é o retorno *out-of-sample* em t e W_t sendo a riqueza ligada ao portfólio em t onde $W_0 = 1$. Assim, o Máximo Drawdown (MDD) é o máximo *drawdown* possível:

$$MDD = \min_{T^{in}+1 \leq t \leq T} DD_t. \quad (3.26)$$

Como por construção são sempre não positivos, o maior valor possível deste é o mais desejado.

4. **Drawdown Médio (DDM):** Considerando as relações introduzidas acima, o DDM é definido como a expectativa de perda *out-of-sample* de um pico observado:

$$DDM = E [DD_t | T^{in} + 1 \leq t \leq T], \quad (3.27)$$

seguindo a mesma lógica acima dado a sua negatividade.

5. **Razão de Calmar:** Considerando novamente as relações destacadas acima, com r_f sendo qualquer valor de retorno definido como objetivo no portfólio (retorno sem risco, por exemplo), é o retorno por unidade de risco quando este é o MDD; isto é:

$$\text{Razão de Calmar} = \frac{E [r_t^{out} - r_f | T^{in} + 1 \leq t \leq T]}{MDD} \quad (3.28)$$

O mesmo pode ser visto como uma versão da razão de *Sharpe* onde se usa a métrica MDD ao invés do desvio-padrão *out-of-sample*, com valores maiores sendo desejados.

6. **Razão de Sortino:** Considerando novamente as relações destacadas acima, é o retorno por unidade de risco quando este é o desvio padrão do *downside*; isto é:

$$\text{Razão de Sortino} = \frac{E [r_t^{out} - r_f | T^{in} + 1 \leq t \leq T]}{\sqrt{E [\min(r_t^{out} - r_f, 0)^2 | T^{in} + 1 \leq t \leq T]}}. \quad (3.29)$$

Note que a interpretação é similar a vista acima.

3.4 OTIMIZAÇÃO CONVEXA *ONLINE*

A otimização convexa *online* representa um quadro de otimização dinâmica, onde decisões sequenciais são tomadas sob a incerteza dos dados futuros. Essa abordagem é cada vez mais aplicada em diversos campos, desde o aprendizado de máquina até o setor financeiro, devido à sua flexibilidade e capacidade de adaptação a novas informações.

No aprendizado de máquina, utiliza-se a otimização convexa *online* para atualizar modelos à medida que novos dados são recebidos. Isso é essencial em áreas como detecção de fraudes e sistemas de recomendação, onde a habilidade de ajustar rapidamente os modelos aos dados emergentes é crítica. No âmbito financeiro, essa metodologia é aplicada na gestão de portfólios e gerenciamento de riscos, adaptando-se continuamente às flutuações do mercado para otimizar decisões de investimento. Formalmente, esse problema pode ser descrito como uma sequência infinita de funções de custo convexo f_1, f_2, \dots e um conjunto de soluções viáveis $S_n \subseteq \mathbb{R}^n$. A cada etapa t , o algoritmo utiliza uma estratégia de alocação de portfólio $\mathbf{b}_t \in S_n$ e, após escolher essa estratégia, avalia o custo em f_t . Como nem todas as informações estão disponíveis antecipadamente, o algoritmo usa o desempenho de $t = 1, \dots, t - 1$ para definir a posição futura. É importante observar que as funções de custo podem ser bastante distintas a cada passo, desde que sejam convexas. Dado que a sequência de funções de custo f_1, f_2, \dots não é conhecida antecipadamente, não é possível escolher \mathbf{b}_t de modo que f_t seja minimizado. Dessa forma, o objetivo não é atingir a solução ideal para cada função de custo específica, já que não têm acesso prévio à sequência completa das funções. Por isso, esses algoritmos focam em minimizar o custo acumulado ao longo do tempo, em vez de buscar a perfeição em cada passo individual. A eficácia desses algoritmos é medida através da função de arrependimento, ou *Regret*. Essa métrica compara o desempenho acumulado do algoritmo *online* com o de um algoritmo hipotético *offline*, que teria acesso a todas as funções de custo desde o início. Assim, o *Regret* quantifica o quanto o algoritmo *online* se desvia do desempenho ideal, ajudando a avaliar sua eficiência em um cenário de informações incompletas.

Na subseção 3.4.1, discute-se a estratégia Constant Rebalanced Portfolio

(CRP). Posteriormente, na subseção 3.4.2, a métrica de *Regret* é examinada para o caso específico de seleção de portfólios. Por fim, na subseção 3.4.3, concluí-se com a exposição do algoritmo *Online Gradient Descent* (OGD) também no contexto da decisão de alocação de ativos.

3.4.1 *Constant Rebalanced Portfolio*

Encontrar o portfólio de máxima riqueza absoluta dentro do mercado é uma tarefa bastante ambiciosa [4], assim, busca-se minimizar a distância entre o portfólio escolhido pelo algoritmo *online* e o *Constant Rebalanced Portfolio* (CRP) tal como proposto por [18].

Devido as movimentações diárias do mercado financeiro, o portfólio apresentará uma alocação diferente da alocação de ativos original. O CRP é uma estratégia de investimentos na qual o portfólio é rebalanceado de tal forma que a proporção destinada a cada um dos ativos na carteira original é preservada.

Dochow [109] constrói a formalização do CRP de maneira bastante didática. Seja $\{A_1, A_2, \dots, A_n\}$ o conjunto de todos os ativos presentes em um portfólio,⁴ Y_i^{inic} a quantidade inicial de ações de cada ativo i , Y_{it} a quantidade de ações do ativo i no tempo t , p_{it} o preço do ativo i no tempo t , a riqueza W , no tempo t , pode ser obtida através da seguinte equação:

$$W_t = \sum_{i=1}^n p_{it} Y_{it}. \quad (3.30)$$

A quantidade de ações em cada ativo i no tempo t pode ser calculada por:

$$Y_{it} = Y_{i(t-1)} + y_{it}, \quad (3.31)$$

onde y_{it} é a quantidade de ações i negociadas em t . Se $y_{it} > 0$ então ações de i foram compradas, se $y_{it} < 0$ ações de i foram vendidas e se $y_{it} = 0$ não houveram transações.

Considerando um horizonte de tempo $\tau = 1, \dots, T$, o número de ações do ativo i presente no portfólio será dado por:

⁴ Considera-se um mercado com n ativos homogêneos e perfeitamente divisíveis.

$$Y_{iT} = Y_i^{inic} + \sum_{t=0}^T y_{it}. \quad (3.32)$$

Então, a riqueza no final do horizonte de tempo T poderá ser calculada como:

$$W_T = \sum_{i=1}^n p_{iT} \left(Y_i^{inic} + \sum_{t=0}^T y_{it} \right). \quad (3.33)$$

O percentual b_{it} do portfólio mantido em cada ativo i no tempo t é definido como:

$$b_{it} = \frac{Y_{i(t-1)} p_{i(t-1)}}{W_{(t-1)}}, \quad (3.34)$$

onde $\sum_{i=1}^n b_{it} = 1$. Combinando as equações 3.32 e 3.34, a proporção destinada a cada ativo $b_{i\tau}$ no instante τ , para o horizonte de tempo $\tau = 1, \dots, T$, pode a ser escrita como:

$$b_{i\tau} = \frac{\left(Y_i^{inic} + \sum_{t=0}^{\tau-1} y_{it} \right) p_{i(\tau-1)}}{W_{(\tau-1)}}. \quad (3.35)$$

Em um portfólio CRP, a cada período de tempo t são realizadas operações de compra e venda de ativos de forma que o percentual de riqueza $b_{it} = b_{i(t+1)} = b_i$ para todo instante de tempo t e todo ativo i pertencente ao portfólio, isto é, o percentual da riqueza W_τ destinado a cada ativo permanece constante ao longo do tempo. Então,

$$\frac{\left(Y_i^{inic} + \sum_{t=1}^{\tau} y_{it} \right) p_{i\tau}}{W_\tau} = \frac{\left(Y_i^{inic} + \sum_{t=0}^{\tau-1} y_{it} \right) p_{i(\tau-1)}}{W_{(\tau-1)}}. \quad (3.36)$$

De acordo com [33], é possível representar uma estratégia baseada no CRP por meio de um vetor $\mathbf{b} = (b_1, b_2, \dots, b_n)^\top$. A partir dessa representação, a riqueza final de um portfólio CRP pode ser calculada como:

$$f_t(\text{CRP}(\mathbf{b})) = \prod_{t=1}^T \mathbf{b}^\top \mathbf{r}_t, \quad (3.37)$$

onde \mathbf{r}_t é o vetor que contém o retorno de cada um dos ativos i no tempo t e \mathbb{S}_n é o conjunto convexo para \mathbf{b} definido pelas restrições $\sum_{i=1}^n b_i = 1$ e $b_i \geq 0 \forall i$. Assim, o CRP ótimo *offline* é definido de maneira formal por:

$$\mathbf{b}^* = \arg \max_{\mathbf{b} \in \mathbb{S}_n} f_t(\text{CRP}(\mathbf{b})) = \arg \max_{\mathbf{b} \in \mathbb{S}_n} \prod_{t=1}^T \mathbf{b}^\top \mathbf{r}_t. \quad (3.38)$$

A função que se deseja maximizar é convexa, podendo então ser resolvida de maneira eficiente. O vetor \mathbf{b}^* é chamado *Best Constant Rebalanced Portfolios* (BCRP). Em seu trabalho seminal, Cover [18] demonstrou que o BCRP é a melhor estratégia em um mercado independente e identicamente distribuído (IID), superando estratégias como *Uniform Constant Rebalanced Portfolio* (UCRP), *Best Stock* (BS) e índices de referência como *Dow Jones Industrial Average Index* (DJIA).

3.4.2 Seleção de Portfólio *Online*

O uso de aprendizado de máquina em problemas de seleção de portfólios se fundamenta em trabalhos como [19, 110, 111, 34]. Estes mostraram que um portfólio constante reequilibrado para um alocação específica (CRP), pode ser mais vantajoso do que selecionar uma carteira baseada em uma medida de desempenho e mantê-la durante todo período. Além disso, [18] demonstrou que existe uma estratégia de alocação constante que proporciona o maior crescimento da riqueza a longo prazo (BCRP).

Nesse contexto, os algoritmos de seleção de portfólio *online* visam aproximar o desempenho do BCRP, adaptando a alocação do portfólio à medida que novos dados de preços se tornam disponíveis. Para avaliar a eficácia desses algoritmos, utiliza-se a função de *Regret*, que mede a diferença entre o desempenho acumulado do algoritmo *online* e o desempenho de um algoritmo *offline* hipotético que teria acesso a todas as informações futuras.

Conforme proposto por [2], a função *Regret* pode ser calculada comparando um algoritmo *online* com a decisão tomada por um algoritmo *offline*, tendo acesso

às primeiras funções de custo f_1, f_2, \dots . A ideia é medir a distância (custo) até uma estratégia ótima. No caso específico da seleção de portfólio *online*, a estratégia de referência é a BCRP. Os algoritmos de otimização de portfólio *online* são baseados no preço atual do ativo para selecionar a alocação futura do portfólio, significando que as informações sobre o preço dos ativos chegam sequencialmente e a decisão de alocação é feita imediatamente. Em cada período de negociação t , para $t = 1, \dots, T$, um investidor observa um vetor de retorno \mathbf{r}_t , onde cada componente r_{it} é o retorno da ação i no tempo t com preço de fechamento até $t - 1$. O vetor \mathbf{b}_t contém a proporção aplicada a cada um dos n ativos. A riqueza de um portfólio investido em \mathbf{b}_t , considerando apenas um período, pode ser calculada pelo produto interno $\mathbf{b}_t^\top \mathbf{r}_t$. Após T períodos, a riqueza acumulada é dada por $\prod_{t=1}^T (\mathbf{b}_t^\top \mathbf{r}_t)$. Tomando o logaritmo da evolução da riqueza, obtém-se a taxa de crescimento $\sum_{t=1}^T \log(\mathbf{b}_t^\top \mathbf{r}_t)$. Da mesma forma, tomando o log da riqueza de um portfólio usando a estratégia de investimento CRP com $\mathbf{b} \in \mathbb{S}_n$, obtém-se a taxa de crescimento logarítmica $\sum_{t=1}^T \log(\mathbf{b}^\top \mathbf{r}_t)$. O *Regret* de um algoritmo *online* que seleciona o \mathbf{b}_t para $t = 1, \dots, T$, é dado por:

$$\text{Regret}(T) = \sum_{t=1}^T \log(\mathbf{b}^{*\top} \mathbf{r}_t) - \sum_{t=1}^T \log(\mathbf{b}_t^\top \mathbf{r}_t). \quad (3.39)$$

Assim, o algoritmo *online* tenta minimizar a distância entre o portfólio \mathbf{b}_t e o CRP ou BCRP ótimo. A lógica por trás disso é que, se um algoritmo tem baixo *Regret*, seu desempenho acumulado ao longo do tempo será próximo do desempenho do melhor CRP possível (ou seja, o BCRP). Em outras palavras, mesmo que o algoritmo não seja sempre ótimo em cada período, seu desempenho geral será semelhante ao do BCRP a longo prazo. Isso ocorre porque o baixo *Regret* implica que a diferença acumulada entre o desempenho do algoritmo e o do BCRP é limitada [2, 3, 24], mesmo quando o número de períodos tende ao infinito. Portanto, espera-se que um portfólio selecionado por um algoritmo de baixo *Regret* tenha o mesmo comportamento assintótico que o BCRP [4].

Na esfera do aprendizado de máquina *online*, conforme delineado por [112], as estratégias empregadas são categorizadas em três grupos principais: *Follow-the-Winner* (FTW), *Pattern-Matching Approaches* (PMA) e *Meta-Learning Algorithms* (MLA). Cada categoria reflete uma abordagem distinta na resolução de problemas.

De acordo com uma análise detalhada realizada por [109], os algoritmos da categoria FTW destacam-se por estabelecer um limite superior para o erro acumulado, denominado *Regret*, em comparação com a melhor solução constante em retrospectiva (BCRP), após um número adequado de iterações. Especificamente no caso do algoritmo *Online Gradient Descent* (OGD), este limite é proporcional à raiz quadrada do número de iterações, evidenciando que o aumento do erro é moderado, mesmo com a expansão do número de decisões tomadas. Tal característica revela-se extremamente benéfica em contextos onde decisões sequenciais são imprescindíveis sob condições de incerteza, tais como na previsão de mercados ou na gestão de recursos em tempo real, aspectos que serão explorados na subsequente subseção.

Ademais, uma vantagem significativa do OGD relaciona-se à sua capacidade de manter um arrependimento total sublinear, mesmo quando a taxa de transação permanece constante.⁵ Conforme descrito por [113], essa propriedade confere uma robustez notável ao OGD, tornando-o adequado para uma ampla gama de cenários de mercado, sem a necessidade de ajustes complexos no parâmetro de taxa de transação.

3.4.3 *Online Gradient Descent* (OGD)

O algoritmo OGD, introduzido por Zinkevich em 2003 [2], representa um marco fundamental no campo da Otimização Convexa *Online*. Projetado para problemas de tomada de decisão sequencial sob incerteza futura, o OGD atualiza soluções utilizando o gradiente da função de perda, que só é revelado após a decisão. Essa abordagem simplificada, porém eficaz, o consagrou como ferramenta essencial em áreas como aprendizado de máquina, finanças e alocação de recursos, onde decisões *online* são essenciais. Desta forma, inicialmente são enumeradas premissas básicas do algoritmo, colocando na prática questões discutidas nas últimas subseções. Depois, a incorporação e aplicabilidade das mesmas são melhor discutidas no restante dessa seção.

i **Ambiente Dinâmico:** O ambiente é considerado dinâmico, onde as condições

⁵ No contexto presente, a taxa de transação refere-se ao custo associado à corretagem, por exemplo, cobrado por corretoras, ou ainda ao custo de oportunidade vinculado à tomada de decisões.

do mercado e as informações disponíveis podem mudar a cada instante. O OGD é desenhado para adaptar-se a essas mudanças, atualizando o portfólio de forma contínua.

- ii **Decisões Iterativas:** O processo de seleção de portfólio \mathbf{b}_t ocorre em etapas (ou rodadas). A cada rodada, o investidor escolhe um portfólio com base nas informações disponíveis e, após observar o desempenho, ajusta sua estratégia para a próxima rodada.
- iii **Função de Perda Convexa:** A função de perda utilizada para avaliar o desempenho do portfólio em cada rodada é convexa, permitindo que o problema de otimização tenha uma estrutura que o OGD possa explorar de maneira eficaz. Vale ressaltar que no contexto de otimização de portfólios, a função de perda geralmente é baseada no logaritmo do retorno do portfólio, garantindo tal convexidade.
- iv **Atualização com Gradiente:** O portfólio é atualizado usando o gradiente da função de perda em relação ao portfólio atual. Isso significa que o ajuste feito a cada rodada é proporcional ao gradiente, movendo o portfólio na direção que minimiza a perda esperada.
- v **Restrições Convexas:** As escolhas de portfólio a cada rodada estão sujeitas a restrições convexas. Por exemplo, a soma das alocações deve ser 1 (a soma das frações investidas em cada ativo) e cada alocação deve ser não-negativa. Isso está relacionado à projeção do vetor de alocação atualizado no conjunto convexo \mathcal{D} , garantindo que a nova alocação \mathbf{b}_{t+1} esteja dentro do espaço permitido.
- vi **Não Expansividade da Projeção:** O operador de projeção $\Pi_{\mathcal{D}}$ é não expansivo, o que significa que a distância entre o vetor projetado e qualquer outro ponto é sempre menor ou igual à distância antes da projeção. Essa propriedade é importante para garantir que as atualizações não causem instabilidade no algoritmo.
- vii **Arrependimento Sublinear:** O OGD visa garantir um arrependimento sublinear ao longo do tempo, ou seja, o arrependimento total $Regret(T)$

crece mais lentamente do que o número de rodadas T , idealmente como $O(\sqrt{T})$. Isso implica que o algoritmo se aproxima do desempenho ótimo ao longo do tempo, minimizando a diferença acumulada entre a perda do portfólio selecionado online e a perda da melhor solução fixa em retrospectiva.

viii **Convergência a uma Solução Ótima:** Sob certas condições, como passos adequados na atualização (η_t bem ajustado), garante-se que o OGD vai convergir para uma solução que seja competitiva com a melhor escolha possível em retrospectiva, ou seja, a solução que teria sido escolhida com conhecimento completo do futuro.

ix **Competitividade de Desempenho:** A performance do OGD é avaliada em termos de arrependimento (*Regret*). O objetivo é minimizar esse arrependimento ao longo do tempo, garantindo que o desempenho do OGD seja competitivo com a melhor solução fixa em retrospectiva.

De fato, o método se baseia na descida de gradiente clássica, adaptada ao contexto *online*. Formalmente, para uma função de perda convexa $f_t(\mathbf{b})$ no tempo t , a atualização pelo OGD é descrita pelo procedimento iterativo:

$$\mathbf{b}_{t+1} = \mathbf{b}_t - \eta_t \nabla f_t(\mathbf{b}_t), \quad (3.40)$$

onde \mathbf{b}_t representa o vetor de parâmetros, $\nabla f_t(\mathbf{b}_t)$ o gradiente da função no ponto \mathbf{b}_t e η_t a taxa de aprendizado no momento t .

O pseudocódigo do OGD é demonstrado abaixo no Algoritmo 1.

Algoritmo 1: Algoritmo *Online Gradient Descent*

Input: $OGD(\eta)$

início $\mathbf{b}_1 \leftarrow \frac{1}{n}$

for $t = 2$ **to** T **do**

$\mathbf{b}_t = \Pi_{\mathcal{D}} \left(\mathbf{b}_{t-1} + \eta \frac{\mathbf{r}_{t-1}}{\mathbf{b}_{t-1}^\top \mathbf{r}_{t-1}} \right)$

end

Onde $\Pi_{\mathcal{D}}(\mathbf{y}) = \arg \min_{\mathbf{b} \in \mathcal{D}} \|\mathbf{y} - \mathbf{b}\|$ e \mathcal{D} convexo

retorna $\mathbf{b}_{(t+1)}$

fim

O algoritmo recebe como entrada a taxa de aprendizado η , que controla o tamanho do passo na direção do gradiente. Na inicialização, o vetor de pesos \mathbf{b}_1 é definido como $\frac{1}{n}$, onde n é a dimensão do espaço de pesos. Em seguida, o algoritmo itera por T rodadas, começando de $t = 2$ até T .

A cada iteração t , o vetor de pesos \mathbf{b}_t é atualizado utilizando a equação de atualização:

$$\mathbf{b}_t = \Pi_{\mathcal{D}} \left(\mathbf{b}_{t-1} + \eta \frac{\mathbf{r}_{t-1}}{\mathbf{b}_{t-1}^\top \mathbf{r}_{t-1}} \right).$$

Nesta equação, \mathbf{b}_{t-1} representa o vetor de pesos da iteração anterior, \mathbf{r}_{t-1} é o vetor de recompensas (ou perdas) observado no tempo $t - 1$, e $\mathbf{b}_{t-1}^\top \mathbf{r}_{t-1}$ é o produto escalar entre o vetor de recompensas e o vetor de pesos, que representa a recompensa (ou perda) total acumulada até $t - 1$. O termo $\frac{\mathbf{r}_{t-1}}{\mathbf{b}_{t-1}^\top \mathbf{r}_{t-1}}$ é o gradiente normalizado da função de perda em relação aos pesos. A atualização dos pesos é realizada tomando um passo na direção desse gradiente, proporcional a η .

Após a atualização, a função $\Pi_{\mathcal{D}}(\mathbf{y})$ é aplicada para projetar o vetor resultante de volta ao conjunto convexo \mathcal{D} , garantindo que os pesos atualizados permaneçam dentro de um espaço válido e limitado. Após T rodadas, o algoritmo retorna o vetor de pesos final \mathbf{b}_{T+1} . Essa forma de projeção é descrita com maior riqueza de detalhes e exemplificada na Seção 4.1.

Este método mostra-se particularmente eficaz em cenários onde funções de perda são expostas sequencialmente, permitindo ajustes contínuos e aprimoramento do desempenho ao longo do tempo.

O *Regret* no algoritmo OGD é uma métrica essencial que mede a diferença entre o custo acumulado pelo algoritmo e o custo acumulado pela melhor decisão fixa em retrospecto, após T rodadas. Um dos pontos fundamentais do OGD, conforme apresentado por Zinkevich, é a garantia de que o arrependimento cresce sublinearmente com o número de iterações, o que é crucial para a sua aplicabilidade em problemas de otimização convexa *online*.

Mais especificamente, o arrependimento $\text{Regret}(T)$ após T rodadas é limitado pela seguinte expressão:

$$\text{Regret}(T) \leq \frac{D^2}{2\eta} + \frac{\eta}{2} \sum_{t=1}^T \|\nabla f_t(\mathbf{b}_t)\|^2,$$

onde η é a taxa de aprendizado, D é uma constante que limita a distância máxima entre qualquer dois pontos no domínio do problema, e $\nabla f_t(\mathbf{b}_t)$ é o gradiente da função de perda no tempo t avaliado no ponto \mathbf{b}_t .

Esta garantia de que o arrependimento cresce sublinearmente, dependendo crucialmente da escolha da taxa de aprendizado e do comportamento dos gradientes ao longo do tempo, bem como da geometria do espaço de decisão, é um pilar central que justifica a eficácia do OGD. A escolha adequada de η garante que o arrependimento aumenta mais lentamente que T , evidenciando que o algoritmo está aprendendo e adaptando-se de forma eficiente ao problema ao longo do tempo.

Essa notação mostra que o arrependimento depende crucialmente da escolha da taxa de aprendizado e do comportamento dos gradientes ao longo do tempo. Com uma escolha adequada de η , o arrependimento cresce mais lentamente do que T , indicando que o algoritmo está aprendendo e se adaptando efetivamente ao problema ao longo do tempo.

Em seu trabalho [2], Zinkevich foca no ajuste do parâmetro de taxa de aprendizado η para garantir uma boa performance do algoritmo em termos de *Regret* cumulativo, que é comparado com uma sequência de melhores respostas estáticas.

A principal formulação de *Regret* apresentada por Zinkevich é expressa como:

$$\text{Regret}(T) \leq \frac{D^2}{2\eta} + \frac{\eta TG^2}{2},$$

onde D é o diâmetro do conjunto viável $\mathcal{D} \subset \mathbb{R}^n$ e G é uma cota superior para a norma do gradiente das funções de custo.⁶ Esta formulação mostra um *trade-off* interessante no ajuste de η : valores maiores de η aumentam o segundo termo proporcional a T , enquanto valores menores aumentam o termo proporcional a D^2 .

A formulação de Zinkevich para o *Regret* em otimização convexa online revela um balanço cuidadosamente calibrado entre o tamanho do passo da descida do gradiente e o horizonte temporal considerado. A escolha da taxa de aprendizado η é um elemento central nesta equação, tendo a capacidade de determinar a rapidez com que o algoritmo se ajusta às novas informações $\frac{\eta TG^2}{2}$ e a estabilidade ou

⁶ $\sum_{t=1}^T \|\nabla f_t(\mathbf{b}_t)\|^2 \leq TG^2$

suavidade deste ajuste $\frac{D^2}{2\eta}$. Ao definir η proporcional a $\frac{1}{\sqrt{T}}$, Zinkevich não somente assegura que o termo relacionado à variação do gradiente das funções de custo e o termo relacionado ao diâmetro do conjunto viável se equilibram, mas também estabelece um crescimento sublinear do *Regret* em relação ao número de iterações T . Esta propriedade sublinear é essencial, pois indica que, em média, a eficácia do algoritmo se aproxima da melhor solução estática possível à medida que T cresce.

Ao escolher $\eta = \frac{1}{\sqrt{T}}$, Zinkevich demonstra que é possível obter um *Regret* da ordem de $\mathcal{O}(\sqrt{T})$:

$$\text{Regret}(T) \leq DG\sqrt{T}.$$

Portanto, mesmo com uma taxa de aprendizado fixa, o algoritmo consegue obter cotas significativas de *Regret* para o caso regular, desde que o valor de η seja escolhido adequadamente.

Concluindo a análise do algoritmo OGD, é importante notar que, como ocorre com muitos algoritmos de otimização *online*, há uma tendência a focar predominantemente na maximização do retorno, sem incorporar explicitamente considerações de risco na formulação do algoritmo. Essa abordagem pode não ser ideal em todos os contextos de investimento, especialmente naqueles onde a volatilidade e os riscos potenciais desempenham um papel crítico.

4 TÉCNICAS DE PROJEÇÃO PARA CONTROLE DE RISCO

No âmbito dos mercados financeiros, os investidores enfrentam o desafio complexo de desenvolver algoritmos eficientes para a seleção de portfólios, com o objetivo primordial de maximizar os retornos potenciais. Frequentemente, os algoritmos de seleção de portfólio *online* propõem alocações que, embora sejam matematicamente ótimas, podem não estar alinhadas com as restrições e particularidades inerentes ao mercado real. Por exemplo, apesar das vendas a descoberto serem permitidas em diversas jurisdições e contextos, limitações operacionais e regulamentações específicas podem impactar sua viabilidade ou atratividade para determinados investidores. Diante desse cenário, torna-se imperativo a aplicação de técnicas de projeção que ajustem essas alocações para um conjunto viável, não apenas considerando as limitações operacionais, mas também controlando o risco de maneira efetiva.

Neste contexto, as técnicas de projeção têm desempenhado um papel crucial nos algoritmos de otimização *online*, sendo o simplex uma abordagem comumente utilizada para ajustar alocações de portfólio. No entanto, além do simplex, esta tese propõe três métodos de projeção adicionais que são especificamente projetados para controlar o risco, oferecendo uma abordagem mais robusta e adaptativa às condições de mercado.

Este capítulo dedica-se a explorar, de forma aprofundada, tanto a técnica de projeção geométrica tradicional do simplex, quanto as técnicas propostas neste trabalho, que têm a capacidade de controlar o risco. Primeiramente, na Seção 4.1, será discutida a projeção em um simplex, incluindo um exemplo de implementação detalhada de um algoritmo simples. Em seguida, na Seção 4.2, será definida uma das projeções utilizadas neste trabalho, que leva em consideração o beta variante no tempo, permitindo uma adaptação contínua ao risco sistemático. Posteriormente, na Seção 4.3, a imposição do *Value at Risk* (VaR) como critério para a escolha do portfólio ótimo é abordada, destacando sua eficácia no controle de risco em cenários adversos. Por fim, na Seção 4.4, o uso do *Conditional Value at Risk* (CVaR) como medida chave para a gestão do risco extremo é analisado, assegurando que o portfólio seja otimizado não apenas para retorno, mas também para robustez em

face de perdas significativas.

Dessa forma, este capítulo não apenas revisa a técnica clássica de projeção em simplex, mas também apresenta e discute métodos inovadores que fortalecem o controle de risco em algoritmos de seleção de portfólio *online*, oferecendo soluções que são ao mesmo tempo viáveis e rigorosamente seguras.

4.1 PROJEÇÃO EM UM SIMPLEX

De acordo com [109], o simplex desempenha um papel crucial na otimização de portfólio. Denotado por \mathfrak{B}_n , o simplex oferece uma representação geométrica precisa das possíveis combinações de alocações da riqueza de um investidor entre n ativos distintos. Em essência, ele ilustra o espaço de alocação em que cada dimensão é representada por um ativo, e todas as combinações viáveis de alocações residem dentro dessa estrutura.

Este conceito é central, pois fornece uma compreensão visual e matemática das restrições e possibilidades inerentes à seleção de portfólio. Matematicamente, o simplex \mathfrak{B}_n pode ser expresso como:

$$\mathfrak{B}_n = \left\{ \mathbf{b} \in \mathbb{R}^n : b_i \geq 0, \sum_{i=1}^n b_i = 1 \right\} \quad (4.1)$$

Aqui, cada componente b_i ilustra a proporção do capital do investidor alocada ao i -ésimo ativo. Esta definição assegura que as alocações sejam sempre não-negativas e que a alocação total seja unitária, refletindo a totalidade do capital investido.

Em geral, existe um número infinito de alocações dentro do simplex \mathfrak{B}_n , porque $b_i \in \mathbb{R}$. No entanto, para fins de simplificação, o simplex contínuo pode ser discretizado, reduzindo assim o número de alocações para um número finito. Seja a um fator multiplicativo com $0 < a \leq 1$ e $\frac{1}{a} \in \mathbb{N}^*$, com \mathbb{N}^* sendo o conjunto de números naturais diferentes de zero $1, 2, \dots$ e $\frac{1}{a}$ uma partição. Além disso, suponha que cada b_i seja um múltiplo de a , ou seja, $\frac{b_i}{a} \in \mathbb{N}^*$. Um investidor deve alocar $\frac{1}{a}$ vezes uma porcentagem de riqueza a em n ativos. Ao final, $\frac{1}{a} \times a = 100\%$ da riqueza é alocada nos n ativos. O investidor pode realizar seleções repetitivas de um

ativo. É uma questão em aberto determinar quantas alocações possíveis existem em um simplex \mathfrak{B}_n quando a e n são dados. Claramente, trata-se de um problema de análise combinatória com repetição em que a ordem de seleção não importa [114]. Desta forma,

$$\Omega = \frac{\left(n + \frac{1}{a} - 1\right)!}{\frac{1}{a}!(n-1)!}$$

quantifica o número de diferentes alocações quando o simplex \mathfrak{B}_n é discretizado. O número de alocações aumenta essencialmente com a redução de a e o aumento de n .

Uma metodologia frequentemente adotada para tipos de algoritmos de otimização consiste em criar uma alocação intermediária $\mathbf{b}'_t = (b'_{1t}, \dots, b'_{nt})^\top$, que inicialmente desconsidera a premissa de não realizar vendas a descoberto ($b_{it} \geq 0$). Ao concluir os cálculos do algoritmo, o resultado de \mathbf{b}'_t é normalizado por uma projeção no simplex \mathfrak{B}_n , conforme expresso na Equação 4.1, garantindo assim que não ocorram vendas a descoberto. Mais detalhadamente, o problema de minimização em:

$$\mathbf{b}_t = \arg \min_{\mathbf{b} \in \mathfrak{B}_n} \|\mathbf{b} - \mathbf{b}'_t\|$$

é resolvido. Em outras palavras, busca-se o valor de \mathbf{b} que possua a menor distância para \mathbf{b}'_t . Para solucionar este problema, em [115] foi proposto um algoritmo, apresentado sob a forma de pseudocódigo no Algoritmo 2. A projeção no simplex é uma etapa crucial em diversos algoritmos de seleção de portfólio *online*, garantindo que os pesos do portfólio permaneçam válidos após as atualizações. Esta técnica é amplamente utilizada em várias abordagens, incluindo estratégias de *Follow the Winner*, como o Exponential Gradient (EG) [22] e o Online Newton Step (ONS) [24]; estratégias de *Pattern Matching*, como o Correlation-Driven Nonparametric Learning (CORN) [40]; e estratégias de *Follow the Loser*, como o Passive Aggressive Mean Reversion (PAMR) [30], o Confidence Weighted Mean Reversion (CWMR) [31], e o Robust Median Reversion (RMR) [32].

No início de seu horizonte de investimento, os algoritmos *online* geralmente têm informações mínimas. Conseqüentemente, eles requerem um ou mais períodos de negociação para gerar alocações viáveis. Nesta fase, a diversificação ingênua, como $\mathbf{b}_1 = \left(\frac{1}{n}, \dots, \frac{1}{n}\right)^\top$, é frequentemente empregada. Este é um passo importante para garantir a estabilidade e viabilidade do algoritmo em seus estágios iniciais. Seguidamente, os elementos de \mathbf{b}'_t são ordenados. A variável auxiliar th é inicializada com zero e é utilizada para armazenar o valor de projeção para normalização. Através de *loops*, os elementos de \mathbf{b}'_t são atualizados conforme a alocação intermediária. A condição $sumDiv(i) > y_{i-1}$ garante que a alocação permaneça dentro do símplice \mathfrak{B}_n .

Algoritmo 2: Algoritmo de Projeção Simplex

Função Projeção da alocação \mathbf{b}_t (b_t):

```

 $y \leftarrow \mathbf{b}'_t$ ;
Ordene  $y$  em ordem crescente  $y_{(1)} \leq y_{(2)} \leq \dots \leq y_{(n)}$ 
 $th \leftarrow 0$ ;
for  $i = n$  to  $2$  do
    for  $j = i$  to  $n - 1$  do
         $sumDiv(i) \leftarrow sumDiv(i) + y_j$ ;
     $sumDiv(i) \leftarrow \frac{sumDiv(i)-1}{n-i}$ ;
    if  $sumDiv(i) > y_{i-1}$  then
         $th \leftarrow sumDiv(i)$ ;
        exit;
for  $i = 1$  to  $n$  do
     $b_i \leftarrow b_i - th$ ;
    if  $y_i < 0$  then
         $b_i \leftarrow 0$ ;
return  $\mathbf{b}_t$ ;

```

Ao final do algoritmo, todos os valores de alocação que resultam negativos são definidos como zero, garantindo que não ocorram vendas a descoberto ($b_{it} \geq 0$) e que não seja investido mais do que a totalidade do capital disponível ($\sum_{i=1}^n b_{it} = 1$).

Recentes avanços na literatura, destacados especialmente nos trabalhos [40, 30, 31, 32], introduziram soluções computacionalmente eficientes que fornecem garantias formais de um *Regret* sublinear. Entretanto, uma lacuna identificada nestas abordagens é a insuficiente consideração da estrutura de risco do portfólio proposto. Apesar de potencialmente alcançarem boas performances em termos de retorno, esses métodos tendem a gerar portfólios que se inclinam a favorecer um conjunto limitado de ativos. Tal característica pode, em diversos contextos, sinalizar uma exposição exacerbada ao risco.

Esta observação é corroborada em estudos como [116, 117], nos quais os algoritmos *Online Newton Step* (ONS) e *Online Gradient Descent* (OGD) foram implementados. Nos experimentos empíricos realizados, restrições relativas ao risco foram impostas, como um teto para investimento em ativos individuais ou um piso no número de ativos no portfólio. A inclusão dessas restrições visa formar portfólios menos concentrados e, conseqüentemente, diminuir o risco inerente. Para ambos, OGD e ONS, foram admitidas apenas posições da carteira não-negativas, garantindo que o portfólio não incorra em posições vendidas e que o somatório de todas as posições não ultrapasse um, prevenindo assim qualquer alavancagem.

Formalmente, o domínio foi descrito por um conjunto de restrições lineares, configurando-se, portanto, como um polítopo convexo. Assim, as projeções exigidas por ambos os algoritmos podem ser eficientemente resolvidas através de programação quadrática. Em ambos os casos, o conjunto viável \mathfrak{B}_n é representado por um polítopo, dentro do qual não se permitem posições vendidas nem alavancagem.

Conforme discutido nas Seções 3.1 e 3.2, a diversificação do portfólio de investimento oferece múltiplos benefícios. No entanto, simplesmente definir um limite superior de investimento em cada ativo pode não necessariamente diminuir o risco da carteira. Como o algoritmo trabalha simplesmente com distâncias do passo de atualização em relação a um vetor de investimentos específicos no conjunto de opções viáveis, limitar o investimento máximo em um ativo necessariamente implica em investimentos em uma quantidade maior de ações, mas nada impede que as demais ações escolhidas sejam por exemplo do mesmo setor.

Na próxima seção, é apresentado formalmente um algoritmo que possibilita

levar em conta e limitar o risco sistemático inerente ao mercado no tempo.

4.2 PROJEÇÃO USANDO BETA VARIANTE NO TEMPO

Antes de discutir a incorporação do beta variante no tempo no contexto OGD para escolha de carteiras, algumas premissas são destacadas. Do contexto do CAPM, vide [5], é pressuposto: (i) ausência de custos de transação; (ii) taxa livre de risco comum a todos no mercado; (iii) investidores com funções utilidade crescentes com relação ao retorno e decrescentes com relação ao risco; (iv) homogeneidade das expectativas de investidores, isto é, todos têm expectativas comuns no mercado, o que é inserido no contexto via equações discutidas na seção 3.2 onde se considera retornos com choques idiossincráticos $(\varepsilon_{i,t})$ que tem distribuição normal com média zero e variância σ_i^2 , o que permite a aplicação do Filtro de Kalman e estimação consistente de seus parâmetros via Máxima Verossimilhança e, assim, a atualização recursiva dos parâmetros à medida que novas informações se tornam disponíveis, o que é otimizado sob a suposição de normalidade dos dados observados. No que tange a incorporação de betas variantes no tempo, é assumido que o coeficiente beta dos ativos $(\beta_{i,t})$ segue um processo de passeio aleatório, o que implica que a melhor estimativa de $\beta_{i,t+1}$ é $\beta_{i,t}$, ajustada por um erro gaussiano $(\varphi_{i,t})$ independente da variável latente. Esta suposição assume que as mudanças no beta são: graduais e imprevisíveis; não utilizam informações futuras para prevê-lo, evitando assim o viés de antecipação; além de facilitar a decomposição e estimação das variáveis através do Filtro de Kalman.

Para incorporar na prática o risco usando o β do portfólio, assume-se intervalos $[\beta_{min}, \beta_{max}]$ que o portfólio ou carteira como um todo deve satisfazer, isto é, o risco sistemático pode ser controlado e ajustado através da manipulação desses limites, alinhando o risco do portfólio às preferências do investidor ou às condições do mercado. Como o modelo CAPM é linear em relação ao β_i dos ativos ($i = 1, 2, \dots, n$), com o β de todo o portfólio sendo uma combinação linear do risco sistemático de cada ativo ponderado pelo respectivo montante investido, o conjunto de alocações possíveis \mathcal{B}_β é definido como:

$$\mathcal{B}_\beta = \left\{ \mathbf{b} \in \mathbb{R}^n : \sum_{i=1}^n b_i \beta_{i,t} \in [\beta_{min}, \beta_{max}], \sum_{i=1}^n b_i = 1, \mathbf{b} \geq 0 \right\}, \quad (4.2)$$

onde $\beta_{i,t}$ é o β variante no tempo do i -ésimo ativo em tempo t . É importante notar que não necessariamente \mathcal{B}_β é não vazio, pois a intersecção dos semiespaços $\{\sum b_i \beta_{i,t} \leq \beta_{max}\}$ e $\{\sum b_i \beta_{i,t} \geq \beta_{min}\}$ pode ser vazia em certas condições. Essa situação pode ocorrer quando não há ativo i no período t tal que $\beta_{i,t} \geq \beta_{min}$ ou não há ativo i no período t tal que $\beta_{i,t} \leq \beta_{max}$.

O algoritmo OGD que incorpora e limita o risco sistemático é apresentado no Algoritmo 3 abaixo. Para evitar qualquer tipo de viés de antecipação, conforme citado acima, a estimativa de β_t^i para o investimento em $t + 1$ pode usar apenas as informações disponíveis desde 1 até t . Portanto, para cada ativo i no tempo t , usa-se apenas as informações dadas no conjunto $\{r_{i,1}, r_{i,2}, \dots, r_{i,t-1}\}$ para usar o filtro de Kalman. Além disso, utilizou-se o filtro de Kalman com iniciação difusa exata disponível no pacote KFAS ([118]) no repositório CRAN para R ([119]), que permite lidar com a questão de inicializar os parâmetros do filtro de maneira eficiente. Para uma discussão sobre algoritmos de filtro de Kalman, ver [95]. Outra consideração importante é a escolha dos limites β_{min} e β_{max} . Esses valores devem ser selecionados de acordo com a tolerância ao risco do investidor e as condições de mercado. Limites muito estreitos podem restringir excessivamente o conjunto de alocações possíveis, potencialmente prejudicando o desempenho do portfólio. Por outro lado, limites muito amplos podem não fornecer o controle de risco desejado. Uma abordagem possível é usar informações históricas sobre os β dos ativos e do mercado para definir esses limites de forma adaptativa ao longo do tempo via uso de janelas deslizantes.

O Algoritmo 3 é iniciado com a alocação UCRP (*Uniform Constant Rebalanced Portfolio*), ou seja, $b_{i,0} = \frac{1}{n}$, alocando uma quantidade igual de riqueza em cada ativo disponível. Essa escolha pela alocação inicial é feita porque ela não toma nenhuma informação que possa criar algum viés na análise, e o UCRP é um *benchmark* forte por si só, pois elimina o risco não sistemático pela diversificação. No entanto, é importante notar que a alocação UCRP pode não ser a escolha ideal em todos os cenários, especialmente quando há uma grande disparidade no

Algoritmo 3: Algoritmo Projeção OGD com beta variante do tempo

input : r = asset returns, b_0 = initial allocation
for $t = 1, \dots, T$ **do**
 Update step $\mathbf{x}_t = \mathbf{b}_{t-1} - \eta \nabla \text{Regret}_{t-1}$;
 for $i = 1, \dots, n$ **do**
 Update $\beta_{i,t-1|t-1} = \beta_{i,t-1} + K_t(r_{i,t-1} - \beta_{i,t} r_{m,t-1} - r_{f,t-1})$
 Project to feasible set $\mathbf{b} = \arg \min_{\mathbf{b} \in \mathcal{B}_\beta} \|\mathbf{b} - \mathbf{x}_t\|$ where \mathcal{B}_β is defined by equation 4.2;
 Update the gradient $\nabla \text{Regret}_t = -\frac{r_t}{\mathbf{b}_t^\top r_t}$;
 for $i = 1, \dots, n$ **do**
 Update β predictions $\beta_{i,t} = \beta_{i,t-1|t-1}$;
 Update Kalman-filter equations (3.9) to (3.12) using $r_{i,t}$;

desempenho esperado dos ativos. Na Seção 6.1 há a apresentação dos resultados da aplicação dessa técnica de projeção. Embora o controle do beta possa ajudar a gerenciar o risco de mercado (risco sistemático), ele não aborda completamente outros tipos de riscos (como o risco não sistemático). Portanto, afirmar que controlar o beta de uma carteira equivale a controlar todo o risco seria uma simplificação excessiva. Com isso em mente, na próxima seção, a análise é expandida para além do controle do risco sistemático através do beta, abordando outras medidas e estratégias importantes para um gerenciamento de risco mais abrangente e eficaz. De fato, o VaR pode complementar o controle do beta, oferecendo uma visão mais completa dos riscos enfrentados por um portfólio.

4.3 VALUE AT RISK COMO RESTRIÇÕES CÔNICAS DE SEGUNDA ORDEM

Na Subseção 3.3.1 foi introduzido o conceito de *Value at Risk* (VaR), uma medida amplamente utilizada para quantificar o risco de mercado em portfólios de investimento. Retomando esse tópico, é importante ressaltar que o VaR oferece aos gestores de risco uma estimativa da perda esperada para um portfólio específico, considerando condições normais de mercado, em um horizonte de tempo predeterminado e com um nível de confiança estabelecido. Essa métrica é essencial para o gerenciamento eficaz do risco, pois permite que as instituições financeiras, fundos de *hedge*, seguradoras e reguladores avaliem e monitorem a exposição ao risco de

mercado de forma padronizada e compreensível. Ao expressar o risco em termos monetários, o VaR oferece uma visão clara do potencial de perdas, auxiliando na tomada de decisões estratégicas e na alocação de capital.

Com isso em mente, inicialmente são enumeradas as premissas que serão utilizadas e discutidas ao longo dessa seção, permitindo a incorporação do VaR no contexto OGD. De fato: (i) o modelo assume que os retornos dos ativos em diferentes períodos são independentes e identicamente distribuídos (IID) com distribuição normal ou t de Student multivariada. Essa suposição facilita a formulação de um problema de otimização convexa e permite o uso eficiente de técnicas de programação cônica, estimativas consistentes dos parâmetros e com capacidade de capturar caudas mais pesada no contexto específico da segunda distribuição citada, o qual representa melhor os eventos extremos no mercado financeiro na segunda distribuição citada; (ii) o uso do VaR no modelo assume que essa métrica é suficiente para limitar a exposição ao risco, conforme os objetivos do investidor; (iii) a formulação matemática do problema de otimização pressupõe que o problema é convexo, permitindo a aplicação de técnicas de programação cônica de segunda ordem. A convexidade é essencial para garantir que o problema seja resolvível de maneira eficiente, e que as soluções obtidas sejam globalmente ótimas dentro do conjunto de soluções viáveis; (iv) supõe-se que a matriz de covariância dos retornos, Σ_t , é positiva semi-definida, garantindo que a matriz possa ser decomposta em uma forma que permite a aplicação de técnicas de programação cônica; (v) pressupõe que a utilização de diferentes níveis de confiança (γ) associados ao VaR é uma forma válida e eficaz de capturar diferentes perfis de risco, pois o mesmo reflete o grau de aversão ao risco do investidor e é tratada como uma variável ajustável no modelo; (vi) assume que o algoritmo de projeção ortogonal utilizado é eficiente e que a solução obtida pela projeção minimiza adequadamente a função objetivo, respeitando as restrições impostas pelo VaR e as restrições cônicas.

4.3.1 VaR e sua relação com a distribuição dos retornos

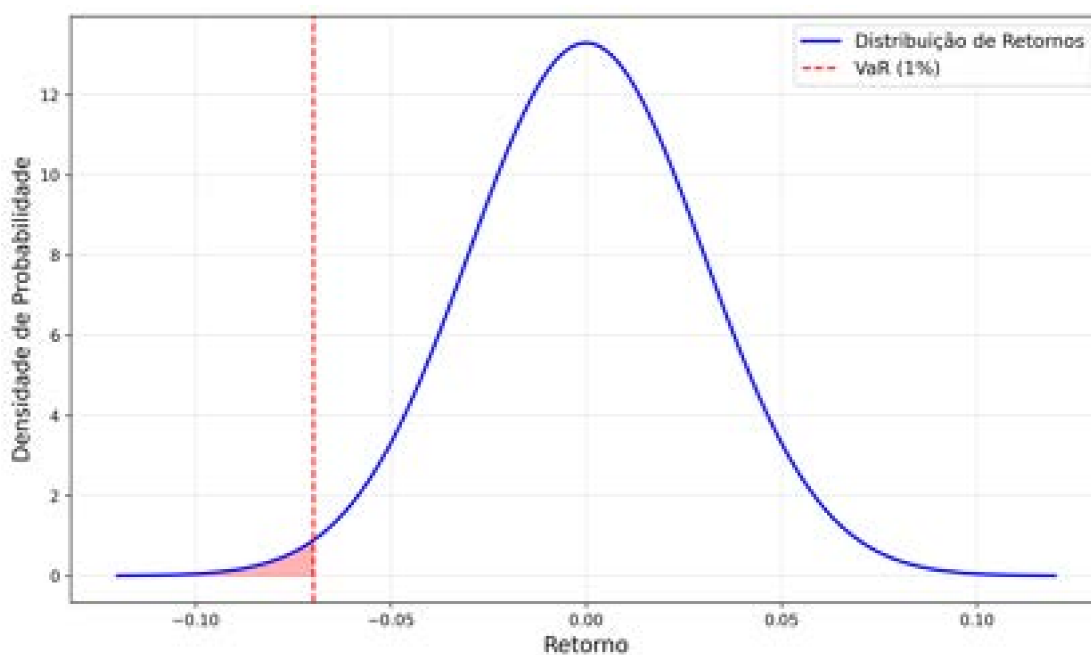
Para compreender melhor o conceito de VaR e sua relação com a distribuição de retornos do portfólio, o Gráfico 2 apresenta uma ilustração do funcionamento. Nesse exemplo, assume-se que os retornos seguem uma distribuição normal com

média 0 e desvio padrão 0,03. A curva azul representa a densidade de probabilidade dos retornos, indicando a probabilidade de ocorrência de cada valor de retorno.

A linha vertical tracejada em vermelho indica o VaR de 1%, que representa a perda máxima esperada com uma probabilidade de 1%. Isso significa que há uma probabilidade de 1% de que os retornos do portfólio sejam iguais ou inferiores ao valor do VaR.

A área sombreada em vermelho à esquerda do VaR de 1% destaca a região correspondente ao quantil de 1% na cauda esquerda da distribuição. Essa área representa a probabilidade de ocorrência de retornos extremamente baixos, iguais ou inferiores ao VaR de 1%.

Gráfico 2 – Distribuição de Retornos do Portfólio e VaR 1%

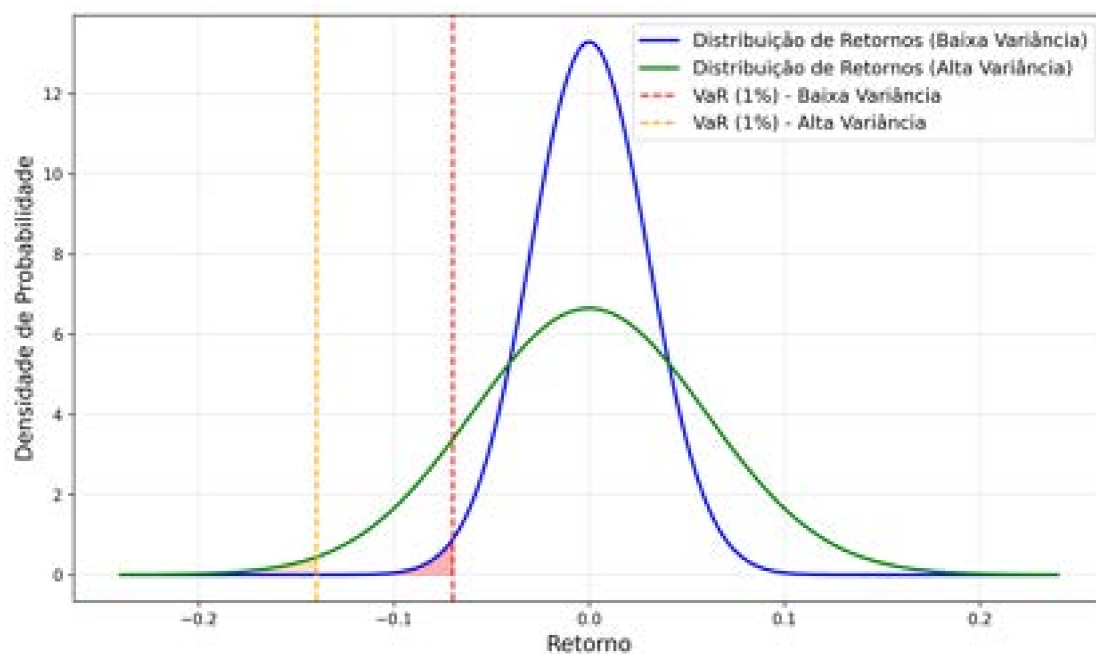


Fonte: Elaborado pelo autor (2024).

A hipótese de normalidade dos retornos em finanças nos proporciona uma base simplificada para análise estatística, assumindo uma distribuição simétrica dos retornos em torno da média, com uma variabilidade padrão. Porém, ao se aprofundar na natureza dos retornos financeiros, nota-se que diferentes níveis de volatilidade podem alterar significativamente o perfil de risco de um portfólio. Expandindo a

compreensão para além das distribuições com variância definida exclusivamente por um único parâmetro, o Gráfico 3, a seguir, ilustra uma comparação entre distribuições de retornos de portfólio que refletem distintos graus de incerteza: a curva azul, com um desvio padrão mais conservador (como a t de Student com mais graus de liberdade), revela uma concentração maior de retornos próximos à média; enquanto a curva verde, representando uma variância maior (como a t de Student com menos graus de liberdade), dispersa mais amplamente os resultados, sugerindo um perfil de risco potencialmente elevado. Ambas são acompanhadas por suas respectivas linhas de VaR a 1%, demonstrando o impacto da volatilidade no cálculo de possíveis perdas. Ao comparar as duas distribuições, fica evidente que a distribuição com alta volatilidade apresenta uma curva mais achatada e dispersa, indicando uma maior probabilidade de ocorrência de retornos extremos, tanto positivos quanto negativos. Consequentemente, o VaR de 1% para a distribuição de alta volatilidade está localizado mais à esquerda em relação à média, implicando em uma perda máxima esperada maior.

Gráfico 3 – Comparação de Distribuições de Retornos com Diferentes Variâncias e seus VaR 1%



Fonte: Elaborado pelo autor (2024).

Essa análise ressalta a importância da volatilidade na avaliação do risco de um portfólio. Uma maior volatilidade significa uma maior dispersão dos retornos em torno da média, o que se traduz em uma maior exposição ao risco. Portanto, portfólios com alta volatilidade tendem a apresentar um VaR mais distante da média, indicando um potencial de perdas mais elevado.

Não menos importante, através da análise comparativa das distribuições com diferentes níveis de volatilidade, é possível estabelecer uma compreensão mais robusta do comportamento dos retornos e do impacto que a volatilidade tem no perfil de risco de um portfólio. A alocação de ativos em um portfólio não é apenas uma questão de maximizar os retornos, mas também de balancear e gerenciar o risco associado a esses retornos. Nesse contexto, o VaR pode ser uma ferramenta crucial na quantificação desse risco. Assim, na próxima subseção, os nuances desse processo são abordados e investigados, destacando como a incorporação do VaR nos modelos de otimização de portfólio pode contribuir para a construção de portfólios que não apenas buscam eficiência em termos de retorno, mas também aderem a um perfil de risco deliberadamente calculado.

4.3.2 Incorporação do VaR como restrição no modelo de otimização

A incorporação de medidas de risco, como o VaR, em modelos de otimização de portfólio é de grande importância para garantir que as alocações sejam não apenas ótimas em termos de retorno esperado, mas também estejam dentro de limites aceitáveis de risco. Nesta subseção, é explorado como o VaR pode ser incorporado no modelo de otimização *online* por meio de restrições cônicas de segunda ordem.

Inicialmente, assume-se que os retornos dos ativos seguem uma distribuição normal multivariada, o que permitirá formular o problema de otimização de forma convexa e eficiente. Em seguida, essa hipótese é relaxada e o caso em que as carteiras seguem uma distribuição *t* de Student é introduzido, proporcionando uma abordagem mais abrangente pois permite incorporar diferentes perfis de risco por meio dos parâmetros distribucionais de maneira simples, com a presença de caudas mais pesadas, usualmente visto no mercado, sendo uma de suas possibilidades.

Ao longo desta seção, demonstra-se como a inclusão do VaR como restrição no modelo de otimização pode ser realizada por meio de manipulações algébricas e da aplicação do teorema de decomposição espectral, como pode ser visto, por exemplo, no capítulo 6 de [120]. Além disso, a formulação epigráfica do problema é apresentada, permitindo resolvê-lo de maneira eficiente utilizando técnicas de programação cônica de segunda ordem. Embora o OGD possa ser usado em um cenário adversário e nenhuma hipótese precisa ser feita em relação a sequência de retornos para garantir *Regret* sublinear, nessa seção é assumido inicialmente que os retornos seguem realizações independentes de uma distribuição normal multivariada. Ao final, tal hipótese é relaxada e será assumido que qualquer combinação linear de ativos segue uma distribuição t de Student. Seja:

$$\mathbf{r}_t \sim N(\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t),$$

o algoritmo OGD que prevê como próxima alocação:

$$\mathbf{b}_t = \Pi_{\mathcal{S}_t} \left(\mathbf{b}_{t-1} + \eta \frac{1}{(\mathbf{b}_{t-1}^\top \mathbf{r}_{(t-1)})} \mathbf{r}_{(t-1)} \right),$$

onde o conjunto de alocações viáveis é definido como:

$$\mathcal{S}_t = \left\{ \mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} \leq \mathbf{b}, \quad \mathbb{P}(\mathbf{x}^\top \mathbf{r}_t \leq \xi) \leq \gamma \right\}.$$

Na definição de \mathcal{S}_t acima, $\mathbf{A}\mathbf{x} \leq \mathbf{b}$ representa restrições lineares, como limites máximos e mínimos de alocação, *shortage* e possivelmente restrições de CAPM definidos na seção anterior. O termo $\mathbb{P}(\mathbf{x}^\top \mathbf{r}_t \leq \xi)$ define que, para cada ξ (VaR), a probabilidade de que o retorno da carteira $\mathbf{x}^\top \mathbf{r}_t$ seja menor do que ξ é de no máximo γ (por exemplo 1%).

Primeiramente, lida-se com a incorporação do Value at Risk (VaR) como uma restrição no modelo de otimização de portfólio. O VaR é utilizado para medir o risco de perda em uma carteira. Aqui, tal característica é incorporada de forma que se busca garantir que a probabilidade de a perda ser maior que um certo valor ξ (o VaR) seja menor ou igual a um nível de confiança γ .

Desta forma, considere a probabilidade:

$$\mathbb{P}(\mathbf{x}^\top \mathbf{r}_t \leq \xi),$$

onde \mathbf{x} é o vetor de pesos da carteira, \mathbf{r}_t é o vetor de retornos dos ativos no tempo t e ξ é o nível de VaR. Isso mede a probabilidade de que o retorno da carteira seja menor do que ξ .

A média da variável aleatória $\mathbf{x}^\top \mathbf{r}_t$ é dada por:

$$\mathbb{E}[\mathbf{x}^\top \mathbf{r}_t] = \mathbf{x}^\top \boldsymbol{\mu}_t,$$

onde $\boldsymbol{\mu}_t$ é o vetor de médias dos retornos dos ativos.

O desvio padrão da variável aleatória $\mathbf{x}^\top \mathbf{r}_t$ é dado por:

$$\sqrt{\text{Var}(\mathbf{x}^\top \mathbf{r}_t)} = \sqrt{\mathbf{x}^\top \boldsymbol{\Sigma}_t \mathbf{x}},$$

onde $\boldsymbol{\Sigma}_t$ é a matriz de covariância dos retornos dos ativos.

A padronização da variável aleatória $\mathbf{x}^\top \mathbf{r}_t$ envolve subtrair a média e dividir pelo desvio padrão. Isso resulta na seguinte expressão padronizada:

$$\frac{\mathbf{x}^\top \mathbf{r}_t - \mathbf{x}^\top \boldsymbol{\mu}_t}{\sqrt{\mathbf{x}^\top \boldsymbol{\Sigma}_t \mathbf{x}}}.$$

A expressão padronizada acima é uma variável aleatória que segue uma distribuição normal padrão (com média 0 e variância 1) caso \mathbf{r}_t seja normalmente distribuído. Tal padronização permite comparar a variável aleatória $\mathbf{x}^\top \mathbf{r}_t$ com a distribuição normal padrão, facilitando o cálculo de probabilidades, como no caso do VaR.

Note que a probabilidade original pode ser reescrita em termos da função de distribuição acumulada da normal padrão, $\Phi(\cdot)$, que mapeia o valor padronizado para a probabilidade cumulativa:

$$\mathbb{P}(\mathbf{x}^\top \mathbf{r}_t \leq \xi) = \Phi\left(\frac{\xi - \mathbf{x}^\top \boldsymbol{\mu}_t}{\sqrt{\mathbf{x}^\top \boldsymbol{\Sigma}_t \mathbf{x}}}\right).$$

Aqui, $\frac{\xi - \mathbf{x}^\top \boldsymbol{\mu}_t}{\sqrt{\mathbf{x}^\top \boldsymbol{\Sigma}_t \mathbf{x}}}$ é o z-score correspondente, que indica quantos desvios padrão ξ está da média $\mathbf{x}^\top \boldsymbol{\mu}_t$.

Agora, a restrição de que essa probabilidade seja menor ou igual a γ é imposta:

$$\mathbb{P}(\mathbf{x}^\top \mathbf{r}_t \leq \xi) \leq \gamma.$$

Substituindo a expressão anterior:

$$\Phi\left(\frac{\xi - \mathbf{x}^\top \boldsymbol{\mu}_t}{\sqrt{\mathbf{x}^\top \boldsymbol{\Sigma}_t \mathbf{x}}}\right) \leq \gamma.$$

Para trabalhar com essa desigualdade de forma mais conveniente, aplica-se a função inversa da distribuição acumulada normal $\Phi^{-1}(\gamma)$ em ambos os lados dado a característica bijetora de tal função. A função $\Phi^{-1}(\gamma)$ é o valor que corresponde ao quantil γ da distribuição normal padrão:

$$\frac{\xi - \mathbf{x}^\top \boldsymbol{\mu}_t}{\sqrt{\mathbf{x}^\top \boldsymbol{\Sigma}_t \mathbf{x}}} \leq \Phi^{-1}(\gamma).$$

Finalmente, rearranja-se essa desigualdade para isolar o termo que contém a matriz de covariância $\boldsymbol{\Sigma}_t$:

$$\sqrt{\mathbf{x}^\top \boldsymbol{\Sigma}_t \mathbf{x}} \leq \frac{\mathbf{x}^\top \boldsymbol{\mu}_t - \xi}{\Phi^{-1}(\gamma)}.$$

Esse é o ponto onde se chegou à restrição desejada, que pode ser usada no modelo de otimização. Essa última expressão vincula o desvio padrão (medido pelo termo $\sqrt{\mathbf{x}^\top \boldsymbol{\Sigma}_t \mathbf{x}}$) ao VaR ξ e ao retorno esperado $\mathbf{x}^\top \boldsymbol{\mu}_t$.

Então, usando que $\boldsymbol{\Sigma}_t$ é positiva semi-definida e aplicando o teorema de decomposição espectral em $\boldsymbol{\Sigma}_t$, escreve-se $\boldsymbol{\Sigma}_t = \mathbf{C}_t \boldsymbol{\Lambda}_t \mathbf{C}_t^\top$ e $\boldsymbol{\Sigma}_t^{\frac{1}{2}} = \mathbf{C}_t \boldsymbol{\Lambda}_t^{\frac{1}{2}} \mathbf{C}_t^\top$, onde \mathbf{C}_t é a matriz formada pelos vetores característicos e $\boldsymbol{\Lambda}_t$ é a matriz diagonal com os valores característicos de $\boldsymbol{\Sigma}_t$. Portanto, dado que $\mathbf{x}^\top \boldsymbol{\Sigma}_t \mathbf{x} = \mathbf{x}^\top \boldsymbol{\Sigma}_t^{\frac{1}{2} \top} \boldsymbol{\Sigma}_t^{\frac{1}{2}} \mathbf{x} = \left\| \boldsymbol{\Sigma}_t^{\frac{1}{2}} \mathbf{x} \right\|^2$, o problema original pode ser reformulado como:

$$\begin{aligned} \min_{\mathbf{x}_t} \quad & \|\mathbf{x}_t - \mathbf{x}_t^*\| \\ \text{sujeito a} \quad & \mathbf{A} \mathbf{x}_t \leq \mathbf{b}, \\ & \left\| \boldsymbol{\Sigma}_t^{\frac{1}{2}} \mathbf{x}_t \right\| \leq \frac{\mathbf{x}_t^\top \boldsymbol{\mu}_t - \xi}{\Phi^{-1}(\gamma)}. \end{aligned} \tag{4.3}$$

Prova da convexidade: Deseja-se provar que o conjunto definido pela restrição

$$\left\| \Sigma_t^{\frac{1}{2}} \mathbf{x}_t \right\| \leq \frac{\mathbf{x}_t^\top \boldsymbol{\mu}_t - \xi}{\Phi^{-1}(\gamma)},$$

é convexo. Para tal, seja dois vetores \mathbf{x}_1 e \mathbf{x}_2 quaisquer onde a restrição em questão é válida. Assim,

$$\left\| \Sigma_t^{\frac{1}{2}} \mathbf{x}_1 \right\| \leq \frac{\mathbf{x}_1^\top \boldsymbol{\mu}_t - \xi}{\Phi^{-1}(\gamma)},$$

$$\left\| \Sigma_t^{\frac{1}{2}} \mathbf{x}_2 \right\| \leq \frac{\mathbf{x}_2^\top \boldsymbol{\mu}_t - \xi}{\Phi^{-1}(\gamma)}.$$

Seja $\lambda \in [0, 1]$ arbitrário onde se define $\mathbf{x}_3 = \lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2$. Para gerar a prova, basta mostrar que \mathbf{x}_3 satisfaz tal restrição. Note que a função norma aplicada à combinação linear de \mathbf{x}_1 e \mathbf{x}_2 pode ser expressa como:

$$\left\| \Sigma_t^{\frac{1}{2}} \mathbf{x}_3 \right\| = \left\| \Sigma_t^{\frac{1}{2}} (\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2) \right\|.$$

Aplicando a desigualdade triangular na norma euclidiana, tem-se:

$$\left\| \Sigma_t^{\frac{1}{2}} (\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2) \right\| \leq \lambda \left\| \Sigma_t^{\frac{1}{2}} \mathbf{x}_1 \right\| + (1 - \lambda) \left\| \Sigma_t^{\frac{1}{2}} \mathbf{x}_2 \right\|.$$

Substituindo as desigualdades dadas para \mathbf{x}_1 e \mathbf{x}_2 :

$$\lambda \left\| \Sigma_t^{\frac{1}{2}} \mathbf{x}_1 \right\| + (1 - \lambda) \left\| \Sigma_t^{\frac{1}{2}} \mathbf{x}_2 \right\| \leq \lambda \frac{\mathbf{x}_1^\top \boldsymbol{\mu}_t - \xi}{\Phi^{-1}(\gamma)} + (1 - \lambda) \frac{\mathbf{x}_2^\top \boldsymbol{\mu}_t - \xi}{\Phi^{-1}(\gamma)}.$$

Como o lado direito é uma combinação convexa, isso pode ser reescrito como:

$$\lambda \frac{\mathbf{x}_1^\top \boldsymbol{\mu}_t - \xi}{\Phi^{-1}(\gamma)} + (1 - \lambda) \frac{\mathbf{x}_2^\top \boldsymbol{\mu}_t - \xi}{\Phi^{-1}(\gamma)} = \frac{(\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2)^\top \boldsymbol{\mu}_t - \xi}{\Phi^{-1}(\gamma)} = \frac{\mathbf{x}_3^\top \boldsymbol{\mu}_t - \xi}{\Phi^{-1}(\gamma)}.$$

Portanto, conclui-se que:

$$\left\| \Sigma_t^{\frac{1}{2}} \mathbf{x}_3 \right\| \leq \frac{\mathbf{x}_3^\top \boldsymbol{\mu}_t - \xi}{\Phi^{-1}(\gamma)},$$

provando o resultado de interesse. \square

Já usando a formulação epigráfica, o problema pode ser reescrito como:

$$\min_{\alpha, \mathbf{x}_t} \alpha \text{ sujeito a } \mathbf{A}\mathbf{x}_t \leq \mathbf{b}, \|\mathbf{x}_t - \mathbf{x}_t^*\| \leq \alpha, \left\| \Sigma_t^{\frac{1}{2}} \mathbf{x}_t \right\| \leq \frac{\mathbf{x}_t^\top \boldsymbol{\mu}_t - \xi}{\Phi^{-1}(\gamma)} \quad (4.4)$$

onde $\alpha \in \mathbb{R}$, $\mathbf{x}_t \in \mathbb{R}^n$, ξ é o VaR dado, e γ é o quantil desejado (1%, 5%, etc.). Tal problema pode ser resolvido de forma eficiente pois é convexo e se trata de uma programação cônica de segunda ordem - *Second-Order Cone Programming* (SOCP).

Para resolver o problema de otimização com as restrições de VaR, propõe-se o uso do Algoritmo 4.

Algoritmo 4: Algoritmo Projeção OGD com restrições VaR Normal

Input: $OGD(\eta, \gamma)$

início $\mathbf{b}_1 \leftarrow \frac{1}{n}$

for $t = 2$ **to** T **do**

$$\left| \begin{array}{l} \mathbf{b}_t = \Pi_{S_t, \boldsymbol{\mu}_t, \gamma} \left(\mathbf{b}_{t-1} + \eta \frac{r_{t-1}}{\mathbf{b}_{t-1}^\top r_{t-1}} \right) \\ \text{Onde } \Pi_{S_t, \boldsymbol{\mu}_t}(\mathbf{y}, \gamma) = \arg \min_{\mathbf{x}_t} \|\mathbf{x}_t - \mathbf{y}\|^2 \text{ sujeito a } \mathbf{A}\mathbf{x}_t \leq \\ \mathbf{b}, \left\| \Sigma_t^{\frac{1}{2}} \mathbf{x}_t \right\| \leq \frac{\mathbf{x}_t^\top \boldsymbol{\mu}_t - \xi}{\Phi^{-1}(\gamma)} \end{array} \right.$$

end

retorna $\mathbf{b}_{(t+1)}$

fm

O funcionamento do algoritmo pode ser descrito da seguinte forma: (i) o algoritmo recebe como entrada os parâmetros η (taxa de aprendizado) e γ (nível de confiança do VaR); (ii) a cada iteração t , o algoritmo atualiza a alocação do portfólio \mathbf{b}_t utilizando a projeção no conjunto viável S_t , que incorpora as restrições lineares e a restrição de VaR. Em suma, a projeção é realizada resolvendo um problema de otimização quadrática com restrições cônicas de segunda ordem, garantindo que a alocação resultante esteja dentro dos limites de risco estabelecidos pelo VaR.

Já sob a hipótese que as carteiras são distribuídas como uma t de Student com variância que depende dos ν graus de liberdade (ver equação 3.18) $\mathbf{x}^\top \Sigma_{\nu,t} \mathbf{x}$ e média $\mathbf{x}_t^\top \boldsymbol{\mu}_t$, pode-se usar a inversa da função de distribuição acumulada de uma t de Student (\mathcal{T}_ν) ao computar o quantil para obter modelos que levam em consideração a possibilidade de caudas mais pesadas para os retornos. Neste caso, o Algoritmo 5 abaixo é visto. Note, porém, que ν acaba sendo considerado um parâmetro de entrada extra nesse caso.

Algoritmo 5: Algoritmo Projeção OGD com restrições VaR t de Student

Input: $OGD(\eta, \gamma, \nu)$

início $\mathbf{b}_1 \leftarrow \frac{1}{n}$

for $t = 2$ **to** T **do**

$$\left| \begin{array}{l} \mathbf{b}_t = \Pi_{S_t, \boldsymbol{\mu}_t, \gamma, \nu} \left(\mathbf{b}_{t-1} + \eta \frac{r_{t-1}}{\mathbf{b}_{t-1}^\top r_{t-1}} \right) \\ \text{Onde } \Pi_{S_t, \boldsymbol{\mu}_t}(\mathbf{y}, \gamma, \nu) = \arg \min_{\mathbf{x}_t} \|\mathbf{x}_t - \mathbf{y}\|^2 \text{ sujeito a } \mathbf{A} \mathbf{x}_t \leq \\ \mathbf{b}, \left\| \Sigma_{\nu,t}^{-\frac{1}{2}} \mathbf{x}_t \right\| \leq \frac{\mathbf{x}_t^\top \boldsymbol{\mu}_t - \xi}{\mathcal{T}_\nu^{-1}(\gamma)} \end{array} \right.$$

end

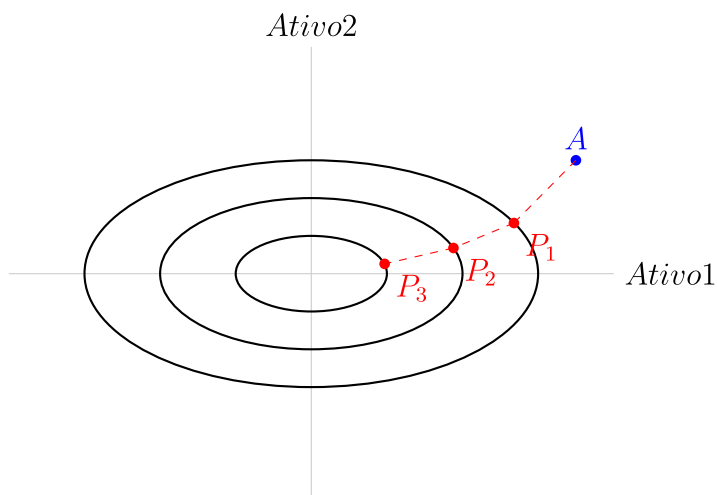
retorna $\mathbf{b}_{(t+1)}$

fim

Para ilustrar o funcionamento dos algoritmos de projeção de VaR, o Gráfico 4 apresenta uma representação simplificada considerando apenas dois ativos. As elipses concêntricas representam diferentes níveis de restrição de VaR, sendo a elipse mais externa (passando por $P1$) associada a um nível menos restritivo e a mais interna (passando por $P3$) a um nível mais restritivo. O ponto A , localizado fora das elipses, indica a alocação inicial do portfólio. A partir desse ponto, são traçadas projeções ortogonais (linhas tracejadas em vermelho) até os pontos $P1$, $P2$ e $P3$, que representam as alocações projetadas para cada nível de restrição de VaR. Portanto, o ponto A é o ponto de partida, antes de aplicar as restrições de risco do VaR. Ele representa a alocação original do portfólio que precisa ser ajustada para atender aos limites de risco estabelecidos pelas restrições cônicas de segunda ordem definidas pelas elipses. Essas projeções são obtidas resolvendo um problema de otimização quadrática com restrições cônicas de segunda ordem, conforme descrito no algoritmo. A curvatura das elipses depende da estrutura da matriz de covariância dos retornos dos ativos. Neste exemplo simplificado, o Ativo1

pode ser interpretado como um ativo livre de risco (por exemplo, atrelado à taxa Selic) e o Ativo2 como um ativo de risco (por exemplo, um ativo corporativo).

Gráfico 4 – Simplificação do modelo de projeção de VaR para dois ativos



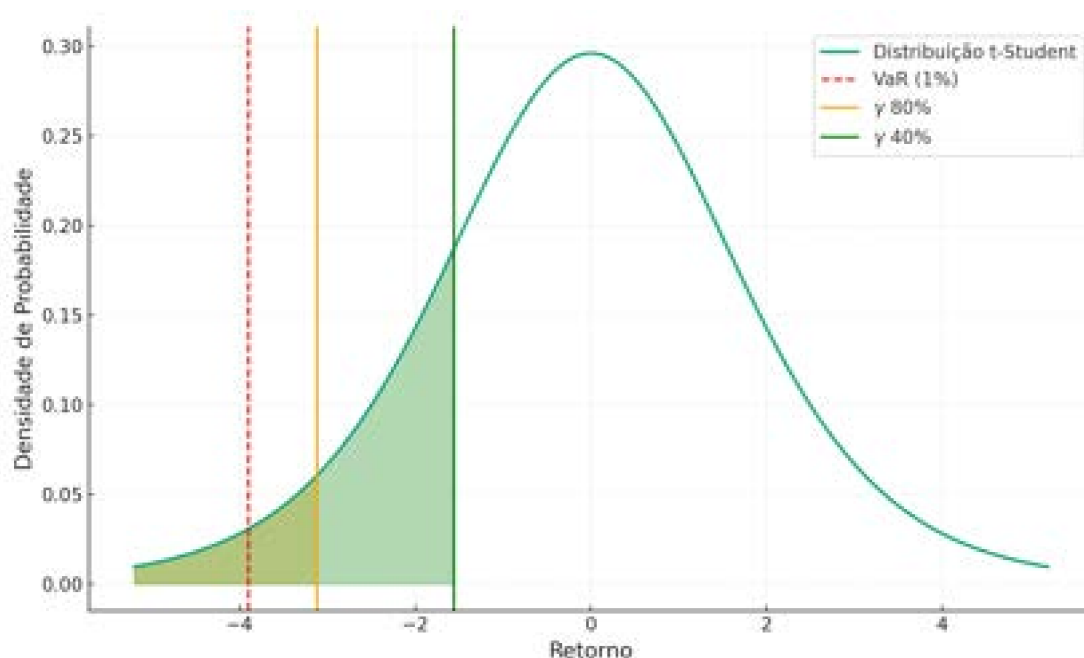
Fonte: Elaborado pelo autor (2024).

Para ilustrar como funcionam essas restrições em termos práticos, no Gráfico 5, plota-se a distribuição de retornos do portfólio hipotético, que neste caso segue uma distribuição t de Student. Essa escolha da distribuição permite capturar de forma mais realista a ocorrência de eventos extremos, uma vez que possui caudas mais pesadas em comparação com a distribuição normal.

A linha tracejada em vermelho indica o VaR de 1%, representando a perda máxima esperada com uma probabilidade de 1%. Essa é a restrição principal desejada perante o portfólio. Além do VaR, o gráfico ilustra diferentes níveis de γ (gamma), que são frações do VaR de 1%. Por exemplo, são destacados os níveis de 80% e 40% do VaR de 1%, representados pelas linhas verticais em amarelo e verde, respectivamente. É importante ressaltar que esses níveis de γ são exemplos ilustrativos e não regras fixas. Eles são utilizados para demonstrar como limites de risco adicionais podem ser incorporados na otimização do portfólio, permitindo ajustar a exposição ao risco de acordo com as preferências individuais do investidor. O princípio geral é que quanto menor o valor de γ escolhido, mais restritiva se torna a região de risco aceitável. No entanto, a escolha específica do valor de γ

deve ser baseada no perfil de risco individual do investidor e não em percentuais predeterminados.

Gráfico 5 – Representação Visual das Restrições de Risco no Projected OGD com VaR



Fonte: Elaborado pelo autor (2024).

O algoritmo 5, utilizado na implementação, busca encontrar a melhor combinação de ativos que satisfaça as restrições de risco impostas pelo VaR e pelo nível de γ selecionado, ao mesmo tempo em que maximiza o retorno esperado do portfólio. As áreas sombreadas abaixo de cada nível de γ destacam as regiões de risco aceitável de acordo com esses limites. Ao explorar diferentes valores de γ , é possível analisar uma variedade de perfis de risco e retorno, desde alocações mais conservadoras até abordagens mais agressivas. Essa flexibilidade permite aos investidores ajustar a exposição ao risco de acordo com suas preferências individuais, objetivos de investimento e tolerância ao risco específica. É fundamental entender que os níveis de γ apresentados são meramente ilustrativos, e cada investidor deve determinar seu próprio nível de γ com base em uma avaliação cuidadosa de sua situação financeira e objetivos pessoais.

Em termos práticos, na Seção 6.2, na qual são apresentados os resultados da

aplicação dessa técnica, utiliza-se somente o quantil de 1% para a restrição de VaR, uma vez que esse nível de confiança é amplamente adotado no mercado financeiro e pelas autoridades regulatórias. De fato, o uso do quantil de 1% permite capturar eventos extremos de perda e fornece uma estimativa mais conservadora do risco.

4.4 *CONDITIONAL VALUE AT RISK* COMO RESTRIÇÃO

Na Subseção 3.3.2 foi introduzido o conceito de *Conditional Value at Risk* (CVaR), uma medida de risco que supera algumas limitações do *Value at Risk* (VaR) e é amplamente utilizada para quantificar o risco de mercado em portfólios de investimento. Retomando esse tópico, é importante ressaltar que o CVaR oferece aos gestores de risco uma estimativa mais robusta da perda esperada para um portfólio específico, considerando não apenas a probabilidade de uma perda exceder um certo limite (como no VaR), mas também a magnitude esperada dessas perdas extremas. O CVaR é calculado como a média das perdas que excedem o VaR, fornecendo assim uma visão mais completa do risco de cauda. Essa métrica é essencial para o gerenciamento eficaz do risco, pois permite que as instituições financeiras, fundos de *hedge*, seguradoras e reguladores avaliem e monitorem a exposição ao risco de mercado de forma mais abrangente e precisa. Ao capturar melhor os eventos extremos, o CVaR oferece uma visão mais conservadora do potencial de perdas, auxiliando na tomada de decisões estratégicas e na alocação de capital em cenários de estresse.

Com isso em mente, inicialmente são enumeradas abaixo as premissas que serão utilizadas e discutidas ao longo dessa seção, permitindo a incorporação do CVaR no contexto de Otimização *Online Gradient Descent* (OGD).

- i O modelo assume que os retornos dos ativos em diferentes períodos são independentes e identicamente distribuídos (IID) com uma distribuição que pode ser adequadamente capturada pela formulação do CVaR. Essa suposição permite a formulação de um problema de otimização convexa, como demonstrado por Rockafellar e Uryasev [9].
- ii O uso do CVaR no modelo assume que essa métrica é suficiente para limitar

a exposição ao risco, conforme os objetivos do investidor, oferecendo uma visão mais completa do risco de cauda em comparação com o VaR.

- iii A formulação matemática do problema de otimização pressupõe que o problema é convexo, permitindo a aplicação de técnicas de programação convexa eficientes. A convexidade é crucial para garantir que o problema seja resolvível de maneira eficiente e que as soluções obtidas sejam globalmente ótimas dentro do conjunto de soluções viáveis.
- iv Supõe-se que a função de perda utilizada no cálculo do CVaR é convexa em relação às variáveis de decisão, garantindo a convexidade global do problema de otimização.
- v Pressupõe-se que a utilização de diferentes níveis de confiança (γ) associados ao CVaR é uma forma válida e eficaz de capturar diferentes perfis de risco, pois reflete o grau de aversão ao risco do investidor e é tratada como uma variável ajustável no modelo.
- vi Assume-se que o algoritmo de otimização utilizado é capaz de lidar eficientemente com a formulação do CVaR, que inclui uma expectativa condicional em sua definição.
- vii O modelo pressupõe que a estimativa empírica do CVaR baseada em dados históricos é uma aproximação adequada do CVaR teórico.

Estas premissas fornecem a base para a incorporação do CVaR como uma restrição de risco no problema de otimização de portfólio, oferecendo uma abordagem mais robusta e conservadora para o gerenciamento de risco em comparação com o uso do VaR.

4.4.1 Incorporação do CVaR como restrição no modelo de otimização

A utilização de medidas de risco é essencial para a construção de portfólios que buscam minimizar a probabilidade de perdas extremas. Entre as diversas métricas disponíveis, o *Value at Risk* (VaR) tem sido amplamente utilizado devido à sua simplicidade e capacidade de quantificar o risco de perdas dentro de um nível

de confiança especificado. No entanto, o VaR apresenta limitações importantes, como a falta de subaditividade e a dificuldade em sua otimização. Para superar essas limitações, foi introduzido o *Conditional Value at Risk* (CVaR), que considera não apenas a probabilidade de exceder uma certa perda, mas também a magnitude das perdas esperadas além desse limite.

Rockafellar e Uryasev [9] demonstraram que a minimização do CVaR pode ser alcançada sem a necessidade de calcular previamente o VaR. Esse resultado é obtido através da minimização da função convexa $F_\gamma(\mathbf{x}_t, \zeta)$, onde ζ é uma variável auxiliar que representa o VaR, γ é o nível de confiança e $F_\gamma(\mathbf{x}_t, \zeta)$ é definida como:

$$F_\gamma(\mathbf{x}_t, \zeta) = \zeta + \frac{1}{\gamma} \mathbb{E}[(-\mathbf{x}_t^\top r_t - \zeta)^+],$$

onde $(\mathbf{x}_t^\top r_t - \zeta)^+$ representa as perdas que excedem o VaR. A função $F_\gamma(\mathbf{x}_t, \zeta)$ é convexa e continuamente diferenciável em relação a ζ , o que permite que o CVaR seja determinado diretamente pela minimização desta função:

$$\text{CVaR}_\gamma(\mathbf{x}_t) = \min_{\zeta \in \mathbb{R}} F_\gamma(\mathbf{x}_t, \zeta).$$

Esse método permite que o CVaR seja minimizado sem a necessidade de calcular o VaR de forma explícita, simplificando o problema de otimização e garantindo que a solução obtida seja globalmente ótima devido à convexidade da função $F_\gamma(\mathbf{x}_t, \zeta)$. Essa abordagem é especialmente útil na construção de portfólios robustos, onde o controle rigoroso do risco é essencial.

Neste trabalho, propõe-se a utilização do CVaR como restrição para controle do risco do portfólio, em substituição ao VaR. Para uma distribuição de retornos dados por uma variável aleatória integrável X , [9] demonstrou que o CVaR pode ser formulado como o seguinte problema de otimização convexo:

$$\text{CVaR}_\gamma(X) = \min_{\zeta} \left[\zeta + \frac{1}{1-\gamma} \mathbb{E}[(-X - \zeta)^+] \right]$$

Onde:

- ζ é uma variável auxiliar que representa o VaR;
- $(X - \zeta)^+$ representa as perdas que ultrapassam o VaR (ou seja, as que contribuem para o cálculo do CVaR).

Quando o CVaR é utilizado no lugar da restrição cônica, o problema de otimização passa a ser:

$$\begin{aligned}
& \min_{\zeta, \mathbf{x}_t, \alpha} \quad \alpha \\
& \text{sujeito a} \quad \mathbf{A}\mathbf{x}_t \leq \mathbf{b}, \\
& \quad \|\mathbf{x}_t - \mathbf{x}_t^*\| \leq \alpha, \\
& \quad \zeta + \frac{1}{1-\gamma} \mathbb{E}[(-\mathbf{x}_t^\top r_t - \zeta)^+] \geq \xi_{\text{cvar}},
\end{aligned} \tag{4.5}$$

onde $\zeta \in \mathbb{R}$, $\mathbf{x}_t \in \mathbb{R}^n$ e ξ_{cvar} é o CVaR máximo ligado ao nível de confiança γ . Note que $\mathbb{E}[(-\mathbf{x}_t^\top r_t - \zeta)^+]$ é convexa, pois é uma composição de uma função convexa não decrescente, $(\cdot)^+$, com uma função linear das variáveis de decisão \mathbf{x}_t e ζ . Nesse caso, o conjunto de soluções viáveis permanece convexo, visto que se está minimizando o CVaR do portfólio em relação ao nível de tolerância ξ_{cvar} .

Para transformar a formulação em um problema quadrático com restrições lineares, considera-se a aproximação da função $\mathbb{E}[(-\mathbf{x}_t^\top r_t - \zeta)^+]$ por sua contrapartida empírica. Nesse caso, substitui-se a esperança por uma média empírica. Denote-se a amostragem empírica dos retornos r_t por y_k , $k = 1, \dots, q$. Tem-se que:

$$\mathbb{E}[(-\mathbf{x}_t^\top r_t - \zeta)^+] \approx \frac{1}{q} \sum_{k=1}^q [(-\mathbf{x}_t^\top y_k - \zeta)^+]. \tag{4.6}$$

O segundo passo é a introdução de variáveis de folga u_k que representam a função $(\cdot)^+$, de forma que $u_k = (-\mathbf{x}_t^\top y_k - \zeta)^+$, e naturalmente u_k precisa satisfazer $u_k \geq 0$ e $u_k + \mathbf{x}_t^\top y_k + \zeta \geq 0$. Isso leva à seguinte formulação:

$$\begin{aligned}
& \min_{\alpha, \mathbf{x}_t} \alpha \\
& \text{sujeito a } \mathbf{A}\mathbf{x}_t \leq \mathbf{b}, \\
& \quad \|\mathbf{x}_t - \mathbf{x}_t^*\| \leq \alpha, \\
& \quad \min_{\zeta, \mathbf{u}} \zeta + \frac{1}{q\gamma} \sum_{k=1}^q u_k \geq \xi_{\text{cvar}}, \\
& \quad u_k + \mathbf{x}_t^\top \mathbf{y}_k + \zeta \geq 0, \\
& \quad u_k \geq 0.
\end{aligned} \tag{4.7}$$

Note que o problema acima ainda não pode ser resolvido de forma eficiente, devido ao problema de otimização presente na restrição. Infelizmente, ao computar o dual do problema interno e juntar as variáveis do problema de otimização leva a uma formulação não linear, que introduz um fator multiplicativo entre x_t e as variáveis duais do problema interno. Portanto, a abordagem utilizada aqui será resolver o problema da (4.7) através do método de gerações de restrições. Note que o problema (4.7) pode ser reescrito como

$$\begin{aligned}
& \min_{\alpha, \mathbf{x}_t} \alpha \\
& \text{sujeito a } \mathbf{A}\mathbf{x}_t \leq \mathbf{b}, \\
& \quad \|\mathbf{x}_t - \mathbf{x}_t^*\| \leq \alpha, \\
& \quad \zeta + \frac{1}{q\gamma} \sum_{k=1}^q u_k \geq \xi_{\text{cvar}}, \quad \forall \zeta, \mathbf{u} \geq 0, \\
& \quad u_k + \mathbf{x}_t^\top \mathbf{y}_k + \zeta \geq 0, \quad \forall \zeta, u_k \geq 0, \quad k = 1, \dots, q.
\end{aligned} \tag{4.8}$$

Naturalmente, não se pode escrever as restrições para todo $\zeta, \mathbf{u} \geq 0$, mas é possível utilizar o método de geração de restrições. Começa-se adicionando a restrição somente para $\mathbf{u} = 0$ e $\zeta = \max\{0, -\min_k \mathbf{x}_t^\top \mathbf{y}_k\}$, que sempre são valores viáveis para ζ, \mathbf{u} . Resolve-se o problema (4.8) com as restrições ativas para somente esses valores de ζ, \mathbf{u} ao invés de todos os valores possíveis. Note que isso é um problema quadrático com restrições lineares que pode ser resolvido de maneira eficiente. Chame a solução ótima obtida de x_{aux} . Depois resolve-se o problema do CVaR original (4.7) com x_{aux} . Ou seja, resolve-se o seguinte problema de Programação Linear:

$$\begin{aligned}
& \min_{\zeta, \mathbf{u}} \quad \zeta + \frac{1}{q\gamma} \sum_{k=1}^q u_k \\
& \text{sujeito a} \quad u_k + \mathbf{x}_{aux}^\top \mathbf{y}_k + \zeta \geq 0, \quad k = 1, \dots, q, \\
& \quad \quad \quad u_k \geq 0, \quad k = 1, \dots, q.
\end{aligned} \tag{4.9}$$

Após obter o valor ótimo, verifica-se se este é maior ou igual que ξ_{cvar} . Se for, então precisa ser o caso em que $\mathbf{x}_t = \mathbf{x}_{aux}$, já que o último é viável e precisa ser ótimo para (4.7), visto que essa é uma versão mais restritiva do que (4.8). Se não for, adiciona-se o novo valor de ζ e \mathbf{u} obtidos no passo anterior como restrições adicionais em (4.8). Esse procedimento é repetido até a convergência.

5 DADOS E APLICAÇÃO EMPÍRICA

Para avaliar o desempenho dos algoritmos propostos, foram utilizados dados do mercado de ações brasileiro. Esses dados foram obtidos a partir das empresas que compunham os portfólios hipotéticos dos índices Bovespa (IBOV) e *Small Cap* (I_SMALL), relativo ao índice *SMLL B3*, e foram coletados por meio do *Yahoo Finance*, utilizando a biblioteca *Pandas DataReader* ([121]) do *Python* 3.11 ([122]).

Ao construir essa base de dados, procurou-se evitar um problema comum em análises financeiras conhecido como "enviesamento de sobrevivência" (*survivorship bias*). Esse enviesamento ocorre quando os *backtests* de estratégias de investimento consideram apenas as empresas que permaneceram ativas no mercado até o presente, excluindo aquelas que falharam ou foram retiradas da bolsa de valores ao longo do período analisado.

Esse tipo de enviesamento pode levar a uma superestimação do desempenho potencial de uma estratégia de investimento, pois ao ignorar as empresas que foram retiradas do mercado, a análise exclui casos de falhas e perdas que teriam impactado negativamente o retorno global. Em outras palavras, ao analisar apenas as "sobreviventes", presume-se incorretamente que a estratégia teria evitado investir em empresas que faliram ou tiveram desempenho ruim, o que não necessariamente seria o caso. Por essa razão, o banco de dados do IBOV consiste em observações dos retornos das ações de 59 empresas. Esse número inclui empresas que já não fazem mais parte do índice, bem como empresas que foram incluídas posteriormente, abrangendo o período de 02 de janeiro de 1998 até 31 de agosto de 2023.

Quanto ao índice *Small Caps* (I_SMALL), não havia informações disponíveis desde o início do período. Portanto, coletamos dados de 64 empresas, abrangendo o período de 05 de janeiro de 2009 a 31 de dezembro de 2023.

O IBOV é o índice mais tradicional do mercado brasileiro, sendo composto por ações de empresas com alto volume de negociação. Por outro lado, o índice *Small Cap* é composto por ações de empresas de baixa capitalização, caracterizadas por baixo volume de negociação e, conseqüentemente, menor liquidez.

Em relação aos resultados apresentados na seção 6.1, focada no beta de

mercado variante no tempo e considerando as ações do IBOV e I_SMALL, adotou-se a seguinte abordagem, usada na avaliação de portfólios em outros artigos da literatura como em [123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135]:

- **Taxa livre de risco:** Utilizo-se a média aritmética ponderada dos rendimentos intrínsecos dos títulos da Dívida Pública Nacional emitidos pelo Tesouro Nacional e detidos pelo Banco Central do Brasil.
- **Retorno de mercado:** Para as ações listadas no *SMLL B3* (I_SMALL), o IBOV foi considerado como o indicador de retorno de mercado. Já para as ações pertencentes ao IBOV, usou-se o *IBrX-100* como referência.
- **Iniciação do algoritmo:** O algoritmo é inicializado com uma carteira inicial *naive* de $\frac{1}{n}$.
- **Período usado para I_Small:** Foram utilizados dados diários de 01 de maio de 2009 até 28 de dezembro de 2017.
- **Período usado para IBOV:** Nesse caso, foram utilizados dados diários de 02 de janeiro de 1998 até 28 de dezembro de 2017.

Já nas seções 6.2 e 6.3, que levam em conta o VaR e o CVaR, respectivamente, há algumas diferenças: (i) adoção da taxa Selic fator diário como *proxy* de retorno de mercado no caso da avaliação focada em dados IBOV; (ii) período usado para os resultados I_Small foi de 06 de janeiro de 2009 até 30 de agosto de 2023; (iii) período usado para os resultados IBOV foi de 05 de julho de 2002 até 30 de agosto de 2023; (iv) um conjunto inicial de dados que corresponde aos 2,5% primeiros retornos observados para as ações do I_Small (10% no caso IBOV), definidos pela quantidade de ações e dados presentes em tais índices,⁷ foram utilizado para a estimação inicial da matriz covariância das carteiras, permitindo, assim, uma carteira inicial já ajustada para os algoritmos.

⁷ Diferentes tamanhos amostrais iniciais foram testados, gerando resultados similares.

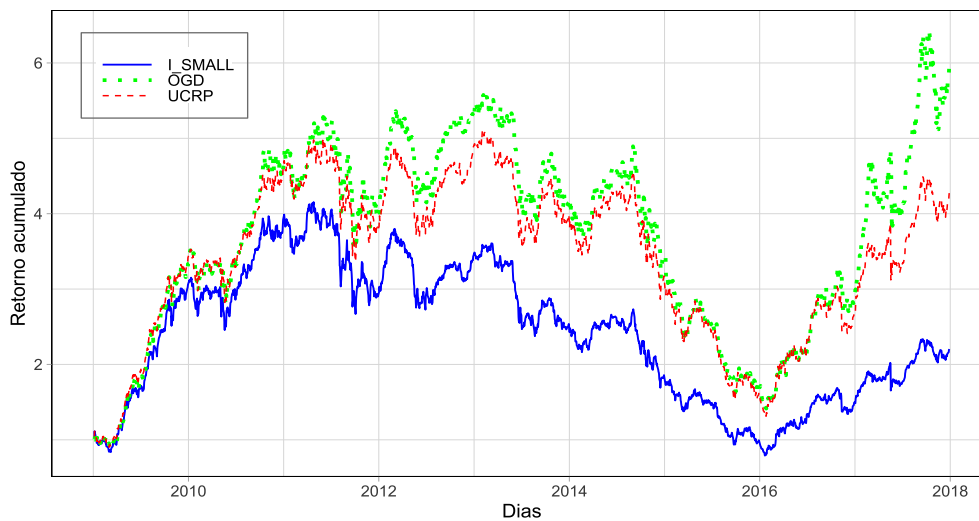
6 RESULTADOS

Esta seção está dividida em 4 partes. Inicialmente, na seção 6.1, são apresentados os resultados do modelo OGD considerando betas variantes no tempo, introduzido na subseção 4.2, com comparações a modelos *benchmarks* como o OGD padrão e o UCRP, tanto para portfólios focados na I_SMALL como na IBOV. Na segunda parte, seção 6.2, uma análise similar é realizada, mas baseada no modelo OGD com restrição VaR considerando tanto retornos com distribuição normal como com distribuição t de Student, conforme detalhado na seção 4.3. Depois, na seção 6.3, o enfoque é no modelo OGD com restrição CVaR mostrado na seção 4.4, mostrando sua superioridade, em particular, na comparação ao modelos OGD Padrão e VaR tanto para dados I_SMALL como para aqueles focados no IBOV. Por fim, na seção 6.4, há um apanhado geral dos resultados mais importantes das três metodologias propostas e uma sucinta comparação com outros trabalhos dentro do escopo de aprendizado de máquina com uso de dados I_SMALL e/ou IBOV.

6.1 RESULTADOS DO OGD COM BETA VARIANTE NO TEMPO

Como uma análise inicial, testou-se o desempenho do algoritmo OGD sem controlar o beta do portfólio e o mesmo é comparado com o índice de *Small Caps* (I_SMALL) e o CRP uniforme (UCRP). Dentre os conjuntos de possíveis ativos, apenas os ativos que pertencem ao índice e o ativo livre de risco estão disponíveis. Da mesma forma, o portfólio UCRP é construído fixando um peso uniforme para cada ativo que pertence ao índice de *Small Caps*, incluindo o ativo livre de risco. Os resultados são mostrados no Gráfico 6.

Gráfico 6 – Resultados do modelo OGD sem controlar o beta - comparação ao índice *Small Caps* e UCRP



(a) Comparação entre o portfólio OGD, índice de *Small Caps* e UCRP



(b) Variação no tempo de beta do portfólio OGD.

Fonte: Elaborado pelo autor (2024).

È possível observar que o algoritmo OGD superou tanto o índice de *Small Caps* quanto o UCRP, enquanto o índice teve o pior desempenho. No entanto, na maioria das vezes, o desempenho do algoritmo OGD é semelhante ao desempenho do UCRP. Além disso, é possível evidenciar que o beta do portfólio está correlacionado com o índice IBrX-100 usado como uma aproximação para o retorno do mercado, variando entre aproximadamente 0,5 e 1,5 durante o período do investimento.

Em termos de métricas de risco, a comparação entre o VaR e o CVaR do portfólio OGD e os *benchmarks* é visto na Tabela 1. O portfólio OGD e o índice de *Small Caps* têm um perfil de risco semelhante, ambos maiores que o portfólio UCRP.

Tabela 1 – Comparação entre métricas de risco para o portfólio OGD sem controlar o beta, índice *Small Caps* e UCRP

	OGD		I_SMALL		UCRP	
	1%	5%	1%	5%	1%	5%
VaR	-0,047	-0,029	-0,048	-0,029	-0,045	-0,028
CVaR	-0,071	-0,042	-0,071	-0,042	-0,067	-0,040

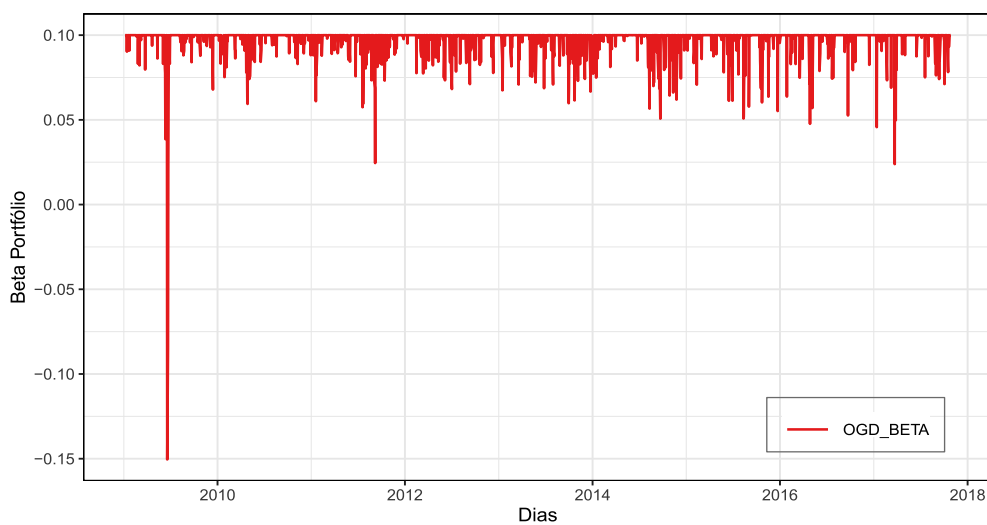
Fonte: Elaborado pelo autor (2024).

No Gráfico 7, enumera-se os resultados para o algoritmo OGD quando controlamos o beta do portfólio. Nesta imagem, o beta foi escolhido para estar entre -0,3 e 0,1, forçando na maioria das vezes uma posição contrária ao índice de *Small Caps*. Como resultado, especialmente entre 2010 e 2012, que é um período de alta valorização para I_SMALL, o algoritmo OGD tomou posições contrárias ao mercado e teve dificuldades para obter ganhos enquanto os valores do índice aumentam rapidamente. No entanto, durante o período de perda entre 2012 e 2016, o portfólio OGD que é contra o mercado é capaz de manter ganhos e tem um retorno acumulado maior para o período total.

Gráfico 7 – Resultados do modelo OGD com beta entre -0,3 e 0,1 - comparação ao índice *Small Caps*



(a) Comparação entre o portfólio OGD com beta entre -0,3 e 0,1 e o índice de *Small Caps*.



(b) Variação no tempo de beta do portfólio OGD com limites.

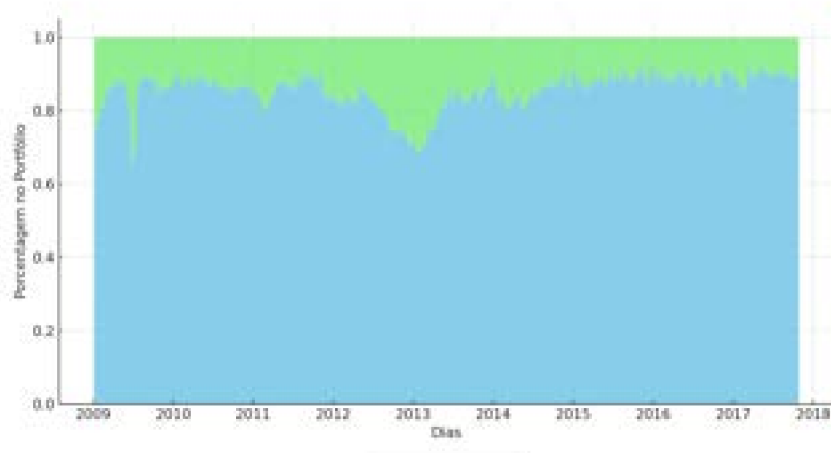
Fonte: Elaborado pelo autor (2024).

Analisando o comportamento do beta na parte b do Gráfico 7, fica claro que a maior parte do tempo o beta é de fato positivo. Uma vez que não se permite que o algoritmo assuma posições vendidas, é difícil construir um portfólio com beta

negativo. A razão disso é que os ativos que pertencem ao índice de *Small Caps* têm uma correlação positiva com o mercado na maioria das vezes, tornando o conjunto de possíveis betas \mathcal{B}_β inviável para alguns períodos se exigirmos beta estritamente negativos.

Portanto, para reduzir a correlação positiva com o mercado, o algoritmo investe a maior parte do dinheiro no ativo livre de risco, como visto no Gráfico 8. Conjectura-se que ao permitir que o algoritmo assuma posições vendidas, deve proporcionar um alcance de resultados ainda melhores durante os mercados de baixa.

Gráfico 8 – Evolução do portfólio OGD_beta - ativos pertencentes ao índice *Small Caps* e/ou livre de risco



Fonte: Elaborado pelo autor (2024).

A Tabela 2 mostra que as métricas de risco são substancialmente menores do que o índice de *Small Caps*, embora os retornos também tenham sido maiores para o período total. Esse resultado é esperado devido à alta concentração do ativo livre de risco no portfólio.

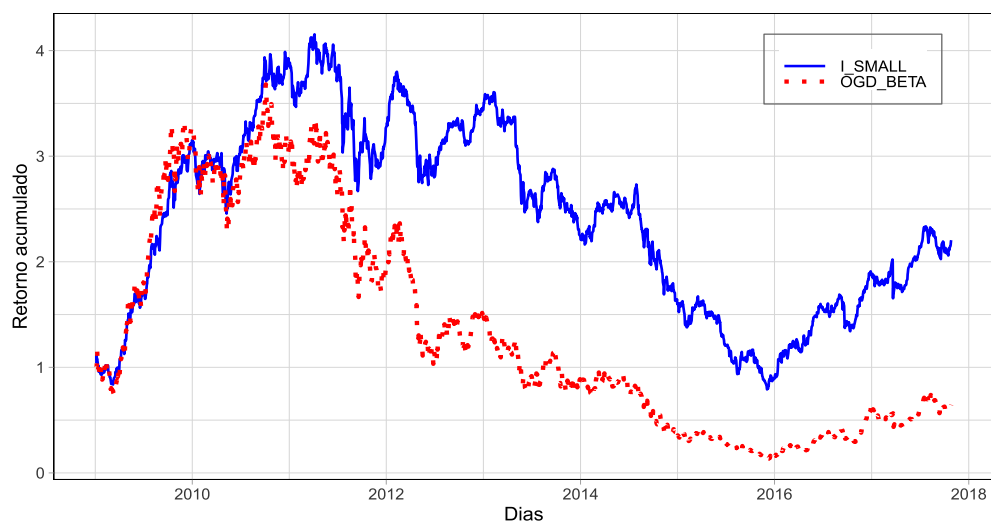
Tabela 2 – Comparação entre métricas de risco para o OGD com beta entre -0,3 e 0,1.

	OGD_beta		I_SMALL	
	1%	5%	1%	5%
VaR	-0,007	-0,004	-0,049	-0,029
CVaR	-0,014	-0,007	-0,071	-0,042

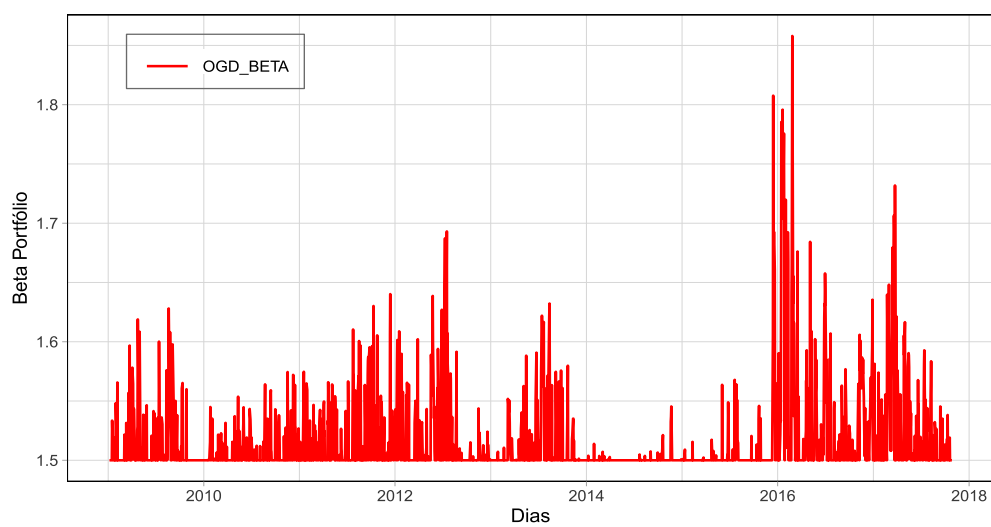
Fonte: Elaborado pelo autor (2024).

Por outro lado, ao forçar uma correlação positiva com o mercado, é possível acompanhar os períodos de crescimento do índice, mas quando o índice cai, o portfólio OGD com alto beta positivo não consegue se adaptar e sofre com o *Bear Market*, como visto na parte (a) do Gráfico 9. Isso sugere que, para obter ganhos substanciais em relação ao mercado, é importante adaptar o intervalo aceito para o beta do portfólio para capturar tanto os períodos de alta quanto os períodos de baixa.

Gráfico 9 – Resultados do modelo OGD com beta entre 1,5 e 3 - comparação ao índice *Small Caps*



(a) Comparação entre o portfólio OGD com beta entre 1.5 e 3 e o índice de *Small Caps*.



(b) Variação no tempo de beta do portfólio OGD com limites.

Fonte: Elaborado pelo autor (2024).

Não menos importante, dado que se impõe um beta maior que um, é possível também esperar um aumento no risco do portfólio, o que de fato ocorreu como evidenciado na Tabela 3.

Tabela 3 – Comparação entre métricas de risco para o OGD com beta entre 1.5 e 3.

	OGD_beta		I_SMALL	
	1%	5%	1%	5%
VaR	-0,077	-0,047	-0,048	-0,029
CVaR	-0,103	-0,067	-0,071	-0,042

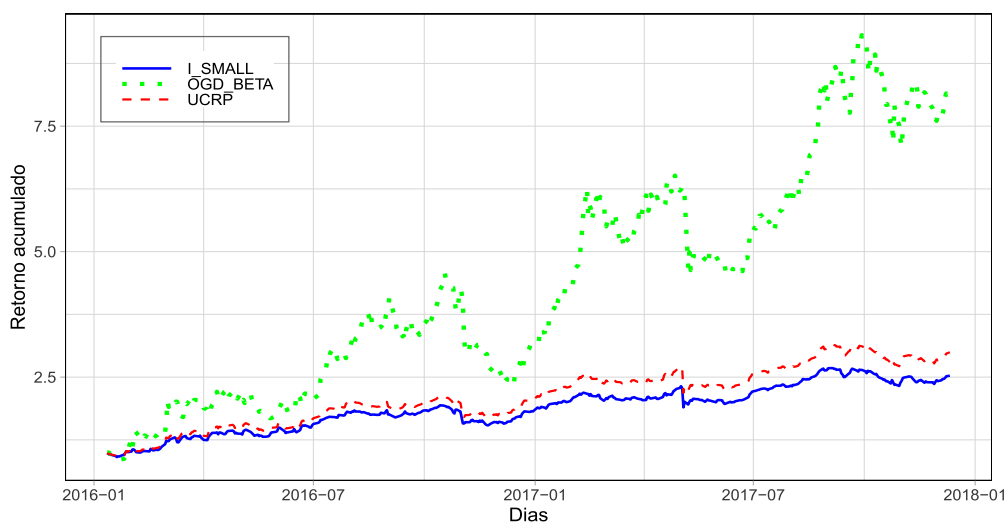
Fonte: Elaborado pelo autor (2024).

6.1.1 OGD com beta variante no tempo para dados *Small Caps* - Resultados para períodos específicos de tempo

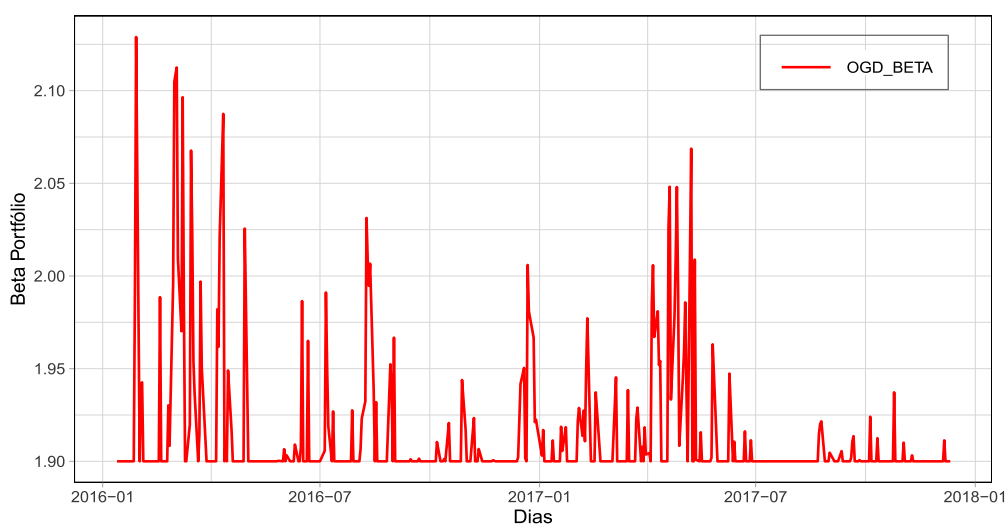
Nesta subseção, será observado que uma escolha adequada do intervalo de beta leva a melhorias no desempenho em períodos de mercado em alta ou em queda.

Na parte (a) do Gráfico 10 abaixo, é observado que o retorno acumulado do OGD ao forçar um beta maior que o mercado leva a melhorias tanto em relação ao índice de *Small Caps* quanto ao *benchmark* UCRP. Este resultado é consistente com a correlação positiva entre o índice I_SMALL e o retorno de mercado usado como referência (IBrX-100) e mostra que um beta maior que um de fato alavanca os retornos do mercado em um mercado em alta.

Gráfico 10 – Resultados do modelo OGD com beta entre 1,9 e 3 - comparação ao índice *Small Caps* e UCRP



(a) Comparação entre o portfólio OGD com beta entre 1,9 e 3, índice de *Small Caps* e UCRP.



(b) Variação no tempo de beta do portfólio OGD com limites.

Fonte: Elaborado pelo autor (2024).

Forçar o beta do portfólio a ser alto durante a etapa de projeção acelera o processo de investimento em empresas mais arriscadas, mas correlacionadas com o mercado; permitindo, assim, que o portfólio alcance retornos maiores do que o portfólio OGD sem controle do beta.

Uma vez que durante todo o tempo o portfólio OGD tem um beta entre 0,5 e 1,5 - parte (b) do Gráfico 6, observa-se que o passo de projeção tende a encontrar portfólios no limite inferior do permitido - parte (b) do Gráfico 10 - e essa alavancagem em relação ao retorno do mercado de fato foi vantajosa durante o mercado em alta, destacando também que o beta variável no tempo para cada ativo de fato capta a correlação com o mercado e é significativo para prever a correlação futura.

De acordo com a teoria CAPM, se o portfólio é uma função linear do retorno do mercado com coeficiente angular β , então, a variância do portfólio, deve ser β^2 vezes a variância do mercado. Assim, dado que é forçando um alto valor absoluto para o beta do portfólio, é esperado um aumento nas métricas de risco utilizadas. Este resultado é exposto na Tabela 4.

Tabela 4 – Comparação entre as métricas de risco para o OGD com beta entre 1,9 e 3.

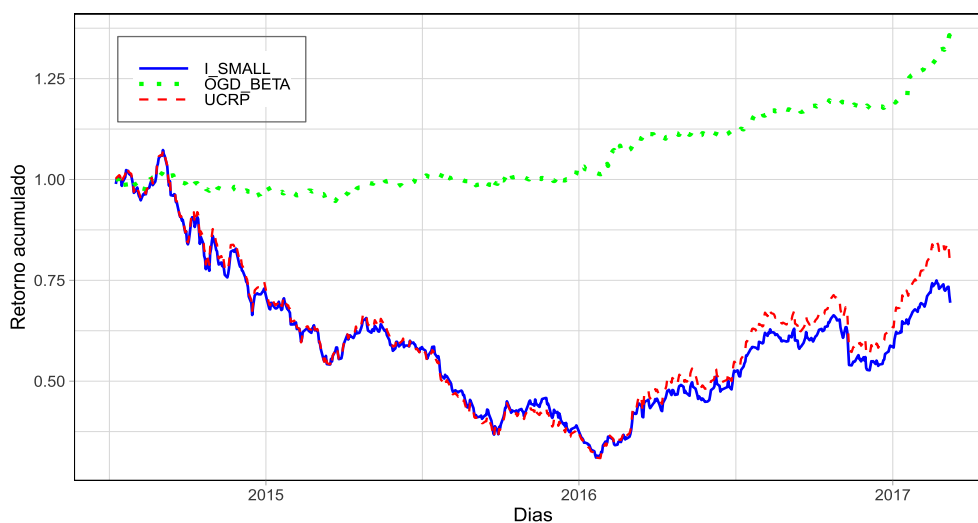
	OGD_beta		I_SMALL		UCRP	
	1%	5%	1%	5%	1%	5%
VaR	-0,095	-0,053	-0,049	-0,024	-0,048	-0,024
CVaR	-0,184	-0,091	-0,100	-0,045	-0,103	-0,045

Fonte: Elaborado pelo autor (2024).

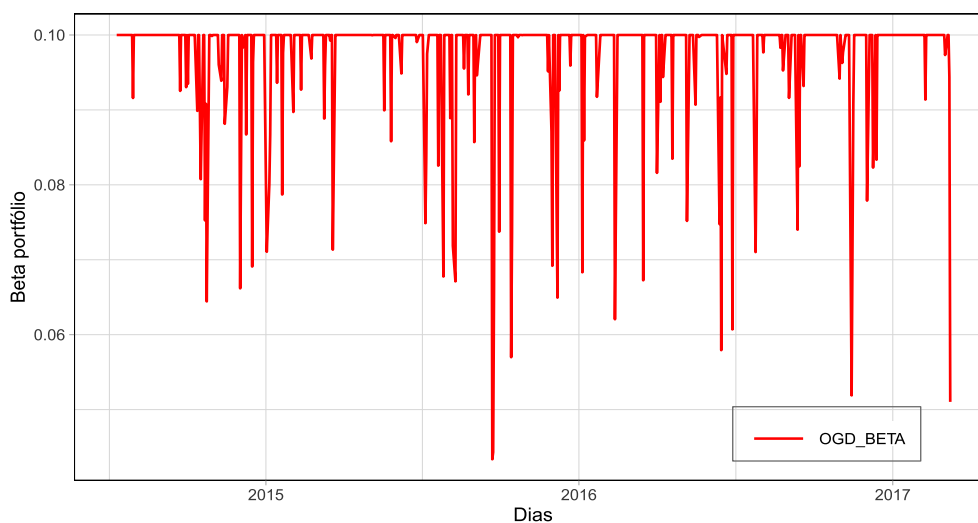
Na parte (a) do Gráfico 11, foi forçada uma correlação negativa com o mercado. Uma vez que o índice I_SMALL e IBOV têm uma correlação positiva, espera-se que o portfólio OGD seja capaz de mover a maior parte do investimento para o ativo livre de risco, evitando o mercado em queda.

Além disso, como discutido anteriormente, uma vez que o algoritmo não pode assumir posições vendidas, para satisfazer as restrições da etapa de projeção do Algoritmo 3, o algoritmo tem que acelerar o processo de sair de posições no mercado e aumentar a posição no ativo livre de risco, reduzindo o risco e evitando o *Bear Market*. Esse comportamento é refletido no beta do portfólio, que tende a zero, conforme mostrado na parte (b) do Gráfico 11.

Gráfico 11 – Resultados do modelo OGD com beta entre -0,3 e 0,1 - comparação ao índice *Small Caps* e UCRP



(a) Comparação entre o portfólio OGD com beta entre -3 e 0,1, índice de *Small Caps* e UCRP.



(b) Variação no tempo de beta do portfólio OGD com limites.

Fonte: Elaborado pelo autor (2024).

Em termos de risco, conforme exposto na Tabela 5, um resultado semelhante foi visto na comparação ao apresentado para o período inteiro. A alocação no ativo livre de risco reduziu consideravelmente o risco do portfólio, superando o índice de *Small Caps* e o portfólio UCRP tanto em retornos quanto em métricas de risco por

várias vezes.

Tabela 5 – Comparação entre métricas de risco para o OGD com beta entre -0,3 e 0,1.

	OGD_beta		I_SMALL		UCRP	
	1%	5%	1%	5%	1%	5%
VaR	-0,008	-0,005	-0,048	-0,034	-0,048	-0,032
CVaR	-0,014	-0,008	-0,072	-0,044	-0,073	-0,044

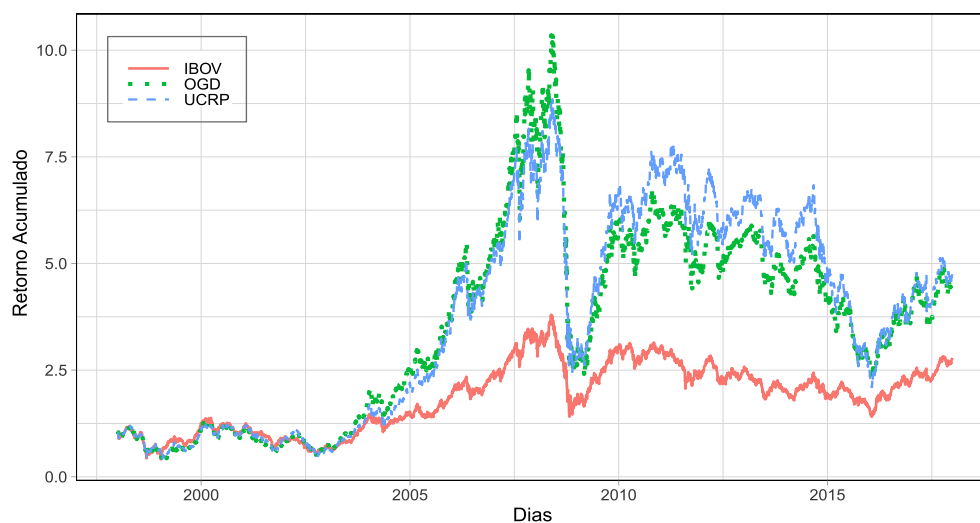
Fonte: Elaborado pelo autor (2024).

6.1.2 OGD com beta variante no tempo para dados IBOV - Resultados Gerais

A seguir, os resultados baseador nas ações do índice Bovespa (IBOV) como possíveis ativos são apresentados. Novamente, usou-se o UCRP e o próprio IBOV como referência; além de manter o IBrX-100 como uma *proxy* para o retorno do mercado.

Na parte (a) do Gráfico 12 é observado que tanto o portfólio UCRP quanto o OGD superaram o índice IBOV. Além disso, o portfólio UCRP teve um desempenho superior ao portfólio OGD. Na maior parte do tempo o portfólio OGD acompanha os retornos do UCRP, o que indica que o método não está sendo capaz de selecionar entre os ativos e, em média, o gradiente tem média zero em cada coordenada; levando, assim, a uma posição quase constante nos ativos disponíveis.

Gráfico 12 – Resultados do modelo OGD sem controlar o beta - comparação ao índice Bovespa e UCRP



(a) Comparação entre o portfólio OGD, Índice Bovespa e UCRP.



(b) Variação no tempo de beta do portfólio OGD.

Fonte: Elaborado pelo autor (2024).

Já na parte (b) do Gráfico 12, fica claro que o beta do portfólio OGD oscila aproximadamente entre 0,5 e 1,3. Em média, o beta do portfólio é próximo de um. Isso se reflete em um risco semelhante entre o portfólio OGD e o próprio mercado. Como o IBOV também é uma aproximação para o retorno do mercado, a Tabela 6

evidencia que o risco é semelhante entre o portfólio OGD e os *benchmarks*.

Tabela 6 – Comparação entre métricas de risco para o portfólio OGD sem controlar o beta índice Bovespa e UCRP

	OGD_beta		IBOV		UCRP	
	1%	5%	1%	5%	1%	5%
VaR	-0,055	-0,033	-0,054	-0,031	-0,058	-0,033
CVaR	-0,085	-0,050	-0,080	-0,047	-0,089	-0,051

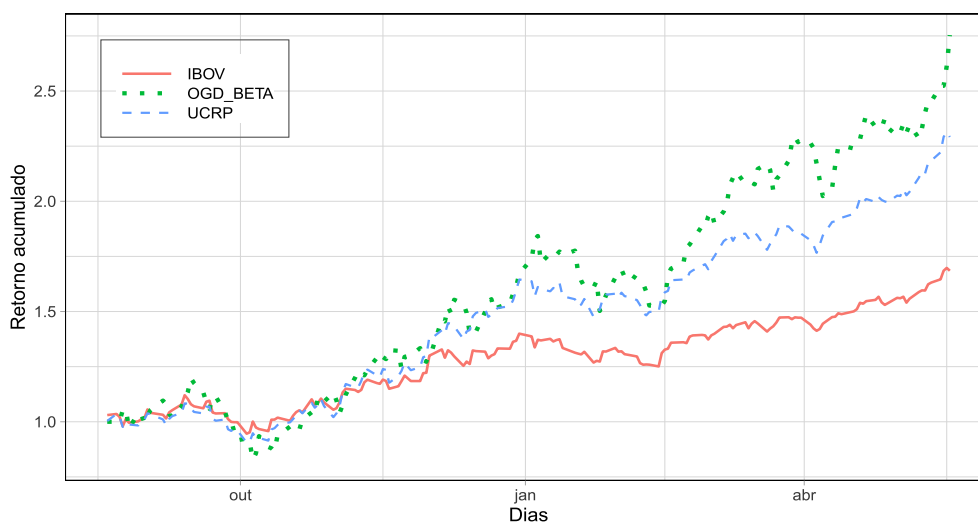
Fonte: Elaborado pelo autor (2024).

A seguir, avaliou-se o desempenho da abordagem proposta controlando o nível de risco sistemático do portfólio para ver se é possível superar o portfólio UCRP apenas com tal controle.

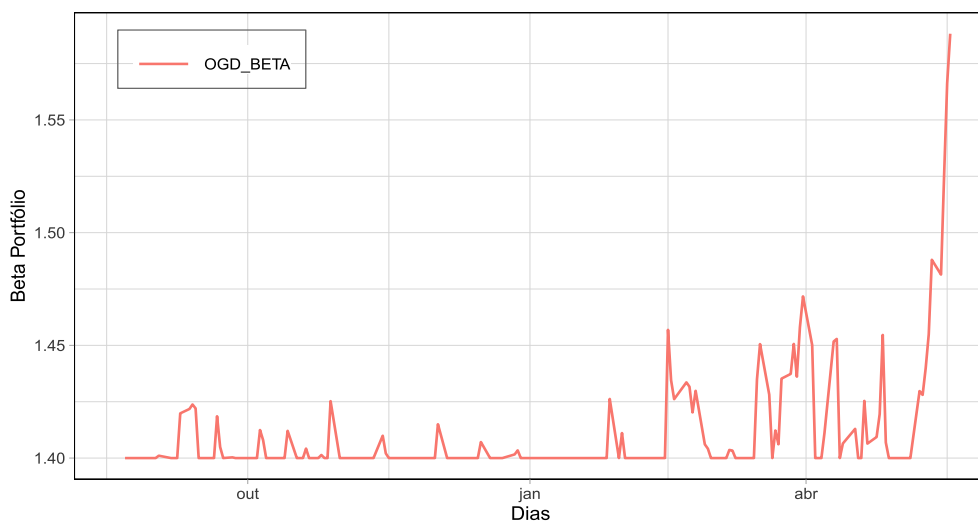
Na parte (a) do Gráfico 13, a eficácia em alavancar o retorno do mercado forçando um alto beta para o algoritmo OGD com controle do mesmo é vista. O algoritmo foi capaz de fornecer retornos maiores do que o UCRP ou o índice IBOV durante o mercado de alta e foi a estratégia dominante durante quase todo o período.

Como esperado, ao analisar a parte (b) do Gráfico 13, tem-se que a maior parte do tempo o passo projetado corresponde ao limite inferior do intervalo de beta usado para as restrições no portfólio.

Gráfico 13 – Resultados do modelo OGD com beta entre 1,4 e 3 - comparação ao índice Bovespa e UCRP



(a) Comparação entre portfólio OGD com beta entre 1,4 e 3, Índice Bovespa e UCRP.



(b) Variação no tempo de beta do portfólio OGD com limites.

Fonte: Elaborado pelo autor (2024).

Como exposto na Tabela 7, forçar o beta a ser superior a 1,4 levou a um aumento nas métricas de risco do portfólio. O portfólio OGD foi aproximadamente duas vezes mais arriscado do que os *benchmarks*. No entanto, como no caso das *Small Caps*, o portfólio apresentou retornos mais altos para todos os períodos no

período de alta selecionado.

Tabela 7 – Comparação entre métricas de risco para o OGD com beta entre 1,4 e 3.

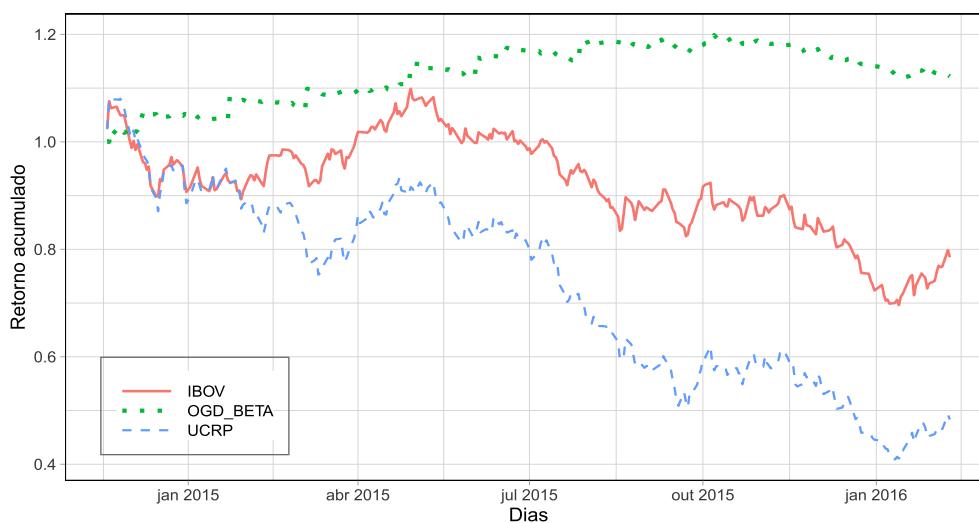
	OGD_beta		IBOV		UCRP	
	1%	5%	1%	5%	1%	5%
VaR	-0,085	-0,052	-0,043	-0,029	-0,052	-0,037
CVaR	-0,139	-0,079	-0,078	-0,043	-0,089	-0,054

Fonte: Elaborado pelo autor (2024).

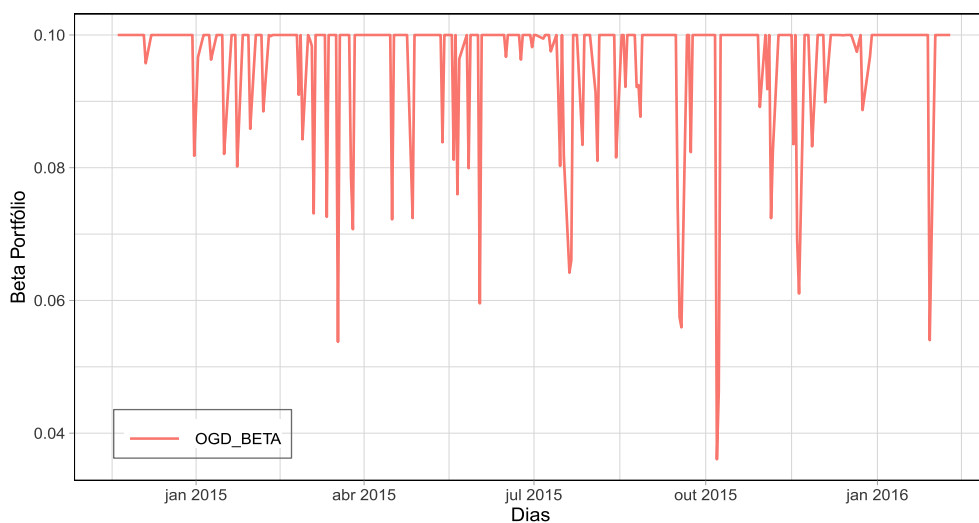
Durante o período de mercado em baixa, restringir o beta do portfólio entre -0,3 e 0,1 faz com que o portfólio evite perdas, investindo principalmente no ativo livre de risco, uma vez que posições vendidas não foram permitidas.

Os resultados com os respectivos betas são expostos no Gráfico 14 abaixo, comprovando a superioridade do OGD com beta com limite focado em período de queda no mercado.

Gráfico 14 – Resultados do modelo OGD com beta entre -0,3 e 0,1 - comparação ao índice Bovespa e UCRP



(a) Comparação entre o portfólio OGD com beta entre -0,3 e 0,1, Índice Bovespa e UCRP.



(b) Variação no tempo de beta do portfólio OGD com limites.

Fonte: Elaborado pelo autor (2024).

Por fim, semelhante ao caso dos *Small Caps*, a principal consequência é a diminuição das métricas de risco do portfólio, conforme é visto na Tabela 8.

Tabela 8 – Comparação entre métricas de risco para o OGD com beta entre -0,3 e 0,1.

	OGD_beta		IBOV		UCRP	
	1%	5%	1%	5%	1%	5%
VaR	-0,008	-0,005	-0,039	-0,026	-0,054	-0,036
CVaR	-0,014	-0,007	-0,070	-0,036	-0,091	-0,050

Fonte: Elaborado pelo autor (2024).

6.2 RESULTADOS DO OGD COM RESTRIÇÃO *VALUE AT RISK*

Nesta seção, apresenta-se os resultados obtidos pela aplicação da segunda metodologia proposta às ações listadas no IBOV e às classificadas como *Small Caps*. O texto busca não apenas expor os achados, mas também aprimorar a compreensão do leitor sobre os processos e decisões envolvidas na análise.

Existem dois modelos/algoritmos principais em relação à forma como o VaR é estimado. O primeiro modelo pressupõe que os retornos seguem uma distribuição normal - apresentado no algoritmo 4. Já o segundo modelo, representado pelo algoritmo 5, estima o VaR supondo que os retornos seguem a distribuição t de Student.

Um importante parâmetro presente nos dois algoritmos estudados nesta seção estabelece um limiar de risco (através do VaR) para definir o nível máximo de perda que o investidor está disposto a tolerar. Para ilustrar diferentes perfis de risco, foram utilizados quatro patamares como exemplo: 80%, 40%, 20% e 10%. É importante ressaltar que esses valores são meramente ilustrativos e não representam regras fixas, mas sim uma forma de representar diferentes disposições de investidores individuais em relação ao risco.

Por exemplo, o parâmetro de 80% configura o modelo para tolerar até 80% do risco associado ao pior cenário estimado pelo quantil de 1% dos retornos passados. Isso significa que o modelo está configurado para aceitar uma perda máxima que é 80% da perda esperada no 1% dos casos mais extremos (pior cenário) observados no histórico de retornos. De maneira similar, ao usar 40%, 20% e 10% como limiares de risco, o modelo está sendo ajustado para tolerar 40%, 20% e 10%

do risco do pior cenário estimado, respectivamente.

Esses diferentes limiares permitem calibrar o nível de aversão ao risco na estratégia de otimização, optando por aceitar uma perda máxima proporcionalmente menor em relação ao pior cenário histórico. É fundamental entender que cada investidor pode ter sua própria tolerância ao risco, que pode ser diferente desses exemplos e deve ser determinada individualmente com base em diversos fatores, incluindo objetivos financeiros, horizonte de investimento e situação pessoal.

A abordagem de usar o quantil de 1% e frações dele busca encontrar um equilíbrio entre a gestão do risco e o potencial de retorno, mantendo o risco aceitável proporcional ao próprio risco dos retornos. Isso permite uma configuração mais granular e conservadora do limiar de risco, ajustando a estratégia de otimização para estar em consonância com uma análise detalhada do perfil de risco dos retornos passados.

A escolha de usar o quantil de 1% dos retornos passados, em vez de um quantil maior (como 5%), foi utilizada como uma forma de evitar "um salto muito grande" no risco. Isso sugere que a utilização de um quantil maior poderia resultar em um limiar de risco que permite perdas excessivamente grandes, potencialmente incompatíveis com a tolerância ao risco desejada.

Ambos modelos exigem a seleção do parâmetro η , responsável por controlar a magnitude dos passos na direção do gradiente negativo durante a otimização. A taxa de aprendizado, $\eta = 0,001$, determina o tamanho dos ajustes nos parâmetros do modelo com o intuito de minimizar a função de custo *Regret*. O parâmetro ν , definindo o nível de confiança para a análise de risco do portfólio, é fixado em $\nu = 0,01$ (1%), calculando o VaR ao nível de 99%. Esta especificação de ν também se aplica tanto aos modelos que adotam a distribuição normal quanto a outros modelos considerados neste estudo.

Para o modelo baseado na distribuição normal, o ν é usado para determinar o valor crítico ϕ , calculado pela função inversa da distribuição cumulativa normal padrão (ver equação 3.17). Similarmente, nos modelos que utilizam a distribuição t de Student, ν e o número de graus de liberdade são empregados para calcular o valor crítico ϕ pela função inversa da distribuição cumulativa t de Student, facilitando a

quantificação precisa do risco (ver equação 3.18).

Assim, ν integra considerações de risco ao processo de otimização de portfólio, ajustando a alocação de ativos para balancear retorno esperado e exposição ao risco. Destaca-se, portanto, a relevância do parâmetro ν tanto na fundamentação teórica quanto na prática de gestão de portfólios.

Para iniciar a otimização com os Algoritmos 4 e 5, selecionou-se um conjunto inicial de dados correspondente aos primeiros 10% dos retornos para ações do IBOV e 2,5% para *Small Caps*. Com base nesse subconjunto, calculou-se a matriz de covariância, essencial para entender as correlações entre os retornos dos ativos. Esta matriz serve como fundamento para uma alocação inicial de ativos, buscando um equilíbrio entre risco e retorno, fundamentado na análise dos dados históricos.

Vale ressaltar que na segunda metodologia, implementada via Algoritmo 5, há duas características extras: (i) por ser baseada na distribuição t de Student, necessita dos graus de liberdade - ajustados via método de máxima verossimilhança de maneira semelhante ao visto em [136]; (ii) a mediana, devido à sua menor sensibilidade a *outliers*, é empregada como medida de tendência central para os graus de liberdade, resultando em 10,70 para *Small Caps* e 10,43 para ações do IBOV, proporcionando estimativas mais estáveis e representativas.

Por fim, define-se o tamanho da janela temporal para o cálculo do VaR deslizante, essencial tanto na teoria quanto na prática de otimização de portfólios *online*, adaptando-se a novas informações. A janela se ajusta a cada novo dado, mantendo o VaR atualizado com as condições de mercado e facilitando a compreensão da volatilidade dos retornos. O período de janela escolhido foi de 44 dias, conforme aplicado nos Algoritmos 4 e 5, detalhados nas subseções 6.2.1 e 6.2.3.

O restante dessa seção está dividido em 4 partes. Nas duas primeiras, com foco nas *small caps*, realizou-se: (i) avaliação do OGD com e sem restrição de VaR assumindo retornos dos ativos com distribuição normal; (ii) avaliação do OGD com e sem restrição de VaR assumindo retornos dos ativos com distribuição t de Student. Já nas duas subseções finais, os dois passos destacados acima foram replicados, mas com enfoque em ativos do IBOV.

6.2.1 OGD com restrição VaR para dados *Small Caps* - Resultados considerando retornos com distribuição normal

Na primeira parte da análise, compara-se o retorno acumulado entre a carteira OGD e outros *benchmarks* tradicionais para avaliar seu desempenho. Conforme antecipado pela teoria dos algoritmos de otimização *online*, observa-se no Gráfico 15 um notável crescimento da carteira OGD, em alguns períodos específicos. No período compreendido entre abril de 2011 e maio de 2013, a performance da carteira OGD demonstrou uma divergência notável em relação aos *benchmarks* de referência, evidenciando uma valorização acentuada e sustentada. Este comportamento destaca uma fase de rendimento extraordinário, em que a carteira OGD superou significativamente os índices I_SMALL, UCRP e Markowitz.

Gráfico 15 – Comparação entre o portfólio OGD, UCRP, I_SMALL e Markowitz



Fonte: Elaborado pelo autor (2024).

Conforme apresentado na Tabela 9, a carteira OGD exibiu uma média de retornos positiva, apresentando a maior Taxa de Crescimento Anual Composta (TCAC), além de ter registrado a maior riqueza terminal.

Tabela 9 – Comparação de Desempenho e Risco: Análise das Carteiras OGD, Markowitz, UCRP e I_SMALL

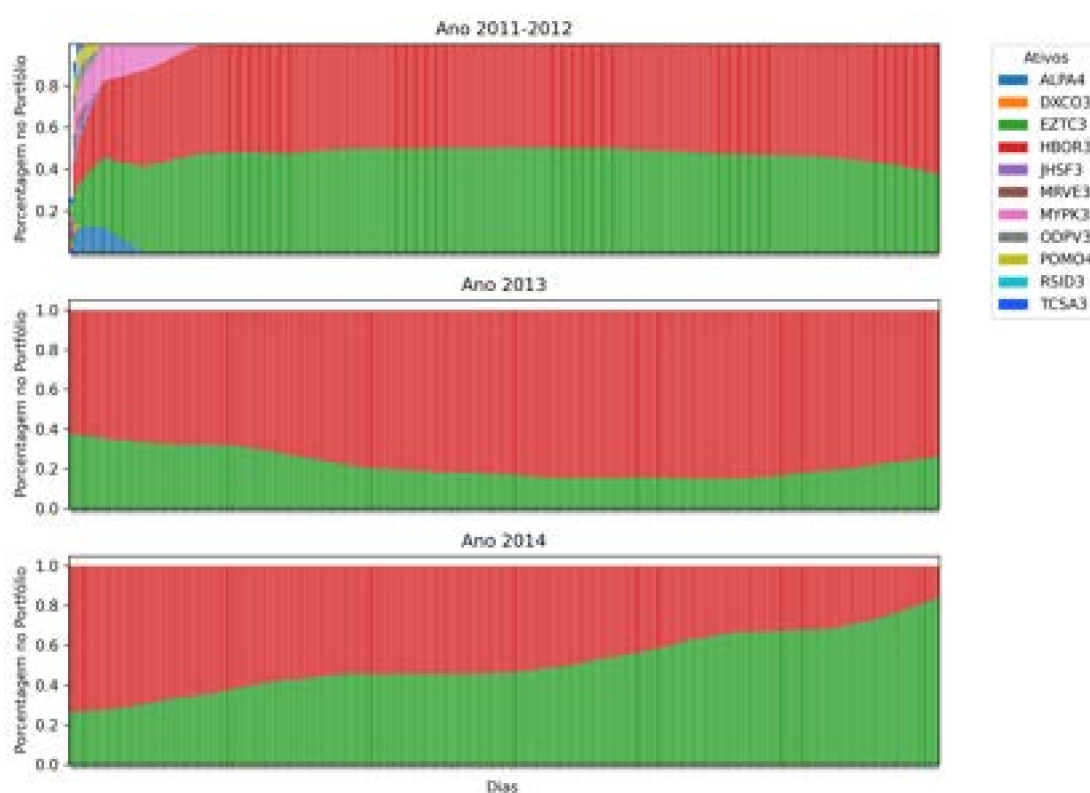
Métricas	OGD	Markowitz	UCRP	I_SMALL
TCAC (%)	11,610	3,234	8,703	8,864
Riqueza Terminal	4,681	1,564	3,231	3,299
Volatilidade Anualizada (%)	41,929	24,490	23,301	23,409
Semi-Desvio	0,036	0,021	0,019	0,020
Máximo <i>Drawdown</i> (%)	-30,464	-22,030	-23,365	-23,459
<i>Drawdown</i> Médio (%)	-11,120	-6,162	-5,447	-6,128
Pior Perda (%)	-28,684	-21,702	-21,222	-22,400
VaR 95 (%)	-6,819	-3,991	-3,718	-3,819
VaR 99 (%)	-10,379	-6,066	-5,651	-5,804
CVaR 95 (%)	-7,918	-4,622	-4,281	-4,337
CVaR 99 (%)	-11,925	-7,351	-7,132	-7,198
Razão de Calmar	0,147	0,064	0,162	0,165
Razão de Sortino	0,021	0,012	0,012	0,012
Recuperação Média	37,654	65,706	77,106	72,832
Assimetria	0,178	-0,563	-1,132	-1,157
Excesso de Curtose	5,921	8,012	14,675	14,221

Fonte: Elaborado pelo autor (2024).

Analisando o percentual de investimento do portfólio OGD destacado no Gráfico 16, no qual se plotou o peso destinado a cada ativo entre o final do ano de 2011 até o ano de 2014, percebe-se que o algoritmo concentrou o investimento em duas ações que apresentavam crescimento durante o período, HBOR3 (Helbor Empreendimentos) e EZTC3 (EZ TEC Empreendimentos e Participações), que foram as duas ações de maior crescimento no período de análise e que apresentam correlação positiva. Essa concentração excessiva em determinadas ações é uma característica comum a todas as estratégias de otimização de portfólios *online*, a tendência de concentrar investimentos em determinados ativos sem considerar adequadamente o risco é, portanto, um aspecto que pode ser observado em algoritmos de otimização de portfólio *online*. Esta característica deriva da estrutura e do objetivo principal desses algoritmos, que focam predominantemente na maximização da riqueza terminal sem integrar explicitamente considerações de risco

na tomada de decisão. Isso pode levar a uma alocação de portfólio que, embora potencialmente otimizada para crescimento sob condições de mercado específicas, pode não ser robusta frente a volatilidades ou adversidades não previstas. A falta de uma abordagem de gerenciamento de risco integrada sugere a necessidade de desenvolvimento e implementação de algoritmos que equilibrem a busca por retornos máximos com a necessidade de mitigar riscos, assegurando assim uma alocação de ativos mais diversificada e resiliente.

Gráfico 16 – Análise Temporal da Alocação de Ativos na Carteira OGD: Concentração e Diversificação de Investimentos de 2011 a 2014 para dados *small caps*



Fonte: Elaborado pelo autor (2024).

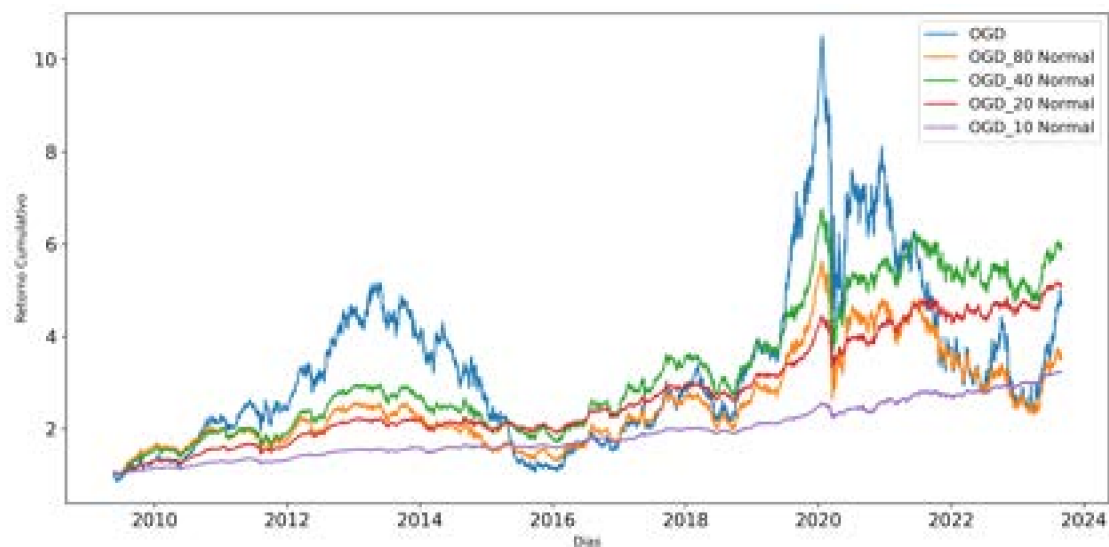
Ao analisar o desempenho e o risco associados ao índice *Small Caps* e à carteira UCRP, observa-se que, embora tenham apresentado um desempenho inferior ao da carteira OGD, ambas mantiveram retornos médios positivos, porém mais modestos. Ainda analisando as informações presentes na Tabela 9, na qual foi apresentada uma série de métricas comparativas de risco e desempenho, a Carteira

OGD apresenta uma maior volatilidade anualizada, indicando flutuações mais significativas nos retornos. Além disso, o maior Semi-Desvio, Máximo *Drawdown* e Pior Perda observados para a carteira OGD sugerem uma propensão para perdas significativas em períodos de baixa, refletindo uma maior vulnerabilidade a condições de mercado adversas.

Enquanto a carteira OGD oferece potencial de crescimento, os investidores devem estar cientes dos riscos associados. As medidas de risco, como o VaR e o CVaR, destacam a possibilidade de perdas extremas, que podem ser preocupantes para investidores com baixa tolerância ao risco. A análise da razão de Calmar e da razão de Sortino, que comparam o retorno ajustado ao risco, sugere que outras carteiras, como UCRP e I_SMALL, podem oferecer uma relação risco-retorno mais equilibrada, apesar de uma TCAC e riqueza terminal menores.

Conforme discutido no capítulo 4, a fusão de técnicas de otimização *online* com metodologias consagradas no campo das finanças apresenta-se como uma via promissora para a construção de estratégias de investimento de alta eficácia. O Gráfico 17 plota o desempenho acumulado de quatro carteiras, configuradas sob distintos graus de aversão ao risco, de acordo com os parâmetros estabelecidos pelo algoritmo 4. O algoritmo, ancorado em fundamentos de otimização *online*, incorpora a flexibilidade necessária para adaptar a composição das carteiras em tempo real, otimizando a relação risco-retorno através do controle da etapa de projeção do algoritmo OGD na forma padrão.

Gráfico 17 – Comparação entre o portfólio OGD, OGD_80 Normal, OGD_40 Normal, OGD_20 Normal e OGD_10 Normal para dados *small caps*



Fonte: Elaborado pelo autor (2024).

Ao definir limiares de risco em 80%, 40%, 20% e 10%, uma visão granular do impacto da aversão ao risco nos resultados financeiros acumulados é fornecido. A superioridade das carteiras OGD_40 Normal e OGD_20 Normal em termos de Riqueza Terminal em relação à implementação padrão do algoritmo OGD é evidente, sugerindo que ajustes nos parâmetros podem amplificar significativamente o desempenho financeiro, como corroborado pelas Tabelas 9 e 10. Estas carteiras não só superaram os *benchmarks* do mercado em TCAC, mas também evidenciaram uma mitigação expressiva no perfil de risco, refletida por métricas como VaR e CVaR.

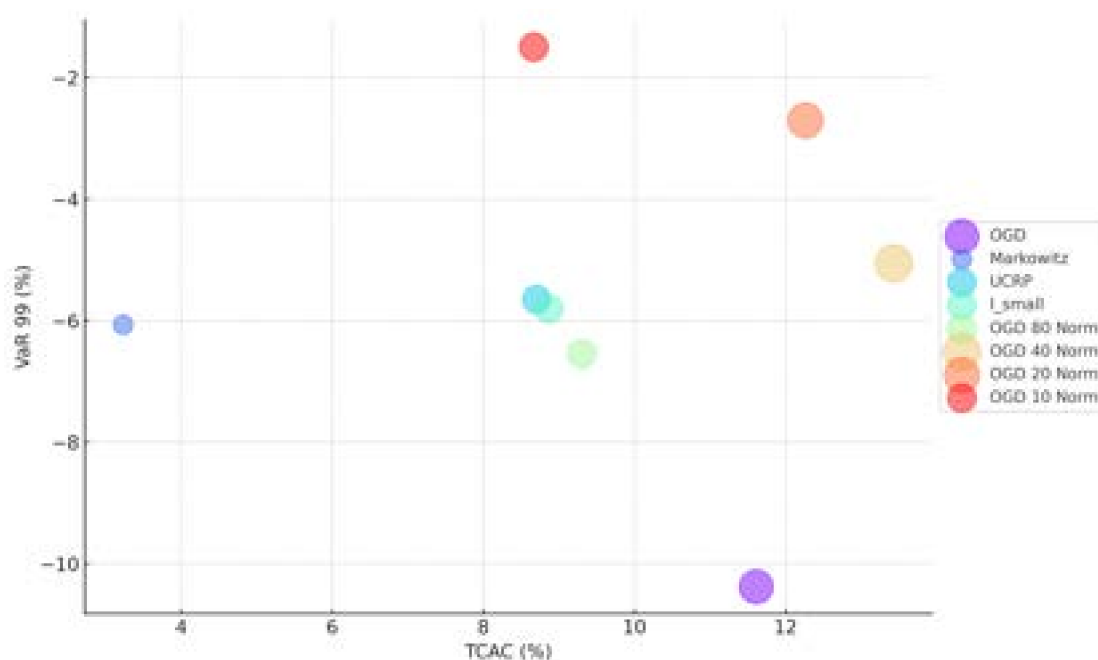
Tabela 10 – Comparação de Desempenho e Risco: Análise das Carteiras s OGD_80 Normal, OGD_40 Normal, OGD_20 Normal e OGD_10 Normal para dados *small caps*

Métricas	OGD 80 Normal	OGD 40 Normal	OGD 20 Normal	OGD 10 Normal
TCAC (%)	9,305	13,429	12,263	8,669
Riqueza Terminal	3,491	5,875	5,081	3,316
Volatilidade Anualizada (%)	26,339	20,474	11,147	6,243
Semi-Desvio	0,023	0,018	0,010	0,006
Máximo <i>Drawdown</i> (%)	-25,028	-20,268	-10,620	-5,334
<i>Drawdown</i> Médio (%)	-5,846	-4,740	-2,696	-1,463
Pior Perda (%)	-22,791	-17,825	-8,605	-4,303
VaR 95 (%)	-4,305	-3,329	-1,784	-0,989
VaR 99 (%)	-6,544	-5,057	-2,708	-1,500
CVaR 95 (%)	-4,953	-3,816	-2,101	-1,204
CVaR 99 (%)	-8,157	-6,298	-3,473	-2,031
Razão de Calmar	0,159	0,292	0,485	0,674
Razão de Sortino	0,013	0,010	0,005	0,003
Recuperação Média	63,180	80,249	143,647	259,687
Assimetria	-0,813	-0,565	-0,542	-0,483
Excesso de Curtose	12,131	12,101	9,351	7,047

Fonte: Elaborado pelo autor (2024).

Adicionalmente, indicadores como Volatilidade Anualizada e Máximo *Drawdown* reforçam a gestão de risco eficiente, alinhando-se ou superando os padrões de referência. A interação dessas variáveis é sintetizada visualmente no Gráfico 18, onde o TCAC (%) é confrontado com o VaR 99 (%) no eixo Y, e o dimensionamento dos círculos quantifica a Riqueza Terminal. Essa representação ressalta a robustez das carteiras OGD_40 Normal e OGD_20 Normal, as quais se distinguem não apenas pela acumulação de capital, mas também por uma alocação estratégica que prioriza a otimização do *trade-off* entre risco-retorno.

Gráfico 18 – Desempenho das carteiras: TCAC (%) vs. VaR 99 (%) com Riqueza Terminal como tamanho do círculo para dados *small caps*



Fonte: Elaborado pelo autor (2024).

Um aspecto interessante, evidenciado pelo Gráfico 19⁸, é que a implementação da restrição do VaR na fase de projeção induz a uma maior diversificação da carteira. Este gráfico ilustra a distribuição dos investimentos que, conjuntamente, representam ao menos 5% do valor total de cada carteira, uma condição imposta para evitar excessiva repetição de cores que poderia comprometer a clareza visual e a interpretação dos dados. Os gráficos apresentam a alocação percentual destinada a cada ativo disponível. Na carteira OGD_80 Normal, por exemplo, observa-se que, apesar de aceitar 80% do risco inerente à carteira OGD padrão, há

⁸ A visualização dos pesos dos ativos em portfólios pode apresentar dificuldades em identificar operações de compra e venda envolvendo pequenas porcentagens, especialmente quando há muitos ativos. Isso ocorre porque a representação gráfica tende a destacar mudanças mais significativas, enquanto alterações menores podem se perder na visualização, sobretudo quando as cores se repetem ou são muito semelhantes. No código utilizado para gerar a imagem, um mapeamento de cores consistente foi aplicado aos ativos, e um limite mínimo de investimento foi definido para filtrar as pequenas alocações.

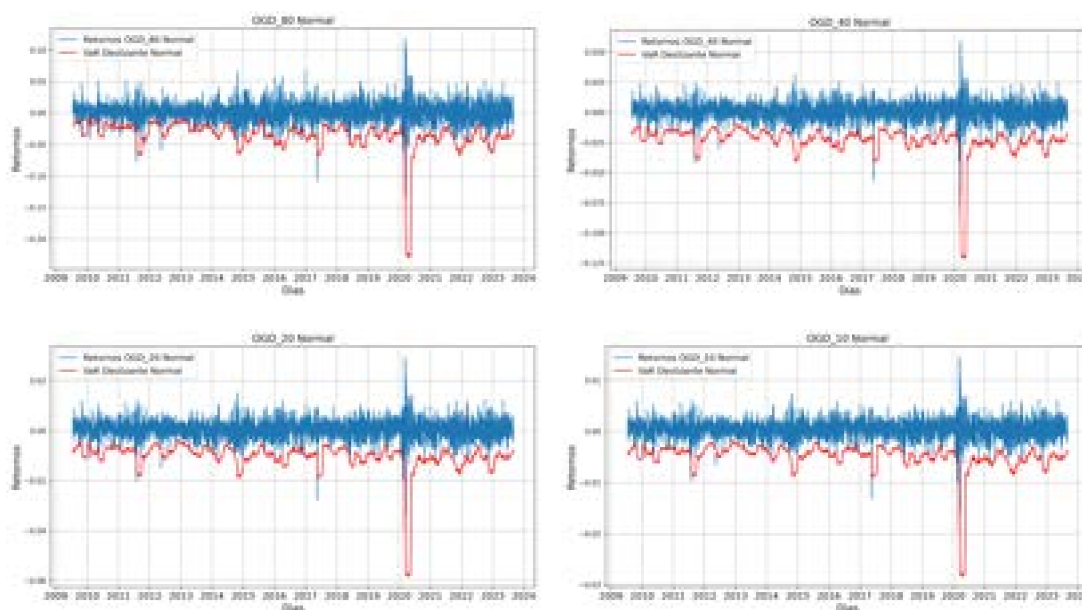
respectivamente, em comparação ao risco da carteira construída pelo OGD. Tal restrição impõe um controle rigoroso sobre o nível de risco aceitável, orientando a alocação de ativos para opções mais seguras, como investimentos em renda fixa, para cumprir com os limiares de risco estabelecidos.

No Gráfico 20, os retornos são plotados, bem como o VaR, para uma janela deslizante de 44 dias, que é o parâmetro de restrição definido para cada atualização dos retornos do portfólio construído. De maneira geral, observou-se que a restrição de VaR foi respeitada, evidenciando a eficácia da metodologia proposta. As ocasiões em que o VaR foi excedido⁹ destacam-se como momentos em que, devido à natureza da otimização *online* e à escolha de investimentos realizada previamente à observação dos retornos, seria particularmente desafiador prever os resultados exatos dos investimentos. Citando como exemplo, ao aplicar a restrição de VaR com base na distribuição normal, identificou-se um total de 66 dias de excedências para a carteira denominada OGD_80 Normal, o que corresponde a aproximadamente 1,87% dos dias de negociação. A Tabela 12, deixada mais a frente para a melhor comparação entre os resultados ligados a retornos com distribuição normal e t de Student, apresenta uma comparação dessas excedências de VaR entre todas as carteiras analisadas.

Embora os resultados até agora sejam promissores, torna-se relevante levar em conta duas métricas estatísticas na avaliação da distribuição dos retornos: a Assimetria e o Excesso de Curtose. É importante destacar que tais métricas, sendo momentos de ordem superior da distribuição dos retornos, não fornecem uma visão completa do risco da carteira, mas sim complementam as medidas de risco mais tradicionais.

A assimetria quantifica o desvio da simetria numa distribuição de probabilidade, com uma distribuição simétrica exibindo uma assimetria igual a zero, denotando que os dados estão uniformemente distribuídos ao redor da média. Na análise das carteiras otimizadas (Tabela 9), observa-se que a carteira OGD apresenta uma assimetria de 0,178, indicando uma leve assimetria positiva; isto é, a probabilidade de valores a esquerda da média são maiores, com presença de

⁹ Número de excedências aqui é o total de dias que um modelo excede em quantificar o risco, no caso, o VaR.

Gráfico 20 – Retornos e VaR Normal para dados *small caps*

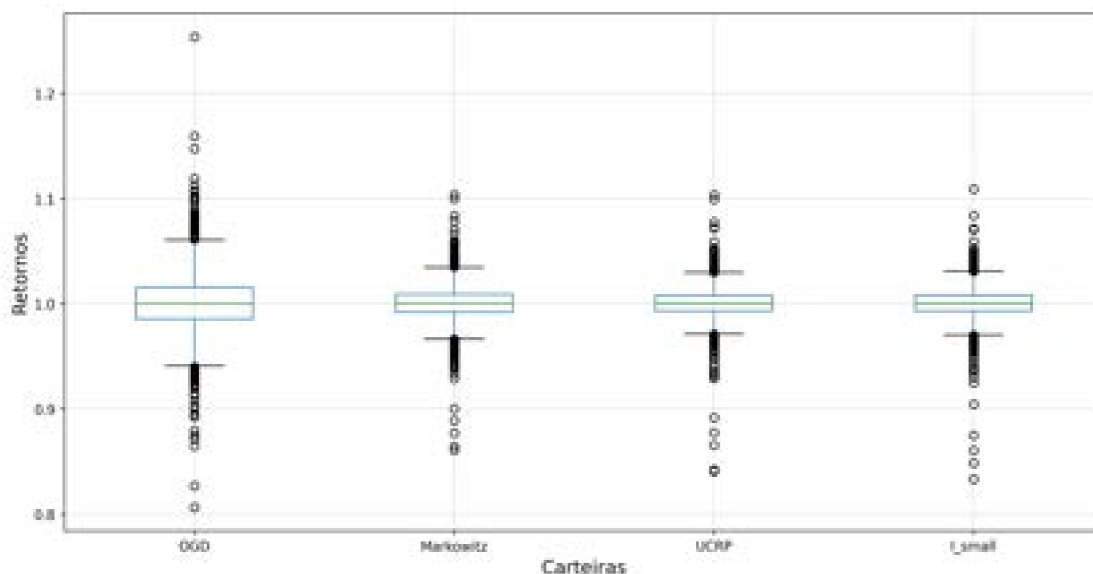
Fonte: Elaborado pelo autor (2024).

valores mais extremos a direita da média. Já as carteiras de referência, como Markowitz, UCRP e I_SMALL, exibem assimetrias negativas, variando de -0,563 a -1,157, sugerindo uma predisposição maior para retornos a direita da média, mas com valores mais extremos a esquerda da média. Vale notar que, apesar destas tenderem a gerar retornos acima da média das mesmas na comparação ao OGD, a distribuição OGD tem valores de quantis mais extremos, conforme pode ser visto no Gráfico 21 abaixo,¹⁰ gerando retornos comparativos maiores (e com maior risco

embutido) conforme discutido na subseção 6.2.1.

¹⁰ Utilizou-se ao longo dos resultados *box-plots* ao invés de histogramas ou mesmo plotar estimações das distribuições (por exemplo, por métodos não paramétricas) pois as distribuições não apresentaram diferentes picos ou inconsistências, além de gerarem uma intuição mais direta e simples. Não menos importante, mesmo com tratamento para possíveis *outliers*, os resultados gerais se mantiveram.

Gráfico 21 – *Box-Plots* dos retornos brutos das carteiras OGD, Markowitz, UCRP e I_SMALL

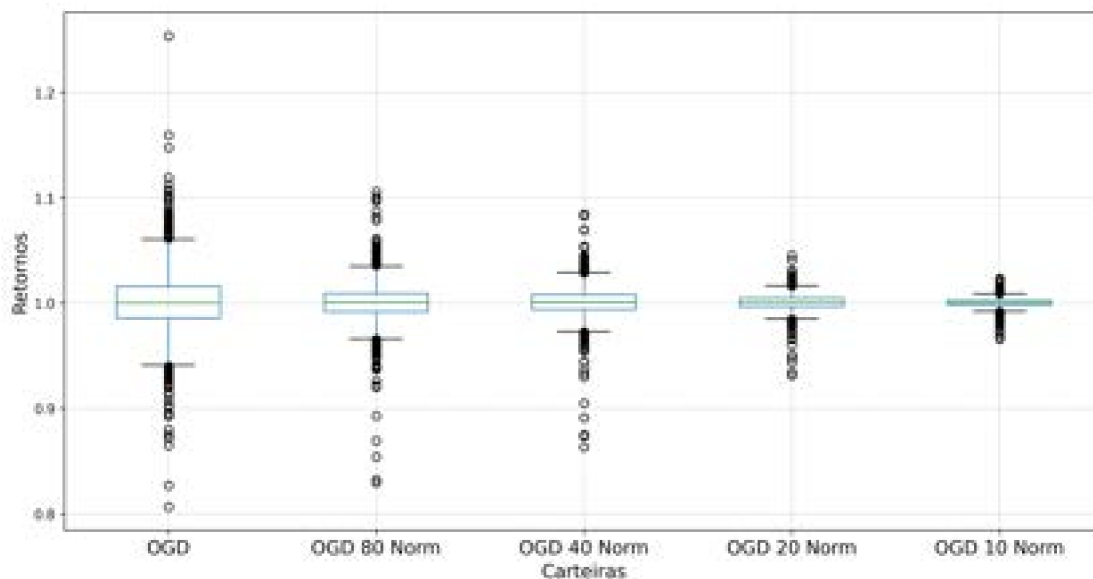


Fonte: Elaborado pelo autor (2024).

Por outro lado, o excesso de curtose apresentou valores elevados para todas as carteiras analisadas. A carteira OGD registrou um excesso de curtose de 5,921, enquanto as carteiras de referência variaram de 8,012 para Markowitz a 14,675 para UCRP. Esta tendência indica uma prevalência acentuada de caudas pesadas na distribuição dos retornos, o que sinaliza uma maior probabilidade de ocorrências extremas, tanto abaixo como acima da média, em comparação com a premissa de uma distribuição normal padrão.

Ao analisar as carteiras com restrição de VaR na Tabela 10, observa-se que a assimetria se torna mais negativa à medida que a restrição de VaR é atenuada, variando de -0,483 para a carteira OGD_10 Normal a -0,565 para a carteira OGD_40 Normal. Este padrão revela uma inclinação para a assimetria negativa, sugerindo uma predisposição para caudas mais extensas à esquerda da média, o que implica a presença de retornos mais aquém da média, apesar da menor probabilidade de ocorrência na comparação aos valores ligados aos quantis acima do central. Isso fica nítido no Gráfico 22 abaixo.

Gráfico 22 – *Box-Plots* dos retornos brutos das carteiras OGD, OGD_80 Normal, OGD_40 Normal, OGD_20 Normal e OGD_10 Normal para dados *small caps*



Fonte: Elaborado pelo autor (2024).

Quanto ao excesso de curtose, as carteiras com restrição de VaR também apresentaram valores elevados, oscilando de 7,047 na carteira OGD_10 Normal a 12,101 na carteira OGD_40 Normal. A identificação de um excesso de curtose elevado nessas carteiras sugere um perfil de risco que vai além da capacidade de representação da distribuição normal, indicando que investidores podem subestimar o risco de eventos de mercado raros e drásticos.

A avaliação das métricas sob a suposição de normalidade ressalta a inadequação desta distribuição para modelar com precisão os riscos e retornos no âmbito financeiro. A evidência de assimetria e de um excesso de curtose aponta para a viabilidade de distribuições alternativas, capazes de acomodar com mais flexibilidade caudas pesadas, como a distribuição t de Student, fornecendo uma representação mais verossímil do comportamento dos retornos financeiros.

A subseção subsequente, 6.2.2, se concentrará na análise dos resultados gerados pelo algoritmo 5, que calcula o VaR sob a premissa de que os retornos seguem a distribuição t de Student.

6.2.2 OGD com restrição VaR para dados *Small Caps* - Resultados considerando retornos com distribuição t de Student

A análise é iniciada no Gráfico 23 do retorno acumulado das carteiras construídas utilizando a restrição do VaR na etapa de projeção, com a suposição de que os retornos seguem a distribuição t de Student. Uma análise visual mostra que o uso dessa estratégia fez com que a carteira OGD_80 t Student apresentasse um desempenho superior, em especial nos períodos compreendidos entre abril de 2011 até maio de 2013 e maio de 2019 até janeiro de 2020.

Gráfico 23 – Comparação entre os portfólios OGD, OGD_80 t Student, OGD_40 t Student, OGD_20 t Student e OGD_10 t Student para dados *small caps*



Fonte: Elaborado pelo autor (2024).

Em contrapartida, quando são analisadas as métricas relacionadas ao risco, nas Tabelas 9 e 11, observa-se que as métricas de risco e desempenho da carteira OGD_80 t Student conseguem apresentar resultados superiores em todos os critérios avaliados na comparação a OGD sem critérios sobre o VaR ou as carteiras de referência, como Markowitz, UCRP e I_SMALL.

De maneira geral, ao comparar as métricas de desempenho e risco presentes nas Tabelas 10 e 11, com controle de risco, nota-se que as carteiras estimadas através da técnica que assume a distribuição t de Student tendem a ter uma TCAC

mais alta em comparação com as estimadas pela distribuição normal, indicando um desempenho superior ao longo do tempo. Esta diferença sugere que assumir uma distribuição t de Student pode resultar em uma melhor performance de crescimento. Similarmente, as carteiras OGD_20 t de Student e OGD_10 t de Student também mostram um desempenho favorável, reforçando a ideia de que a modelagem com a distribuição t de Student pode ser mais eficaz em capturar retornos superiores, mesmo em níveis de alocação mais conservadores. Outro ponto a se ressaltar é que houve, em geral, uma melhora na Volatilidade Anualizada, especialmente notável nas alocações OGD_20 t Student e OGD_40 t Student. Essa melhora pode indicar que essa estratégia oferece uma abordagem que apresenta um perfil de risco moderadamente mais baixo.

Tabela 11 – Comparação de Desempenho e Risco: Análise das Carteiras OGD_80 t Student, OGD_40 t Student, OGD_20 t Student, OGD_10 t Student para dados *small caps*

Métricas	OGD_80 t Student	OGD_40 t Student	OGD_20 t Student	OGD_10 t Student
TCAC (%)	12,152	14,086	12,128	10,794
Riqueza Terminal	5,011	6,380	5,010	4,223
Volatilidade Anualizada (%)	29,659	17,479	8,631	4,197
Semi-Desvio	0,030	0,016	0,008	0,004
Máximo <i>Drawdown</i> (%)	-27,785	-17,013	-8,795	-4,443
<i>Drawdown</i> Médio (%)	-8,737	-4,355	-2,212	-1,100
Pior Perda (%)	-25,808	-14,688	-7,617	-3,853
VaR 95 (%)	-5,637	-2,841	-1,398	-0,680
VaR 99 (%)	-8,574	-4,315	-2,122	-1,031
CVaR 95 (%)	-6,537	-3,282	-1,633	-0,797
CVaR 99 (%)	-9,811	-5,279	-2,641	-1,298
Razão de Calmar	0,165	0,372	0,6212	1,098
Razão de Sortino	0,017	0,009	0,004	0,002
Recuperação Média	45,787	91,296	184,224	379,939
Assimetria	-0,093	-0,746	-0,773	-0,824
Excesso de Curtose	5,929	9,302	9,664	10,407

Fonte: Elaborado pelo autor (2024).

Ao comparar o Máximo *Drawdown* das carteiras que utilizaram a distribuição Normal e a t de Student na etapa de projeção, observou-se que as carteiras que adotaram a t de Student apresentam valores ligeiramente maiores (em valor absoluto) em comparação com as carteiras que utilizaram a Normal. Por exemplo, para a carteira OGD_80, o Máximo *Drawdown* é de -27,785% para a t de Student e -25,028% para a Normal. Esse padrão se repete para as demais carteiras, com a diferença diminuindo à medida que a alocação em renda variável diminui.

O Máximo *Drawdown* representa a maior perda observada em um período, medida do pico ao vale. Um Máximo *Drawdown* maior indica que a carteira experimentou uma queda mais acentuada em seu valor. Portanto, os resultados sugerem que as carteiras que utilizaram a *t* de Student na etapa de projeção apresentaram perdas máximas ligeiramente maiores em comparação com as carteiras que adotaram a Normal.

No entanto, ao analisar a Razão de Calmar, que mede o retorno em excesso por unidade de Máximo *Drawdown*, fica evidenciado que as carteiras que adotaram a *t* de Student apresentam valores consistentemente superiores em comparação com as carteiras que utilizaram a Normal. Isso indica que, apesar de terem Máximos *Drawdowns* maiores, essas carteiras conseguiram gerar um retorno mais alto por unidade de perda máxima.

Esse resultado sugere que, embora a utilização da *t* de Student na etapa de projeção possa levar a perdas máximas ligeiramente maiores, isso é compensado por retornos mais elevados. A Razão de Calmar mais alta para as carteiras que adotaram a *t* de Student indica que essas carteiras obtiveram um melhor desempenho ajustado ao Máximo *Drawdown* em comparação com as carteiras que utilizaram a Normal.

Passando para a Razão de Sortino, que mede o retorno em excesso por unidade de risco de queda (*Downside Risk*), utilizando o Semi-Desvio como medida de risco, observa-se um padrão semelhante. As carteiras que adotaram a *t* de Student apresentam Razões de Sortino mais altas em comparação com as carteiras que utilizaram a Normal. Isso sugere que, mesmo considerando o risco de queda, as carteiras que utilizaram a *t* de Student na etapa de projeção obtiveram um melhor desempenho ajustado ao risco.

Antes de seguir com a análise das demais métricas apresentadas, é importante reconhecer que a escolha da distribuição para modelar os retornos na etapa de estimação do VaR tem um impacto significativo na estimativa do risco e, conseqüentemente, na alocação dos ativos na carteira. No entanto, quando se opta pela distribuição *t* de Student, um parâmetro adicional é inserido o problema, graus de liberdade, que deve ser selecionado a priori e que tem a capacidade de interferir no desempenho de cada carteira.

Assim como já exposto na seção 4.3, o grau de liberdade controla a forma da distribuição, especialmente a espessura das caudas. Um grau de liberdade menor implica caudas mais pesadas e uma maior probabilidade de retornos extremos, o que pode levar a uma superestimação do risco e a uma alocação excessivamente conservadora. Por outro lado, um grau de liberdade maior aproxima a distribuição t de Student da distribuição normal, o que pode subestimar o risco e levar a uma alocação mais agressiva. Para os dados em questão, ações pertencentes a *Small Caps*, utilizou-se os graus de liberdade como 10,7, uma vez que este representava a mediana para o conjunto de treinamento.

Ao comparar o VaR 99% das carteiras estimadas utilizando a distribuição Normal e a t de Student na etapa de projeção, observa-se que as carteiras que adotam a t de Student apresentam valores de VaR 99% consistentemente maiores (em valor absoluto) do que aquelas que utilizaram a Normal. Por exemplo, para a carteira OGD_80, o VaR 99% é de -8,574% para a t de Student e -6,544% para a Normal. Esse padrão se mantém para todas as outras carteiras.

Esse resultado pode ser explicado pela diferença nas características das distribuições utilizadas durante a etapa de projeção. A distribuição t de Student, devido às suas caudas mais pesadas, atribui maiores probabilidades a eventos extremos em comparação com a Normal. Isso significa que, ao utilizar a t de Student na projeção, o algoritmo de otimização considera um risco maior de perdas extremas, o que pode levar a uma alocação mais conservadora dos ativos.

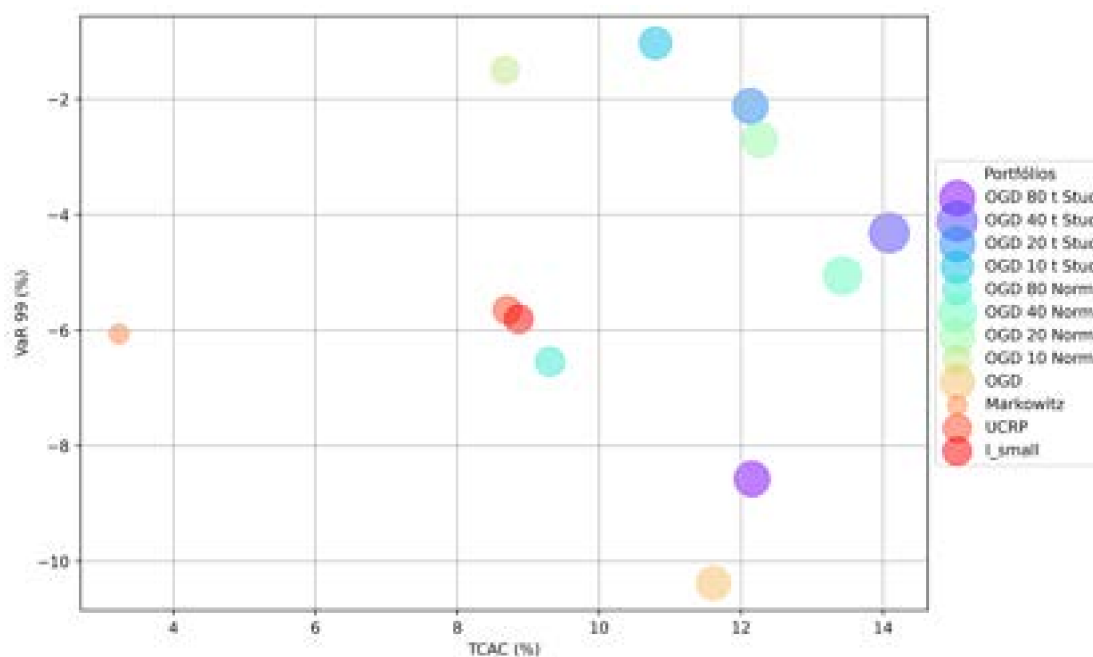
Apesar do VaR 99% ser maior para as carteiras que utilizaram a t de Student na projeção, é interessante notar que essas carteiras apresentaram métricas de desempenho superiores, como o TCAC e a Riqueza Terminal, em comparação com as carteiras que adotaram a Normal. Isso sugere que a consideração de uma distribuição com caudas mais pesadas na etapa de projeção, embora possa resultar em um VaR maior, não necessariamente implica em um desempenho inferior. Pelo contrário, neste caso, as carteiras que utilizaram a t de Student obtiveram retornos mais atrativos.

Outro ponto relevante é que a diferença entre o VaR 99% das carteiras que utilizaram as duas distribuições é mais acentuada para as carteiras com maior

alocação em renda variável (OGD_80 e OGD_40). Isso pode ser atribuído ao fato de que ativos de renda variável geralmente apresentam caudas mais pesadas em suas distribuições de retornos, fazendo com que a escolha da distribuição na etapa de projeção tenha um impacto maior na estimativa do risco.

Essas análises podem ser observadas no Gráfico 24 onde se plotou todas as carteiras construídas utilizando dados da *Small Caps*. Observa-se que as carteiras que utilizaram a distribuição t de Student na etapa de projeção conseguem obter maior riqueza terminal e TCAC, embora apresentem maior risco avaliado pelo VaR 99%. De maneira geral, as duas formas de projeção conseguem apresentar desempenho superior aos *benchmarks*; apresentando, assim, serem boas alternativas de investimento.

Gráfico 24 – Desempenho das carteiras: TCAC (%) vs. VaR 99 (%) com Riqueza Terminal como tamanho do círculo para dados *small caps*

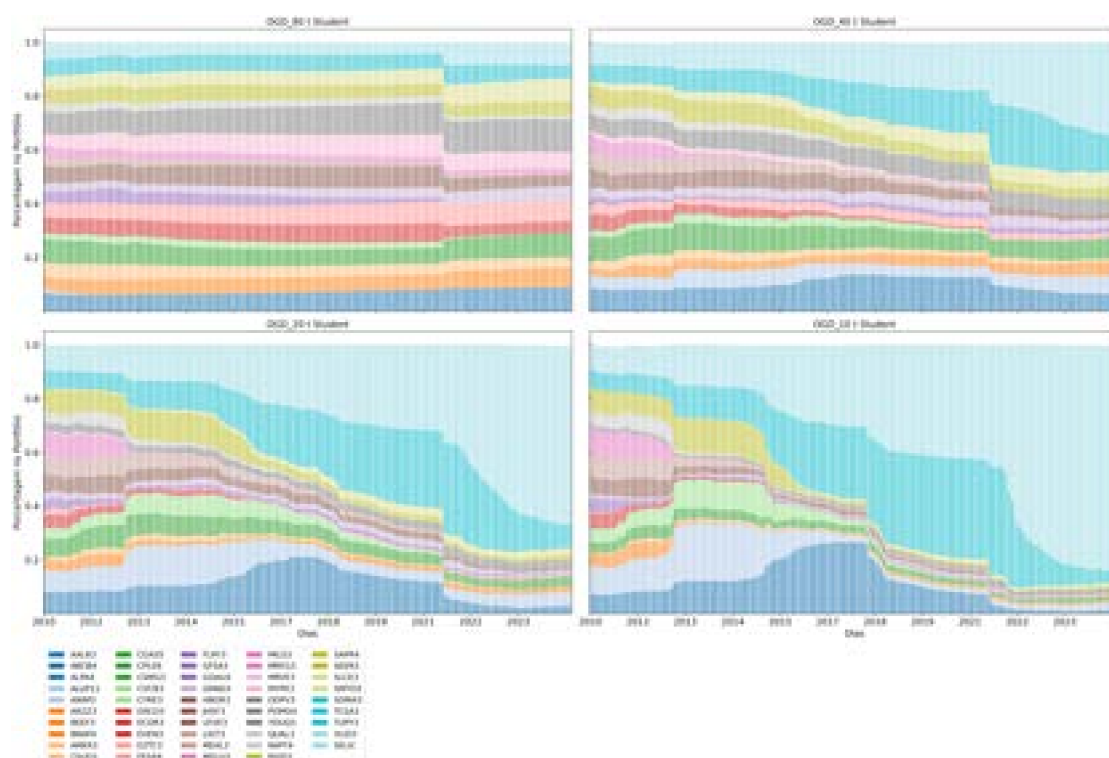


Fonte: Elaborado pelo autor (2024).

Ao analisar as alocações de investimento das carteiras, conforme ilustrado nos Gráficos 19 e 25, identificou-se discrepâncias significativas nos padrões de alocação. O Gráfico 19 evidencia que as carteiras com orientação mais conservadora tendem

a alocar uma parcela substancial dos investimentos na taxa livre de risco, a SELIC, enquanto possuem uma parcela menor do portfólio em ativos corporativos. Em contraste, a abordagem baseada na distribuição t de Student, conforme demonstrado no Gráfico 25, favorece uma maior diversificação de ativos, notadamente nas carteiras OGD_80 t Student e OGD_40 t Student. Inclusive, observa-se que as carteiras consideradas mais conservadoras, especificamente OGD_20 t Student e OGD_10 t Student, mantêm uma proporção menor de investimentos em SELIC.

Gráfico 25 – Comparação dos pesos investidos nos portfólios OGD_80 t Student, OGD_40 t Student, OGD_20 t Student e OGD_10 t Student para dados *small caps*



Fonte: Elaborado pelo autor (2024).

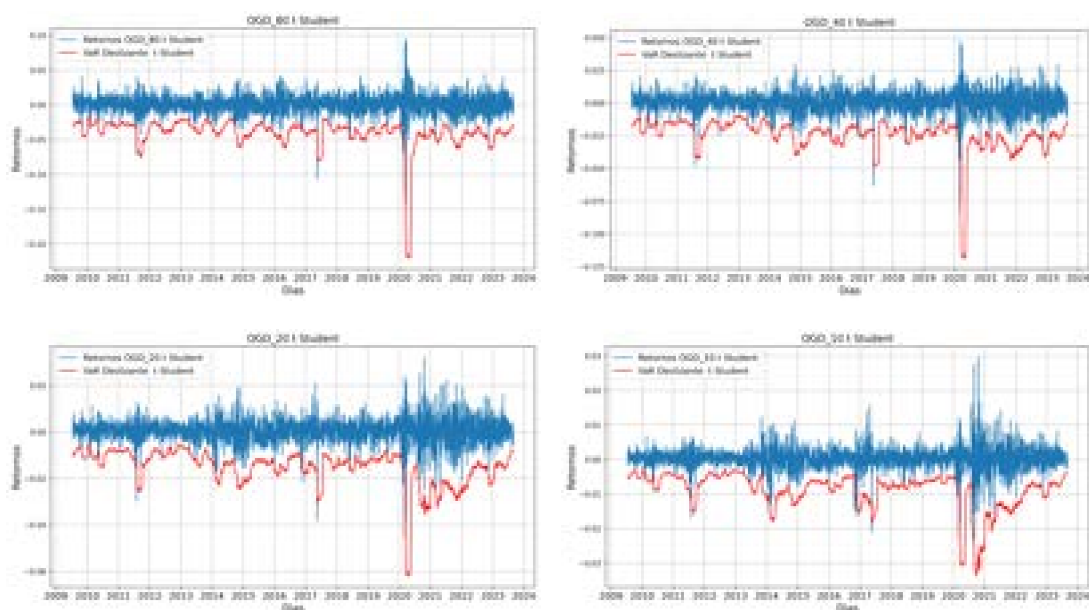
Este padrão divergente de alocação parece ser, em grande medida, reflexo das premissas distintas acerca da distribuição dos retornos. A adoção da distribuição t de Student, que confere maior probabilidade a ocorrências de eventos extremos, induz o modelo de otimização a abarcar uma gama mais extensa de ativos na busca por mitigar o risco de perdas expressivas. Este fenômeno parece conduzir a

uma diversificação ampliada das carteiras, inclusive nas abordagens conservadoras. Essa estratégia acaba promovendo um incremento médio nos retornos, evidenciado tanto pela TCAC quanto pela Riqueza Terminal. Esta dinâmica sugere que a contemplação de uma diversificação robusta de ativos e a incorporação de eventos extremos na estimação do VaR contribuem para uma otimização da relação risco-retorno, compensando adequadamente o risco adicional inerente.

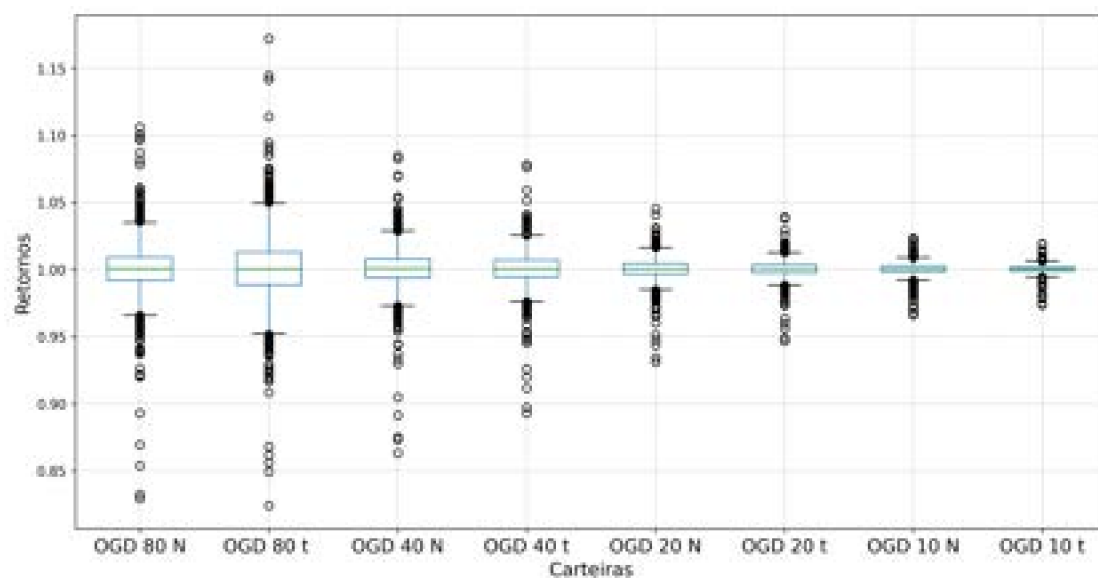
Adicionalmente, ao considerar indicadores como a Razão de Calmar e a Razão de Sortino, verifica-se que as carteiras fundamentadas na distribuição t de Student superam, ainda que marginalmente, aquelas baseadas na distribuição Normal em termos de geração de retorno por unidade de risco assumido. Esta observação reitera a eficácia da distribuição t de Student não apenas na diversificação de carteira, mas também na capacidade de produzir retornos mais atrativos, ajustados ao risco.

No Gráfico 26 foi plotado os retornos e o VaR 99% para cada uma das carteiras construídas. É importante ressaltar que, conforme mencionado no texto, as carteiras que utilizam a distribuição t de Student apresentam uma menor participação em ativos livres de risco, mesmo nas carteiras mais conservadoras. Essa observação é relevante, pois sugere que a distribuição t de Student leva a uma maior alocação em ativos de risco, o que pode influenciar o comportamento das excedências do VaR.

Ao analisar novamente as Tabelas 10 e 11, observa-se que as carteiras baseadas na distribuição t de Student apresentam valores de assimetria e excesso de curtose consistentemente menores em comparação com as carteiras baseadas na distribuição Normal. De fato, avaliando o Gráfico 27 abaixo, fica claro a menor discrepância de dados tanto a direita como a esquerda do quantil central das distribuições. Em especial para o OGD_80 t Student, a distribuição é quase simétrica, com os valores extremos em ambos os lados da distribuição bastante similares com relação a média dos retornos da carteira. Já com relação aos outros níveis, tais efeitos são diminuídos, mas ainda estão presentes.

Gráfico 26 – Retornos e VaR t de Student para dados *small caps*

Fonte: Elaborado pelo autor (2024).

Gráfico 27 – *Box-Plots* comparativos dos retornos brutos das carteiras OGD_80, OGD_40, OGD_20 e OGD_10 considerando distribuições Normal e t de Student para dados *small caps*

Fonte: Elaborado pelo autor (2024).

Esses resultados sugerem que, embora a distribuição t de Student possua caudas mais pesadas que a Normal, as carteiras resultantes apresentam uma distribuição de retornos menos assimétrica e com menor excesso de curtose. Isso pode ser atribuído à maior diversificação de ativos proporcionada pela abordagem baseada na t de Student, conforme evidenciado pela menor participação em ativos livres de risco.

No entanto, é importante notar que, mesmo com uma distribuição de retornos aparentemente mais favorável, as carteiras baseadas na t de Student apresentam um maior número de excedências do VaR, especialmente para as carteiras mais conservadoras (OGD_20 t Student e OGD_10 t Student), conforme mostrado na Tabela 12. Isso sugere que, apesar da menor assimetria e excesso de curtose, a distribuição t de Student pode ainda estar subestimando a probabilidade de perdas extremas nessas carteiras.

Tabela 12 – Comparação do número de excedências para os modelos VaR Normal e VaR t de Student para dados *small caps*

Disposição ao Risco	Distribuição Normal	Distribuição t de Student
OGD_80	66	36
OGD_40	63	78
OGD_20	58	135
OGD_10	42	187

Fonte: Elaborado pelo autor (2024).

Uma possível explicação para esse fenômeno é que, embora a distribuição t de Student capture melhor as caudas pesadas em comparação com a Normal, ela ainda assume simetria em torno da média. No entanto, em carteiras mais conservadoras, mesmo aquelas com uma menor participação em ativos livres de risco, a distribuição dos retornos pode apresentar uma assimetria negativa mais acentuada devido à menor exposição a ativos de risco. Assim, a distribuição t de Student pode não estar capturando adequadamente essa assimetria, levando a uma subestimação do VaR e, conseqüentemente, a um maior número de excedências.

Além disso, é importante considerar que a escolha dos graus de liberdade da distribuição t de Student pode ter um impacto significativo na estimativa do VaR.

Como mencionado anteriormente, se os graus de liberdade não forem ajustados adequadamente para cada nível de disposição ao risco, o VaR estimado pode não refletir corretamente a probabilidade de perdas extremas, contribuindo para o maior número de excedências em carteiras mais conservadoras.

Em suma, levando em conta as Tabelas 10 e 11, mais especificamente os valores de assimetria e excesso de curtose apresentados para as carteiras construídas, além dos Gráficos 22 e 27, algumas observações interessantes são destacadas:

- i Para todas as carteiras construídas usando a distribuição Normal, observou-se uma assimetria negativa, variando de -0,483 a -1,813. Isso indica que essas carteiras apresentam retornos com caudas esquerdas mais pesadas. Além disso, o excesso de curtose é alto, variando de 7,047 a 12,131; evidenciando, assim, a presença de caudas pesadas em relação à distribuição Normal. Esses resultados sugerem que considerar uma distribuição Normal não está capturando adequadamente as características dos retornos dessas carteiras. A assimetria negativa e o excesso de curtose indicam que a Normal está subestimando as perdas potenciais e que uma distribuição com caudas mais pesadas seria mais apropriada para modelar esses retornos.
- ii Já para as carteiras construídas usando a distribuição t de Student, observa-se uma melhora significativa na assimetria, com valores variando de -0,093 a -0,824. Isso indica uma distribuição mais simétrica dos retornos. O excesso de curtose também é menor em comparação com as carteiras baseadas na Normal, ficando entre 5,929 e 10,407. Esses valores indicam que a distribuição t de Student está capturando melhor as características dos retornos em comparação com a Normal. A menor assimetria e o menor excesso de curtose sugerem que a t de Student está proporcionando uma representação mais adequada das caudas pesadas presentes nos dados.

Fazendo um apanhado geral, a carteira OGD original apresenta uma leve assimetria positiva de 0,178 e um excesso de curtose de 5,921, indicando caudas pesadas. As variações do OGD com restrição de VaR baseadas na distribuição Normal apresentam assimetria negativa e excesso de curtose elevado, sugerindo

que a Normal não é a melhor escolha para modelar esses retornos. Por outro lado, as variações do OGD com restrição de VaR baseadas na distribuição t de Student mostram uma melhora na assimetria e no excesso de curtose, indicando que a t de Student é mais adequada para capturar as características dos retornos.

No entanto, é importante ressaltar que mesmo com a distribuição t de Student, ainda observamos uma assimetria e um excesso de curtose não desprezíveis, em especial nas carteiras mais conservadoras, sugerindo que pode haver espaço para explorar outras distribuições que capturem ainda melhor as características específicas dos retornos dessas carteiras.

6.2.3 OGD com restrição VaR para dados IBOV - Resultados considerando retornos com distribuição normal

Agora o foco está na análise de desempenho das ações que fazem parte do índice Bovespa. Ao analisar o Gráfico 28, no qual o retorno acumulado da carteira OGD em comparação com os *benchmarks* do mercado (Índice Bovespa, UCRP e Markowitz) é plotado, observam-se dois aspectos. Primeiramente, a carteira OGD exibe um desempenho substancialmente superior ao longo do período em análise. Em segundo lugar, nota-se uma queda expressiva no início de 2020, atribuída ao impacto da pandemia da COVID-19.

Analisando as métricas comparativas dos portfólios presentes na Tabela 13, identifica-se que a carteira OGD se destaca por seu retorno superior, evidenciado pela TCAC de 19,841%, consideravelmente maior do que os *benchmarks*. Além disso, a Riqueza Terminal alcançada pela estratégia OGD é substancialmente mais alta, indicando que a aplicação dessa abordagem resultou em um montante final expressivamente maior.

Gráfico 28 – Comparação entre o portfólio OGD, IBOV, UCRP e Markowitz



Fonte: Elaborado pelo autor (2024).

Tabela 13 – Comparação de Desempenho e Risco: Análise das Carteiras OGD, Markowitz, UCRP e IBOV

Métricas	OGD	Markowitz	UCRP	I_bovespa
TCAC(%)	19,841	6,051	11,698	9,453
Riqueza Terminal	29,881	3,012	7,976	5,449
Volatilidade Anualizada (%)	26,179	21,508	23,657	27,042
Semi-Desvio	0,022	0,0173	0,0202	0,023
Máximo <i>Drawdown</i> (%)	-27,043	-19,294	-25,114	-25,673
<i>Drawdown</i> Médio (%)	-11,699	-7,247	-9,909	-10,967
Pior Perda (%)	-25,899	-12,690	-24,299	-24,432
VaR 95 (%)	-3,228	-2,668	-3,030	-3,470
VaR 99 (%)	-6,395	-4,575	-5,050	-5,565
CVaR 95 (%)	-4,790	-3,927	-4,407	-5,020
CVaR 99 (%)	-7,923	-6,500	-7,146	-7,204
Razão de Calmar	0,354	0,105	0,208	0,158
Razão de Sortino	0,013	0,010	0,011	0,013
Recuperação Média	62,759	81,093	67,918	58,854
Assimetria	0,003	-0,515	-0,322	-0,138
Excesso de Curtose	12,359	10,512	11,454	8,533

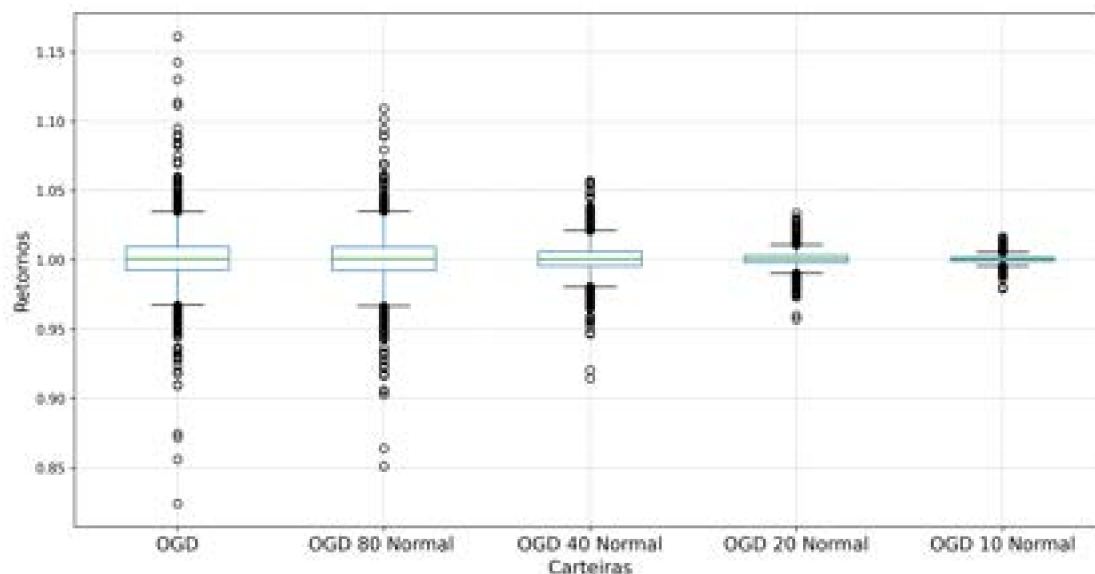
Fonte: Elaborado pelo autor (2024).

A carteira OGD apresenta uma Volatilidade Anualizada de 26,179%, ligeiramente inferior à do IBOV (27,042%), mas superior aos demais *benchmarks*. Esse resultado sugere que a variabilidade dos retornos da estratégia OGD é comparável à do mercado acionário brasileiro. Ademais, a análise do Máximo *Drawdown* revela que a carteira OGD experimentou perdas significativas de -27,043% em determinados períodos, valor próximo ao observado para o IBOV (-25,673%) e o UCRP (-25,114%).

Com relação ao VaR 99%, observa-se que a carteira OGD apresenta o maior risco de perda, com um valor de -6,395%, seguida pelo IBOV (-5,565%), UCRP (-5,050%) e Markowitz (-4,575%). No entanto, ao avaliar o retorno ajustado ao risco por meio da Razão de Calmar, constata-se que a carteira OGD supera os *benchmarks*, sugerindo que a estratégia proporciona um melhor equilíbrio entre retorno e risco de perda.

Analisando as métricas de assimetria e excesso de curtose, verifica-se que a carteira OGD exibe uma assimetria positiva (0,003), indicando que a distribuição dos retornos é quase simétrica, com uma leve tendência para valores extremos acima da média - ver gráfico 29 abaixo. Por outro lado, o excesso de curtose de 12,359 sugere que a distribuição dos retornos possui caudas mais pesadas em relação a uma distribuição normal.

Gráfico 29 – *Box-Plots* dos retornos brutos das carteiras OGD, OGD_80 Normal, OGD_40 Normal, OGD_20 Normal e OGD_10 Normal para dados *IBOV*



Fonte: Elaborado pelo autor (2024).

A análise dos resultados obtidos a partir da aplicação da restrição de VaR, calculada com base na suposição de normalidade dos retornos, revela aspectos interessantes sobre a relação entre risco e retorno: o Gráfico 30 evidencia que a carteira OGD padrão, sem restrições adicionais de VaR, superou todas as demais estratégias em termos de retorno acumulado ao longo do período analisado. Esse resultado sugere que, embora a imposição de restrições de VaR possa ser eficaz na redução do risco, ela também pode limitar o potencial de retorno da carteira.

Gráfico 30 – Comparação de Desempenho e Risco: Análise das Carteiras s OGD_80 Normal, OGD_40 Normal, OGD_20 Normal e OGD_10 Normal para dados IBOV



Fonte: Elaborado pelo autor (2024).

Ao examinar as métricas de desempenho e risco apresentadas na Tabela 14, observa-se que as carteiras com maior disposição ao risco (80% e 40%) apresentam os maiores valores de TCAC e Riqueza Terminal. Conforme a disposição ao risco é reduzida para 20% e 10%, esses valores diminuem consideravelmente, reforçando a noção de que um maior controle de risco pode restringir os retornos potenciais. Em contrapartida, as métricas de volatilidade anualizada, semi-desvio, máximo *drawdown*, *drawdown* médio e pior perda exibem uma clara tendência de redução à medida que a disposição ao risco das carteiras é diminuída. Essa constatação indica que a imposição de restrições de VaR efetivamente controla o risco das carteiras, limitando a volatilidade e as perdas potenciais.

Tabela 14 – Comparação de Desempenho e Risco: Carteiras OGD com suposição de distribuição Normal para dados IBOV

Métricas	OGD_80 Normal	OGD_40 Normal	OGD_20 Normal	OGD_10 Normal
TCAC(%)	16,558	16,819	14,043	11,687
Riqueza Terminal	17,742	18,503	11,780	7,961
Volatilidade Anualizada (%)	24,863	14,852	7,539	3,635
Semi-Desvio	0,021	0,013	0,006	0,003
Máximo <i>Drawdown</i> (%)	-23,274	-13,444	-7,423	-3,681
<i>Drawdown</i> Médio (%)	-8,622	-4,892	-2,897	-1,445
Pior Perda (%)	-21,559	-12,918	-6,706	-3,339
VaR 95 (%)	-3,304	-1,991	-1,030	-0,493
VaR 99 (%)	-5,253	-3,118	-1,624	-0,798
CVaR 95 (%)	-4,609	-2,789	-1,436	-0,697
CVaR 99 (%)	-7,228	-4,3604	-2,2655	-1,0995
Razão de Calmar	0,275	0,421	0,721	1,648
Razão de Sortino	0,012	0,007	0,003	0,002
Recuperação Média	64,121	106,716	210,071	435,997
Assimetria	-0,393	-0,334	-0,215	-0,179
Excesso de Curtose	6,742	6,087	5,314	5,117

Fonte: Elaborado pelo autor (2024).

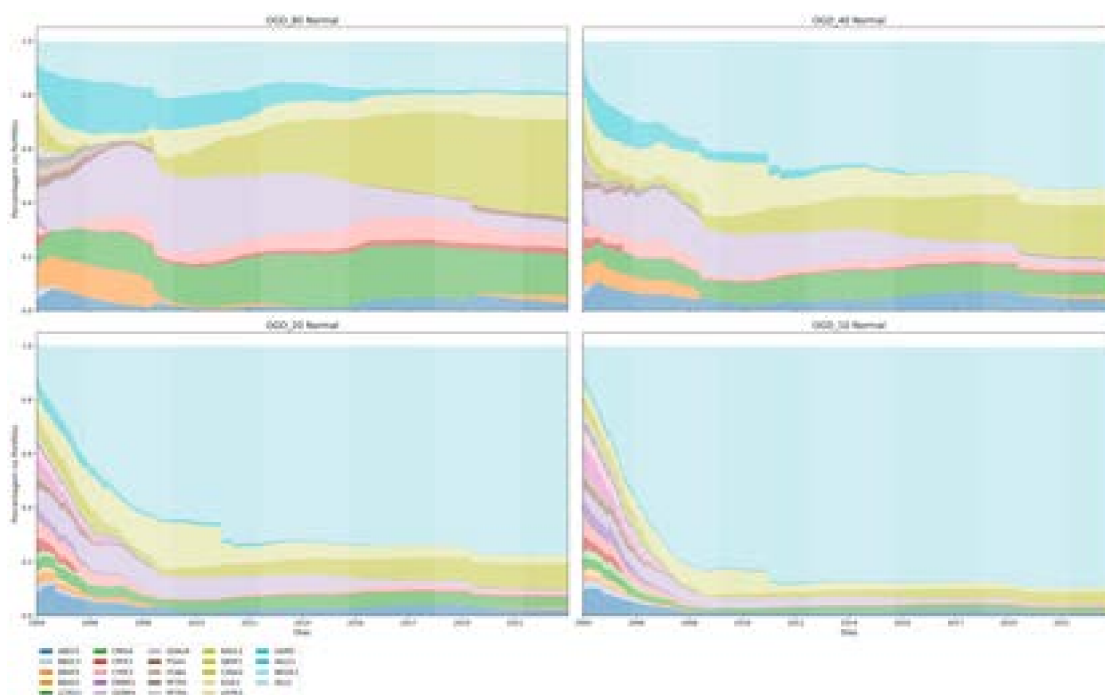
Os valores de VaR diminuem consistentemente conforme a disposição ao risco das carteiras é reduzida. Esse comportamento é esperado, uma vez que as restrições mais rígidas de VaR são impostas durante a etapa de projeção do algoritmo OGD.

Adicionalmente, a análise das Razões de Calmar e Sortino revela uma interação dinâmica entre o ajuste de retorno frente ao *drawdown* máximo e a sensibilidade ao risco das carteiras. A Razão de Calmar, indicador do retorno ajustado ao *drawdown* máximo, demonstra um incremento progressivo à medida que a disposição ao risco decresce. Esse fenômeno sugere uma capacidade superior das carteiras de menor risco em equilibrar o retorno obtido frente às perdas máximas experimentadas. Contrariamente, a Razão de Sortino, que avalia exclusivamente os retornos inferiores à média, apresenta uma redução conforme a aversão ao risco se intensifica. Este padrão sugere que, para carteiras mais conservadoras, o retorno ajustado para riscos específicos de baixa tende a ser menor.

Uma interpretação possível para a divergência observada entre as Razões de Calmar e Sortino encontra-se ao examinar detalhadamente o Gráfico 31. A análise sugere que, com o aumento das restrições ao risco, há uma inclinação para um maior investimento no ativo livre de risco. Este, agindo como um elemento

de estabilização, assegura um retorno garantido, diminuindo simultaneamente a exposição do portfólio à volatilidade do mercado. Embora tal alocação contribua significativamente para a mitigação do risco, ela implica, por outro lado, uma limitação ao potencial de obtenção de retornos extremamente positivos. Essa configuração resulta em uma Razão de Calmar mais elevada, especialmente notável nos portfólios com menor disposição ao risco (20% e 10%). Paralelamente, a preferência por uma maior alocação em ativos livres de risco conduz a uma redução no risco geral de baixa, dada a contribuição do retorno garantido desses ativos para contrabalançar possíveis perdas oriundas de outras classes de ativos. Contudo, esta estratégia de redução de risco tem como consequência retornos potenciais mais baixos, impactando diretamente na Razão de Sortino.

Gráfico 31 – Comparação dos pesos investidos nos portfólios OGD, OGD_80 Normal, OGD_40 Normal, OGD_20 Normal e OGD_10 Normal para dados IBOV

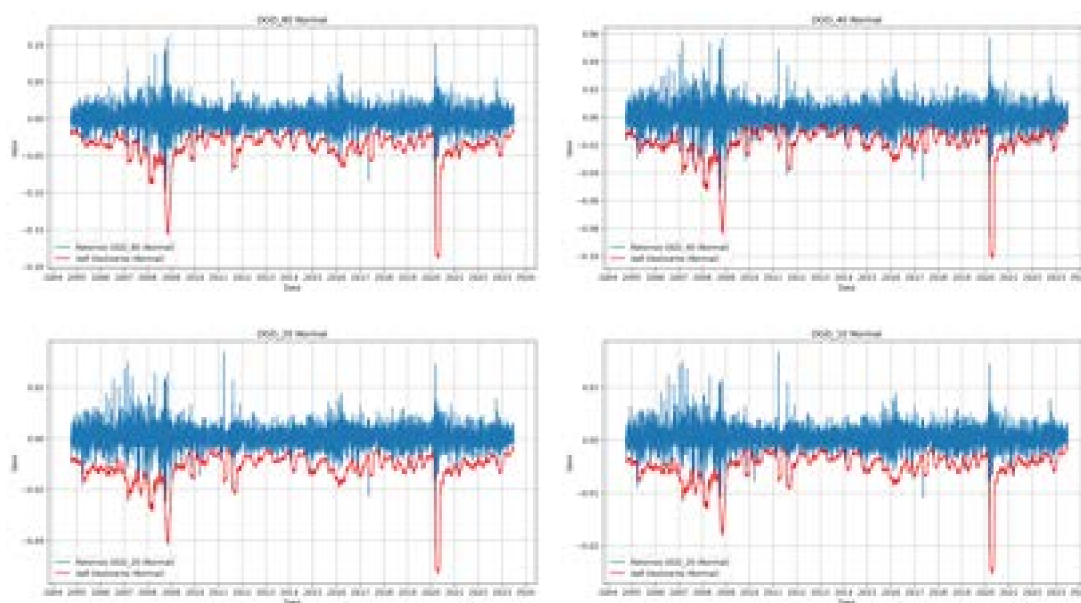


Fonte: Elaborado pelo autor (2024).

No Gráfico 32, apresenta-se os retornos juntamente com o VaR estimado, que serve como um parâmetro de restrição para cada atualização dos retornos do portfólio analisado. De maneira geral, observa-se que a restrição de VaR foi

respeitada, um padrão similar ao visto para dados *small caps* e que se repete na análise focada no IBOV. Similarmente, as instâncias em que o VaR foi ultrapassado coincidem com períodos de crise, eventos que seriam complexos de antecipar.

Gráfico 32 – Retornos e VaR Normal para dados IBOV



Fonte: Elaborado pelo autor (2024).

Analisando os dados relacionados à assimetria e ao excesso de curtose para as ações do IBOV na Tabela 14, percebe-se que, assim como observado para as *small caps*, a imposição da restrição do VaR usando a distribuição normal introduz uma distorção na assimetria dos retornos. Todas as carteiras com restrição apresentam assimetria negativa, variando de -0,393 (80%) a -0,179 (10%) - isso é nítido ao analisar novamente o Gráfico 29, com dados mais extremos nos quantis abaixo dos 25% da distribuição na comparação a aqueles acima dos 75%. No entanto, é interessante notar que a magnitude dessa distorção é menor para as ações do IBOVESPA em comparação com as *small caps* (ver Tabela 9). Isso pode ser atribuído à maior estabilidade e liquidez das ações que compõem o índice, o que as torna menos propensas a eventos extremos e retornos assimétricos. Quanto ao excesso de curtose, observa-se que ele diminui à medida que a restrição do VaR se torna mais rígida, variando de 6,742 (80%) a 5,117 (10%). Apesar dessa redução, o

excesso de curtose ainda é alto para todas as carteiras com restrição, indicando que os retornos do IBOVESPA apresentam caudas mais pesadas do que a distribuição normal. Esses resultados sugerem que, embora a distribuição normal possa ser mais adequada para modelar os retornos das ações do IBOVESPA em comparação com as *small caps*, ela ainda pode não capturar completamente as características desses retornos, especialmente na presença de caudas pesadas.

De maneira geral, embora todas as métricas diretas de risco analisadas, como Volatilidade Anualizada, Semi-Desvio, Máximo *Drawdown*, *Drawdown* Médio, Pior Perda, VaR 95%, VaR 99%, CVaR 95% e CVaR 99%, tenham apresentado substancial queda no risco com a imposição de restrições mais rígidas no algoritmo de projeção, a análise da assimetria e do excesso de curtose indica que outras distribuições podem ser mais aplicáveis. Distribuições como a t de Student são capazes de acomodar com mais flexibilidade caudas pesadas, fornecendo uma representação mais verossímil do comportamento dos retornos financeiros.

A próxima subseção se concentrará na análise dos resultados gerados pelo algoritmo 5, que calcula o VaR sob a premissa de que os retornos seguem a distribuição t de Student para as ações do IBOVESPA. Ela ajudará a corroborar com o exposto até aqui.

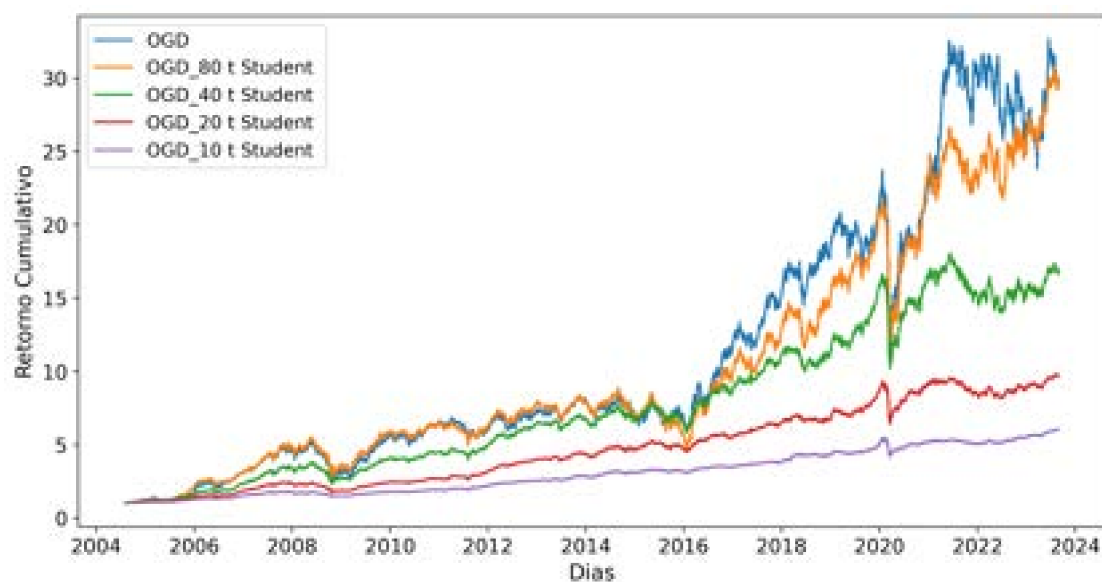
6.2.4 OGD com restrição VaR para dados IBOV - Resultados considerando retornos com distribuição t de Student

Ao analisar os resultados obtidos pela aplicação da restrição de VaR calculada com base na distribuição t de Student, em comparação com a carteira OGD padrão, observam-se importantes diferenças no desempenho e no perfil de risco das estratégias.

Começando pela análise visual do Gráfico 33, nota-se que a carteira OGD padrão apresenta a maior riqueza acumulada ao longo do período, seguida de perto pela carteira OGD_80 t Student. Conforme a disposição ao risco é reduzida nas carteiras 40%, 20%, 10%, observa-se uma diminuição gradual no retorno acumulado, sugerindo que a imposição de restrições mais rígidas de VaR limita o potencial de crescimento da carteira, ajustando risco-retorno de acordo com as preferências do

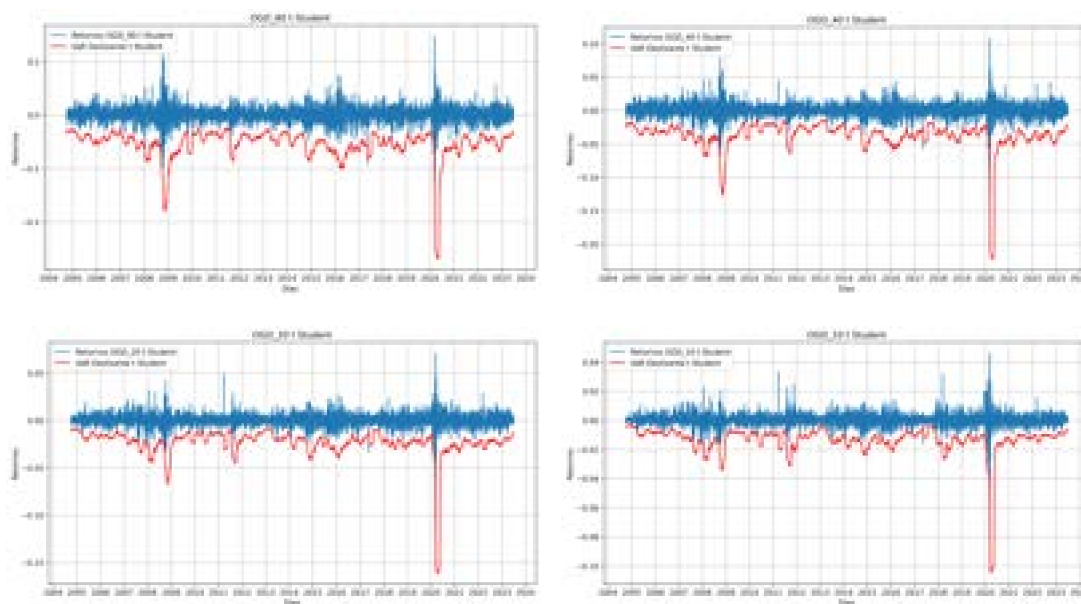
investidor. Além disso, ao avaliar o Gráfico 34, fica claro que a restrição de VaR foi novamente respeitada, assim como outros casos já avaliados, com o VaR tendo sido sobrepujado apenas em períodos de crise, cuja antecipação é mais restrita.

Gráfico 33 – Comparação entre os portfólios OGD, OGD_80 t Student, OGD_40 t Student, OGD_20 t Student e OGD_10 t Student para dados IBOV



Fonte: Elaborado pelo autor (2024).

Gráfico 34 – Retornos e VaR t de Student para dados IBOV



Fonte: Elaborado pelo autor (2024).

Ao examinar as métricas de desempenho e risco na Tabela 15, constata-se que a carteira OGD_80 t Student apresenta um TCAC de 19,838% e uma riqueza terminal de 29,867, valores muito próximos aos da carteira OGD padrão (19,841% e 29,881, respectivamente, conforme visto na Tabela 13). Isso indica que, mesmo com a imposição da restrição de VaR baseada na distribuição t de Student, a carteira 80% consegue capturar grande parte do potencial de retorno da estratégia original, mas apresentando métricas de risco mais comportadas, como volatilidade anualizada de 21,085% (26,179% no OGD padrão), um máximo *drawdown* de -22,708% (-27,043% no OGD padrão) e VaR 99% de -5,317 (-6,395 no OGD padrão).

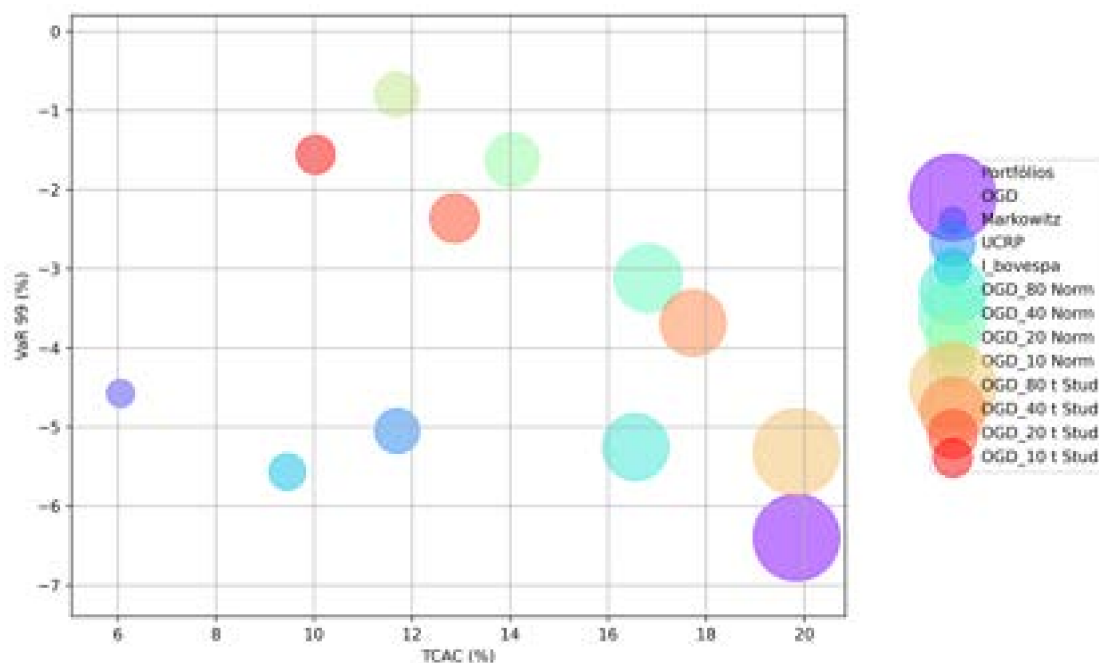
Tabela 15 – Comparação de Desempenho e Risco: Análise das Carteiras OGD_80 t Student, OGD_40 t Student, OGD_20 t Student, OGD_10 t Student para dados IBOV

Métricas	OGD_80 t Student	OGD_40 t Student	OGD_20 t Student	OGD_10 t Student
TCAC (%)	19,838	17,733	12,866	10,033
Riqueza Terminal	29,867	17,422	9,6962	6,017
Volatilidade Anualizada (%)	21,085	17,671	12,524	7,272
Semi-Desvio	0,020	0,015	0,010	0,006
Máximo <i>Drawdown</i> (%)	-22,708	-18,032	-15,650	-10,430
<i>Drawdown</i> Médio (%)	-8,328	-6,717	-4,316	-2,848
Pior Perda (%)	-15,831	-13,032	-11,650	-7,127
VaR 95%(%)	-3,077	-2,298	-1,468	-0,904
VaR 99 (%)	-5,317	-3,686	-2,361	-1,557
CVaR 95 (%)	-4,569	-3,273	-2,138	-1,362
CVaR 99 (%)	-7,557	-5,362	-3,587	-2,334
Razão de Calmar	0,393	0,374	0,410	0,432
Razão de Sortino	0,018	0,014	0,012	0,009
Recuperação Média	66,617	91,953	143,756	235,101
Assimetria	-0,256	-0,381	-1,306	-1,325
Excesso de Curtose	4,764	7,224	12,737	15,338

Fonte: Elaborado pelo autor (2024).

O Gráfico 35 permite analisar o desempenho das duas estratégias de projeção de maneira mais visual. Começando pela carteira OGD padrão, observa-se um alto retorno (TCAC de 19,841%) associado a um maior risco (VaR 99% de -6,395%). Essa combinação resulta na maior riqueza terminal (29,881) entre todas as carteiras analisadas, conforme indicado pelo tamanho do círculo no gráfico. Ao analisar as carteiras otimizadas pela distribuição Normal, nota-se uma redução gradual no TCAC e no VaR 99% à medida que a disposição ao risco diminui (de 80% para 10%). Essa relação é claramente visível no gráfico, com as carteiras dispostas em uma linha descendente da esquerda para a direita. A redução no risco (VaR 99%) é acompanhada por uma diminuição na riqueza terminal, conforme indicado pelos tamanhos decrescentes dos círculos. Por outro lado, as carteiras otimizadas pela distribuição t de Student apresentam um comportamento distinto. Embora o TCAC e o VaR 99% também diminuam à medida que a disposição ao risco é reduzida, a magnitude dessa redução é menor em comparação às carteiras otimizadas pela Normal. Isso pode ser observado no gráfico, onde as carteiras otimizadas pela t de Student estão posicionadas mais próximas da carteira OGD padrão.

Gráfico 35 – Desempenho das carteiras: TCAC (%) vs. VaR 99 (%) com Riqueza Terminal como tamanho do círculo para dados IBOV



Fonte: Elaborado pelo autor (2024).

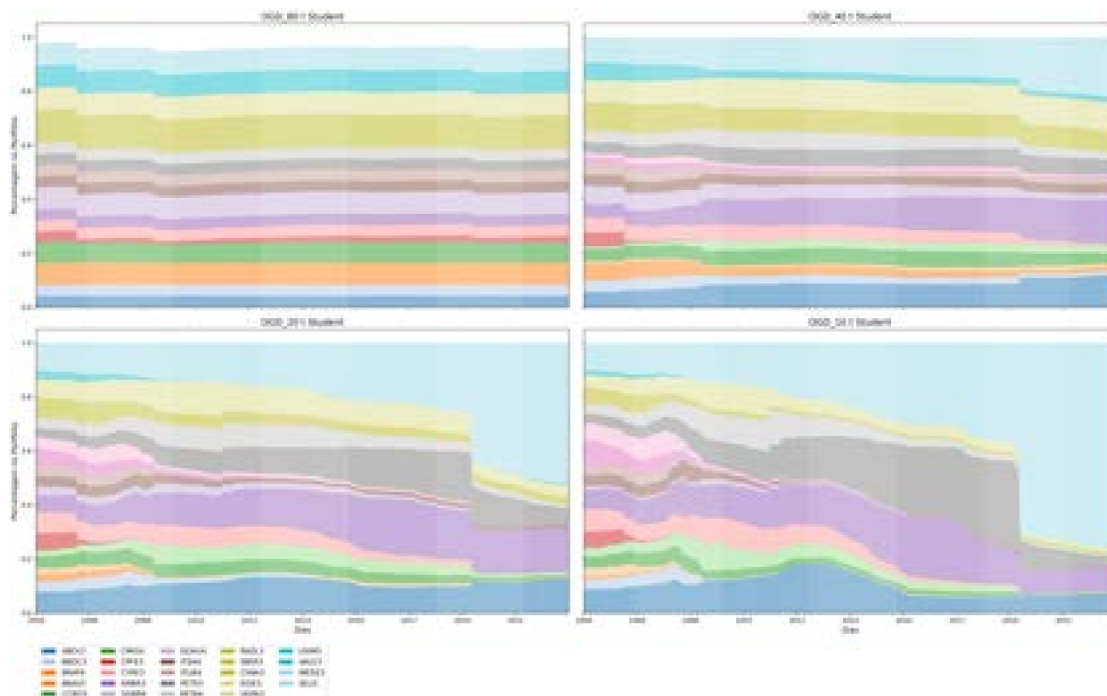
Analisando agora as medidas de desempenho ajustado ao risco, vemos que a razão de Sortino apresenta valores decrescentes conforme a disposição ao risco diminui. Essa redução sugere que, embora as carteiras com menor disposição ao risco apresentem uma menor exposição a perdas, elas também podem ter um desempenho inferior em termos de retorno ajustado ao *downside risk*. Por outro lado, a Razão de Calmar, que mede o retorno em excesso em relação ao máximo *drawdown*, apresenta uma tendência crescente, ou seja, os resultados sugerem que, à medida que a disposição ao risco diminui, as carteiras com restrição de VaR baseada na distribuição t de Student oferecem um melhor equilíbrio entre o retorno e o máximo *drawdown*, indicando uma maior eficiência na geração de retornos em relação ao risco de perda máxima.

Esse comportamento divergente entre a razão de Sortino e de Calmar foi similar nas duas distribuições testadas. Uma possível explicação para a divergência reside nas características da distribuição dos retornos das carteiras. As carteiras

com menor disposição ao risco 20% e 10% apresentaram assimetria negativa mais acentuada e excesso de curtose mais elevado em comparação às carteiras com maior disposição ao risco 80% e 40%. A Razão de Sortino, ao considerar apenas o *downside risk*, é mais sensível à assimetria negativa, penalizando as carteiras com valores mais extremos a esquerda da média. Já a Razão de Calmar, ao levar em conta o máximo *drawdown*, é mais influenciada pela magnitude das perdas extremas, que tendem a ser maiores em distribuições com caudas mais pesadas. Conseqüentemente, as carteiras com menor disposição ao risco, apesar de apresentarem uma Razão de Sortino mais baixa, podem exibir uma Razão de Calmar mais alta devido à sua maior capacidade de limitar perdas extremas.

Outro ponto que deve ser levado em consideração é a comparação da composição das carteiras, plotada no Gráfico 36. Conforme a disposição ao risco diminui, o algoritmo de otimização tende a selecionar ativos com menor correlação entre si, visando reduzir o risco da carteira. Essa maior diversificação pode resultar em retornos mais estáveis, porém potencialmente menores, o que se refletiria em uma Razão de Sortino mais baixa.

Gráfico 36 – Comparação dos pesos investidos nos portfólios OGD_80 t Student, OGD_40 t Student, OGD_20 t Student e OGD_10 t Student para dados IBOV



Fonte: Elaborado pelo autor (2024).

No entanto, a maior diversificação também pode contribuir para a redução do máximo *drawdown*, uma vez que a presença de ativos menos correlacionados tende a suavizar as perdas da carteira durante períodos de estresse do mercado. Isso explicaria o aumento da Razão de Calmar nas carteiras com menor disposição ao risco, indicando uma melhor proteção contra perdas extremas. Comparando os valores da Razão de Sortino para as duas distribuições (Tabelas 14 e 15), observa-se que as carteiras otimizadas com base na distribuição t de Student apresentam valores consistentemente mais altos em comparação às carteiras correspondentes baseadas na distribuição Normal. Isso sugere que, ao considerar a distribuição t de Student, as carteiras otimizadas pelo algoritmo OGD são mais eficientes em termos de retorno ajustado ao *downside risk*.

Essa diferença pode ser atribuída às características das distribuições utilizadas. A distribuição t de Student possui caudas mais pesadas em comparação à

distribuição Normal, o que significa que ela atribui maior probabilidade a eventos extremos, tanto positivos quanto negativos. Ao incorporar essa informação na otimização das carteiras, o algoritmo OGD pode selecionar ativos que oferecem um melhor equilíbrio entre retorno e *downside risk*, levando a uma Razão de Sortino mais alta.

Além disso, é interessante notar que, para ambas as distribuições, a Razão de Sortino diminui conforme a disposição ao risco diminui. Isso indica que, independentemente da distribuição utilizada, as carteiras com menor disposição ao risco tendem a apresentar uma menor eficiência em termos de retorno ajustado ao *downside risk*. No entanto, a magnitude dessa diminuição é menor para as carteiras baseadas na distribuição t de Student, sugerindo que essa abordagem pode oferecer uma melhor proteção contra o *downside risk*, mesmo em carteiras com menor disposição ao risco.

Ao considerar a assimetria e o excesso de curtose tanto para os testes realizados para Small Caps (Tabela 11 para o caso t de Student e Tabela 10 para o caso com distribuição normal) quanto aqueles ligados ao Ibovespa (Tabelas 15 e 14, respectivamente), algumas similaridades são constatadas. De fato:

- i A assimetria é negativa e seu valor decai de acordo com o aumento da restrição, ou seja, quando quanto mais restrita é a exigência, mas negativa se torna a assimetria; além disso, a assimetria antes positiva no OGD padrão (ver Tabela 9 para o caso *small caps* e Tabela 13 para o caso IBOV), passa a ser negativa para ambas estratégias.
- ii Para a restrição de 80% vemos uma diminuição do excesso de curtose, no caso do Ibovespa essa redução é significativa, passando de 12,359 na carteira OGD padrão para 4,764 na carteira de 80% no caso t de Student (6,742 para o a distribuição normal). Além disso, esse valor começa a aumentar na medida que impomos restrições mais rígidas.
- iii As carteiras com restrições de VaR baseadas na distribuição t de Student apresentam assimetrias menos acentuadas em comparação com as carteiras baseadas na distribuição Normal. Isso sugere que a distribuição t de Student

pode ser mais adequada para capturar as características dos retornos, especialmente a presença de caudas pesadas, resultando em uma representação mais precisa da assimetria.

Todavia, algumas diferenças também foram constatadas:

- i Apesar da assimetria dos retornos sob a distribuição t de Student no caso IBOV também ter sido negativo, variando de -0,256 (OGD_80 t Student) a -1,325 (OGD_10 t Student), esses valores são comparáveis aos observados sob a suposição de normalidade, sugerindo que a mudança na distribuição não teve um impacto significativo na assimetria dos retornos do IBOVESPA na comparação aos vistos para o *small caps*.
- ii Já o excesso de curtose foi menor ao utilizar a distribuição t de Student, variando de 4,764 (OGD_80 t Student) a 15,338 (OGD_10 t Student), em comparação com os valores obtidos sob a suposição de normalidade para o caso IBOV. Esse resultado indica que a distribuição t de Student é mais apropriada para modelar as caudas pesadas dos retornos do IBOVESPA, embora a redução no excesso de curtose seja menos acentuada do que a observada para as *small caps*.

Em resumo, a análise das métricas de assimetria e curtose dos retornos das carteiras do IBOVESPA e das *small caps*, considerando a distribuição t de Student, reforça a importância de se utilizar uma distribuição de probabilidade que seja capaz de acomodar as características específicas dos ativos, como a presença de caudas pesadas e assimetria. Embora a distribuição t de Student tenha se mostrado mais adequada do que a distribuição normal, especialmente para as *small caps*, é fundamental avaliar continuamente a adequação das distribuições utilizadas e considerar outras alternativas que possam fornecer uma representação ainda mais precisa do comportamento dos retornos financeiros.

6.3 RESULTADOS DO OGD COM RESTRIÇÃO *CONDITIONAL VALUE AT RISK*

Dada a clara relação entre as métricas VaR e CVaR, além das similaridades entre os algoritmos propostos para ambas, esta seção utiliza dados e objetivos similares aos vistos na seção 6.2. Não menos importante, uma análise comparativa entre a estratégia OGD padrão e as estratégias de investimento adotadas como *benchmarks* revelou que a estratégia OGD apresenta superioridade em termos de desempenho geral. Notou-se também que as estratégias de VaR para a distribuição normal, especialmente aquelas com menor exposição ao risco, demonstraram uma relação risco-retorno mais favorável do que a estratégia OGD padrão. Além disso, a suposição de que os retornos seguem uma distribuição t de Student resultou em uma melhoria significativa na relação risco-retorno em comparação com as estratégias VaR baseadas na distribuição normal e OGD padrão.

Nesta seção, apresentamos os detalhes da implementação da metodologia que utiliza o CVaR como critério de restrição na otimização de portfólios. O objetivo é ilustrar os procedimentos e as decisões envolvidas na aplicação dessa abordagem, tanto para ações listadas no IBOV quanto para aquelas classificadas como *Small Caps*. Assim como na metodologia baseada no VaR, o CVaR é utilizado para estabelecer um limiar de risco que define a perda máxima tolerável pelo investidor. No entanto, enquanto o VaR considera apenas um ponto específico da distribuição dos retornos, o CVaR oferece uma medida mais robusta ao capturar as perdas esperadas além desse ponto crítico.

A inexistência de uma distribuição específica na implementação do CVaR é uma característica importante que distingue essa medida de risco das abordagens que dependem de tal pressuposto, como o VaR baseado, por exemplo, na distribuição normal ou na distribuição t de Student. De fato, o VaR requer a suposição de uma distribuição específica dos retornos para calcular o valor crítico e, conseqüentemente, o risco associado. Enquanto isso, o CVaR é calculado de forma que considera todas as perdas além de um certo limiar (o VaR) e calcula a média dessas perdas, o que permite capturar o risco de cauda sem precisar fazer suposições sobre a forma da distribuição dos retornos. Isso é particularmente útil em mercados financeiros, onde

os retornos frequentemente exibem caudas pesadas e assimetrias que não são bem modeladas por distribuições padrão como a normal.

Para simplificar a comparação, também foram adotados quatro patamares de tolerância ao risco: 80%, 40%, 20% e 10%, de forma análoga ao procedimento adotado com o VaR. Esses patamares são aplicados para ajustar o nível de aversão ao risco na estratégia de otimização, permitindo uma adaptação mais precisa às preferências individuais dos investidores.

Assim como na implementação do VaR, os intervalos de tempo dos dados utilizados na aplicação do CVaR são exatamente os mesmos, com uma janela temporal deslizante de 44 dias, essencial tanto na teoria quanto na prática de otimização de portfólios *online*. Essa janela se ajusta a cada novo dado, mantendo o CVaR atualizado com as condições de mercado e facilitando a compreensão da volatilidade dos retornos. Adicionalmente, os modelos exigem a seleção do parâmetro η , responsável por controlar a magnitude dos passos na direção do gradiente negativo durante a otimização. A taxa de aprendizado, $\eta = 0,001$, determina o tamanho dos ajustes nos parâmetros do modelo com o intuito de minimizar a função de custo *Regret*. O parâmetro ν , definindo o nível de confiança para a análise de risco do portfólio, é fixado em $\nu = 0,01$ (1%), calculando o VaR ao nível de 99%.

Diante dos resultados mencionados e com o objetivo de proporcionar uma análise mais detalhada, optou-se por apresentar exclusivamente os resultados comparando as metodologias CVaR perante o OGD Padrão e o VaR t de Student. Essa escolha permite análises distintas e gráficos mais detalhados, que serão apresentados nas subseções a seguir. Na seção 6.3.1, são comparados os modelos em questão usando dados *small caps*. Por fim, na seção 6.3.2, o mesmo é efetuado para carteiras formadas pelo IBOV.

6.3.1 OGD com restrição CVaR para dados *Small Caps*

Antes de começar, é importante destacar que todas as métricas relevantes para a comparação com o modelo VaR baseado na distribuição t de Student estão devidamente apresentadas na Tabela 11.

Assim como apresentado nos resultados anteriores, são plotados no Gráfico 37 abaixo os retornos acumulados das diferentes estratégias de investimento utilizando a restrição de CVaR. Na análise dos gráficos de retorno acumulado, observa-se que a estratégia de projeção utilizando CVaR exibe um comportamento de crescimento mais estável e controlado ao longo do tempo, em comparação com a estratégia baseada no VaR utilizando a distribuição t de Student. A estratégia CVaR, como esperado, prioriza a mitigação de perdas extremas, resultando em uma curva de retorno mais suave, especialmente evidente nos períodos de alta volatilidade, como a crise de 2020. Essa abordagem, embora menos agressiva em termos de retorno absoluto, demonstra uma capacidade superior de preservação de capital durante períodos de estresse de mercado.

Gráfico 37 – Comparação entre os portfólios OGD, OGD_80 CVaR, OGD_40 CVaR, OGD_20 CVaR e OGD_10 CVaR para dados *small caps*



Fonte: Elaborado pelo autor (2024).

Comparativamente, ao observar o Gráfico 23, a estratégia baseada no VaR com t Student, embora apresente um retorno acumulado mais elevado em alguns momentos, como nos períodos de crescimento acelerado antes de 2020, exibe uma maior volatilidade e uma queda mais acentuada durante os períodos de crise. Essa diferença reflete a natureza das duas metodologias: enquanto o CVaR está mais focado na redução do risco de cauda, o VaR com t Student permite maior exposição

ao risco em busca de retornos mais elevados, mas com maior suscetibilidade a *drawdowns* significativos.

Essa análise é suportada por métricas presentes nas tabelas 11 e 16. A estratégia de projeção utilizando o CVaR demonstra um comportamento de crescimento mais estável e controlado. Essa estabilidade é suportada por métricas como a Volatilidade Anualizada, que é significativamente menor para as carteiras CVaR em todos os níveis de disposição ao risco, variando de 19,78% para OGD_80 CVaR a apenas 3,25% para OGD_10 CVaR. Em comparação, as carteiras otimizadas com VaR t Student apresentam maior volatilidade, alcançando até 29,66% para OGD_80 t Student. Isso sugere que a estratégia CVaR é mais eficaz na mitigação das flutuações extremas de mercado, resultando em uma curva de retorno mais suave, especialmente evidente durante períodos de alta volatilidade, como a crise de 2020.

Tabela 16 – Comparação de Desempenho e Risco: Análise das Carteiras OGD_80 CVaR, OGD_40 CVaR, OGD_20 CVaR, OGD_10 CVaR para dados *small caps*

Métricas	OGD_80 CVaR	OGD_40 CVaR	OGD_20 CVaR	OGD_10 CVaR
TCAC (%)	17,532	16,218	13,034	11,293
Riqueza Terminal	9,879	8,464	5,594	4,497
Volatilidade Anualizada (%)	19,782	12,047	6,185	3,245
Semi-Desvio	0,017	0,010	0,005	0,003
Máximo <i>Drawdown</i> (%)	-38,244	-21,686	-12,963	-8,849
<i>Drawdown</i> Médio (%)	-8,992	-4,716	-1,567	-0,522
Pior Perda (%)	-13,583	-8,008	-4,080	-2,123
VaR 95 (%)	-2,187	-1,313	-0,657	-0,328
VaR 99 (%)	-3,351	-2,023	-1,022	-0,520
CVaR 95 (%)	-2,853	-1,691	-0,845	-0,427
CVaR 99 (%)	-5,234	-3,105	-1,571	-0,808
Razão de Calmar	0,363	0,512	0,768	1,276
Razão de Sortino	0,009	0,006	0,003	0,001
Recuperação Média	85,283	136,118	264,068	502,213
Assimetria	-1,200	-1,131	-1,074	-1,084
Excesso de Curtose	15,303	12,724	12,670	12,683

Fonte: Elaborado pelo autor (2024).

A Pior Perda e o Máximo *Drawdown* são outras métricas que suportam essa observação. A Pior Perda para as carteiras CVaR é consistentemente menor, com -13,58% para OGD_80 CVaR, comparado a -25,81% para OGD_80 t Student. De forma similar, o Máximo *Drawdown* para a carteira OGD_80 CVaR é de -48,24%,

enquanto para a carteira OGD_80 t Student é de -27,78%, refletindo a natureza da metodologia CVaR em priorizar a preservação do capital durante períodos de estresse de mercado, mesmo que isso implique uma menor agressividade no crescimento dos retornos.

Por outro lado, a estratégia baseada no VaR com t Student, enquanto apresenta um TCAC competitivo em certos momentos (12,15% para OGD_80 t Student comparado a 17,53% para OGD_80 CVaR) mostra uma maior suscetibilidade a perdas significativas durante crises. Isso é evidenciado pelo *Drawdown* Médio, que é maior nas carteiras t Student (ex: -8,74% para OGD_80 t Student comparado a -8,99% para OGD_80 CVaR), e pela maior Recuperação Média para as carteiras CVaR, que indica a eficiência na recuperação após períodos de perda.

Além disso, o foco do CVaR em mitigar perdas extremas é corroborado pelas métricas de VaR 95% e 99% e CVaR 95% e 99%, onde os valores são consistentemente menores nas carteiras CVaR, sugerindo uma menor exposição a riscos de cauda em comparação com as carteiras t Student.

Ainda comparando os dados das duas tabelas, focando nas métricas de desempenho ajustado ao risco, observou-se um comportamento interessante ao comparar as carteiras otimizadas com restrição de CVaR e aquelas baseadas no VaR sob a suposição de distribuição t de Student. A Razão de Sortino apresenta valores consistentemente mais baixos nas carteiras otimizadas com CVaR, independentemente do nível de disposição ao risco. Essa tendência indica que, embora o CVaR seja eficaz em limitar as perdas extremas, ele também pode resultar em um retorno ajustado ao risco inferior, especialmente no que tange ao *downside risk*. A penalização da Razão de Sortino pode ser atribuída à natureza conservadora do CVaR, que tende a restringir a exposição a riscos de cauda, potencialmente limitando o *upside* da carteira.

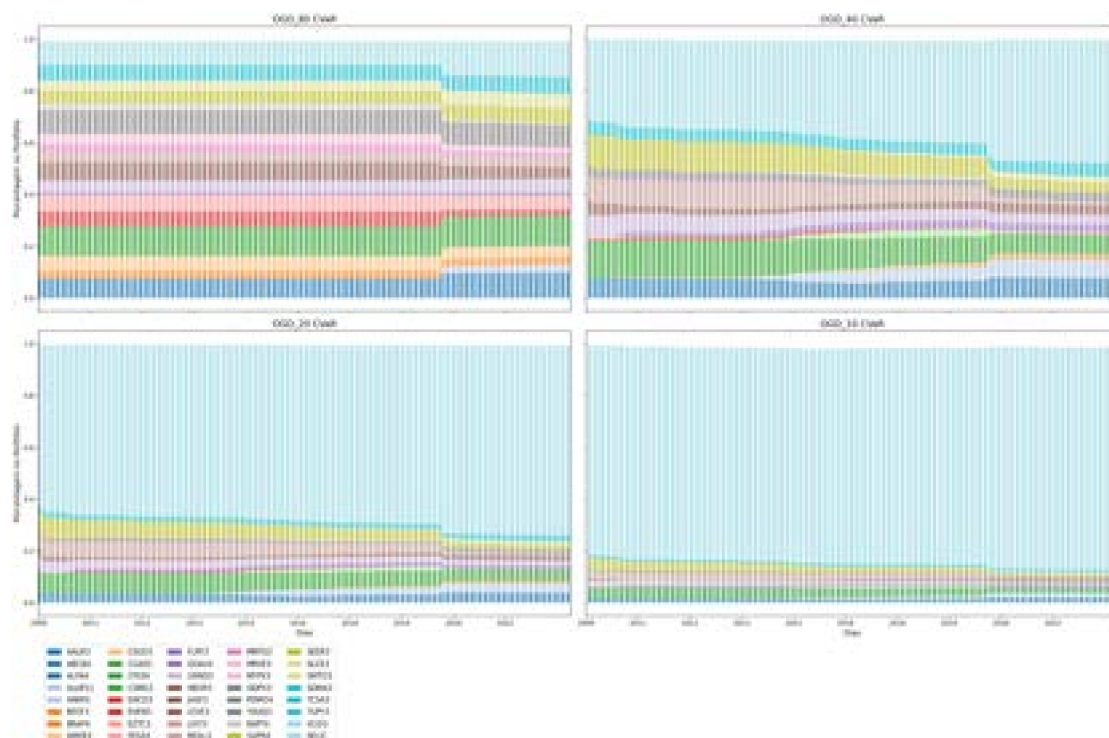
Por outro lado, a Razão de Calmar mostra uma melhoria significativa nas carteiras otimizadas com CVaR, especialmente à medida que a disposição ao risco diminui. Isso sugere que, mesmo que o CVaR leve a uma redução no retorno ajustado ao *downside risk*, ele compensa ao proporcionar uma proteção superior contra o máximo *drawdown*. Dessa forma, as carteiras CVaR parecem oferecer

um equilíbrio mais eficiente entre o retorno e a proteção contra grandes quedas, especialmente em ambientes de mercado voláteis.

A divergência entre a Razão de Sortino e a Razão de Calmar pode ser explicada pela forma como cada métrica interage com as características dos retornos das carteiras. As carteiras otimizadas com CVaR, que mostraram maior assimetria negativa e excesso de curtose, indicam uma distribuição de retornos mais concentrada e com eventos extremos. A Razão de Sortino, sendo mais sensível à assimetria negativa, penaliza essas carteiras, refletindo o impacto das perdas extremas à esquerda da média. Em contrapartida, a Razão de Calmar, ao considerar o máximo *drawdown*, favorece as carteiras CVaR pela sua capacidade superior de limitar perdas significativas, o que é especialmente relevante em contextos de maior volatilidade.

Além disso, outro ponto que deve ser levado em consideração (ver Gráfico 38) é a proporção investida em cada ativo ajustada de forma a minimizar o risco de cauda, o que tende a priorizar ativos com retornos mais estáveis e menor exposição a perdas extremas. Isso resulta em uma carteira que, embora possa ter retornos mais moderados, apresenta uma menor volatilidade negativa e *drawdowns* controlados, favorecendo uma Razão de Calmar mais alta.

Gráfico 38 – Comparação dos pesos investidos nos portfólios OGD_80 CVaR, OGD_40 CVaR, OGD_20 CVaR e OGD_10 CVaR para dados *small caps*

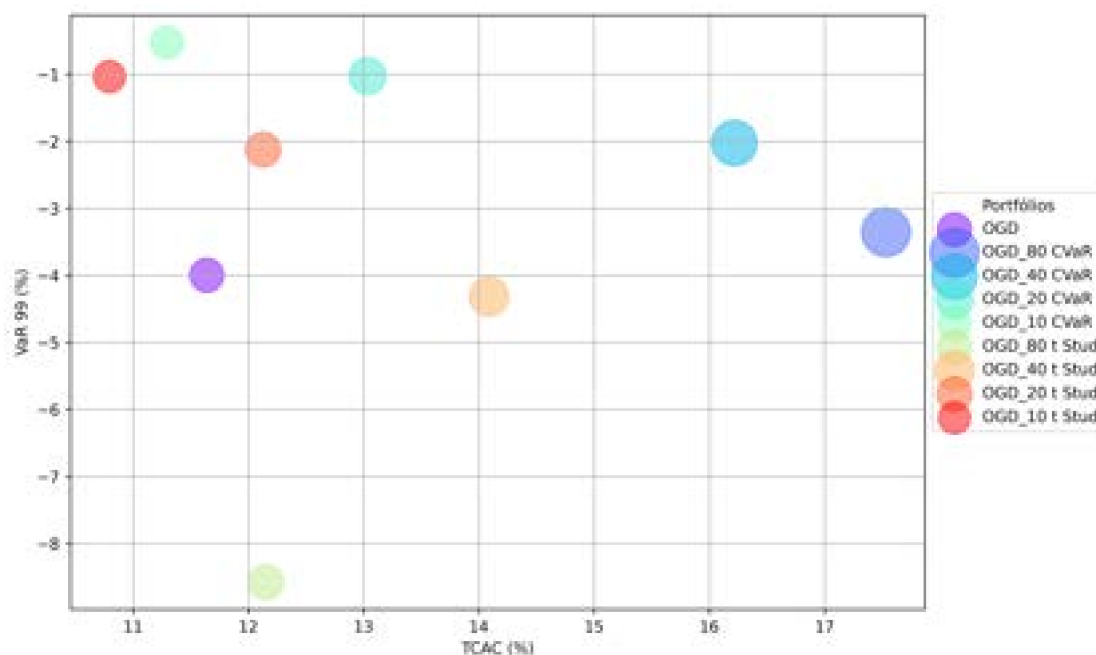


Fonte: Elaborado pelo autor (2024).

Por outro lado, a metodologia VaR t Student pode atribuir pesos maiores a ativos com maior potencial de retorno, mesmo que isso venha acompanhado de maior volatilidade e risco de cauda. Essa abordagem pode melhorar a Razão de Sortino, especialmente se os ativos com maior peso tiverem um bom desempenho durante períodos de alta volatilidade. No entanto, a maior exposição ao risco também pode aumentar o máximo *drawdown*, o que pode comprometer a Razão de Calmar se as perdas extremas se materializarem.

Seguindo a mesma estratégia utilizada na Seção 6.2 e o enfoque na comparação de modelos, plotou-se o Gráfico 39 que ilustra a comparação entre diferentes estratégias de otimização de portfólio. Neste gráfico de bolhas, três dimensões críticas de desempenho são novamente representadas: a Taxa de Crescimento Anual Composta (TCAC) no eixo x, o VaR 99% no eixo y, e a Riqueza Terminal através do tamanho das bolhas.

Gráfico 39 – Desempenho das carteiras: TCAC (%) vs. VaR (%) com Riqueza Terminal como tamanho do círculo para dados *small caps*



Fonte: Elaborado pelo autor (2024).

Observa-se, com grande margem, a superioridade da estratégia de minimização do CVaR, particularmente nas carteiras OGD_80 CVaR e OGD_40 CVaR. Essas carteiras não apenas apresentam TCAC significativamente superiores, mas também conseguem alcançar um nível de Riqueza Terminal mais elevado. Além disso, essas estratégias mantêm um perfil de risco mais controlado, conforme indicado pelos menores valores de VaR 99% em comparação com as carteiras otimizadas usando a distribuição t de Student.

Especificamente, as carteiras OGD_80 CVaR e OGD_40 CVaR destacam-se no gráfico com as maiores TCACs, respectivamente 17,53% e 16,22%, combinadas com as maiores Riquezas Terminais, de 9,879 e 8,464. Isso sugere que a metodologia de minimização do CVaR permite capturar um crescimento substancial do portfólio, ao mesmo tempo em que gerencia de maneira eficaz os riscos de cauda.

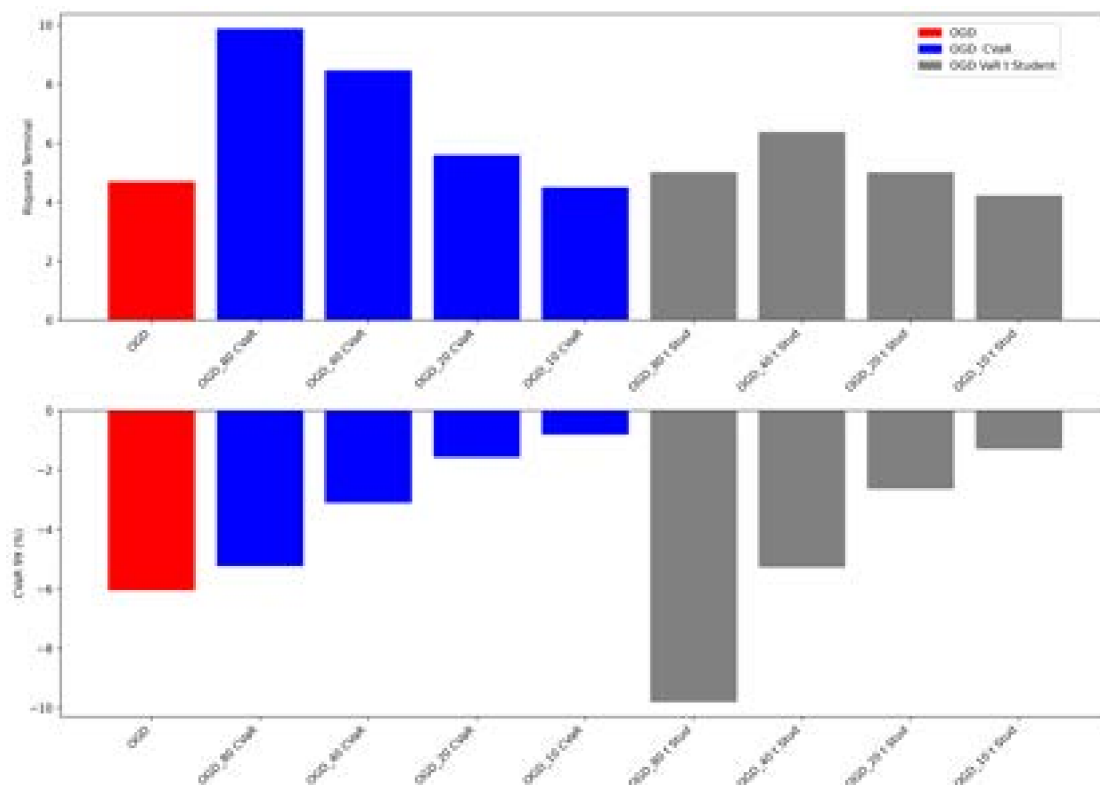
No que tange ao controle de risco, as estratégias baseadas no CVaR mostram uma clara vantagem. Os valores de VaR 99% para as carteiras OGD_80 CVaR

e OGD_40 CVaR são significativamente menores (-3,351% e -2,023%, respectivamente) em comparação com as carteiras otimizadas pela metodologia t Student, que apresentam valores de VaR 99% variando entre -8,574% e -1,031%. Este controle aprimorado de riscos extremos é crucial para investidores que buscam limitar perdas severas em períodos de maior volatilidade.

Por outro lado, as carteiras otimizadas com base no VaR t Student, embora ofereçam uma maior exposição ao risco, resultam em menores TCACs e menores Riquezas Terminais, além de um VaR 99% consideravelmente mais alto. Isso indica que, apesar de uma potencial recompensa em termos de retorno, essas carteiras estão mais sujeitas a perdas extremas, o que pode ser menos atraente para investidores avessos ao risco.

Uma outra comparação, trazida especificamente para esta parte do trabalho, faz uma análise entre o OGD em sua implementação na forma padrão, a projeção do CVaR, e o VaR t de Student. Essa comparação, plotada no Gráfico 40 através de duas imagens, mostra na parte superior a Riqueza Terminal e, na imagem inferior, o nível de CVaR 99% associado a cada estratégia.

Gráfico 40 – Análise Comparativa de Riqueza Terminal e CVaR entre a carteira OGD e estratégias de minimização de CVaR, e minimização de VaR com Distribuição t Student para dados *small caps*



Fonte: Elaborado pelo autor (2024).

Observa-se que o OGD em sua forma padrão, representado pela barra vermelha, alcança uma Riqueza Terminal de 4,681, que é significativamente inferior quando comparada às estratégias baseadas em CVaR, especialmente para os níveis de risco mais elevados (80% e 40%). As estratégias OGD_80 CVaR e OGD_40 CVaR apresentam Riquezas Terminais de 9,879 e 8,464, respectivamente, demonstrando um desempenho superior em termos de crescimento de capital.

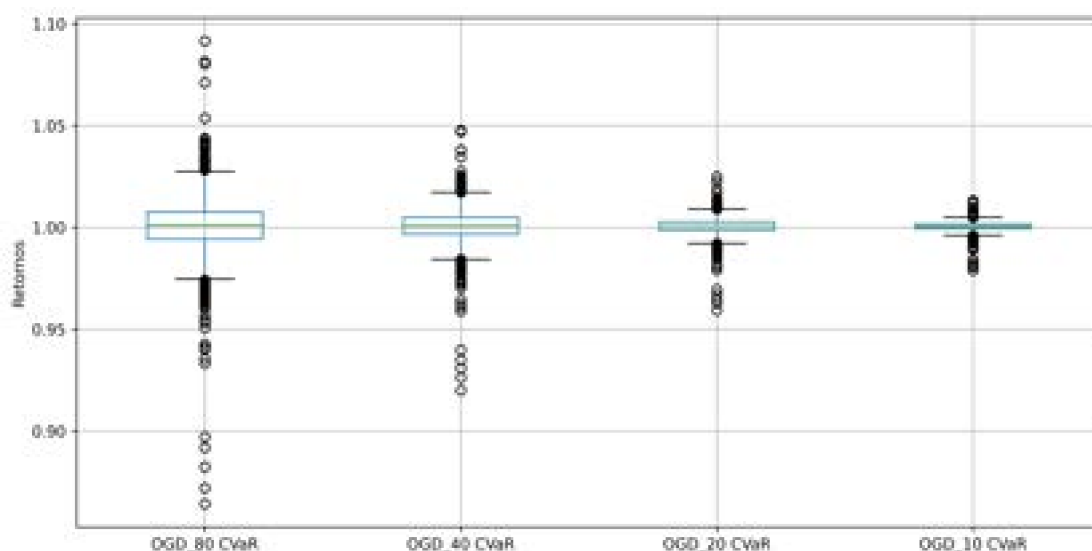
Por outro lado, ao analisar o CVaR 99%, vemos que a implementação padrão do OGD registra um valor de -11,925, o mais elevado entre todas as estratégias analisadas, indicando uma maior exposição ao risco de cauda. Em contraste, as estratégias baseadas em CVaR mantêm um controle mais rigoroso sobre o risco extremo, com o CVaR 99% variando de -5,234 (OGD_80 CVaR) a -0,808 (OGD_10

CVaR), evidenciando a eficácia dessa abordagem em proteger contra grandes perdas.

A comparação com as estratégias baseadas no VaR t de Student, representadas pelas barras cinzas, também reflete a superioridade do CVaR na etapa de projeção. Apesar da estratégia t de Student apresentar uma relação risco-retorno melhor que o OGD padrão, ela ainda fica aquém das estratégias CVaR, tanto em termos de Riqueza Terminal quanto na mitigação do risco extremo.

Uma última análise diz respeito às métricas de distribuição. Realizou-se uma comparação entre a distribuição do OGD na forma padrão (Gráfico 22), a distribuição VaR t de Student (Gráfico 27) e o CVaR (Gráfico 41) apresentados como *Box-Plots*.

Gráfico 41 – *Box-Plots* dos retornos brutos das carteiras OGD_80 CVaR, OGD_40 CVaR, OGD_20 CVaR e OGD_10 CVaR para dados *small caps*



Fonte: Elaborado pelo autor (2024).

Observa-se que o OGD exibe valores de quantis mais extremos. A mediana dos retornos para ambas as estratégias de projeção permanece próxima de 1,00, indicando uma tendência central similar. No entanto, a estratégia CVaR mostra uma redução mais acentuada na dispersão dos retornos conforme o nível de tolerância ao risco diminui, evidenciada pela contração das caixas dos *box-plots* e o encurtamento das linhas de extensão.

A volatilidade anualizada confirma essa observação, com a estratégia CVaR apresentando uma queda significativa de 19,782% (OGD_80) para 3,245% (OGD_10), comparada à redução de 29,659% para 4,197% na estratégia t de Student.

Adicionalmente, a estratégia CVaR revela uma assimetria negativa mais pronunciada, variando de -1,200 a -1,084, em comparação com -0,093 a -0,824 na estratégia t de Student, sugerindo que o CVaR é mais eficaz na penalização das perdas extremas, conforme esperado para mitigar riscos de cauda.

O excesso de curtose é significativamente maior na estratégia CVaR (15,303 a 12,683) em comparação com a t de Student (5,929 a 10,407), indicando uma distribuição mais leptocúrtica e com caudas mais pesadas na abordagem CVaR.

Baseado nesses resultados, observa-se que a implementação do CVaR demonstrou eficácia em proporcionar uma proteção robusta contra os riscos de cauda, sendo especialmente eficiente na contenção de perdas extremas. Essa característica revela-se fundamental para investidores que buscam minimizar a exposição a cenários adversos severos. As métricas associadas ao CVaR, como a Razão de Calmar e o Semi-Desvio, indicam que essa abordagem não apenas reduz a magnitude das perdas máximas, mas também limita a volatilidade negativa. A distribuição dos retornos das carteiras CVaR evidencia uma concentração em torno da média e uma redução dos outliers negativos, reforçando a eficácia do CVaR em ambientes voláteis.

Entretanto, considera-se que essa abordagem mais conservadora apresenta certas desvantagens. A concentração maior dos retornos em torno de uma média, observada nos histogramas e boxplots, sugere que o CVaR pode limitar o potencial de ganhos extremos. Além disso, a maior assimetria negativa nas distribuições dos retornos das carteiras CVaR pode indicar uma tendência a sacrificar possíveis retornos elevados para evitar perdas, o que pode ser menos atraente para investidores que buscam maior exposição ao risco em busca de retornos mais altos.

6.3.2 OGD com restrição CVaR para dados IBOV

Conforme destacado no início dessa seção, o foco novamente se dará na comparação entre o CVaR, OGD padrão e nos resultados da estratégia VaR t

de Student para dados IBOV, tendo como base a capacidade do CVaR superar estratégias de investimento que são consideradas como *benchmarks* do mercado, assim como visto nas subseções anteriores. Logo, grande parte da análise se baseará na comparação com as métricas vistas nas Tabelas 13 e 15.

A Figura 42 apresenta o retorno cumulativo das carteiras otimizadas com base em diferentes níveis de risco usando a metodologia CVaR, em comparação com a estratégia padrão OGD. Ao longo do período analisado, é evidente que a carteira OGD_80 CVaR, representada pela linha laranja, supera consistentemente as outras estratégias em termos de retorno acumulado, alcançando os níveis mais elevados ao final do horizonte temporal. Essa superioridade reflete a capacidade da estratégia de capturar ganhos substanciais durante os ciclos de alta do mercado, enquanto ainda se beneficia do controle de risco oferecido pelo CVaR.

Gráfico 42 – Comparação entre os portfólios OGD, OGD_80 CVaR, OGD_40 CVaR, OGD_20 CVaR e OGD_10 CVaR para dados IBOV



Fonte: Elaborado pelo autor (2024).

A carteira OGD, em sua forma padrão, também demonstra um desempenho robusto, mas menos expressivo quando comparado às carteiras CVaR com maior tolerância ao risco, como OGD_80 CVaR (todo período) e OGD_40 CVaR (neste caso, inicialmente até 2017; depois entre 2020 e 2021). Este comportamento sugere

que, embora a estratégia padrão OGD seja eficaz, a inclusão do CVaR como uma restrição adicional de risco permite uma alocação mais agressiva em condições de mercado favoráveis, resultando em um retorno acumulado superior.

Por outro lado, as carteiras OGD_20 CVaR e OGD_10 CVaR, que impõem restrições de risco mais rigorosas, exibem retornos acumulados mais conservadores. Embora estas estratégias ofereçam menor volatilidade e protejam o portfólio contra perdas extremas, como indicado pela suavidade das curvas ao longo do tempo, elas também limitam a participação nos grandes movimentos de alta do mercado, resultando em um crescimento de capital menos acentuado.

Essa análise é suportada pelas métricas presentes na Tabela 17, que comparam o desempenho das carteiras otimizadas com base em diferentes níveis de CVaR para os dados do IBOV. É evidente que a estratégia OGD_80 CVaR destaca-se por oferecer o maior retorno cumulativo e a maior Riqueza Terminal, ao custo de uma volatilidade um pouco mais elevada. Essa relação risco-retorno é reforçada pela Razão de Calmar, que demonstra a robustez da estratégia em equilibrar o crescimento do capital com a contenção de perdas severas ao longo do tempo.

Tabela 17 – Comparação de Desempenho e Risco: Análise das Carteiras OGD_80 CVaR, OGD_40 CVaR, OGD_20 CVaR, OGD_10 CVaR para os dados IBOV

Métricas	OGD_80 CVaR	OGD_40 CVaR	OGD_20 CVaR	OGD_10 CVaR
TCAC (%)	20,497	17,641	13,368	11,710
Riqueza Terminal	32,718	20,887	10,455	7,936
Volatilidade Anualizada (%)	21,079	13,466	7,346	4,044
Semi-Desvio	0,018	0,012	0,006	0,004
Máximo <i>Drawdown</i> (%)	-36,508	-29,986	-17,732	-8,732
<i>Drawdown</i> Médio (%)	-6,725	-3,403	-1,522	-0,541
Pior Perda (%)	-13,857	-9,418	-5,579	-2,998
VaR 95 (%)	-2,312	-1,462	-0,784	-0,415
VaR 99 (%)	-3,550	-2,253	-1,215	-0,653
CVaR 95 (%)	-2,957	-1,859	-1,009	-0,542
CVaR 99 (%)	-5,064	-3,458	-1,920	-1,058
Razão de Calmar	0,431	0,552	0,677	1,203
Razão de Sortino	0,010	0,006	0,003	0,002
Recuperação Média	76,906	125,434	232,649	424,860
Assimetria	-0,197	-0,253	-0,107	-0,101
Excesso de Curtose	12,387	20,233	24,665	23,024

Fonte: Elaborado pelo autor (2024).

Comparando essas estratégias com o OGD padrão, cujas métricas foram

apresentadas na Tabela 13, observa-se que a inclusão de uma restrição baseada no CVaR permite uma gestão de risco mais eficaz, especialmente em contextos de alta volatilidade, sem sacrificar significativamente o crescimento do capital. Embora a volatilidade anualizada das estratégias CVaR seja menor que a do OGD padrão (26,179%), o controle adicional de risco oferecido pelo CVaR resulta em um Semi-Desvio e um Máximo *Drawdown* mais favoráveis (0,022 e -27,043% para o OGD padrão, respectivamente), proporcionando um investimento mais estável e menos sujeito a choques extremos.

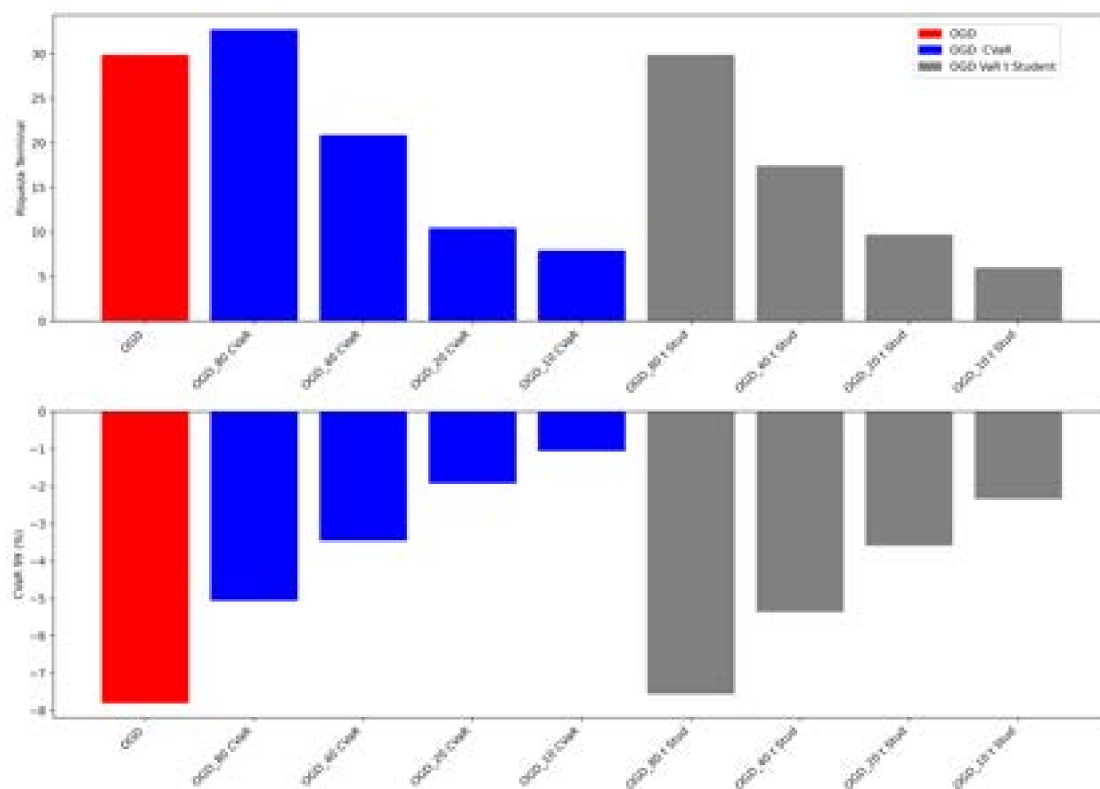
Ao comparar as métricas de Razão de Sortino e Razão de Calmar entre as estratégias OGD padrão, VaR t de Student (ver Tabela 15) e CVaR, é possível identificar diferenças significativas na forma como cada abordagem equilibra risco e retorno. A Razão de Sortino, que mede o retorno ajustado ao risco penalizando apenas a volatilidade negativa, apresenta valores semelhantes entre o OGD padrão (0,013) e as estratégias baseadas no VaR t de Student (entre 0,09 e 0,018), refletindo uma capacidade razoável dessas abordagens em gerar retornos consistentes ajustados ao risco de queda. No entanto, as estratégias baseadas em CVaR, apesar de exibirem um controle mais rigoroso sobre as perdas extremas, resultam em uma Razão de Sortino ligeiramente inferior. Isso ocorre porque o foco intenso na mitigação de riscos de cauda pode limitar o potencial de ganhos elevados, resultando em um menor retorno ajustado ao risco quando comparado às outras abordagens.

Por outro lado, a Razão de Calmar, que avalia o retorno em relação ao máximo *drawdown*, evidencia uma superioridade das estratégias baseadas em CVaR, especialmente nos níveis de maior tolerância ao risco. Essas estratégias conseguem não apenas limitar as perdas extremas, mas também capturar um crescimento significativo do capital ao longo do tempo, o que se reflete em uma Razão de Calmar mais elevada. Comparativamente, o OGD padrão e as estratégias VaR t de Student, embora competitivas, demonstram uma Razão de Calmar inferior (a maior vista dentre elas foi de 0,432 no caso t de Student com menor propensão ao risco), sugerindo que essas abordagens estão mais expostas a *drawdowns* maiores, o que pode comprometer a resiliência do portfólio em períodos de alta volatilidade.

Essa análise é corroborada pelo Gráfico 43, onde é possível observar que as estratégias que minimizam o CVaR destacam-se tanto em termos de Riqueza

Terminal quanto na redução dos riscos extremos. As carteiras OGD_80 CVaR e OGD_40 CVaR, por exemplo, demonstram um equilíbrio superior entre crescimento de capital e controle de risco, resultando em uma maior Riqueza Terminal comparada ao OGD padrão, sem aumentar excessivamente o CVaR. Isso mostra que as abordagens que utilizam o CVaR conseguem limitar os impactos negativos em situações de mercado adversas, ao mesmo tempo em que capturam de forma eficiente as oportunidades de crescimento.

Gráfico 43 – Análise Comparativa de Riqueza Terminal e CVaR entre a carteira OGD e estratégias de minimização de CVaR, e minimização de VaR com Distribuição t Student para dados IBOV



Fonte: Elaborado pelo autor (2024).

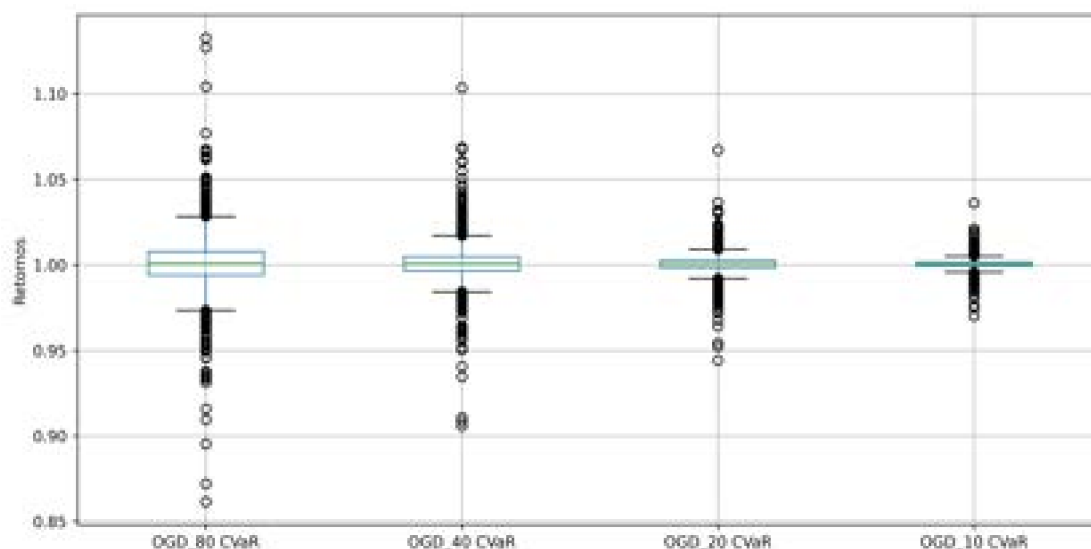
Por outro lado, as estratégias que empregam a minimização do VaR com a distribuição t de Student, apesar de apresentarem boas performances em termos de Riqueza Terminal, ainda ficam atrás das estratégias CVaR no quesito controle de risco. Isso é evidente ao observar que essas estratégias têm um CVaR 99% mais

elevado, indicando maior vulnerabilidade a perdas extremas. Este comportamento reflete a maior sensibilidade dessas carteiras a eventos de cauda, o que pode se traduzir em uma menor resiliência durante períodos de alta volatilidade.

Analisando o *Box-Plot* apresentado no Gráfico 29, observa-se uma dispersão significativa nos retornos, caracterizada por um maior Semi-Desvio e pela presença de outliers mais pronunciados nas caudas dos boxplots. Isso sugere que, embora a estratégia OGD possa oferecer oportunidades de ganho, ela também está mais exposta a riscos de cauda e volatilidade elevada, conforme refletido pela maior volatilidade anualizada (26,179%) e pela Pior Perda (-25,899%) em comparação com as variantes ajustadas para CVaR.

Por outro lado, as estratégias ajustadas para CVaR, ilustradas no Gráfico 44, apresentam uma redução na dispersão dos retornos à medida que o nível de CVaR é reduzido. Esse comportamento é evidenciado pelas menores volatilidades anualizadas e Semi-Desvios observados nas estratégias CVaR, particularmente nos níveis de risco mais baixos, como OGD_20 CVaR e OGD_10 CVaR, que exibem Semi-Desvios de 0,006 e 0,004, respectivamente. Esse padrão reflete a principal função do CVaR, que é mitigar os riscos extremos ao limitar as perdas potenciais além de um certo quantil.

Gráfico 44 – *Box-Plots* dos retornos brutos das carteiras OGD_80 CVaR, OGD_40 CVaR, OGD_20 CVaR e OGD_10 CVaR para dados IBOV



Fonte: Elaborado pelo autor (2024).

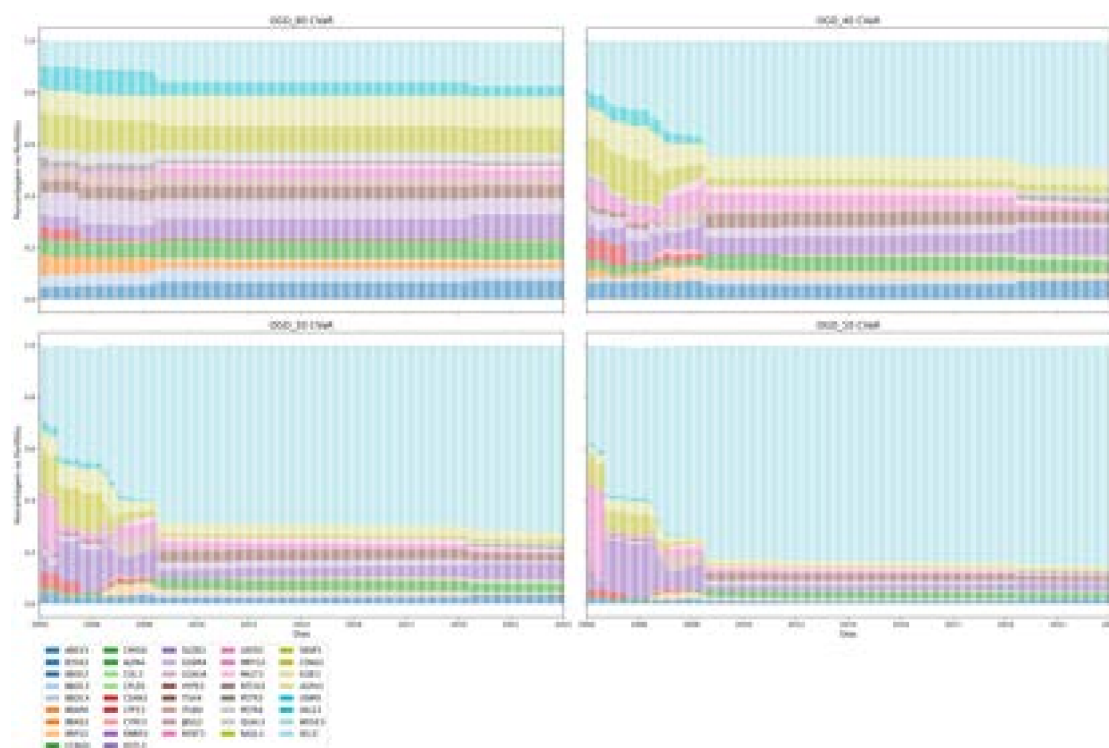
Ao observar os primeiros 3 quartis dos retornos nos gráficos de boxplot, nota-se que as estratégias ajustadas para CVaR mantêm valores centrais semelhantes à mediana. Isso indica que, apesar da mitigação do risco de cauda (*extreme tail risk*), o retorno central da estratégia não é significativamente comprometido. Além disso, a menor dispersão das estratégias CVaR sugere uma performance mais estável e previsível na comparação ao OGD padrão, o que pode ser preferível para investidores com maior aversão ao risco.

Ao analisar os gráficos de alocação de ativos das carteiras otimizadas com as estratégias OGD_CVaR (Gráfico 45) e OGD_t Student (Gráfico 36), observam-se diferenças marcantes na distribuição dos pesos ao longo do tempo, refletindo as abordagens distintas na gestão de risco, particularmente no tratamento de riscos extremos e na estabilização da volatilidade.

As carteiras otimizadas com CVaR demonstram uma tendência a manter uma diversificação mais uniforme e persistente ao longo do tempo, evidenciada pela distribuição mais estável dos pesos. Isso é especialmente perceptível nas carteiras OGD_20 CVaR e OGD_10 CVaR, onde a menor concentração em ativos

individuais e a menor volatilidade anualizada (*e.g.*, 7,346% para OGD_20 CVaR e 4,044% para OGD_10 CVaR) indicam uma gestão de risco mais eficiente. Essa característica de estabilidade é uma consequência direta da formulação do CVaR, que busca minimizar o risco de cauda ao limitar as perdas potenciais além de um certo quantil, resultando em uma alocação de ativos que reage menos drasticamente a flutuações de mercado.

Gráfico 45 – Comparação dos pesos investidos nos portfólios OGD_80 CVaR, OGD_40 CVaR, OGD_20 CVaR e OGD_10 CVaR para dados IBOV



Fonte: Elaborado pelo autor (2024).

Em contraste, as carteiras otimizadas utilizando o VaR t Student apresentam uma alocação de ativos mais volátil, com ajustes significativos nos pesos ao longo do tempo. Esse comportamento é particularmente evidente nas carteiras OGD_40 t Student e OGD_10 t Student, onde a alocação tende a se concentrar em um número reduzido de ativos. Essa concentração eleva o risco específico da carteira, o que é corroborado pela maior volatilidade anualizada registrada (*e.g.*, 17,671%

para OGD_40 t Student e 7,272% para OGD_10 t Student), além de uma maior Pior Perda (*e.g.*, -13,032 % para OGD_40 t Student).

Além disso, as carteiras baseadas em CVaR demonstram uma performance superior em termos de mitigação de perdas, conforme evidenciado pelos valores mais baixos de CVaR 95% e 99%. Por exemplo, a carteira OGD_10 CVaR apresenta valores de CVaR 99% de -1,058%, indicando uma maior eficiência na contenção de perdas extremas em comparação com as carteiras baseadas em VaR t Student. Essa capacidade de limitar perdas faz da abordagem CVaR uma escolha robusta para investidores que buscam preservar o capital em ambientes de alta volatilidade, garantindo uma alocação de ativos mais equilibrada e menos suscetível a grandes oscilações.

No que tange as métricas distribucionais, resultados um pouco distintos aos vistos para *small caps* foram evidenciados, apesar de no caso IBOV, o OGD padrão ter assimetria quase nula: (i) a inserção da restrição CVaR, apesar de levar a uma assimetria negativa, estes valores são significativamente inferiores (entre -0,197 para 80% e -0,101 para 10%) aos da modelagem VaR t de Student (entre -0,256 e -1,325), além de haver no primeiro momento um aumento de assimetria negativa entre a propensão 80% e 40% (de -0,197 para -0,253), com queda e constância nos graus seguintes (-0,107 e -0,101, respectivamente), contra um aumento constante do caso VaR t de Student. Isso sugere a presença de valores mais (menos) extremos a direita (esquerda) da mediana na comparação OGD Padrão, evitando mais perdas, além de uma propensão maior em suavizar o retorno do investimento na comparação ao Var t de Student, em especial nas menores propensões ao risco; (ii) Um excesso de curtose similar ao visto no OGD padrão (12, 359) e modelo CVaR 80% (12,387), mas com constante elevação nos níveis seguintes (entre 20,233 e 23,024), mas em patamares bastante superiores aos evidenciados no caso VaR t de Student (entre 4,675 e 15,358), indicando uma propensão maior por carteiras nos quantis centrais do retorno, corroborando com a suavização dos retornos dos investimentos.

Em suma, o CVaR para dados IBOV, assim como no caso *small caps*, mostrou-se mais robusto em evitar riscos de cauda, evitando perdas extremas e volatilidade negativa, mas também presa por carteiras com diversificação gerando retornos mais próximos da média, em especial para perfis de risco mais conservadores

e em momentos de grande volatilidade do mercado. Dito isso, pode não ser a estratégia mais indicada para perfis mais propensos a tomar riscos em prol de um maior retorno. Vale ressaltar que elevar a propensão ao risco para um nível de 80% mostrou que o retorno acumulado final do período foi ligeiramente superior ao OGD Padrão, mas com riqueza terminal maior e níveis de risco significativamente menores tanto contra o OGD como perante o VaR t de student, gerando robustez a tal estratégia. Já na comparação mais direta com o VaR t de Student, para todos os níveis de disposição, o CVaR mostrou melhores relações retorno x risco, enaltecendo ainda mais a abordagem focada no CVaR. Entretanto, os resultados se mostraram menos expressivos na comparação aos dados *small caps*, algo possivelmente explicado pela natureza mais volátil dos ativos que formam ambos índices. Com isso em mente, futuras avaliações em mercados mais e menos voláteis poderão gerar ainda mais robustez para a modelagem proposta.

6.4 RESUMO DE RESULTADOS E COMPARAÇÕES COM OUTROS TRABALHOS

Esta seção tem por objetivo resumir os principais resultados dos modelos propostos e fazer uma comparação sucinta com outros poucos trabalhos que tiveram como enfoque modelos de aprendizado de máquinas ou similares com uso de dados do IBOV e/ou *Small Caps*. Abaixo há um apanhado dos resultados ligados as metodologias propostas nesse trabalho.

1. **OGD Beta para dados *Small Caps*:** A estratégia de manter um beta variante no tempo e limitado entre -3 e 0,1 no período de 1998 a 2015, considerado de baixa no mercado para os ativos que compõem o índice, onde o algoritmo define uma maior parcela do ativo livre de risco na composição da carteira ótima ao longo do tempo; além de entre 1,9 e 3 para os anos restantes, um período de alta no mercado, com maior predisposição a escolha de ativos de maior risco e correlacionados com o mercado de ações, geraram - (i) carteiras com maior retorno acumulado em todo período na comparação ao índice *Small Caps*, o OGD padrão e ao UCRP; (ii) menor (maior) risco total captado via VaR e CVaR na comparação a tais modelos no caso de

baixa (alta) do mercado.

2. **OGD Beta para dados IBOV:** A estratégia de manter um beta variante no tempo e limitado entre 1,4 e 3 gerou maiores retornos acumulados em períodos de alta do mercado, mas com um risco mais elevado observado via VaR e CVaR na comparação ao IBOV, OGD padrão e ao UCRP. Já restringir o beta entre -0,3 e 0,1 na baixa (como entre 2015 e início de 2016) gerou novamente maior retorno acumulado, mas ligado a um menor risco na comparação aos modelos concorrentes, de maneira similar ao visto para dados *Small Caps*.

3. **Modelo VaR para dados *Small Caps*:** Levando em conta os modelos *benchmark* OGD padrão, Markowitz, UCRP e o Índice *Small Caps* (I_SMALL), foi observado a superioridade do OGD em termos de retorno (Riqueza Terminal e taxa de crescimento anualizado do retorno - TCAC), mas com risco definido pelas métricas Volatilidade anualizada, Semi-Desvio, Máximo *Drawdown*, *Drawdown* Médio, VaR, CVaR, Razões de Calmar e Sortino com ordem decrescente de relevância definida por OGD, Markowitz, UCRP e I_SMALL, respectivamente. Ao considerar os modelos OGD VaR com perda máxima que são 80%, 40%, 20% ou 10% da perda esperada no 1% dos casos mais extremos (pior cenário) observados no histórico de retornos, também chamados de graus de disposição ao risco, foi constatado que tanto considerando a hipótese de retornos com distribuição normal ou aquelas com t de Student - (i) os valores ligados ao risco foram negativamente correlacionados a disposição ao risco, isto é, quanto menor a disposição ao risco, melhores valores ligados as métricas de interesse foram constatadas, com valores destes melhorando na comparação ao OGD padrão segundo a diminuição da disposição ao risco, gerando valores comparáveis aos modelos *benchmark* de menor risco comparativo Markowitz, UCRP e I_SMALL, respectivamente; (ii) já com relação a valorização do retorno, foram vistos resultados superiores ligados as métricas de retorno nos modelo OGD VaR com disposição a 40% e 20% (40% ligeiramente acima dos 20%) na comparação ao OGD, além de valores relativamente similares e abaixo dos modelos 80% e 10%, mas comparáveis nesse quesito ao OGD padrão. Desta forma, fica claro a melhor combinação risco x retorno entre disposições médias ao risco, apesar de haver retornos similares dos graus 10%

e 80% na comparação ao OGD padrão, mas a um risco menor. No que tange a comparação dos modelos supondo retornos com distribuição normal e t de Student, foram vistos valores superiores ligados aos retornos considerando a distribuição t de Student em toda disposição ao risco considerada (mais eficaz em capturar retornos superiores, algo que vai de encontro com as caudas mais pesadas ligadas a tal distribuição). Já levando em consideração as métricas de risco, algumas considerações importantes foram vistas - (i) Máximo *Drawdown* (Semi-Desvio) ligeiramente superior e Razões de Calmar (Razão de Sortino) consistentemente superiores na distribuição t de Student, indicando que nesse caso é possível maior retorno esperado por unidade de perda máxima tomada (risco tomado); (ii) valores consistentemente superiores do VaR para a distribuição t de Student, ligado a maior probabilidade que tal distribuição da a valores extremos (a composição da carteira com maior gama de ativos de maior retorno e risco, ao mesmo tempo que com maior diversificação e menor alocação em ativos livre de risco, deixa isso claro). Por fim, com relação a assimetria e curtose, houve uma melhora significativa da assimetria negativa e excesso de curtose na distribuição t de Student para todos níveis de disposição ao risco considerados, além de tal melhoria estar negativamente correlacionada ao grau de de disposição, indicando que a distribuição em questão capturou de maneira mais adequada o efeito de caudas pesadas visto nos dados, apesar de serem valores ainda significantes. Em suma, a distribuição t de Student gerou uma combinação mais eficaz de risco x retorno na comparação tanto ao modelo VaR Normal, como perante todos modelos *benchmark*, mas ainda com espaço para melhora, em especial na capacidade de capturar efeitos de ativos extremos.

4. **Modelo VaR para dados IBOV:** Assim como ocorreu no caso *Small Caps* detalhado acima, foi visto superioridade de métricas de retornos para o OGD padrão na comparação aos outros *benchmark* Markowitz, UCRP e IBOV. Porém, nem todas as métricas de risco ligadas aos modelos Markowitz, UCRP e IBOV foram superiores ao OGD padrão, com vantagem deste com relação as Razões de Calmar e Sortino, indicando que ele tende ter um melhor equilíbrio risco x retorno. Outra diferença é que o OGD aplicado a dados IBOV mostrou

métrica de assimetria e curtose indicando uma distribuição quase simétrica e com uma leve tendência para valores extremos acima da média. Considerando novamente os modelos OGD VaR com graus de disposição ao risco de 80%, 40%, 20% ou 10%, foi constatado que tanto para a hipótese de retornos com distribuição normal ou aquelas com t de Student - (i) os valores ligados as métricas de risco tiveram caminhos similares aos vistos em *Small Caps*; (ii) já com relação aos valores ligados a retorno, diferente do exposto acima, nenhum grau de aversão gerou valores superiores ao OGD Padrão, mas com o caso t de Student 80% tendo valores um pouco abaixo do OGD (comparável caso se considere o TCAC), mas com clara vantagem no quesito risco para este, mostrando que a estratégia com menor controle de risco se mostrou mais plausível nesse caso, o que pode também pode ser explicado em parte pela característica de maior estabilidade e liquidez dos ativos presentes no IBOV. No que tange a comparação dos modelos supondo retornos com distribuição normal e t de Student, foram vistos valores superiores ligados a proporção retornos x risco considerando a distribuição t de Student para 40% e 80%, mas abaixo nas menores disposições ao risco na comparação a normal, algo distinto do visto para dados *Small Caps*. Já levando em consideração as métricas de risco, resultados similares ao caso *Small Caps* foram evidenciados. Por fim, com relação a assimetria e curtose - (i) assim como no caso *Small Caps*, tanto a a assimetria negativa como o excesso de curtose para ambas distribuições está negativamente correlacionado ao grau de de disposição; (ii) apesar de no caso IBOV as assimetrias vistas para a t de Student terem sido menos acentuadas, elas ainda são comparáveis ao visto no caso normal, sugerindo que a mudança na distribuição não teve um impacto significativo na assimetria dos retornos do IBOV na comparação aos vistos para o Small Caps; (iii) o excesso de curtose foi menor no caso IBOV usando t de Student na comparação a distribuição normal, mas em menor escala na comparação ao modelo *Small Caps*. Em suma, a distribuição t de Student gerou uma melhor combinação de risco x retorno para a maior de disposição ao risco na comparação tanto ao modelo VaR Normal como todos *benchmarks*, mas a escolha da distribuição teve resultados abaixo do *Small Caps*, deixando claro a necessidade de um estudo mais apurado da adequação de outras

distribuições para o caso IBOV.

5. **Modelo CVaR para dados *Small Caps***: Dados os resultados presentes na avaliação do modelo baseado no VaR e dos *benchmark* OGD Padrão, Markowitz, UCRP e I_Small, o foco aqui foi na comparação dos modelos OGD CVaR com graus de disposição ao risco de 80%, 40%, 20% ou 10% contra os com melhor desempenho comparativo, isto é, o OGD Padrão e o VaR baseado no distribuição t de Student para os mesmos níveis de aversão ao risco destacados acima. No quesito retorno cumulativo, diferente do VaR t de Student onde apenas o modelo com disposição ao risco 40% gerou efetivamente maiores valores finais, aqui todos os níveis acima de 20% (80% o principal) tiveram tal característica, além de contar com um crescimento mais estável e controlado ao longo do tempo devido a sua natureza de priorizar a mitigação de perdas extremas em períodos de crise, priorizando ativos com retornos mais estáveis e menor exposição a perdas extremas. A análise das métricas de risco Volatilidade anualizada, Semi-Desvio, Máximo *Drawdown*, *Drawdown* Médio, VaR, CVaR, Razões de Calmar e Sortino corroboraram com tal resultado, mas também indicam que a tendência em evitar perdas de cauda do CVaR pode também gerar um retorno ajustado ao risco menor, em especial em períodos de grande alta de mercado, mas que, ao mesmo tempo, proporcionam retornos mais eficientes contra grandes perdas, comum em momentos de crise, tendo valores mais significativos nesse quesito com o aumento da aversão ao risco do investidor. No geral, a proporção retornos x risco, tanto capturada pela relação Taxa de Crescimento Anual Composta (TCAC) x métrica VaR 99% quanto por riqueza terminal x métrica CVaR 99%, demonstrou evidente vantagem para os modelos CVaR nas comparações a todos graus de aversão ao risco, por exemplo: o CVaR com disposição de 80% gerou risco relativamente similar ao OGD padrão e ao VaR com mesma disposição, mas com larga vantagem em termos de retorno para ambos. Já no que tange a aspectos distribucionais: (i) correlação entre menor dispersão de retornos e menor disposição ao risco, mais acentuado na comparação aos modelos VaR, os quais tem a mesma característica; (ii) assimetria negativa mais acentuada no CVaR, indicando que penaliza de maneira mais eficaz

as perdas extremas, algo esperado por evitar mais o risco de cauda; (iii) significativo aumento no excesso de curtose na comparação ao VaR, indicando que tende a dar mais pesos a retornos mais centrais em detrimento aos mais extremos.

6. **Modelo CVaR para dados IBOV:** Assim como no caso *Small Caps*, aqui os resultados foram mais baseados na comparação do modelo CVaR com disposição ao risco de 80%, 40%, 20% ou 10% com o OGD Padrão (superior a todos outros *benchmark*) e em menor escala com o VaR t de Student (que acabou gerando retornos acumulados e riqueza terminal abaixo do OGD padrão). No quesito retorno cumulativo, apenas o CVaR 80% gerou resultados finais maiores que o OGD padrão, com o de disposição de 40% tendo se mostrado a frente em alguns períodos, em particular quando a volatilidade foi menor. Já no que tange riqueza terminal x risco, além da clara superioridade com relação ao VaR em todos os níveis considerados (risco efetivamente menor na comparação ao OGD padrão em todas análises), foi observado significativa superioridade do CVaR 80% na comparação ao OGD Padrão. A análise das métricas de risco Volatilidade anualizada, Semi-Desvio, Máximo *Drawdown*, *Drawdown* Médio, VaR, CVaR, Razões de Calmar e Sortino indicam: níveis de risco mais satisfatórios capturados pelas métricas VaR e CVaR, o foco mais intenso da estratégia CVaR em evitar perdas extremas, ao mesmo tempo que consegue capturar um crescimento mais robusto e suave que as outras estratégias, em especial em níveis de disposição ao risco maiores, com a diversificação de ativos corroborando com isso. Em relação as métricas distribucionais: (i) introduziu uma assimetria negativa, antes positiva e quase nula no OGD Padrão, com valores inferiores aos da modelagem VaR t de Student, sugerindo a presença de valores mais (menos) extremos a direita (esquerda) da mediana na comparação OGD Padrão, evitando mais perdas e suavizando o retorno do investimento de maneira mais eficaz na comparação ao VaR t de Student, em especial nas menores propensões ao risco; (ii) excesso de de curtose similar ao visto no OGD padrão, mas com constante crescimento nos níveis seguintes, porém em níveis superiores aos evidenciados no caso VaR t de Student, indicando uma

propensão maior por carteiras nos quantis centrais do retorno, corroborando com a suavização dos retornos dos investimentos. No geral, foi superior ao OGD Padrão na relação retorno x risco para a maior disposição ao risco de 80%, independente das métricas usadas (parte do do período menos volátil no caso de 40%), com clara vantagem em termos de risco ao OGD Padrão em todas as disposições; além de superior ao VaR t de student em todos os quesitos, independente da perspectiva tomada. Porém, vale ressaltar que foram gerados resultados inferiores aos dados *Small Caps*, talvez pela natureza intrínseca de tais dados na comparação ao IBOV, gerando indicativos que análises baseadas em dados de outros mercados irão ajudar a gerar maior robustez a metodologia proposta.

Por fim, há abaixo uma tabela resumo dos principais resultados de outros estudos recentes inseridos no contexto de *machine learning* para dados do Brasil, com o atual trabalho tendo tido resultados em nível similar ou superior tanto com relação aos retornos como perante o risco, em especial ao se considerar o CVaR, mesmo com um tempo de avaliação superior a quase todos eles, o que permitiu um estudo mais robusto levando em conta diferentes períodos de aquecimento e queda no mercado:

Tabela 18 – Resumo de resultados de outros trabalhos focados em dados brasileiros e modelagem *machine learning*

Artigo	Dados	Modelo	Período	Controle para Risco?	Resultados Gerais
Granja (2004) [123]	IBOV	Redes Neurais - Algoritmo Backpropagation	1999 até 2003	não	Lucro acumulado de 128% no período contra 28% do IBOV e 107% no modelo MV com risco medido pelo desvio padrão similar

Artigo	Dados	Modelo	Período	Controle para Risco?	Resultados Gerais
Thommé Neto et al. (2011) [127]	IBOV	Carteira Naive (1/n) e Mínima Variância Global	1998 até 2018	não	Modelo Média Variância Global gerou estatisticamente retornos médios e acumulados maiores e com desvio padrão menor na comparação ao IBOV e Naive (1/N) quando os pesos de cada ativo foi limitado em no máximo 0,1
Santos e Tessari (2012) [124]	IBOV	Carteira Naive (1/n), MV, Mínima Variância e baseada em Risk-Metrics	2009 a 2011	sim, via técnicas RiskMetrics	Modelagem MV e Mínima Variância com estimadores não paramétricos para a matriz covariância dos retornos gerou maiores retornos ajustados ao risco na comparação aos outros modelos

Artigo	Dados	Modelo	Período	Controle para Risco?	Resultados Gerais
Ferreira et al. (2018) [133]	IBOV	Parallel multiobjective evolutionary algorithms	2010 até 2015	não	Distribuição do Máximo Drawdown e do CVaR deslocado a esquerda de todos algoritmos testados na comparação ao IBOV, com a distribuição de retornos a direita numa comparação similar.
Silva de Souza et al. (2018) [135]	<i>Small Caps</i>	Automated Trading System (ATS) baseada em modelos de médias móveis dos retornos	2007 e 2016	não	Portifólios baseados em <i>Small Caps</i> geraram retornos médios similares aos com estratégias baseadas em bluechips.
Duarte e Castro (2019) [130]	IBOV	Algoritmo Fuzzy Clustering, baseado na construção hierárquica de portfólios	2009 até 2017	não	Estratégia focada no algoritmo teve retornos médios superiores e volatilidade similar a visto no IBOV, com o retorno médio aumentando com o número de grupos de ações considerados para gerar o cluster de interesse.

Artigo	Dados	Modelo	Período	Controle para Risco?	Resultados Gerais
Paiva et al (2019) [131]	IBOV	Método Support Method Machine (SVM) em conjunto com MV	2001 até 2016	não	Estratégia conjunta SVM e MV gerou maiores retornos acumulados na comparação ao IBOV e ao modelo SVM com naive (1/N) ao descartar a possibilidade de custos de transação, acima do IBOV em metade do período levando custos intermediarios e sempre abaixo do IBOV considerando altos custos de transação.

Artigo	Dados	Modelo	Período	Controle para Risco?	Resultados Gerais
Gotardelo (2020) [128]	IBOV e <i>Small Caps</i>	Carteiras Convexas e Algoritmos Evolutivos Multiobjetivos (AEM)	1994 a 2020 para IBOV; 2008 até 2020 para <i>Small Caps</i>	sim, focado no Maximum Draw-down	Carteira convexa aplicada a dados IBOV com rentabilidade acumulada superior entre os anos de 1998 até 2017 na comparação a CAPM, OCAPM e Naive (1/n), ficando abaixo do CAPM entre 2018 e 2020; Carteira convexa aplicada a dados <i>Small Caps</i> com rentabilidade acumulada abaixo do CAPM em todos os 5 períodos, mas acima do Naive (1/n) em todos períodos e apenas um acima do OCAPM; 4 das 12 carteiras baseadas em estratégias AEM tiveram retornos acumulados superiores as carteiras CAMP, OCAPM, Naive e Carteira convexa no caso IBOV, sendo sempre abaixo da CAPM e CDI no caso <i>Small Caps</i> .

Artigo	Dados	Modelo	Período	Controle para Risco?	Resultados Gerais
Conegundes e Pereira (2020) [134]	IBOV	Modelagem Deep Reinforcement Learning (DRL)	2015 até 2019	não	Modelos DRL analisados com janelas flutuantes menores tiveram retornos acumulados bem acima do IBOV, com a modelagem focada na janela maior com valores similares para metade do período e acima nos anos restantes. Já o Máximo Drawdown médio foi acima também na comparação ao IBOV.
Freitas e Junior (2023) [132]	IBOV	Stock Network Portfolio Allocation (SNPA)	2018 até 2019	sim, via minimização de risco	Modelo resultou em retorno acumulado muito acima dos modelos MV, CVaR, IBOV e Hierarchical Risk Parity (HRP); além da superioridade levando em conta a razão de sortido o o Maximo Drawdown.

Fonte: Elaborado pelo autor (2024).

7 CONCLUSÕES

Neste trabalho, foram propostas três metodologias para aprimorar a gestão de risco e a adaptabilidade na alocação dinâmica de ativos em portfólios. O objetivo principal foi incorporar métricas de risco econométricas, como o coeficiente beta (β), *Value at Risk* (VaR) e *Conditional Value at Risk* (CVaR), no processo de otimização de portfólios utilizando algoritmos de aprendizado *online*.

Inicialmente, investigou-se a viabilidade de utilizar o coeficiente beta variante no tempo, estimado por meio do Filtro de Kalman, para capturar a dinâmica temporal da correlação entre os ativos e o mercado. Esta técnica mostrou-se eficiente, com reduções consideráveis dos valores estimados para as métricas VaR e CVaR em comparação aos modelos OGD sem limites ao beta, o UCRP e os próprios índices *Small Caps* e Ibovespa.

Ao definir um beta negativo, a estratégia tende a aumentar a alocação em ativos de menor risco, como a renda fixa, minimizando a exposição às flutuações do mercado. Esta abordagem é vantajosa em períodos de baixa, onde a proteção contra perdas pode ser mais crítica. Por outro lado, estabelecer um beta elevado aumenta a volatilidade do portfólio e amplifica a exposição a movimentos de alta do mercado, o que pode ser benéfico em ambientes de crescimento, mas arriscado durante as correções do mercado.

Um dos principais desafios na utilização dessa metodologia é a previsão do comportamento futuro do mercado. A escolha de um beta específico baseia-se frequentemente em expectativas que podem não se concretizar, tornando a projeção uma ferramenta de risco considerável. Por exemplo, um portfólio configurado para um alto beta em antecipação a um mercado em alta pode sofrer perdas substanciais se o mercado inesperadamente declinar.

A eficácia da projeção do beta também foi funcional na análise, evidenciando que, apesar das incertezas, é possível alcançar retornos superiores ajustando dinamicamente a correlação do portfólio com o mercado. Esta estratégia, entretanto, requer uma gestão ativa e um monitoramento constante das condições de mercado para ajustar as expectativas e realocar recursos conforme necessário, a fim de

maximizar os retornos e minimizar os riscos.

Portanto, o controle do beta na imposição de limites oferece uma ferramenta valiosa para o gerenciamento do risco sistemático, mas deve ser aplicado com cautela, integrado a uma abordagem de investimento mais abrangente que considere tanto a análise técnica quanto a avaliação contínua do ambiente econômico e financeiro. Além disso, é importante reconhecer que o risco sistemático não é a única fonte de risco.

Avaliou-se também um modelo inicial que aborda o risco de maneira direta, estabelecendo limites de risco baseados no VaR. Esta técnica mostrou-se eficiente tanto na geração de retornos quanto no controle de risco, superando as métricas tradicionais de avaliação de risco em comparação com outros modelos *benchmark*. Nesse contexto, a inclusão das restrições de VaR nas carteiras OGD Normal e OGD t de Student pode ter um impacto significativo na exposição a ativos com potencial de altos retornos, o que, por sua vez, pode afetar a assimetria e o excesso de curtose da distribuição dos retornos da carteira.

O algoritmo OGD com restrições de VaR busca maximizar o retorno da carteira estando sujeito a restrições lineares e à restrição de VaR. Esta restrição limita a exposição da carteira a ativos com alta volatilidade e baixos retornos esperados, a fim de manter o VaR abaixo do nível especificado. No entanto, ao fazer isso, a restrição de VaR também pode limitar a exposição a ativos com potencial de altos retornos, uma vez que esses ativos geralmente apresentam maior volatilidade.

Além disso, a limitação da exposição a ativos com potencial de altos retornos pode aumentar o excesso de curtose, que mede a quantidade de observações extremas (tanto positivas quanto negativas) em relação a uma distribuição normal, da distribuição dos retornos da carteira. Ao limitar a possibilidade de retornos extremos positivos, a restrição de VaR pode fazer com que a distribuição dos retornos da carteira tenha caudas mais pesadas do que o esperado em uma distribuição normal, aumentando, assim, o excesso de curtose.

Embora a restrição de VaR vise controlar o risco da carteira, sua utilização pode ter o efeito colateral de limitar a possibilidade de retornos extremos positivos. Isso pode ser uma desvantagem para investidores que buscam maximizar os retornos,

mesmo que isso implique em um maior risco.

É crucial encontrar um equilíbrio entre o controle de risco e a busca por retornos. A inclusão da restrição de VaR no algoritmo OGD permite ajustar o nível de risco de acordo com as preferências e tolerância ao risco do investidor, ao mesmo tempo em que se busca uma carteira ótima. Com isso em mente, foi proposta e avaliada uma extensão deste modelo, baseada no CVaR, uma métrica que, além de permitir o ajuste para diferentes níveis de risco e ter propriedades similares ao VaR, é mais robusta por ser coerente, ou seja, permite gerar retorno enquanto mitiga risco por meio da diversificação em mercados com distribuições de retornos, como as com caudas mais pesadas.

A implementação do CVaR demonstrou eficácia significativa em proporcionar uma proteção robusta contra os riscos de cauda, sendo particularmente eficiente na contenção de perdas extremas. Esta característica é fundamental para minimizar a exposição a cenários adversos severos. As métricas associadas ao CVaR, como a Razão de Calmar e o Semi-Desvio, indicaram que essa abordagem não apenas reduz a magnitude das perdas máximas, mas também limita a volatilidade negativa. A distribuição dos retornos das carteiras CVaR evidenciou uma concentração em torno da média e uma redução dos *outliers* negativos, reforçando a eficácia do CVaR em ambientes voláteis.

Entretanto, é importante reconhecer que essa abordagem mais conservadora apresenta certas desvantagens. A maior concentração dos retornos em torno de uma média, observada nos histogramas e *boxplots*, sugere que o CVaR pode limitar o potencial de ganhos extremos. Além disso, a maior assimetria negativa nas distribuições dos retornos das carteiras CVaR pode indicar uma tendência a sacrificar possíveis retornos elevados para evitar perdas, o que pode ser menos atraente para investidores que buscam maior exposição ao risco em busca de retornos mais altos.

Em suma, a estratégia de otimização de portfólio baseada no CVaR oferece uma abordagem equilibrada para o gerenciamento de riscos, especialmente em mercados voláteis ou com distribuições de retornos não-normais. Além disso, proporciona uma proteção eficaz contra perdas extremas, mantendo um nível

aceitável de retorno. No entanto, sua natureza conservadora pode limitar o potencial de ganhos excepcionais em períodos de alta do mercado.

Portanto, a escolha entre as diferentes estratégias de otimização — seja baseada em beta, VaR ou CVaR — deve ser feita considerando cuidadosamente o perfil de risco do investidor, as condições de mercado e os objetivos de investimento. Uma abordagem híbrida, que combine elementos dessas diferentes estratégias, pode oferecer uma solução mais adaptável para diferentes cenários de mercado, permitindo um equilíbrio dinâmico entre a busca por retornos e a gestão prudente de riscos.

7.1 TRABALHOS FUTUROS

Com base nos resultados e limitações identificados neste estudo, algumas direções de pesquisas futuras são propostas para aprimorar a eficácia e a aplicabilidade prática dos modelos de otimização de portfólios. Uma área de investigação futura envolve a implementação de um parâmetro η dinâmico. Na análise atual, esse parâmetro, que controla a taxa de aprendizagem no algoritmo OGD, foi mantido constante. No entanto, a introdução de um η dinâmico poderia melhorar o desempenho do modelo, permitindo uma adaptação mais eficaz às mudanças nas condições de mercado. Uma análise observada ao longo deste estudo indica que o algoritmo OGD, ao incluir proteção contra o risco utilizando VaR ou CVaR, tende a alterar a composição do portfólio para ativos menos arriscados, como esperado. No entanto, uma característica observada foi a permanência dessa alocação em ativos conservadores mesmo após o retorno à normalidade nas condições de mercado. Uma das possíveis causas para essa rigidez no comportamento pode estar associada ao uso de um η fixo no algoritmo. Esse parâmetro, responsável por controlar a taxa de aprendizagem, ao ser mantido constante, não permite que o modelo se adapte de forma eficiente às mudanças dinâmicas do mercado, levando a uma composição excessivamente conservadora do portfólio mesmo quando a volatilidade diminui.

Uma possível abordagem seria desenvolver um mecanismo que ajuste automaticamente o valor de η com base no desempenho recente do portfólio. Isso poderia incluir, por exemplo, um sistema que reduza gradualmente o valor de η ao

longo do tempo, mantendo um valor mínimo para assegurar que o modelo responda de forma mais dinâmica a novas informações. Outra possibilidade seria vincular o valor de η à volatilidade do mercado, permitindo ajustes mais agressivos em períodos de baixa volatilidade e mais conservadores em períodos de alta volatilidade. Além disso, técnicas de otimização bayesiana poderiam ser empregadas para encontrar o valor ótimo de η em diferentes janelas de tempo, capturando padrões sazonais ou cíclicos no mercado.

Adicionalmente, as projeções propostas neste trabalho são aplicáveis a outros métodos de otimização *online*. Assim, uma tarefa importante seria testar essas metodologias em outros métodos, como *Online Newton Step* e *Exponential Gradient*, para avaliar a robustez dessas abordagens em diferentes algoritmos de otimização.

Outra área de pesquisa futura é a incorporação de taxas de transação na otimização de portfólio, uma consideração crucial para uma representação mais realista dos custos associados ao rebalanceamento do portfólio. O desenvolvimento de um modelo de custo de transação que incorpore tanto custos fixos quanto proporcionais é um aspecto relevante dessa linha de pesquisa. Os custos fixos poderiam representar taxas de corretagem ou custos administrativos associados a cada transação, enquanto os custos proporcionais refletiriam os custos que aumentam com o volume da transação, como o impacto no mercado de grandes negociações. Além disso, seria pertinente a implementação de uma restrição adicional no problema de otimização para limitar o *turnover* do portfólio, refletindo as limitações práticas enfrentadas pelos gestores de fundos, que precisam equilibrar a necessidade de rebalanceamento com os custos associados a transações frequentes. Uma análise detalhada de como diferentes níveis de taxas de transação afetam a estratégia ótima de negociação poderia gerar *insights* sobre a liquidez do mercado e o tamanho ideal do fundo.

Uma linha de pesquisa promissora seria a integração do η dinâmico e das taxas de transação. O desenvolvimento de um *framework* que otimize simultaneamente o η dinâmico e as decisões de rebalanceamento, considerando as taxas de transação, poderia resultar em estratégias de investimento mais eficientes. Isso envolveria uma análise cuidadosa do *trade-off* entre a frequência de rebalanceamento (impactada pelo η dinâmico) e os custos de transação, com o objetivo de encontrar

um equilíbrio que maximize os retornos ajustados ao risco, considerando os custos reais de implementação. A validação dessas abordagens integradas exigiria *backtests* abrangentes em diferentes regimes de mercado, permitindo avaliar a robustez da estratégia em diversas condições econômicas.

Em conclusão, as propostas de pesquisa mencionadas visam aprimorar a sofisticação e a aplicabilidade prática dos modelos de otimização de portfólio baseados em OGD e outros métodos *online*. Ao abordar questões como a adaptabilidade do algoritmo às mudanças nas condições de mercado e a incorporação realista dos custos de implementação, essas linhas de investigação podem fornecer percepções relevantes tanto para a academia quanto para a prática da gestão de investimentos.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] MARKOWITZ, Harry. Portfolio selection. **The journal of finance**, Wiley Online Library, v. 7, n. 1, p. 77–91, 1952.
- [2] ZINKEVICH, Martin. Online convex programming and generalized infinitesimal gradient ascent. In: PROCEEDINGS of the 20th International Conference on Machine Learning (ICML-03). [S.l.: s.n.], 2003. P. 928–936.
- [3] AGARWAL, Amit et al. Algorithms for portfolio management based on the newton method. In: ACM. PROCEEDINGS of the 23rd international conference on Machine learning. [S.l.: s.n.], 2006. P. 9–16.
- [4] HAZAN, Elad; ARORA, Sanjeev. **Efficient algorithms for online convex optimization and their applications**. [S.l.]: Princeton University, 2006.
- [5] SHARPE, William F. Capital asset prices: A theory of market equilibrium under conditions of risk. **The journal of finance**, Wiley Online Library, v. 19, n. 3, p. 425–442, 1964.
- [6] SHARPE, William F. Mutual fund performance. **The Journal of business**, JSTOR, v. 39, n. 1, p. 119–138, 1966.
- [7] SHARPE, William F. The sharpe ratio. **The journal of portfolio management**, Institutional Investor Journals, v. 21, n. 1, p. 49–58, 1994.
- [8] JP MORGAN, REUTERS. Riskmetrics–technical document. **New York. Morgan Guaranty Trust Company of New York, Forth edition. December**, 1996.
- [9] ROCKAFELLAR, R Tyrrell; URYASEV, Stanislav. Optimization of conditional value-at-risk. **Journal of risk**, v. 2, p. 21–42, 2000.
- [10] ROCKAFELLAR, R Tyrrell; URYASEV, Stanislav. Conditional value-at-risk for general loss distributions. **Journal of banking & finance**, Elsevier, v. 26, n. 7, p. 1443–1471, 2002.
- [11] LINTON, Oliver. **Financial econometrics**. [S.l.]: Cambridge University Press, 2019.

- [12] ANKUDINOV, Andrei; IBRAGIMOV, Rustam; LEBEDEV, Oleg. Heavy tails and asymmetry of returns in the Russian stock market. **Emerging Markets Review**, Elsevier, v. 32, p. 200–219, 2017.
- [13] DE JESÚS, Raúl; ORTIZ, Edgar. Risk in emerging stock markets from Brazil and Mexico: Extreme value theory and alternative value at risk models. **Frontiers in Finance and Economics**, v. 8, n. 2, p. 49–88, 2011.
- [14] BERNARDI, Mauro; PETRELLA, Lea. Interconnected risk contributions: A heavy-tail approach to analyze US financial sectors. **Journal of Risk and Financial Management**, MDPI, v. 8, n. 2, p. 198–226, 2015.
- [15] STRAETMANS, Stefan; CHAUDHRY, Sajid M. Tail risk and systemic risk of US and Eurozone financial institutions in the wake of the global financial crisis. **Journal of International Money and Finance**, Elsevier, v. 58, p. 191–223, 2015.
- [16] ANYIKWA, Izunna; PHIRI, Andrew. Quantile connectedness amongst BRICS equity markets during the COVID-19 pandemic and Russia–Ukraine war. **Cogent Economics & Finance**, Taylor & Francis, v. 11, n. 2, p. 2251300, 2023.
- [17] MORKŪNAITĖ, Indrė; CELOV, Dmitrij; LEIPUS, Remigijus. Evaluation of Value-at-Risk (VaR) using the Gaussian Mixture Models. **Research in Statistics**, Taylor & Francis, v. 2, n. 1, p. 2346075, 2024.
- [18] COVER, Thomas M. Universal portfolios. **Mathematical finance**, Wiley Online Library, v. 1, n. 1, p. 1–29, 1991.
- [19] KELLY, John L. A new interpretation of information rate. **Bell Labs Technical Journal**, Wiley Online Library, v. 35, n. 4, p. 917–926, 1956.
- [20] BREIMAN, Leo et al. Optimal gambling systems for favorable games. **The Kelly Capital Growth Investment Criterion**, p. 47–60, 1961.
- [21] FINKELSTEIN, Mark; WHITLEY, Robert. Optimal strategies for repeated games. **Advances in Applied Probability**, Cambridge University Press, v. 13, n. 2, p. 415–428, 1981.

- [22] HELMBOLD, David P et al. On-Line Portfolio Selection Using Multiplicative Updates. **Mathematical Finance**, Wiley Online Library, v. 8, n. 4, p. 325–347, 1998.
- [23] VOVK, Volodya; WATKINS, Chris. Universal portfolio selection. In: PROCEEDINGS of the eleventh annual conference on Computational learning theory. [S.l.: s.n.], 1998. P. 12–23.
- [24] HAZAN, Elad et al. Logarithmic regret algorithms for online convex optimization. In: SPRINGER. COLT. [S.l.: s.n.], 2006. v. 4005, p. 499–513.
- [25] SINGER, Yoram. Switching portfolios. **International journal of neural systems**, World Scientific, v. 8, n. 04, p. 445–455, 1997.
- [26] LEVINA, Tatsiana; SHAFER, Glenn. Portfolio selection and online learning. **International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems**, World Scientific, v. 16, n. 04, p. 437–473, 2008.
- [27] KOZAT, Suleyman S; SINGER, Andrew C. Universal switching portfolios under transaction costs. In: IEEE. 2008 IEEE International Conference on Acoustics, Speech and Signal Processing. [S.l.: s.n.], 2008. P. 5404–5407.
- [28] KOZAT, Suleyman S; SINGER, Andrew C. Universal semiconstant rebalanced portfolios. **Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics**, Wiley Online Library, v. 21, n. 2, p. 293–311, 2011.
- [29] BORODIN, Allan; EL-YANIV, Ran; GOGAN, Vincent. Can we learn to beat the best stock. **Advances in Neural Information Processing Systems**, v. 16, 2003.
- [30] LI, Bin et al. PAMR: Passive aggressive mean reversion strategy for portfolio selection. **Machine learning**, Springer, v. 87, p. 221–258, 2012.
- [31] LI, Bin et al. Confidence weighted mean reversion strategy for online portfolio selection. **ACM Transactions on Knowledge Discovery from Data (TKDD)**, ACM New York, NY, USA, v. 7, n. 1, p. 1–38, 2013.

- [32] HUANG, Dingjiang et al. Robust Median Reversion Strategy for On-Line Portfolio Selection.(2013). In: PROCEEDINGS of the Twenty-Third International Joint Conference on Artificial Intelligence: IJCAI 2013. [S.l.: s.n.], 2012.
- [33] LI, Bin; HOI, Steven Chu Hong. **Online Portfolio Selection: Principles and Algorithms**. [S.l.]: CRC Press, 2015.
- [34] ALGOET, Paul H; COVER, Thomas M. Asymptotic optimality and asymptotic equipartition properties of log-optimum investment. **The Annals of Probability**, JSTOR, p. 876–898, 1988.
- [35] GYÖRFI, L.; SCHÄFER, D. Nonparametric prediction. In: SUYKENS, J.; HORVATH, G.; BASU, S. (Ed.). **Advances in Learning Theory: Methods, Models and Applications**. Amsterdam, The Netherlands: IOS Press, 2003. P. 339–354.
- [36] GYÖRFI, László; LUGOSI, Gábor; UDINA, Ferran. Nonparametric kernel-based sequential investment strategies. **Mathematical Finance**, v. 16, n. 2, p. 337–357, 2006.
- [37] GYÖRFI, László; UDINA, Ferran; WALK, Harald. Nonparametric nearest neighbor based empirical portfolio selection strategies. **Statistics and Decisions**, v. 26, n. 2, p. 145–157, 2008.
- [38] GYÖRFI, László; URBÁN, András; VAJDA, István. Kernel-based semi-log-optimal empirical portfolio selection strategies. **International Journal of Theoretical and Applied Finance**, World Scientific, v. 10, n. 03, p. 505–516, 2007.
- [39] OTTUCSÁK, Gábor; VAJDA, Imre. An asymptotic analysis of the mean-variance portfolio selection. **Statistics and Decisions**, v. 25, p. 63–88, 2007.
- [40] LI, Bin; HOI, Steven CH; GOPALKRISHNAN, Vivekanand. Corn: Correlation-driven nonparametric learning approach for portfolio selection. **ACM Transactions on Intelligent Systems and Technology (TIST)**, ACM New York, NY, USA, v. 2, n. 3, p. 1–29, 2011.

- [41] GYÖRFI, László; VAJDA, Imre. Growth optimal investment with transaction costs. In: PROCEEDINGS of the International Conference on Algorithmic Learning Theory. Budapest, Hungary: [s.n.], 2008. P. 108–122.
- [42] VOVK, Volodimir G. Aggregating strategies. In: PROCEEDINGS of the third annual workshop on Computational learning theory. [S.l.: s.n.], 1990. P. 371–386.
- [43] AKCOGLU, Karhan; DRINEAS, Petros; KAO, Ming-Yang. Fast universalization of investment strategies. **SIAM Journal on Computing**, SIAM, v. 34, n. 1, p. 1–22, 2004.
- [44] DAS, Puja; BANERJEE, Arindam. Meta optimization and its application to portfolio selection. In: PROCEEDINGS of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining. [S.l.: s.n.], 2011. P. 1163–1171.
- [45] HAZAN, Elad; SESHADHRI, Comandur. Efficient learning algorithms for changing environments. In: PROCEEDINGS of the 26th annual international conference on machine learning. [S.l.: s.n.], 2009. P. 393–400.
- [46] SALAS, Arnold. **Bayesian online learning for online portfolio selection**. 2020. Tese (Doutorado) – University of Oxford.
- [47] KHEDMATI, Majid; AZIN, Pejman. An online portfolio selection algorithm using clustering approaches and considering transaction costs. **Expert Systems with Applications**, Elsevier, v. 159, p. 113546, 2020.
- [48] GUO, Sini; GU, Jia-Wen; CHING, Wai-Ki. Adaptive online portfolio selection with transaction costs. **European Journal of Operational Research**, Elsevier, v. 295, n. 3, p. 1074–1086, 2021.
- [49] ZHANG, Yifan et al. Cost-sensitive portfolio selection via deep reinforcement learning. **IEEE Transactions on Knowledge and Data Engineering**, IEEE, v. 34, n. 1, p. 236–248, 2020.
- [50] DAI, Hong-Liang et al. An online portfolio strategy based on trend promote price tracing ensemble learning algorithm. **Knowledge-Based Systems**, Elsevier, v. 239, p. 107957, 2022.

- [51] DAI, Hong-Liang et al. Novel online portfolio selection algorithm using deep sequence features and reversal information. **Expert Systems with Applications**, Elsevier, v. 255, p. 124565, 2024.
- [52] ZHANG, Wenyu et al. Effective online portfolio selection for the long-short market using mirror gradient descent. **IEEE Signal Processing Letters**, IEEE, 2023.
- [53] HE, Jin'an; YIN, Shicheng; PENG, Fangping. Weak aggregating specialist algorithm for online portfolio selection. **Computational Economics**, Springer, v. 63, n. 6, p. 2405–2434, 2024.
- [54] HE, Hongliu; LI, Hua. A new boosting algorithm for online portfolio selection based on dynamic time warping and anti-correlation. **Computational Economics**, Springer, v. 63, n. 5, p. 1777–1803, 2024.
- [55] MICHAUD, Richard O. The Markowitz optimization enigma: Is 'optimized' optimal? **Financial analysts journal**, Taylor & Francis, v. 45, n. 1, p. 31–42, 1989.
- [56] MANDELROT, Benoit. The Variation of Certain Speculative Prices. **The Journal of Business**, v. 36, n. 4, p. 394–419, 1963.
- [57] FAMA, Eugene F. The behavior of stock-market prices. **The journal of Business**, JSTOR, v. 38, n. 1, p. 34–105, 1965.
- [58] CONT, Rama. Empirical properties of asset returns: stylized facts and statistical issues. **Quantitative finance**, IOP Publishing, v. 1, n. 2, p. 223, 2001.
- [59] MERTON, Robert C. Lifetime portfolio selection under uncertainty: The continuous-time case. **The review of Economics and Statistics**, JSTOR, p. 247–257, 1969.
- [60] BAJEUX-BESNAINOU, Isabelle; JORDAN, James V; PORTAIT, Roland. Dynamic asset allocation for stocks, bonds, and cash. **The Journal of Business**, JSTOR, v. 76, n. 2, p. 263–287, 2003.
- [61] JOBSON, J David; KORKIE, Robert M. Putting Markowitz theory to work. **The Journal of Portfolio Management**, Institutional Investor Journals, v. 7, n. 4, p. 70–74, 1981.

- [62] ELTON, Edwin J et al. **Modern portfolio theory and investment analysis**. [S.l.]: John Wiley & Sons, 2009.
- [63] BLACK, Fischer; JENSEN, Michael C.; SCHOLLES, Myron. The Capital Asset Pricing Model: Some Empirical Tests. In: JENSEN, Michael C. (Ed.). **Studies in the Theory of Capital Markets**. New York: Praeger, 1972.
- [64] FAMA, Eugene F; FRENCH, Kenneth R. The capital asset pricing model: Theory and evidence. **Journal of economic perspectives**, American Economic Association, v. 18, n. 3, p. 25–46, 2004.
- [65] LINTNER, John. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. **The review of economics and statistics**, JSTOR, p. 13–37, 1965.
- [66] CAMPBELL, John Y et al. Título do Capítulo 5. In: THE econometrics of financial markets. [S.l.]: Cambridge University Press, 1998. cap. 5, p. 181–217.
- [67] ROSENBERG, Barr; GUY, James. Prediction of beta from investment fundamentals: part one. **Financial Analysts Journal**, Taylor & Francis, v. 32, n. 3, p. 60–72, 1976.
- [68] ROSENBERG, Barr; GUY, James. Prediction of Beta from Investment Fundamentals: Part Two. **Financial Analysts Journal**, Taylor & Francis, v. 32, n. 4, p. 62–70, 1976.
- [69] SCHLUETER, Tobias; SIEVERS, Soenke. Determinants of market beta: the impacts of firm-specific accounting figures and market conditions. **Review of Quantitative Finance and Accounting**, Springer, v. 42, p. 535–570, 2014.
- [70] JACOB, Nancy L. The measurement of systematic risk for securities and portfolios: Some empirical results. **Journal of Financial and Quantitative Analysis**, Cambridge University Press, v. 6, n. 2, p. 815–833, 1971.
- [71] BLUME, Marshall E. Betas and their regression tendencies. **The Journal of Finance**, JSTOR, v. 30, n. 3, p. 785–795, 1975.

- [72] FABOZZI, Frank J; FRANCIS, Jack Clark. Beta as a random coefficient. **Journal of Financial and Quantitative Analysis**, Cambridge University Press, v. 13, n. 1, p. 101–116, 1978.
- [73] HARVEY, Campbell R. Time-varying conditional covariances in tests of asset pricing models. **Journal of Financial economics**, Elsevier, v. 24, n. 2, p. 289–317, 1989.
- [74] FERSON, Wayne E; HARVEY, Campbell R. The risk and predictability of international equity returns. **Review of financial Studies**, Oxford University Press, v. 6, n. 3, p. 527–566, 1993.
- [75] HARVEY, Campbell R. Predictable risk and returns in emerging markets. **The review of financial studies**, Oxford University Press, v. 8, n. 3, p. 773–816, 1995.
- [76] JAGANNATHAN, Ravi; WANG, Zhenyu. The conditional CAPM and the cross-section of expected returns. **The Journal of finance**, Wiley Online Library, v. 51, n. 1, p. 3–53, 1996.
- [77] FAMA, Eugene F; FRENCH, Kenneth R. Industry costs of equity. **Journal of financial economics**, Elsevier, v. 43, n. 2, p. 153–193, 1997.
- [78] FERSON, Wayne E; HARVEY, Campbell R. Conditioning variables and the cross section of stock returns. **The Journal of Finance**, Wiley Online Library, v. 54, n. 4, p. 1325–1360, 1999.
- [79] ADRIAN, Tobias; FRANZONI, Francesco. **Learning about beta: a new look at CAPM tests**. [S.l.], 2004.
- [80] ADRIAN, Tobias; FRANZONI, Francesco. Learning about beta: Time-varying factor loadings, expected returns, and the conditional CAPM. **Journal of Empirical Finance**, Elsevier, v. 16, n. 4, p. 537–556, 2009.
- [81] GROENEWOLD, Nicolaas; FRASER, Patricia. Time-varying estimates of CAPM betas. **Mathematics and Computers in Simulation**, Elsevier, v. 48, n. 4-6, p. 531–539, 1999.
- [82] AVRAMOV, Doron; CHORDIA, Tarun. Asset pricing models and financial market anomalies. **The Review of Financial Studies**, Oxford University Press, v. 19, n. 3, p. 1001–1040, 2006.

- [83] PETKOVA, Ralitsa; ZHANG, Lu. Is value riskier than growth? **Journal of Financial Economics**, Elsevier, v. 78, n. 1, p. 187–202, 2005.
- [84] JOSTOVA, Gergana; PHILIPPOV, Alexander. Bayesian analysis of stochastic betas. **Journal of Financial and Quantitative Analysis**, Cambridge University Press, v. 40, n. 4, p. 747–778, 2005.
- [85] ENGLE, Robert F; BOLLERSLEV, Tim. Modelling the persistence of conditional variances. **Econometric reviews**, Taylor & Francis, v. 5, n. 1, p. 1–50, 1986.
- [86] BAILLIE, Richard T; DEGENNARO, Ramon P. Stock returns and volatility. **Journal of financial and Quantitative Analysis**, Cambridge University Press, v. 25, n. 2, p. 203–214, 1990.
- [87] BRAILSFORD, Timothy J; FAFF, Robert W. Testing the conditional CAPM and the effect of intervaling: a note. **Pacific-Basin Finance Journal**, Elsevier, v. 5, n. 5, p. 527–537, 1997.
- [88] LEDOIT, Olivier; SANTA-CLARA, Pedro; WOLF, Michael. Flexible multivariate GARCH modeling with an application to international stock markets. **Review of Economics and Statistics**, MIT Press, v. 85, n. 3, p. 735–747, 2003.
- [89] GLOSTEN, Lawrence R; JAGANNATHAN, Ravi; RUNKLE, David E. On the relation between the expected value and the volatility of the nominal excess return on stocks. **The journal of finance**, Wiley Online Library, v. 48, n. 5, p. 1779–1801, 1993.
- [90] ZAKOIAN, Jean-Michel. Threshold heteroskedastic models. **Journal of Economic Dynamics and control**, Elsevier, v. 18, n. 5, p. 931–955, 1994.
- [91] FAFF, Robert W; HILLIER, David; HILLIER, Joseph. Time varying beta risk: An analysis of alternative modelling techniques. **Journal of Business Finance & Accounting**, Wiley Online Library, v. 27, n. 5-6, p. 523–554, 2000.
- [92] CHOUDHRY, Taufiq; WU, Hao. Forecasting the weekly time-varying beta of UK firms: GARCH models vs. Kalman filter method. **The European Journal of Finance**, Taylor & Francis, v. 15, n. 4, p. 437–444, 2009.

- [93] MESTRE, Roman. A wavelet approach of investing behaviors and their effects on risk exposures. **Financial Innovation**, SpringerOpen, v. 7, n. 1, p. 1–37, 2021.
- [94] COSTA, LEANDRO SANTOS DA et al. Conditional pricing model with heteroscedasticity: Evaluation of Brazilian funds. **Revista de Administração de Empresas**, SciELO Brasil, v. 59, p. 225–241, 2019.
- [95] DURBIN, James; KOOPMAN, Siem Jan. **Time series analysis by state space methods**. [S.l.]: OUP Oxford, 2012. v. 38.
- [96] BRYZGALOVA, Svetlana. Spurious factors in linear-asset pricing models. **Working paper, Stanford University**, 2017.
- [97] GUIISO, Luigi; JAPPELLI, Tullio; TERLIZZESE, Daniele. Income risk, borrowing constraints, and portfolio choice. **The American economic review**, JSTOR, p. 158–172, 1996.
- [98] JORION, Philippe. Risk2: Measuring the risk in value at risk. **Financial analysts journal**, Taylor & Francis, v. 52, n. 6, p. 47–56, 1996.
- [99] JORION, Philippe. **Value at risk: the new benchmark for managing financial risk**. [S.l.]: The McGraw-Hill Companies, Inc., 2007.
- [100] SUPERVISION, BASLE COMMITTEE ON BANKING. Amendment to the capital accord to incorporate market risks. **Basle, Switzerland, jan**, 1996.
- [101] ABAD, Pilar; BENITO, Sonia; LÓPEZ, Carmen. A comprehensive review of Value at Risk methodologies. **The Spanish Review of Financial Economics**, Elsevier, v. 12, n. 1, p. 15–32, 2014.
- [102] ANTONELLI, Sabrina; IOVINO, Maria Gabriella. Optimization of Monte Carlo procedures for value at risk estimates. **Economic Notes**, Wiley Online Library, v. 31, n. 1, p. 59–78, 2002.
- [103] ARTZNER, Philippe et al. Coherent measures of risk. **Mathematical finance**, Wiley Online Library, v. 9, n. 3, p. 203–228, 1999.
- [104] ACERBI, Carlo; TASCHE, Dirk. On the coherence of expected shortfall. **Journal of banking & finance**, Elsevier, v. 26, n. 7, p. 1487–1503, 2002.

- [105] BASEL COMMITTEE ON BANKING SUPERVISION. **Basel III Monitoring Report**. Basel, Switzerland, 2013. Accessed: 08/05/2024. Disponível em: <<https://www.bis.org/publ/bcbs289.pdf>>.
- [106] STEINKI, Oliver; MOHAMMAD, Ziad. Common metrics for performance evaluation: Overview of popular performance measurement ratios. **Available at SSRN 2662054**, 2015.
- [107] CESARONE, Francesco; MARTINO, Manuel L; TARDELLA, Fabio. Mean-Variance-VaR portfolios: MIQP formulation and performance analysis. **OR Spectrum**, Springer, v. 45, n. 3, p. 1043–1069, 2023.
- [108] BACON, Carl; CHAIRMAN, S. How sharp is the Sharpe ratio? Risk-adjusted Performance Measures. **Statpro**, nd, 2009.
- [109] DOCHOW, Robert. **Online algorithms for the portfolio selection problem**. [S.l.]: Springer, 2016.
- [110] BELL, Robert M; COVER, Thomas M. Competitive optimality of logarithmic investment. **Mathematics of Operations Research**, INFORMS, v. 5, n. 2, p. 161–166, 1980.
- [111] BELL, Robert; COVER, Thomas M. Game-theoretic optimal portfolios. **Management Science**, INFORMS, v. 34, n. 6, p. 724–733, 1988.
- [112] LI, Bin; HOI, Steven CH. Online portfolio selection: A survey. **ACM Computing Surveys (CSUR)**, ACM, v. 46, n. 3, p. 35, 2014.
- [113] RESTELLI, MARCELLO; VITTORI, EDOARDO; TROVO, FRANCESCO et al. Online gradient descent for online portfolio optimization with transaction costs. Politecnico di Milano, 2018.
- [114] ROSEN, Kenneth H. **Discrete mathematics and its applications**. [S.l.]: The McGraw Hill Companies, 2007.
- [115] CHEN, Yunmei; YE, Xiaojing. Projection onto a simplex. **arXiv preprint arXiv:1101.6081**, 2011.

- [116] YAMIM, J. D. M. et al. Seleção de Portfólio baseado em Otimização Convexa Online. In: INFORMATION, Missing editor (Ed.). **Proceedings of the XXXVIII Ibero-Latin American Congress on Computational Methods in Engineering**. Florianopolis: [s.n.], 2017. v. XXXVII.
- [117] YAMIM, J. D. M. **Um modelo de seleção de carteiras de ações baseado em otimização convexa online**. 2018. Diss. (Mestrado) – Universidade Federal de Juiz de Fora (UFJF), Brasil. Primeiro Orientador: Borges, Carlos Cristiano Hasenclever. Co-orientador: Fonseca Neto, Raul.
- [118] HELSKE, Jouni. KFAS: Exponential Family State Space Models in R. **Journal of Statistical Software**, v. 78, n. 10, p. 1–39, 2017. DOI: [10.18637/jss.v078.i10](https://doi.org/10.18637/jss.v078.i10).
- [119] R CORE TEAM. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2019. Disponível em: [<http://www.R-project.org/>](http://www.R-project.org/).
- [120] SEBER, George AF. **A matrix handbook for statisticians**. [S.l.]: John Wiley & Sons, 2008. v. 15.
- [121] MCKINNEY, Wes et al. Data structures for statistical computing in python. In: AUSTIN, TX. PROCEEDINGS of the 9th Python in Science Conference. [S.l.: s.n.], 2010. v. 445, p. 51–56.
- [122] VAN ROSSUM, Guido; DRAKE, Fred L. **Python 3 Reference Manual**. Scotts Valley, CA: CreateSpace, 2009. ISBN 1441412697.
- [123] GRANJA, Daniel de Moraes et al. **Modelo de inferência não linear para alocação de carteira**. 2004. Tese (Doutorado) – Universidade de São Paulo, São Paulo, SP. Disponível em: [<https://www.teses.usp.br/teses/disponiveis/92/92131/tde-05042022-101739/en.php>](https://www.teses.usp.br/teses/disponiveis/92/92131/tde-05042022-101739/en.php).
- [124] SANTOS, André Alves Portela; TESSARI, Cristina. Técnicas quantitativas de otimização de carteiras aplicadas ao mercado de ações brasileiro. **Revista Brasileira de Finanças**, Sociedade Brasileira de Finanças, v. 10, n. 3, p. 369–393, 2012.

- [125] YAMIM, J. D. M.; BORGES, Carlos Cristiano Hasenclever; NETO, Raul Fonseca. Online Portfolio Optimization with Risk Control. **Trends in Computational and Applied Mathematics**, SciELO Brasil, v. 22, p. 475–493, 2021.
- [126] YAMIM, J. D. M.; BORGES, CCH; NETO, RF. Portfolio Optimization Via Online Gradient Descent and Risk Control. **Computational Economics**, Springer, v. 62, n. 1, p. 361–381, 2023.
- [127] THOMÉ NETO, Cesar; LEAL, Ricardo Pereira Câmara; ALMEIDA, Vinício de Souza. Um índice de mínima variância de ações brasileiras. **Economia Aplicada**, SciELO Brasil, v. 15, p. 535–557, 2011.
- [128] GOTARDELO, Davi Riani. **Otimização multiobjetivo utilizando algoritmos evolutivos em seleção de carteiras: uma abordagem envolvendo ômega, assimetria e antifragilidade**. 2020. Tese (Doutorado) – Universidade Federal de Juíz de Fora, Juíz de Fora, MG. Disponível em:
<<https://repositorio.ufjf.br/jspui/handle/ufjf/12525>>.
- [129] CIVILETTI, Fabio; CAMPANI, Carlos Heitor; ROQUETE, Raphael. Equally weighted portfolios and “momentum effect”: an interesting combination for unsophisticated investors? **BBR. Brazilian Business Review**, SciELO Brasil, v. 17, p. 506–522, 2020.
- [130] DUARTE, Flavio Gabriel; CASTRO, Leandro Nunes de. A fuzzy clustering algorithm for portfolio selection. In: IEEE. 2019 IEEE 21st Conference on Business Informatics (CBI). [S.l.: s.n.], 2019. v. 1, p. 414–418.
- [131] PAIVA, Felipe Dias et al. Decision-making for financial trading: A fusion approach of machine learning and portfolio selection. **Expert Systems with Applications**, Elsevier, v. 115, p. 635–655, 2019.
- [132] FREITAS, Washington Burkart; JUNIOR, João Roberto Bertini. Random walk through a stock network and predictive analysis for portfolio optimization. **Expert Systems with Applications**, Elsevier, v. 218, p. 119597, 2023.

- [133] FERREIRA, Fernando GDC et al. Parallel moeas for combinatorial multiobjective optimization model of financial portfolio selection. In: IEEE. 2018 IEEE Congress on Evolutionary Computation (CEC). [S.l.: s.n.], 2018. P. 1–8.
- [134] CONEGUNDES, Leonardo; PEREIRA, Adriano C Machado. Beating the stock market with a deep reinforcement learning day trading system. In: IEEE. 2020 International Joint Conference on Neural Networks (IJCNN). [S.l.: s.n.], 2020. P. 1–8.
- [135] SILVA DE SOUZA, Matheus José et al. Do small caps generate above average returns in the Brazilian stock market? **Review of development finance**, AfricaGrowth Institute, v. 8, n. 1, p. 18–24, 2018.
- [136] ULUDAG, G et al. ML Estimation of Distribution Parameters for VaR Calculation Using Evolutionary Algorithm. **WSEAS Transaction on Business and Economics**, 2005.