BRUNA MAIA BARBOSA
DIEGO MANZO DE ARRUDA
LIVIA LEITE DE ALMEIDA

OTIMIZAÇÃO E PERFORMANCE DE PORTFÓLIOS DE INVESTIMENTOS POR METÓDOS DE PROGRAMAÇÃO MATEMÁTICA

São Paulo
2022

BRUNA MAIA BARBOSA
DIEGO MANZO DE ARRUDA
LIVIA LEITE DE ALMEIDA

OTIMIZAÇÃO E PERFORMANCE DE PORTFÓLIOS DE INVESTIMENTOS POR METÓDOS DE PROGRAMAÇÃO MATEMÁTICA

Trabalho apresentado à Escola Politécnica da Universidade de São Paulo para obtenção do Título de Engenheiro Eletricista com ênfase em Controle e Automação

São Paulo
2022

BRUNA MAIA BARBOSA
DIEGO MANZO DE ARRUDA
LIVIA LEITE DE ALMEIDA

OTIMIZAÇÃO E PERFORMANCE DE PORTFÓLIOS DE INVESTIMENTOS POR METÓDOS DE PROGRAMAÇÃO MATEMÁTICA

Trabalho apresentado à Escola Politécnica da Universidade de São Paulo para obtenção do Título de Engenheiro Eletricista com ênfase em Controle e Automação

Orientador:
Oswaldo Luiz do Valle Costa

São Paulo
2022

Catalogação-na-publicação

Barbosa, Bruna Maia
Otimização e performance de portfólios de investimento por métodos de programação matemática / B. M. Barbosa, D. M. Arruda, L. L. Almeida -- São Paulo, 2022.
53 p.

Trabalho de Formatura - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Telecomunicações e Controle.

1.Portfólios de investimento 2.Programação Matemática 3.Otimização I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Telecomunicações e Controle II.t. III.Arruda, Diego Manzo IV.Almeida, Livia Leite de

AGRADECIMENTOS

Agradecemos ao professor Oswaldo Luiz, nosso orientador, pelo acompanhamento e mentoria durante o ano e à nossa família, amigos e parceiros pelo suporte constante durante o curso.

RESUMO

O projeto tem como objetivo compor uma carteira representativa do benchmark e usar métodos de otimização e dados históricos para rastrear seu comportamento.

São estudados e aplicados sete métodos de otimização e o Ibovespa é o benchmark escolhido. Dentre os métodos, quatro deles são métodos de programação linear e três são métodos computacionais. Os métodos de programação linear minimizam o maior desvio absoluto entre o retorno do portfólio e do benchmark ou a soma dos desvios absolutos. Já os métodos computacionais, minimizam o erro quadrático, o erro não sistêmico ou a variância do erro.

Por fim, é feita uma análise de performance dos resultados através do Índice de Sharpe, do β de cada otimização e da capacidade da carteira de se aproximar do retorno do índice.

Palavras-Chave – Rastreamento, Ibovespa, Otimização, Retorno, Portfólio.

ABSTRACT

The project aims to compose a representative portfolio of the benchmark and use optimization methods and historical data to track its behavior.

Seven optimization methods are studied and applied and Ibovespa is the chosen benchmark. Among the methods, four of them are linear programming methods and three are computational methods. Linear programming methods minimize the largest absolute deviation between the portfolio and benchmark return or the sum of the absolute deviations. Computational methods minimize the quadratic error, the non-systemic error or the error variance.

At last, a performance analysis of the results is made through the Sharpe Ratio, the β of each optimization and the portfolio's ability to approach the index return.

Keywords – Tracking, Ibovespa, Optimization, Return, Portfolio.

LISTA DE FIGURAS

LISTA DE TABELAS

SUMÁRIO

1 INTRODUÇÃO AO TEMA

A Bolsa de Valores coleta e organiza informações sobre os negócios realizados em cada pregão e os divulga por meio de índices que mostram o comportamento do mercado. Entre os índices divulgados, existe o Ibovespa (Índice de Bolsa de Valores de São Paulo), que foi criado em 1968 e é o indicador de desempenho das ações negociadas na B3, sigla que representa a bolsa de valores oficial do Brasil, sediada na cidade de São Paulo, e faz referência às letras iniciais de Brasil, Bolsa, Balcão.

Esse indicador reúne as empresas mais relevantes do mercado de capitais brasileiros e serve como referência para os investidores. O Ibovespa também é uma forma de visualizar o desempenho da economia brasileira à longo prazo:

Fonte: TradingView.

Figura 1: Histórico do índice BOVESPA de 1995 até 2022

A partir dessa visão histórica, é possível identificar como grandes eventos afetam a economia, o que pode ser visto no gráfico acima, no ano de 2020, quando as cotações do Ibovespa caíram cerca de 44% por conta da pandemia da COVID-19, sendo um dos piores momentos na história do índice.

Por ter essa característica de indicador de mercado, é muito comum que investidores tenham interesse em rastreá-lo, fazendo com que sua carteira acompanhe o índice e tenha um retorno semelhante. Em certos fundos, o administrador é avaliado de acordo com a capacidade de sua carteira seguir ou superar o comportamento de um índice. Os riscos envolvidos são sistêmicos, uma vez que o índice não obrigatoriamente vai crescer com o tempo, mas é uma forma interessante de se expor ao mercado e conseguir retornos maiores do que os retornos oferecidos por investimentos em renda fixa, por exemplo.

Um desafio dessa estratégia é o fato de os índices serem compostos por um número grande de ativos e muitos serem de difícil negociação, devido à sua liquidez. Dessa forma, é necessária uma otimização da carteira, decidindo quais ativos escolher e em qual proporção, de forma a se aproximar o máximo possível do benchmark.

Ter um parâmetro para comparação, tanto na renda fixa quanto na renda variável pode ajudar a definir se o investimento está sendo positivo ou não. Em geral, na renda fixa, o benchmark costuma ser o CDI, enquanto na renda variável, costuma ser o Ibovespa. Por isso, muitos investidores que possuem carteira de ações e querem avaliar a sua performance costumam fazer a comparação de seus investimentos com o Ibovespa.

A principal motivação para a realização desse projeto de rastreamento é definir qual é o melhor método para conseguir um bom desempenho em uma carteira de ações quando comparada ao índice.

2 TRABALHOS RELACIONADOS

Essa sessão traz uma breve explicação de alguns temas necessários para entender os modelos escolhidos e as simulações de otimização em seguida.

2.1 Fundo Passivo

Um fundo passivo é um fundo que tem a estratégia de evitar correr riscos não sistêmicos, de forma a obter retornos que se aproximem do benchmark escolhido.

Para seguir com essa estratégia, a solução mais intuitiva é montar uma carteira com as mesmas proporções de ativos que a composição do benchmark. Entretanto, observa-se que os índices no geral tem uma quantidade grande de ativos em sua composição e que nem sempre esses ativos possuem a liquidez necessária para compor um fundo, dificultando negociações.

Diante desse cenário, faz-se necessário o uso de métodos de otimização da carteira que, por exemplo, minimizem o erro entre o retorno da carteira e do benchmark.

2.2 O Modelo de Markowitz

O modelo proposto por H. Markowitz em 1952, amplamente conhecido, permite que o investidor componha uma carteira com a relação entre risco e retorno que faça mais sentido para o seu objetivo a partir de uma fronteira eficiente. É um problema de otimização quadrática com restrições lineares.

O modelo assume que os investidores avaliam sua carteira de acordo com o retorno esperado e com a variância das taxas de retorno no período analisado e parte da premissa de que se os investidores precisassem escolher entre duas carteiras com retornos iguais, escolheriam a de menor risco e entre duas carteiras de risco igual, escolheriam a de maior retorno.

Para o caso de uma carteira com retorno P, que apresenta n ativos de risco, com retornos $R_1, ...R_n$, retornos esperados $r_1, ...r_n$ e matriz de covariância Σ, uma proporção

ω_j é investida no ativo que possui retorno R_j:

$$P = \sum_{j=1}^{n} \omega_j R_j$$

Usando uma notação vetorial para todos os valores de 1 à n e sabendo que:

$$r_i = E(R_i)$$

$$\Sigma = cov(R) = E((R - r)(R - r)')$$

Sendo μ a média do retorno P, dada por:

$$\mu = E(P) = E(\omega' R) = \omega' r$$

E sendo σ^2 a variância de P, dada por :

$$\sigma^2 = \omega' \Sigma \omega$$

Usa-se a notação vetorial a seguir:

$$e = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \in \Re$$

$$\omega = \begin{bmatrix} \omega_1 \\ \vdots \\ \omega_n \end{bmatrix} \in R^n.$$

Tem-se o problema de otimização abaixo para uma rentabilidade μ almejada:

$$min \omega' \Sigma \omega$$

$$sujeito \quad a$$

$$\omega' r = \mu,$$

$$\omega' e = 1,$$

$$\omega \in R^n.$$

2.3 Índice de Sharpe - IS

O índice de Sharpe, desenvolvido por William Sharpe, em 1966, é uma das formas de avaliação de desempenho mais conhecidas e amplamente usada para avaliar fundos de investimentos. O índice se relaciona com a teoria de seleção de carteira, principalmente no modelo CAPM, que defende que nenhuma carteira deve ter seu IS maior do que a carteira de mercado [1]. A utilização correta do índice depende da estimativa correta dos seus parâmetros e de como será aplicado. O indice definido a seguir:

$$IS = \frac{R_e - R_f}{\sigma}$$

Sendo R_e o retorno esperado, R_f o retorno livre de risco e σ a volatilidade. A volatilidade representa o risco da carteira, e o retorno livre de risco pode ser representado pelo CDI, por exemplo. O índice deve ser usado pelo investidor para analisar a sua carteira. Caso a carteira não possua investimentos com risco, seleciona-se aquela com maior IS.

2.4 CAPM - Modelo de Precificação de Ativos de Capital

O CAPM é um modelo que analisa a relação entre o risco sistêmico e o retorno de um investimento para determinar uma precificação de ativos de risco.

O método assume algumas premissas, como a que os investidores fazem uso da otimização por média variância para alocar seus recursos. O modelo avalia o ativo da seguinte forma:

$$E(R) = R_f + \beta(R_m - R_f)$$

Sendo $E(r)$ o retorno esperado do investimento, R_f o retorno livre de risco, R_m a rentabilidade oferecida pelo mercado e $(R_m - R_f)$ o chamado no mercado como "prêmio pelo risco".

A mensuração de risco é alterada de σ para β, fazendo uma comparação do risco tomado com o risco tomado pela carteira de mercado, ou seja, um $\beta = 1$ significa correr um risco igual ao mercado. O β de uma carteira pode ser calculado utilizando a seguinte expressão:

$$\beta = \frac{cov(R_M, R)}{(\sigma_M)^2}$$

Sabendo que R_M é o retorno da carteira de mercado e que σ_M a variância do retorno da carteira de mercado. O risco de um ativo pode ser separado em dois tipos, os quais são associados pelo β: sistêmico, que é inerente ao mercado, e não-sistêmico, que pode ser reduzido pela diversificação da carteira.

3 MODELOS DE RASTREAMENTO SELECIONADOS

3.1 Modelos de Programação Linear

Para a otimização de carteiras, é comum rastrear o erro quadrático por possuir melhores qualidades estatísticas e por sua facilidade de implementação computacional. Entretanto, em várias situações o erro linear ou o desvio absoluto entre o resultado do portfólio e do benchmark podem se mostrar mais relevantes ou serem resultados mais fáceis e intuitivos de interpretar, mostrando-se melhores escolhas.

Para os modelos de programação linear mostrados a seguir, faz-se a minimização do erro de rastreamento linear. Sendo ω a carteira a ser rastreada, ω_B o benchmark, P o retorno, e R a taxa de retorno, tem-se que o erro de rastreamento é descrito por:

$$P_e = P - P_B = (\omega - \omega_B)'R$$

A partir de uma série histórica com T dados sobre o índice escolhido para o rastreamento e de p ativos básicos que compõem o índice, y sendo o vetor T-dimensional com as observações do índice que se deseja rastrear, e Γ a matriz T $\times$ p com T dados sobre os p ativos básicos que se considera:

$$\Gamma = \begin{bmatrix} R_1(1) & \cdots & R_p(1) \\ \vdots & \ddots & \vdots \\ R_1(T) & \cdots & R_p(T) \end{bmatrix}$$

Os modelos de rastreamento abaixo foram propostos [2]:

I. MAD (Mean Absolute Deviation): O modelo MAD faz a minimização dos desvios absolutos da média. Os pesos do portfólio são determinados de forma que a soma dos desvios absolutos entre os retornos do benchmark e os retornos da carteira seja minimizada:

$$min \quad e'(|y - \Gamma\omega|)$$

$$com \quad e = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \in \Re^T$$

Em modelos que utilizam o erro quadrático, os desvios da média são ressaltados por conta de serem elevados ao quadrado, portanto são modelos mais sensíveis do que o modelo MAD.

II. Min-Max: Minimiza o maior desvio entre os retornos da carteira e os do benchmark. A estratégia do modelo é ser uma proteção para os piores casos possíveis. O modelo Min-Max é menos robusto em relação à desvios do que o MAD.

$$min \quad (max \quad |y - \Gamma\omega|)$$

Uma percepção intuitiva de risco é observar se o retorno do portfólio está abaixo do retorno do benchmark, por isso foram criadas as variações "Downside" dos modelos MAD e MinMax, que possuem a restrição de fazer a minimização apenas se o retorno representar um risco. O traço nos vetores $\overline{\omega}$ e $\overline{y}$ indica que só estão contidos os valores em que o retorno do benchmark se mostrou menor do que o do portfolio:

III. MADD (Mean Absolute Downside Deviation): É uma variação do modelo MAD que penaliza apenas quando o retorno está abaixo do benchmarking.

$$min \quad e'(|\overline{y - \Gamma\omega}|)$$

IV. DMin-Max (Downside Min-Max): É uma variação do modelo Min-Max que penaliza apenas quando o retorno está abaixo do benchmarking. O maior desvio negativo é minimizado.

$$min \quad (max \quad |\overline{y - \Gamma\omega}|)$$

3.2 Métodos Computacionais

Os modelos de rastreamento abaixo foram propostos [3]:

3.2.1 Mínima Variância do Erro

O modelo de mínima variância do erro tem por objetivo minimizar a volatilidade existente entre o retorno da carteira e o do benchmark.

Ele possui o retorno esperado da carteira de mínima variância global $\mu_g = 0$ e número de ativos $p < n$. Para esse caso, a carteira ω que contém a distribuição otimizada dos ativos possui a mesma dimensão n do benchmark, porém, apenas os p ativos escolhidos são não-nulos. O seu modelo matemático é representado a seguir:

$$min \quad (\omega - \omega_B)' \sum (\omega - \omega_B)$$

$$sujeito \quad a$$

$$(\omega - \omega_B)' r = 0$$

$$\omega \geq 0$$

$$\omega' e = 1$$

$$\omega_i = 0, i = n - p + 1, ..., n$$

$$\omega \in R^n$$

3.2.2 Mínimo Erro Não Sistêmico

Antes de apresentar o modelo de mínimo risco não-sistêmico, é necessário relembrar o conceito de carteira de mercado: é uma carteira eficiente (ou seja, se encontra na hipérbole do modelo de Markowitz) que representa o mercado, como por exemplo a composição do índice Ibovespa. Essa carteira assume como premissa que o mercado faça otimização por média-variância (por isso a carteira se encontra na hipérbole do modelo de Markowitz) e assume também que não há arbitragem (mercado em equilíbrio).

Outro conceito que deve ser relembrado antes de prosseguir para o modelo propriamente dito é o conceito de β . Este muda a mensuração de risco do σ para β, fazendo uma comparação do risco tomado com o risco tomado pela carteira de mercado, ou seja, um β = 1 significa correr um risco igual ao mercado. O β de uma carteira pode ser calculado utilizando a seguinte expressão:

$$\beta = \frac{cov(P_M, P)}{(\sigma_M)^2}$$

Com P sendo o retorno da carteira e P_M o retorno da carteira de mercado. De posse destes conceitos, assumindo uma carteira com retorno P, retorno esperado μ e variância

σ^2 Temos que:

$$\sigma^2 = Var(P) = \beta^2 Var(P_M) + Var(Z) = \beta^2(\sigma_M)^2 + Var(Z)$$

A parcela $\beta^2(\sigma_M)^2$ representa o erro sistêmico, que não pode ser evitado, enquanto que a $Var(Z)$ representa o risco não-sistêmico, que pode ser evitado com uma boa diversificação.

Para o caso de rastrear a carteira de mercado, queremos $\beta = 1$, então podemos enunciar o modelo da seguinte forma:

$$min \quad \omega' \sum \omega$$

$$sujeito \quad a$$

$$\omega \geq 0$$

$$\omega' e = 1$$

$$\sum_{i=1}^{p} \beta_i \omega_i = 1$$

$$\omega \in R^p$$

3.2.3 Mínimo Erro Quadrático

O modelo de mínimo erro quadrático tem por objetivo encontrar a carteira que minimize a diferença ao quadrado entre seus retornos e os retornos do benchmark. O primeiro passo necessário para o desenvolvimento do método é definir os p ativos que serão considerados na otimização. Dado um benchmark com n ativos, selecionam-se p ativos, tal que p $\leq$ n.

O critério de escolha dos ativos que serão utilizados na otimização é, em geral, o percentual de participação no índice que será rastreado. O problema consiste, então, na minimização do seguinte erro:

$$err = (y - \Gamma\omega)$$

Sendo ω a composição da carteira que será escolhida ao fim da otimização. O seu modelo matemático é representado a seguir:

$$min(y - \Gamma\omega)'(y - \Gamma\omega)$$

$$sujeito \quad a$$

$$\omega \geq 0$$

$$\omega' e = 1$$

$$\omega \in R^p$$

4 AQUISIÇÃO DOS DADOS

Uma etapa crucial para a simulação dos modelos é garantir os dados necessários para alimentar o modelo e realizar a simulação. Para isso é necessário encontrar uma fonte de dados adequada e confiável para fazer a extração. No caso deste trabalho, foi utilizado o site yahoo.finance, site que possui dados dos mais diversos ativos disponíveis no mercado financeiro e tem a facilidade de fornecê-los nos períodos necessários e através de scripts em Python.

Para obter os dados foi utilizada a biblioteca "web.get_data_yahoo"em um script em Python, em que todos os dados foram coletados. Após isso, foram necessários alguns tratamentos para os datasets ficassem no formato adequado para o código. Sobre o tratamento, pode-se citar a seleção e formatação do nome das colunas a serem utilizadas e o cálculo da variação, por exemplo. Através desse script, para a coleta e o tratamento dos dados, basta que seja definido o período e os ativos que serão selecionados.

De inicio foram escolhidos 5 ativos: VALE3 (Vale), PETR4 (Petrobras), ITUB4 (Itaú Unibanco), BBDC4 (Banco Bradesco) e B3SA3 (B3), por serem os mais representativos do Ibovespa, índice utilizado como benchmark. Contudo, vale ressaltar que, como a aquisição e transformação dos dados foi automatizada, é possível mudar tanto a quantidade quanto os ativos escolhidos.

4.1 Criação de repositório Online

Com o surgimento de um script em Python para a aquisição de dados, veio à tona uma questão muito importante para qualquer projeto que envolva o desenvolvimento de software: o controle de versões. Por isso, se tornou necessário utilizar um software que ajudasse nesse controle. O software escolhido foi o GitHub, pois com ele é possível criar um repositório na nuvem para armazenar os códigos e também branchs para organizar as versões do código.

5 IMPLEMENTAÇÃO

5.1 Automação da Solução

Com a complexidade do projeto e sendo necessário executar diversos scripts para obter os resultados, naturalmente, o processo começa a se tornar mais demorado e mais complicado, fazendo com que seja necessário ter muito claro toda a ordenação de passos para que os dados sejam armazenados e as simulações fiquem corretas. Por isso, foi desenvolvida uma automação para que todos os processos de execução dos scripts sejam feitos automaticamente, desde a aquisição, até a validação dos resultados. Foi utilizada a linguagem Python por meio de um script.

Esse script é responsável pela orquestração de diversos outros scripts, tanto em Python como em Matlab, sendo este utilizado apenas como um motor para execução da otimização, enquanto que os demais processos seriam feitos todos na linguagem Python.

Além da aquisição de dados e da simulação dos scripts do Matlab para o cálculo da otimização, também é necessário realizar a análise dos resultados obtidos. Para isso, em substituição do Matlab, foram utilizados notebooks Jupyter, os quais são muito utilizados nas áreas de análise de dados, por utilizarem a linguagem Python em um ambiente onde o usuário tem um controle melhor das variáveis e de quais trechos do script Python estão sendo executados.

Neste instante do fluxo de execução do projeto é que se percebe a importância da automação dos processos de aquisição e simulação, uma vez que, com estes processos automatizados, o usuário pode se preocupar apenas com as análises que deseja realizar, sem precisar alterar parâmetros dentro dos comandos para gerar os dados que necessita, fazendo o processo de geração dos dados da análise muito mais rápido e menos suscetível a erros.

A seguir há uma representação gráfica dos scripts que são executados pela automação.

Na Figura 2, é possível ver que o script principal, responsável por executar a automação, é o chamado "simulation tool". Os scripts dentro do script principal são responsáveis por executar os processos de aquisição e simulação da automação. No lado esquerdo da figura há três arquivos Python que serão utilizados na aquisição de dados.

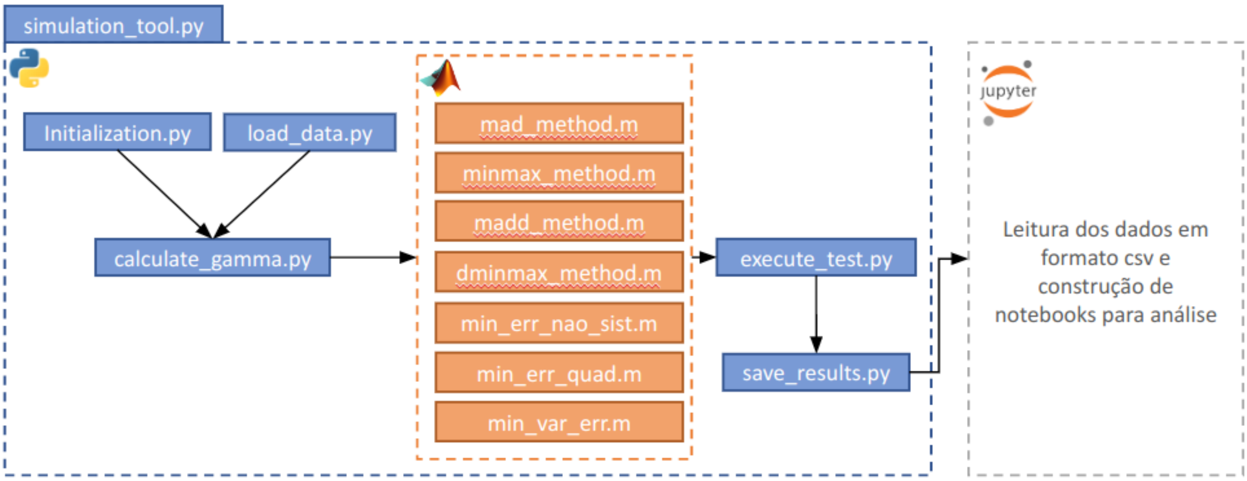

Figura 2: Fluxo de execução da automação

O primeiro deles, "initialization", é o responsável por adquirir os dados para a simulação desejada. Estes dados são parâmetros como as datas e os ativos que serão utilizados durante a simulação e serão carregados via prompt de comandos em tempo de execução.

O segundo script a ser executado é o "load data", responsável por interagir com a API que fornece os dados dos ativos e do benchmark selecionado e depois salvá-los em formato csv.

Já o terceiro script, "calculate gamma", tem o papel de transformar os dados da aquisição nas matrizes que serão utilizadas durante a simulação no Matlab.

De posse de todos os parâmetros necessários para a simulação, o script principal executa as funções no Matlab, que retornam o valor encontrado para a otimização, ou seja, o peso de cada ativo na carteira.

Uma vez que os pesos dos ativos na carteira são adquiridos, é executada uma etapa de validação dos resultados obtidos, ou seja, os pesos são fixados e é simulado um período para avaliar o rastreamento.

Finalmente, o processo de automação acaba com o script "save results"sendo executado para que os resultados dos métodos sejam salvos em formato csv.

Após a obtenção dos resultados dos métodos, é feita a execução do Jupyter Notebook, com o objetivo de gerar, automaticamente, todos os gráficos que serão utilizados para as análises. Assim, ao fim das simulalções, apenas com as informações sobre as datas de otimização e validação, é possível gerar as visualizações e fazer quaisquer análises que se desejar sobre os resultados.

5.2 Implementação das Simulações dos Modelos de Programação Linear

Na seção a seguir há uma breve explicação sobre a implementação de cada um dos modelos de otimização escolhidos. Foram usadas as funções *linprog* ou *intlinprog* do Matlab para fazer as implementações. A minimização é realizada da seguinte forma:

$$min_x f^T x$$

$$sujeito \quad a$$

$$A.x \leq B,$$

$$Aeq.x = beq$$

$$lb \leq x$$

I. MAD (Mean Absolute Deviation): O método é descrito por:

$$min \quad e'(|y - \Gamma\omega|)$$

$$com \quad e = (1...1) \in \Re^T$$

Para a implementação, chegou-se aos seguintes valores para as matrizes e vetores da função:

$$var = 2.row + n_{fund};$$

sendo n_{fund} o número de ativos usados;

$$A = [];$$

$$b = [];$$

e as seguintes variáveis auxiliares:

$$I = eye(row);$$

$$e = ones(n_{fund}, 1);$$

$$et = ones(row, 1);$$

$$z = zeros(1, row);$$

$$z1 = zeros(1, n_{fund});$$

e por fim, define-se:

$$f = [et' \quad et' \quad z1];$$

$$Aeq = [I \quad -I.\Gamma; z \quad z \quad e'];$$

$$beq = [y \quad 1];$$

$$lb = zeros(var, 1);$$

II. Min-Max: o modelo tem o objetivo de realizar a seguinte otimização:

$$min \quad (max|y - \Gamma\omega|)$$

Para a implementação, chegou-se aos seguintes valores para as matrizes e vetores da função:

$$var = n_{fund} + 1;$$

sendo n_{fund} o número de ativos usados. Tem-se as variáveis auxiliares a seguir:

$$e = ones(n_{fund}, 1);$$

$$et = ones(row, 1);$$

$$z1 = zeros(1, n_{fund});$$

e por fim, define-se:

$$A = [-et \quad \Gamma; -et \quad -\Gamma];$$

$$b = [y \quad -y];$$

$$Aeq = [0 \quad e'];$$

$$beq = [1];$$

$$lb = zeros(var, 1);$$

$$f = [1 \quad z1];$$

III. MADD (Mean Absolute Downside Deviation): o método é definido por:

$$min \quad e'(|\overline{y - \Gamma\omega}|)$$

Para a implementação, chegou-se aos seguintes valores para as matrizes e vetores da função:

$$var = 2.row + n_{fund};$$

sendo n_{fund} o número de ativos usados. Tem-se as variáveis auxiliares a seguir:

$$I = eye(row);$$

$$e = ones(n_{fund}, 1);$$

$$et = ones(row, 1);$$

$$z = zeros(1, row);$$

$$z1 = zeros(1, n_{fund});$$

e por fim, define-se:

$$A = [-I \quad -\Gamma];$$

$$b = [-y];$$

$$Aeq = [z \quad e'];$$

$$beq = [1];$$

$$f = [et' \quad z1];$$

$$lb = zeros(var, 1);$$

IV. DMin-Max (Downside Min-Max):o modelo realiza a otimização:

$$min \quad (max|\overline{y - \Gamma\omega}|)$$

Para a implementação, chegou-se aos seguintes valores para as matrizes e vetores da função:

$$var = n_{fund} + 1;$$

sendo n_{fund} o número de ativos usados. Tem-se as variáveis auxiliares a seguir:

$$e = ones(n_{fund}, 1);$$

$$et = ones(row, 1);$$

$$z1 = zeros(1, n_{fund});$$

e por fim, define-se:

$$A = [-et \quad \Gamma];$$

$$b = [-y];$$

$$Aeq = [0 \quad e'];$$

$$beq = [1];$$

$$lb = zeros(var, 1);$$

$$f = [1 \quad z1];$$

5.3 Implementação das Simulações dos Métodos Computacionais

Como os modelos computacionais não são lineares, não é possível utilizar as funções *linprog* ou *intlinprog* do Matlab para fazer as suas resoluções, conforme foi visto nos modelos anteriores.

Portanto, foi utilizada a função *quadprog*, também presente no Matlab, mas utilizada para resolver funções quadráticas com variáveis lineares. Essa função realiza a minimização para um problema nos seguintes moldes:

$$min_x \frac{1}{2} x^T H x + f^T x$$

$$sujeito \quad a$$

$$A.x \leq B,$$

$$Aeq.x = beq$$

Para a resolução dos modelos a seguir, as equações de minimização foram adaptadas para corresponder a esse molde e realizar a otimização seguindo cada um dos métodos.

5.3.1 Mínima Variância do Erro

O modelo de mínima variância do erro, tem por objetivo realizar a seguinte otimização:

$$min(\omega - \omega_B)'\Sigma(\omega - \omega_B)$$

$$sujeito \quad a$$

$$(\omega - \omega_B)'r = 0$$

$$\omega \geq 0$$

$$\omega' e = 1$$

$$\omega_i = 0, i = n - p + 1, ..., n$$

$$\omega \in R^n$$

Conforme já mencionado anteriormente, a matriz Σ é a matriz de covariância e o vetor ω_B é a composição da carteira do benchmark que, no caso em questão, é o Ibovespa.

Em todos os modelos de simulação realizados só foi necessário coletar a variação dos fundos que compunham a carteira de rastreamento e que foram utilizados na otimização. Porém, nesse caso, diferentemente dos outros, precisou-se coletar toda a composição do índice Ibovespa e as respectivas variações de cada um dos ativos.

Para realizar isso de uma maneira mais otimizada, foi feito um código em Python que, a partir de uma composição do índice, coleta os dados de todos os ativos presentes.

Ao final dessa coleta, todos os dados necessários para a otimização já estavam presentes e, portanto, foi possível realizar a equivalência dos valores da função *quadprog* do Matlab com a equação de minimização do modelo.

Após algumas manipulações, chegou-se nos seguintes valores para as matrizes da função:

$$H = 2\Sigma$$

$$f = -2\Sigma\omega_B$$

E, para as matrizes que definem as restrições, obte-se o seguinte:

$$A = -I_{(nxn)}$$

$$b = zeros_{(nx1)}$$

$$Aeq = \begin{bmatrix} \multicolumn{2}{c}{ones'_{(nx1)}} \\ \multicolumn{2}{c}{ones_{(1xn)}} \\ zeros_{(n-p,p)} & I_{(n-p)x(n-p)} \end{bmatrix}$$

$$beq == \begin{bmatrix} \omega'_B.ones_{(nx1)} \\ 1 \\ zeros_{(n-p,1)} \end{bmatrix}$$

O valor n corresponde a todos os ativos que compõem e o índice Ibovespa e o valor p corresponde apenas aos ativos utilizados na otimização. As matrizes I, *ones* e *zeros* representam uma matriz identidade, uma matriz de uns e uma matriz de zeros, respectivamente. E os valores subescritos representam as dimensões de cada matriz.

Com todos os valores definidos, foi possível realizar o cálculo da função *quadprog*:

$$[x, z] = quadprog(H, f, A, b, Aeq, beq)$$

Como apenas os p primeiros valores de ω são os que representam a distribuição da carteira otimizada, o resultado é definido da seguinte forma:

$$\omega = x(1 : p, 1)$$

Por fim, para definir o *z_otimo* é necessário somar uma constante ao valor z encontrado pela função *quadprog*. Isso se deve ao fato de que, após a transformação da equação do modelo nas matrizes da função do Matlab, sobrou um termo constante, que deve ser adicionado ao resultado final da otimização. Portanto, o cálculo feito é:

$$z_{otimo} = z + \omega'_B.\Sigma.\omega_B$$

Assim, a otimização é finalizada com o valor da distribuição da carteira representado por ω e o valor da Mínima Variância do Erro representado por *z_otimo*.

5.3.2 Mínimo Erro Não Sistêmico

Para o modelo de mínimo erro não sistêmico, assim como no modelo anterior, foi necessário construir a matriz de covariância Σ e, além disso, foi necessário calcular os valores de β para cada ativo.

Esses valores de β serão utilizados na restrição de β da carteira. Inicialmente, será assumido $\beta = 1$, ou seja, um risco igual ao risco de mercado. Recaptulando o modelo:

$$min \omega' \sum \omega$$

$$sujeito \quad a$$

$$\omega \geq 0$$

$$\omega' e = 1$$

$$\sum_{i=1}^{P} \beta_i \omega_i = 1$$

$$\omega \in R^P$$

Para implementar este modelo no Matlab também foi utilizada a função quadprog com os seguintes parâmetros.

$$H = 2\Sigma$$

$$f = zeros_{(nx1)}$$

As matrizes que definem as restrições foram preenchidas da seguinte forma.

$$A = -I_{(nxn)}$$

$$b = zeros_{(nx1)}$$

$$Aeq = \begin{bmatrix} ones'_{(nx1)} \\ betas \end{bmatrix}$$

$$beq = ones_{(2x1)}$$

Em que *betas* é um vetor contendo o valor de β_i do ativo i.

5.3.3 Mínimo Erro Quadrático

Para a implementação do modelo do mínimo erro quadrático, assim como nos anteriores, foi feita a equivalência da equação do modelo com a função *quadprog* do Matlab. A partir da equação do método, dada por:

$$min(y - \Gamma\omega)'(y - \Gamma\omega)$$

$$sujeito \quad a$$

$$\omega \geq 0$$

$$\omega' e = 1$$

$$\omega \in R^p$$

As equivalências encontradas para a função quadprog foram as seguintes:

$$H = 2\Gamma'\Gamma$$

$$f = -2\Gamma' y$$

As matrizes que definem as restrições foram preenchidas da seguinte forma:

$$A = -I_{(nxn)}$$

$$b = zeros_{(nx1)}$$

$$Aeq = ones_{(1xn)}$$

$$beq = 1$$

6 RESULTADOS DAS SIMULAÇÕES

Nesta seção estão os resultados obtidos. O processo de simulação é separado em duas partes: otimização e validação do modelo. Para cada modelo, foi necessário passar por estes processos para obter os resultados.

Primeiramente, é necessário, a partir de uma série histórica de retornos, construir a matriz Γ que é posteriormente utilizada na simulação dos modelos. Este processo de construção de Γ é o que chamamos de otimização.

Para o processo de otimização foram utilizados dados de períodos diferentes, de janeiro de 2019 à dezembro de 2019, de janeiro de 2020 à dezembro de 2020 e de janeiro de 2021 à dezembro de 2021.

Essa abordagem foi definida porque o índice utilizado como benchmark (Ibovespa) teve uma variação bastante brusca nos anos de 2020 e 2021 devido aos impactos da pandemia da Covid-19, sendo interessante utilizar dados tanto do período pré pandemia quanto dados do período de maior impacto no índice.

Na primeira parte dos resultados apresentados, a otimização foi feita apenas uma vez, não sendo acrescentados dados conforme os meses se passam, ou seja, os percentuais definidos inicialmente para cada ativo foram mantidos ao longo de toda a validação. Em seguida, apresenta-se resultados com reotimizações realizadas trimestralmente, de forma a analisar a influência do período de treino na assertividade da previsão.

Por último, no processo de validação, é o momento que cada modelo é implementado levando em consideração sua filosofia. A seguir estão os resultados e algumas análises do que foi obtidos nas simulações.

6.1 Modelos de Programação Linear

6.1.1 Otimização e validação com os mesmos dados

De início pode não parecer intuitivo otimizar e validar o modelo com os mesmos dados, mas com essa combinação de dados é possível observar se os modelos de fato estão entregando o que se propõem. A seguir, estão os gráficos e uma análise dos resultados obtidos das simulações executadas com os dados de 2019 e 2020.

- Otimização e validação com dados do ano de 2019

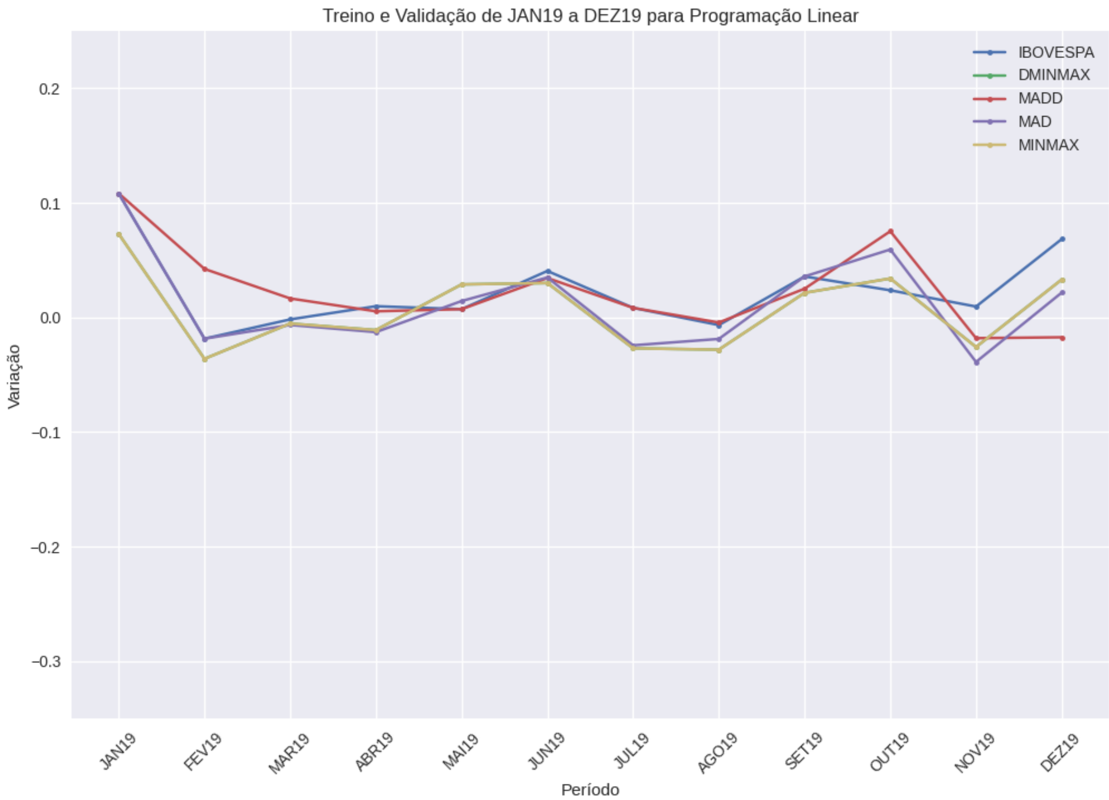

Figura 3: Resultado da simulação de otimização e validação com dados de 2019

Do gráfico é possível perceber que os modelos conseguiram acompanhar o benchmark, no entanto, não é possível fazer uma análise muito aprofundada da eficácia dos métodos, por isso, foi criada a tabela abaixo.

A tabela ilustra o desempenho de cada um dos métodos perante quatro indicadores de desempenho que são os indicadores utilizados como função objetivo nos métodos. Dessa forma, é possível analisar cada um dos métodos para constatar se de fato este foi o melhor no indicador que utiliza, ou seja, o que se espera como resultado é um melhor desempenho do método quando comparado aos outros no indicador utilizado como sua função objetivo.

Na tabela, os valores em negrito representam o menor valor dentre os métodos.

Tabela 1: Validação métodos de programação linear para o período de 2019

Metodologia utilizada	MAD	MinMax	MADD	DMinMax
Soma dos desvios absolutos	**21,57%**	26,14%	26,76%	26,14%
Máximo desvio absoluto	4,84%	**3,53%**	8,57%	3,53%
Soma dos desvios negativos	17,30%	22,95%	**13,46%**	22,95%
Máximo desvio negativo	4,84%	3,53%	8,57%	**3,53%**

- Otimização e validação com dados do ano de 2020

  Para as simulações utilizando os dados de 2020, os resultados foram muito semelhantes. Assim como no caso anterior, também foi construída uma tabela para análise da efetividade dos métodos.

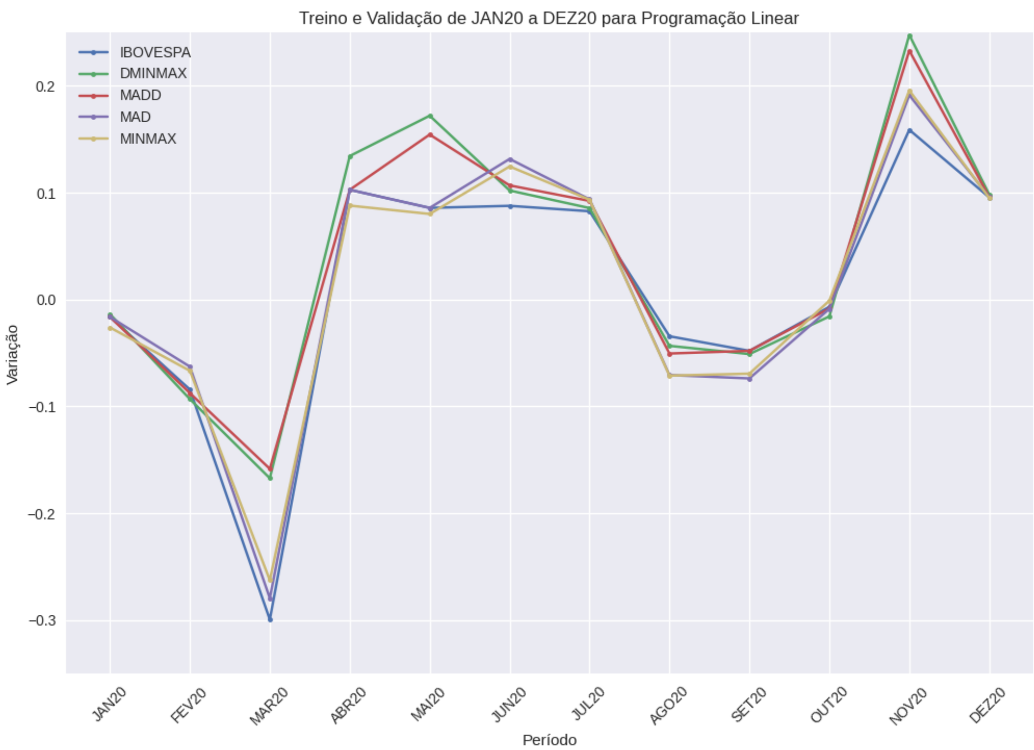

Figura 4: Resultado da simulação de otimização e validação com dados de 2020

Tabela 2: Validação métodos de programação linear para o período de 2020

Metodologia utilizada	MAD	MinMax	MADD	DMinMax
Soma dos desvios absolutos	**19,32%**	23,37%	33,41%	39,16%
Máximo desvio absoluto	4,39%	**3,68%**	14,11%	13,23%
Soma dos desvios negativos	6,40%	8,91%	**2,11%**	2,99%
Máximo desvio negativo	3,63%	3,68%	1,61%	**0,90%**

Assim como era esperado, os métodos cumpriram com suas metodologias, entregando os menores valores nos seus respectivos indicadores.

6.1.2 Otimização e validação com dados de períodos diferentes

Agora serão mostradas as simulações que representam a realidade. Foram feitas otimizações utilizando dados do passado e as porcentagens obtidas da otimização foram utilizadas em uma "carteira teórica", simulando uma compra dos ativos no início do período de validação e levando-os até o fim desse período. Fez-se treinos para prever resultados anuais e treinos para prever resultados trimestralmente

- Otimização com dados do ano de 2019 e validação com dados do ano de 2021

  A seguir estão os resultados das simulações. Assim como anteriormente, foi construída uma tabela para análise do desempenho dos métodos.

Tabela 3: Otimização por programação linear com dados de 2019 e validação com dados de 2021

Metodologia utilizada	MAD	MinMax	MADD	DMinMax
Soma dos desvios absolutos	33,63%	37,22%	**30,77%**	37,22%
Máximo desvio absoluto	8,04%	7,29%	**5,48%**	7,29%
Soma dos desvios negativos	31,41%	35,09%	**28,52%**	35,09%
Máximo desvio negativo	8,04%	7,29%	**5,48%**	7,29%

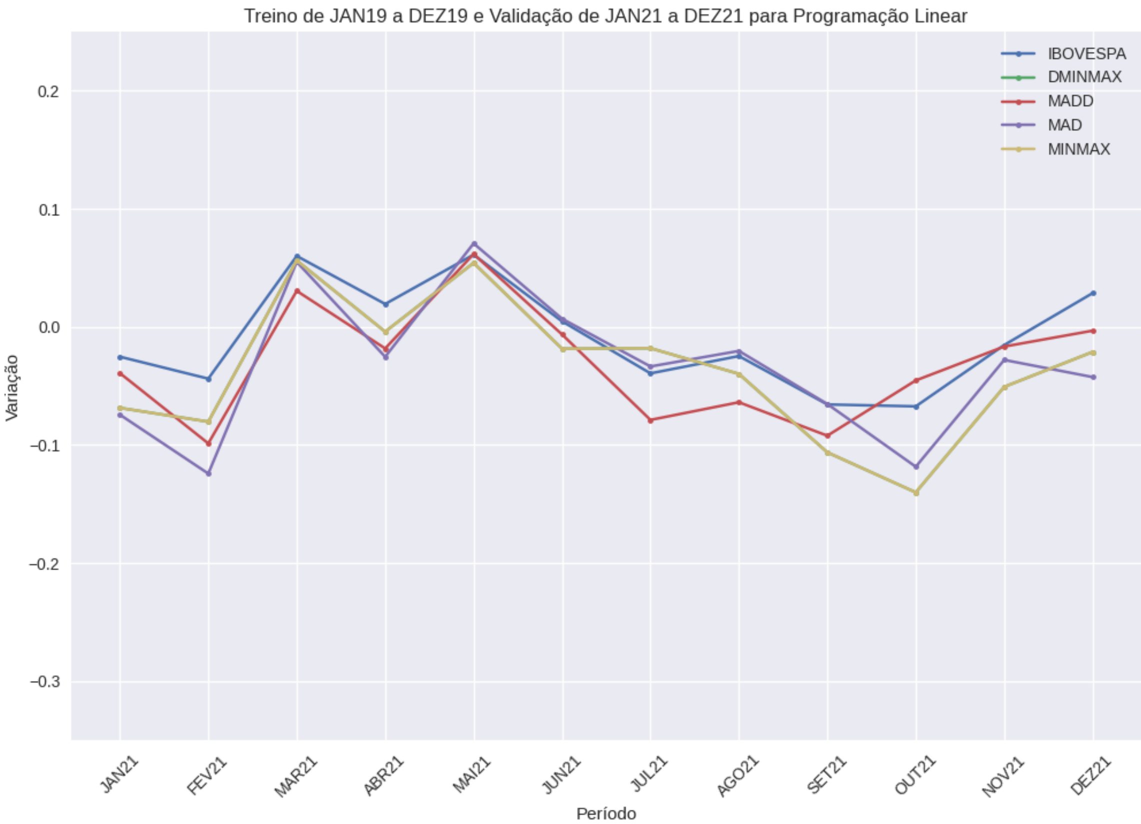

Figura 5: Resultado da simulação de otimização com dados de 2019 e validação com dados de 2021

- Otimização com dados do ano de 2020 e validação com dados do ano de 2021

Tabela 4: Otimização por programação linear com dados de 2020 e validação com dados de 2021

Metodologia utilizada	MAD	MinMax	MADD	DMinMax
Soma dos desvios absolutos	36,26%	35,37%	32,80%	**31,26%**
Máximo desvio absoluto	**4,85%**	5,10%	9,13%	8,26%
Soma dos desvios negativos	32,69%	31,54%	26,25%	**23,83%**
Máximo desvio negativo	**4,85%**	5,10%	9,13%	8,26%

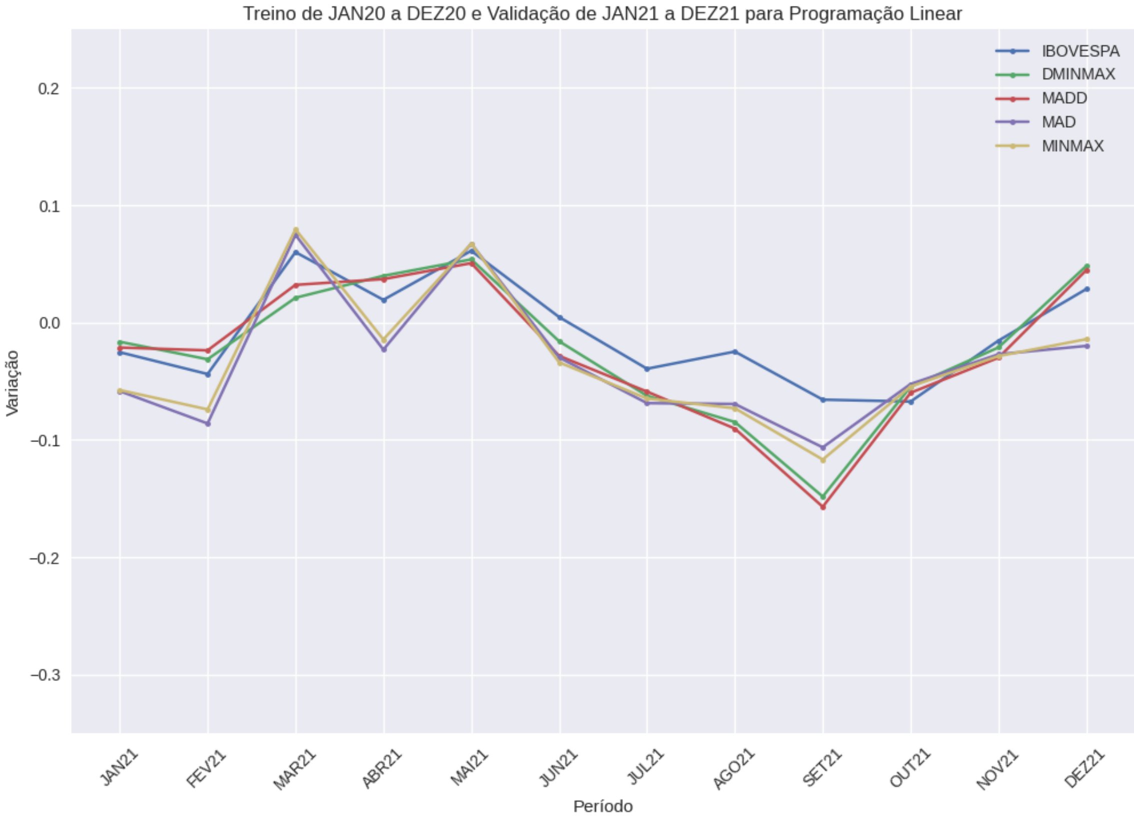

Figura 6: Resultado da simulação de otimização com dados de 2020 e validação com dados de 2021

Observando os resultados, é possível ver que, novamente, pelos gráficos, os métodos conseguiram acompanhar a movimentação do benchmark, no entanto não é possível extrair informações muito detalhadas.

Já nas tabelas, houve uma grande diferença entre os casos com validação utilizando dados diferentes dos dados da otimização e os casos com validação utilizando dados iguais aos dados da otimização. Desta vez, os métodos não se destacaram nos seus respectivos indicadores, o que não é um problema, visto que, o retorno futuro de um índice é uma variável aleatória que não pode ser prevista, apenas estimada.

Por último, apresenta-se uma simulação com reotimização trimestral, de forma a avaliar se a otimização com um menor período de tempo tem um resultado mais assertivo. A cada 3 meses o Γ foi recalculado e foi aplicada uma carteira com novos percentuais de cada ativo. Na sessão "Análise dos Resultados"é apresentada uma análise sobre os resultados dessa diferença de período na otimização.

- Reotimização com dados trimestrais de outubro de 2020 à setembro de 2021 e validação trimestral de janeiro à dezembro de 2021

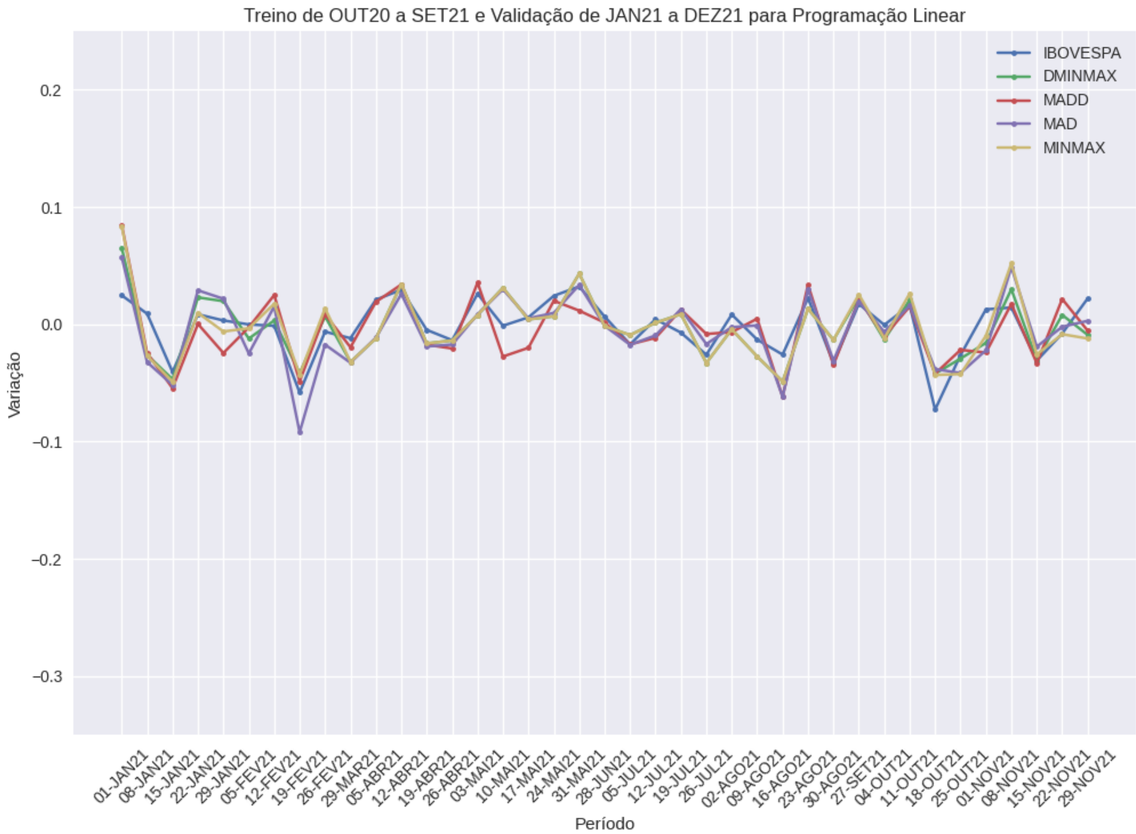

Figura 7: Resultado da simulação de reotimização com dados semestrais para prever o ano de 2021

6.2 Métodos Computacionais

A seguir é feita a validação dos métodos de Mínimo Erro Quadrático, Mínimo Erro Não Sistêmico e Mínima Variância do Erro de forma a verificar se os métodos seguem seus fundamentos teóricos. Para isso, são comparados os dados reais do benchmark com a carteira simulada para o mesmo ano.

6.2.1 Otimização e validação com os mesmos dados

A seguir, estão os gráficos e uma análise dos resultados obtidos das simulações executadas com os dados de 2020 e 2021.

- Otimização e validação com dados do ano de 2020

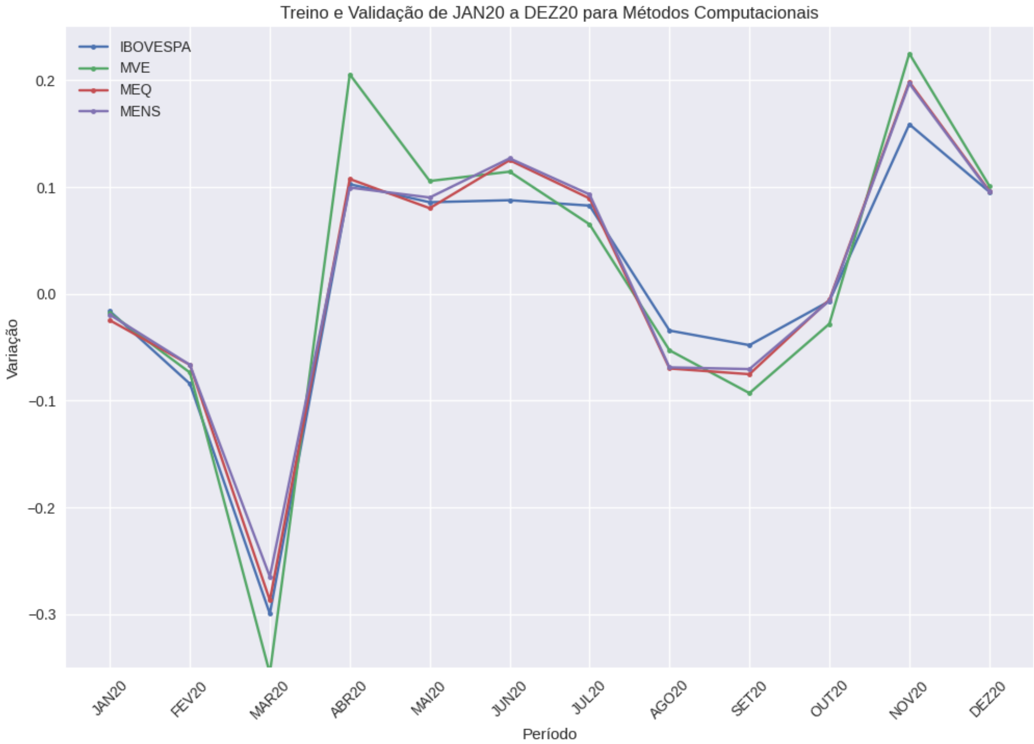

Figura 8: Resultado da simulação de otimização e validação com dados de 2020

A partir do gráfico, é possível perceber que o método de Mínima Variância do Erro apresentou uma maior variação em relação ao Ibovespa quando comparado aos outros dois. Isso ocorreu, principalmente, pois, para realizar a otimização nesse período, foi necessário diminuir as restrições implementadas no Matlab, ou seja, o erro permitido na simulação foi maior do que nos outros métodos. Essa alteração foi necessária, pois, como foi utilizada uma composição fixa do índice Ibovespa - do dia 26/09/2022 - há muitos ativos que hoje fazem parte do índice, mas que não apresentam dados no ano de 2020. Esse fato acontece, pois, periodicamente, a composição do Ibovespa se altera, retirando ativos antigos e adicionando novos. Desse modo, uma otimização muito restrita não estava sendo capaz de encontrar um valor ótimo e, então, essa alteração foi necessária.

Entretanto, mesmo com esse fato, é possível perceber que os modelos acompanham as tendências do benchmark. Por outro lado, não é possível fazer uma análise aprofundada da eficácia dos métodos, para isso usa-se a tabela abaixo.

A tabela busca ilustrar o desempenho de cada um dos métodos perante três indicadores de desempenho que são utilizados como função objetivo nos métodos.

Dessa forma, é possível analisar cada um deles para constatar se de fato este foi o melhor no indicador que utiliza, ou seja, o que se espera como resultado é um melhor desempenho do método quando comparado aos outros no indicador utilizado como sua função objetivo.

Na tabela, os valores em negrito representam o menor valor dentre os métodos.

Tabela 5: Otimização e validação dos métodos computacionais no ano de 2020

Metodologia utilizada	MEQ	MENS	MVE
Mínimo erro quadrático	**0,56%**	0,64%	1,55%
Mínimo erro não sistêmico	1,69%	**1,56%**	2,64%
Mínima Variância do Erro	0,32%	0,34%	**0,12%**

- Otimização e validação com dados do ano de 2021

  Para as simulações utilizando os dados de 2021, os resultados foram muito semelhantes. Porém, como pode-se perceber, como os dados são mais atualizados, é possível observar um melhor desempenho do método de Mínima Variância do Erro.

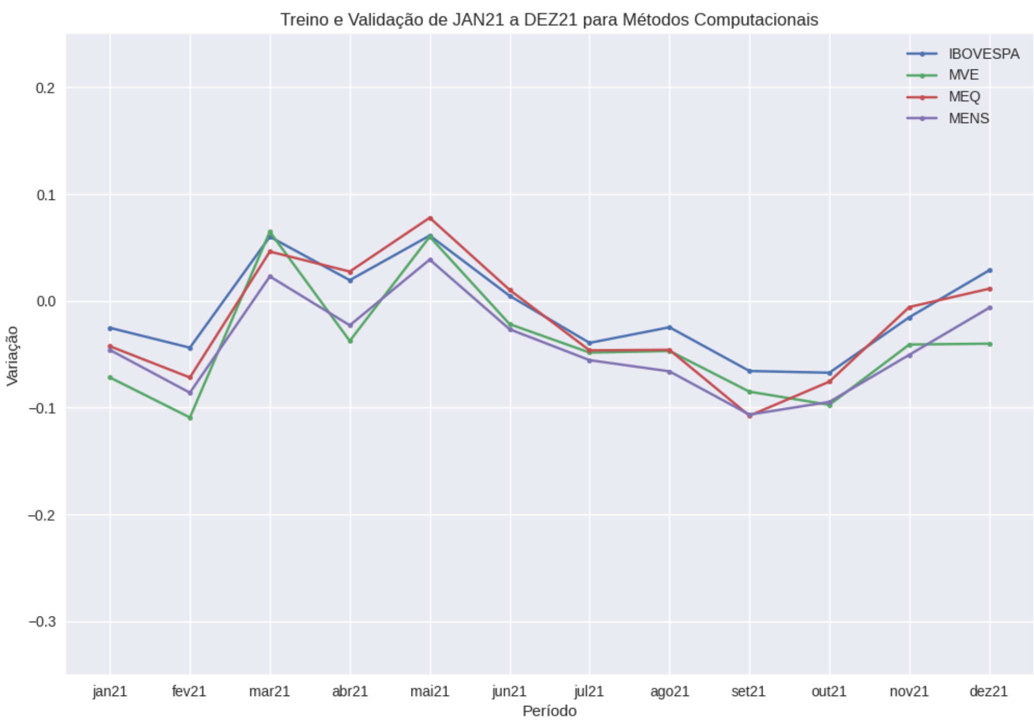

Figura 9: Resultado da simulação de otimização e validação com dados de 2021

Assim como no caso anterior, também foi construída uma tabela para análise da efetividade dos métodos.

Tabela 6: Otimização e validação dos métodos computacionais no ano de 2021

Metodologia utilizada	MEQ	MENS	MVE
Mínimo erro quadrático	**0,44%**	1,37%	0,51%
Mínimo erro não sistêmico	0,31%	**0,20%**	0,30%
Mínima Variância do Erro	0,17%	0,08%	**0,03%**

Conforme era esperado, os métodos cumpriram com suas metodologias, entregando os menores valores nos seus respectivos indicadores.

6.2.2 Otimização e validação com dados de períodos diferentes

Agora os métodos serão usados para prever resultados de um período subsequente. Foram feitas otimizações utilizando dados do passado e as carteiras obtidas na otimização foram utilizadas em uma "carteira teórica", simulando uma compra dos ativos no início do período de validação e levando-os até o fim desse período.

- Otimização com dados do ano de 2019 e validação com dados do ano de 2020

  A seguir estão os resultados das simulações.

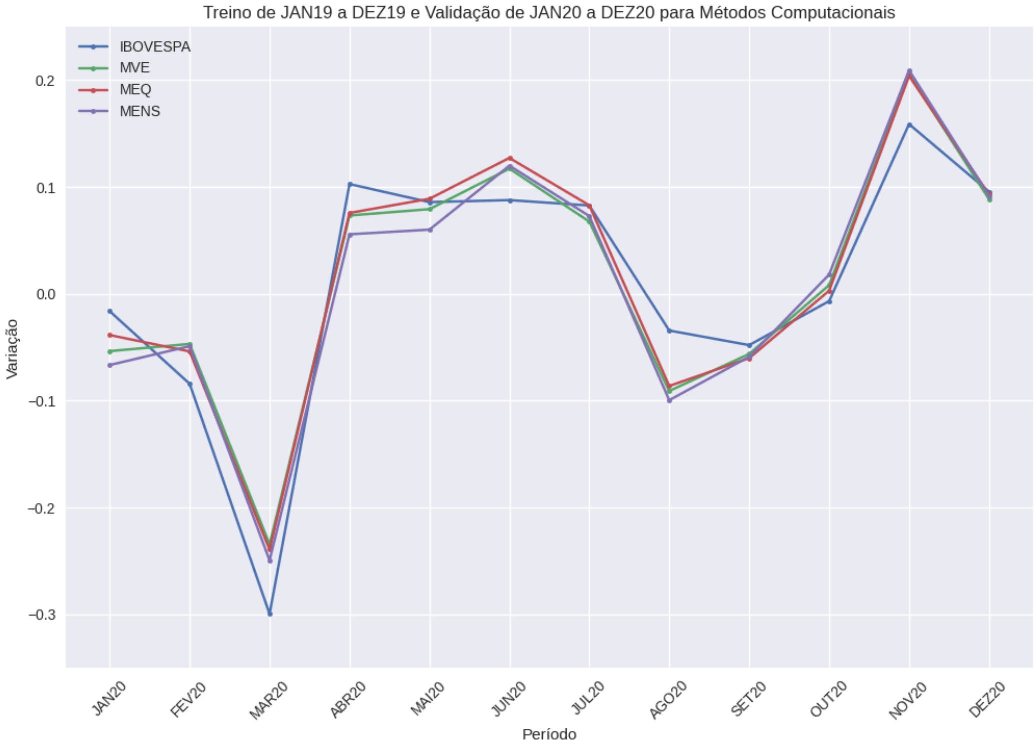

Figura 10: Resultado da simulação de otimização com dados de 2019 e validação com dados de 2020

- Otimização com dados do ano de 2020 e validação com dados do ano de 2021

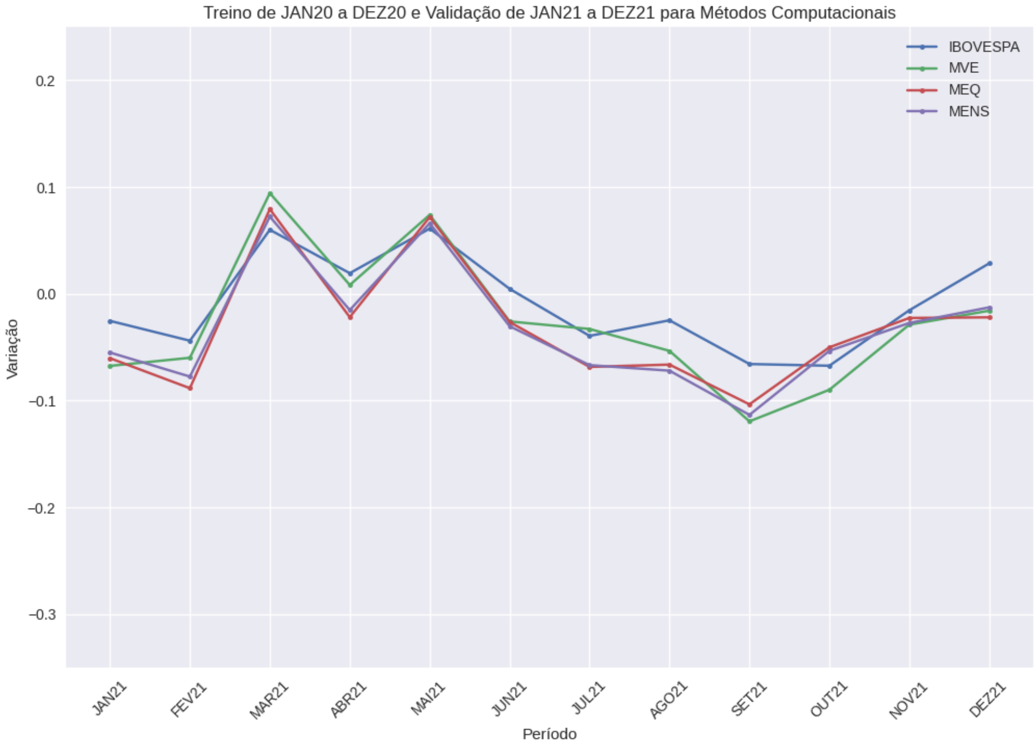

Figura 11: Resultado da simulação de otimização com dados de 2020 e validação com dados de 2021

Observando os resultados, concluímos que os métodos conseguiram acompanhar a movimentação do benchmark.

Em seguida, foram realizadas outras otimizações, com períodos menores, em busca de resultaddos mais assertivos.

- Reotimização com dados trimestrais de outubro de 2020 à setembro de 2021 e validação trimestral de janeiro à dezembro de 2021

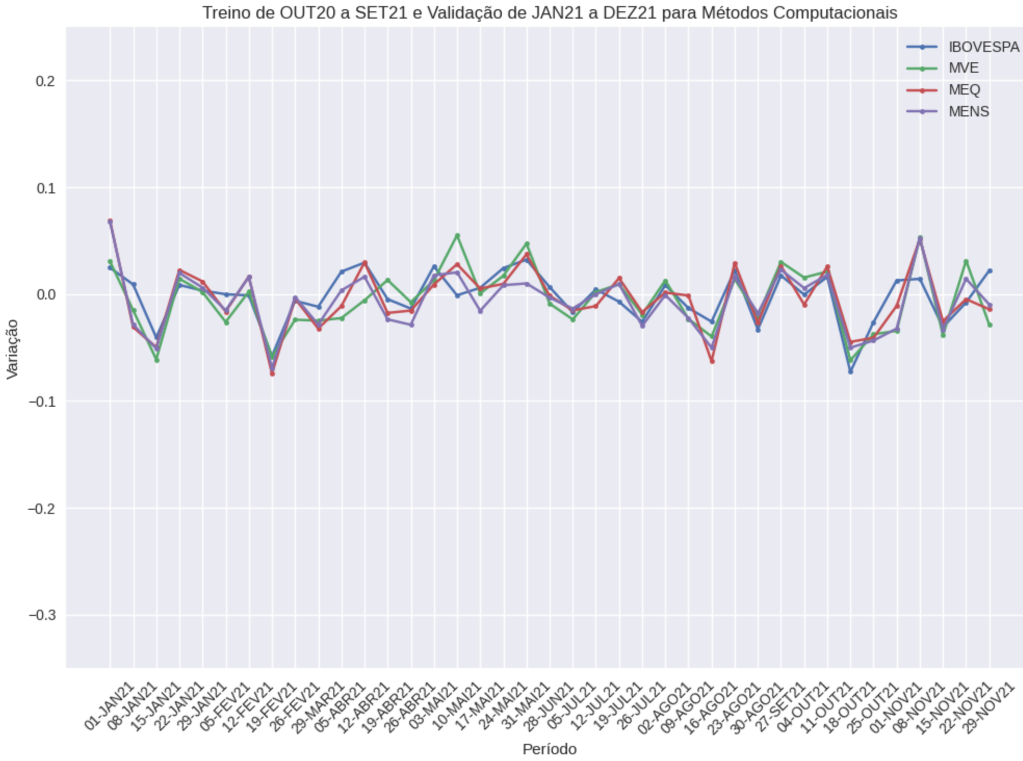

Figura 12: Resultado da simulação de reotimização com dados semestrais para prever o ano de 2021

7 ANÁLISE DOS RESULTADOS

Nesta seção há uma análise comparativa dos resultados obtidos pelas simulações dos diferentes métodos. Para realizar esta análise foram utilizados três métricas de comparação: o índice de Sharpe, o β da carteira e o módulo do erro. Cada uma delas avalia um aspecto da carteira. O índice de Sharpe serve como forma de medir se o risco tomado por uma determinada composição da carteira justifica seu retorno se comparado a um ativo livre de risco, o β é uma forma de mensurar o risco assumido pela carteira, e o módulo do erro diz o quão distante o modelo ficou do benchmark em termos de retorno.

A seguir é feita a análise das três métricas separadamente. Para isso, foram utilizadas duas formas de otimização, primeiro é feita a otimização utilizando dados mensais do ano de 2020 e validação com dados de 2021 e depois é feita a otimização com dados semanais em uma janela de três meses anteriores ao período de validação até atingir um período de um ano, sendo a validação feita no ano de 2021.

7.1 Análise do índice de Sharpe

Nesta seção é feita a análise do índice de Sharpe de cada carteira.

Tabela 7: Índices de Sharpe obtidos para cada método no ano de 2021

Método	IS
MAD	-0,67
MADD	-0,47
MINMAX	-0,63
DMINMAX	**-0,44**
MENS	-0,65
MEQ	-0,62
MVE	-0,48

Em negrito está destacado o melhor valor obtido dentro os métodos. Neste caso, o melhor valor é o mais alto dentre os métodos, ou seja, o menos negativo.

É possível ver que nenhuma carteira obteve um valor positivo, isso indica que o retorno livre de risco conseguiu ser melhor do que o retorno de todas as carteiras obtidas nas simulações. Esse resultado pode ser explicado pelo alto retorno que o ativo livre de risco possui, então uma carteira precisa ter uma boa performance no período para superá-lo, algo difícil de ser alcançado em um período de muitas incertezas e busca do mercado por opções de investimentos mais seguros, como foi o ano de 2021.

Outro ponto interessante de analisar é o fato de que os dois métodos que obtiveram os melhores resultados foram justamente os métodos de programação linear que não penalizam a carteira quando possuem um retorno maior do que o benchmark, isso mostra que a filosofia utilizada para construção dos métodos mostrou resultado.

Os resultados obtidos com a otimização utilizando períodos trimestrais são mostrados a seguir.

Tabela 8: Índice de Sharpe obtido para cada método no ano de 2021 com otimização trimestral

Método	IS
MAD	-0,29
MADD	-0,28
MINMAX	**-0,23**
DMINMAX	-0,27
MENS	-0,35
MEQ	-0,25
MVE	-0,29

Assim como no outro formato de otimização, todos os métodos obtiveram um índice de Sharpe negativo. No entanto, neste caso o método que performou melhor foi o MINMAX. Os valores obtidos nesta simulação se mostraram ligeiramente melhores que os obtidos com otimização anual, ou seja, o retorno das carteiras se mostrou um pouco mais próximo do retorno do ativo livre de risco.

7.2 Análise do β das carteiras

Agora é feita a análise do β das carteiras.

Tabela 9: β da carteira para cada método no ano de 2021

Método	β
MAD	1,143
MADD	1,172
MINMAX	1,188
DMINMAX	**1,135**
MENS	1,158
MEQ	1,161
MVE	1,319

Novamente, está destacado em negrito o melhor valor obtido dentro os métodos. Para esta métrica, o melhor valor considerado é o menor dentre os valores de todos os métodos dado que este indica menor volatilidade. Também poderia ser escolhido o valor mais alto de β como melhor valor por indicar que a carteira pode ter um retorno maior do que o de mercado, no entanto, como o foco está no rastreamento e não necessariamente no retorno final da carteira, foi optado por utilizar o valor mais baixo, de menor risco.

Mais uma vez, todos os métodos apresentaram uma característica em comum, já que obtiveram β maior do que um, ou seja, todas as carteiras estão assumindo um risco maior do que o mercado. Olhando do ponto de vista do investidor, possuir um risco maior do que um ativo que se deseja rastrear pode significar assumir um risco maior do que o desejado.

Assim como na métrica anterior, para a otimização anual é possível ver que o método com o melhor resultado é o DMINMAX.

Para a otimização com período trimestral, os resultados obtidos foram os seguintes.

Tabela 10: β da carteira para cada método no ano de 2021 com otimização trimestral

Método	β
MAD	0,958
MADD	0,867
MINMAX	0,868
DMINMAX	**0,811**
MENS	0,895
MEQ	0,956
MVE	0,853

Diferentemente da otimização anual, a otimização trimestral obteve, para todos os casos, um β menor do que um, indicando que todas as carteiras possuem um risco menor do que o mercado. Isso é importante, pois, se o retorno do ativo for semelhante ao retorno do benchmark (que nesse caso representa o mercado), seria possível obter o mesmo retorno com um risco menor, sendo essa a melhor situação para um investidor.

7.3 Análise do módulo do erro

Por último é feita a análise do módulo do erro. Esta métrica é calculada obtendo o módulo da diferença entre o retorno da carteira e o retorno do benchmark.

Tabela 11: Módulo do erro para cada método no ano de 2021

Método	Módulo do erro
MAD	23,3%
MADD	17,0%
MINMAX	22,4%
DMINMAX	**14,4%**
MENS	22,3%
MEQ	21,9%
MVE	17,8%

Esta métrica é do interesse de um investidor que deseja fazer o rastreamento de um ativo, pois, o mais importante em um rastreamento é o retorno que foi obtido no final do período do investimento, ou seja, quando o investidor faz a venda do ativo e observa a relação lucro/prejuízo. Por isso, o melhor valor é o menor dentre os métodos.

E mais uma vez, o método que se saiu melhor com otimização anual é o DMINMAX, pois foi capaz de obter resultados significativamente melhores do que os outros métodos. Estes resultados são surpreendentes, uma vez que, o método MEQ, por exemplo, foca em diminuir justamente o erro entre a carteira e o benchmark, mas essa falha em obter o melhor resultado mostra que quando se tenta prever o comportamento de um ativo financeiro, o resultado obtido não tem nenhum compromisso com a realidade, fazendo com que o resultado esperado não seja alcançado.

Tabela 12: Módulo do erro para cada método no ano de 2021 com otimização trimestral

Método	Módulo do erro
MAD	10,3%
MADD	7,4%
MINMAX	**2,6%**
DMINMAX	3,5%
MENS	13,4%
MEQ	5,9%
MVE	9,7%

Neste formato de otimização, o erro obtido se mostrou muito mais baixo do que na otimização anual. Neste caso, o método que melhor performou foi o MINMAX. O método obteve erro baixo, o risco sendo um dos mais baixos dentre todos e risco menor do que um, sendo uma boa alternativa para o investidor que deseja ter uma carteira que acompanha o Ibovespa com um risco menor.

8 CONCLUSÕES

No início do projeto foi definido como objetivo construir um algoritmo que fosse capaz de compor uma carteira representativa de um benchmark utilizando dados históricos e diversos métodos de otimização. Para isso, foram definidos sete métodos, sendo quatro de programação linear e três métodos computacionais para alcançar o objetivo de rastrear o índice que representa o mercado brasileiro, o Ibovespa.

Para alcançar estes objetivos foram desenvolvidas sete funções no Matlab, cada uma implementando um método, e mais alguns scripts em Python responsáveis por fornecer os elementos necessários para simulação e análise dos resultados. O conjunto composto pelas funções do Matlab e os scripts de Python formaram uma automação que pode ser vista como um produto completo para geração de carteiras para rastreamento de benchmarks.

Desde o ínicio do projeto esta visão de um produto final foi muito reforçada, pois, facilita o processo do investidor na hora de gerar carteiras, não sendo necessário preparar um novo script para cada otimização que desejar realizar. Além disso, essa visão de criação de um produto trouxe um aprendizado muito grande na desenvolvimento de projetos, visto que foi necessário utilizar frameworks de metodologia ágil bastante conhecidos no mercado e técnicas de desenvolvimento de software, como a utilização de repositório em nuvem para versionamento e armazenamento de dados.

Analisando os resultados, é possível ver que os valores obtidos utilizando otimização trimestral se mostraram superiores aos resultados obtidos na otimização anual, visto que em termos de β obtiveram valores mais baixos, ou seja, risco mais baixo; em termos do índice de Sharpe obtiveram valores menos negativos, isto é, mais próximos do retorno do ativo livre de risco e em termos de módulo do erro, obtiveram valores significativamente mais baixos, ou seja, conseguiram ao final do período entregar uma carteira com retorno mais próximo do benchmark. Esse resultado pode ser explicado pelo fato de a otimização estar usando dados mais atualizados e com otimizações feitas para uma previsão mais curta do que o caso anual.

Outro resultado importante é o fato de os métodos de programação linear, mesmo que sendo mais simples, terem produzido resultados superiores aos métodos computacionais. Em particular, o método MINMAX e seu variante DMINMAX que ganharam em todos as métricas apontadas, com destaque para o método DMINMAX que se saiu melhor em todas as simulações com otimização anual e também superou os outros métodos no β para otimização trimestral. Já os métodos computacionais não tiveram resultados muito inferiores aos resultados dos métodos de programação linear mas não mostraram que sua complexidade trouxe resultado. Particularmente, o método MVE se mostrou bastante difícil de ser implementado, uma vez que foi necessário abrir mão de restrições mais fortes para obter um resultado para as simulações.

REFERÊNCIAS

1 VARGA, G. Índice de sharpe e outros indicadores de performance aplicados a fundos de ações brasileiros. *Revista de Administração Contemporânea*, SciELO Brasil, v. 5, p. 215–245, 2001.

2 RUDOLF, M.; WOLTER, H.-J.; ZIMMERMANN, H. A linear model for tracking error minimization. *Journal of Banking & Finance*, Elsevier, v. 23, n. 1, p. 85–103, 1999.

3 COSTA, O. L. d. V.; ASSUNÇÃO, H. G. V. Análise de risco e retorno em investimentos financeiros. 2005.

4 ROLL, R. A mean/variance analysis of tracking error. *Journal of portfolio management*, v. 18, n. 4, p. 13–22, 1992.

5 PAULO, W. L. de; OLIVEIRA, E. M. de; COSTA, O. L. do V. Enhanced index tracking optimal portfolio selection. *Finance Research Letters*, Elsevier, v. 16, p. 93–102, 2016.

6 PADULA FELIPE; PACHECO, G. J. M. *Comparação de métodos de programação matemática para otimização de portfólios*. [S.l.]: Tese de monografia Universidade de São Paulo, 2021.

7 YAHOO!, F. *Dados históricos do Ibovespa*. Disponível em: ⟨https://finance.yahoo.com/⟩.

ANEXO A

1. Planilha com os cálculos realizados para criar as tabelas apresentadas na sessão "Resultados das Simulações" para métodos computacionais: Planilha teste de teoria métodos computacionais

2. Planilha com os cálculos realizados para criar as tabelas apresentadas na sessão "Resultados das Simulações" para métodos de programação linear: Planilha teste de teoria programação linear

3. Repositório com todos os códigos desenvolvidos no trabalho: Repositório TCC