

CESAR SCHOOL

PÓS-GRADUAÇÃO EM ANÁLISE E ENGENHARIA DE DADOS

ALDO FERREIRA DE SOUZA MONTEIRO

**AVALIAÇÃO DE CARTEIRAS DE INVESTIMENTOS
AUTOMATIZADAS POR OTIMIZAÇÃO
MULTIOBJETIVO E TEORIA DO PORTFÓLIO:
COMPARAÇÃO COM BENCHMARK EM SIMULAÇÕES
COMPUTACIONAIS**

RECIFE

2023

ALDO FERREIRA DE SOUZA MONTEIRO

**AVALIAÇÃO DE CARTEIRAS DE INVESTIMENTOS
AUTOMATIZADAS POR OTIMIZAÇÃO
MULTIOBJETIVO E TEORIA DO PORTFÓLIO:
COMPARAÇÃO COM BENCHMARK EM SIMULAÇÕES
COMPUTACIONAIS**

Trabalho de Conclusão apresentado ao Curso de Pós-Graduação do Centro de Estudos e Sistemas Educacionais do Recife – CESAR School, como requisito para conclusão da especialização em Análise e Engenharia de Dados.

Orientação: Prof. Domingos Sávio de Oliveira Santos Júnior

RECIFE

2023

Avaliação de Carteiras de Investimentos Automatizadas por Otimização Multiobjetivo e Teoria do Portfólio: Comparação com Benchmark em Simulações Computacionais

Resumo. Este estudo investiga a alocação de ativos no mercado financeiro brasileiro usando a teoria moderna do portfólio com otimização multiobjetivo, por meio de simulações computacionais. Inicialmente, definem-se conceitos fundamentais em análise de séries temporais financeiras, abordando retorno e risco. Em seguida, descrevemos a teoria do portfólio e como ela pode ser aplicada na construção de carteiras de investimento. Desenvolvemos um simulador computacional que permite a realização de experimentos para avaliar diferentes estratégias de alocação de ativos e otimização de portfólio. Realizamos 23 simulações com parâmetros variados, incluindo janelas temporais, frequência de realocação e estimadores de matriz de covariância. Os resultados mostram que algumas configurações superam um benchmark tradicional em termos de retorno, enquanto mantêm a volatilidade em níveis mais baixos. A configuração mais eficaz atingiu um retorno acumulado de 147,04% com um desvio padrão de 1,56%. No entanto, este estudo também destaca as limitações e desafios na aplicação da teoria do portfólio e sugere áreas para pesquisas futuras.

Palavras-chave

Teoria Moderna do Portfólio; Otimização Multiobjetivo; Séries Temporais Financeiras.

Abstract. This study investigates asset allocation in the Brazilian financial market using modern portfolio theory with multi-objective optimization through computational simulations. Initially, we define fundamental concepts in financial time series analysis, addressing return and risk. We then describe portfolio theory and how it can be applied in constructing investment portfolios. We develop a computational simulator that allows for experiments to evaluate different asset allocation and portfolio optimization strategies. We conducted 23 simulations with varied parameters, including time windows, reallocation frequency, and covariance

matrix estimators. The results show that some configurations outperform a traditional benchmark in terms of return while keeping volatility at lower levels. The most effective configuration achieved a cumulative return of 147.04% with a standard deviation of 1.56%. However, this study also highlights limitations and challenges in applying portfolio theory and suggests areas for future research.

Keywords

Modern Portfolio Theory; Multi-Objective Optimization; Financial Time Series.

1. Introdução

O mercado financeiro brasileiro, com sua complexidade e volatilidade, apresenta desafios únicos para investidores. A alocação de ativos, essencial para a gestão de portfólio, desempenha um papel crítico na busca por retornos sólidos e riscos controlados. Este estudo investiga estratégias de alocação de ativos no mercado brasileiro, aplicando a teoria moderna do portfólio com otimização multiobjetivo em ambiente simulado.

1.1. Justificativa

A importância deste estudo reside na necessidade de desenvolver estratégias de alocação mais sofisticadas e adaptáveis ao contexto do mercado financeiro brasileiro. Ao simular a aplicação da teoria do portfólio de Markowitz no contexto do mercado financeiro brasileiro, com a utilização de otimização multiobjetivo, buscamos verificar a aplicabilidade desses métodos quantitativos neste mercado no processo de construção de carteiras de investimento que maximizem o retorno e minimizem o risco.

1.2. Problema

Diante da complexidade e volatilidade do mercado financeiro brasileiro, surge o seguinte problema de pesquisa: como podemos aprimorar a alocação de ativos no mercado financeiro brasileiro, considerando a maximização do retorno e a minimização do risco como objetivos simultâneos? Para resolver esse problema, exploramos diferentes estratégias de alocação de ativos baseadas na teoria moderna do portfólio,

visando identificar configurações eficazes que possam superar benchmarks tradicionais em termos de retorno e volatilidade.

1.3. Objetivos

Este trabalho tem como objetivo geral investigar estratégias de alocação de ativos no mercado financeiro brasileiro, aplicando a teoria moderna do portfólio com otimização multiobjetivo em ambiente simulado computacionalmente para avaliar a eficácia de tais estratégias. Para atingir este objetivo, podemos elencar os seguintes objetivos específicos:

- Desenvolver um simulador computacional para conduzir experimentos e avaliar diferentes estratégias de alocação de ativos;
- Realizar simulações com variações de parâmetros, incluindo janelas temporais, frequência de realocação e estimadores de matriz de covariância;
- Analisar e comparar os resultados das simulações, identificando configurações eficazes que superem o benchmark IBovespa em termos de retorno e volatilidade.

1.4. Estrutura do Artigo

O artigo está dividido em 6 seções, onde a primeira introduz o contexto da pesquisa destacando sua importância e seus objetivos, a segunda fornece os conceitos fundamentais necessários para compreender o estudo, a terceira descreve a metodologia empregada neste trabalho, bem como as tecnologias utilizadas, a quarta foca no tratamento dos dados utilizados nas simulações e a quinta descreve os algoritmos de simulação e otimização do portfólio além dos parâmetros testados e os resultados obtidos. Por fim, a sexta e última seção traz um resumo dos resultados obtidos, elencando as limitações e sugerindo abordagens para trabalhos futuros.

2. Fundamentação Teórica

Nesta seção serão introduzidos alguns conceitos fundamentais utilizados no restante do trabalho. A seção está dividida em duas partes: a primeira define alguns conceitos básicos para análise de séries temporais financeiras e a segunda traz um breve resumo da teoria de Harry Markowitz para seleção de portfólio.

2.1. Conceitos Básicos

Nesta seção, serão definidos alguns conceitos básicos em análise de séries temporais financeiras que serão utilizados no decorrer deste artigo. Tipicamente, uma série temporal financeira consiste em uma sequência de preços de um ativo financeiro no decorrer do tempo. Entretanto, a variação absoluta do preço isoladamente não é relevante para determinar a rentabilidade de um investimento: uma ação que passou de R\$100,00 para R\$110,00 (ganho de R\$10,00) não rendeu mais que uma que passou de R\$1,00 para R\$2,00 (ganho de R\$1,00), apesar do ganho absoluto ter sido bem superior. É necessário considerar o ganho em relação ao valor investido, para isso utiliza-se o conceito de retorno.

Seja p_t o preço de um determinado ativo financeiro no instante t , o retorno no instante t (r_t) pode ser definido pela equação 1.

$$r_t = \frac{p_t - p_{t-1}}{p_{t-1}} \quad (1)$$

Apesar de muito útil e fácil de interpretar, o retorno pode não ser a melhor opção quando se precisa realizar certas operações aritméticas. Por exemplo, um retorno de +50% seguido de um de -50% não resulta num retorno total de 0. O retorno acumulado não é a soma dos retornos. Para facilitar os cálculos, o chamado log-retorno é muito usado ao invés do retorno. O log-retorno no tempo t pode ser definido pela equação 2.

$$r_t = \ln\left(\frac{p_t}{p_{t-1}}\right) \quad (2)$$

Outro aspecto muito importante nos investimentos é o risco, que pode ser definido como o nível de incerteza acerca da rentabilidade de um investimento (O QUE É RISCO, [2017-]). Uma forma muito usada para mensurar o risco de um investimento é a variância dos retornos (ou log-retornos) ao longo do tempo. Quanto maior a variância, mais incerto e imprevisível é o resultado futuro do investimento.

Neste trabalho, foram utilizados o log-retorno e a variância como métricas de retorno e risco respectivamente.

2.2. Teoria Moderna do Portfólio

A teoria do portfólio, introduzida por Markowitz (1952) em seu artigo “Portfolio Selection”, prega que o investidor racional deve maximizar o retorno esperado e minimizar o risco. Dessa forma, a construção de carteiras de investimento se torna um problema de otimização multiobjetivo com dois objetivos conflitantes: retorno x risco. O resultado desse processo de otimização é a chamada fronteira eficiente, que é formada pelo conjunto de portfólios que possuem o maior retorno para um dado nível de risco ou o menor risco para um dado retorno esperado. Também nesse trabalho, é formulado como deve ser calculado o retorno esperado e variância da carteira a partir, respectivamente, dos retornos e covariâncias entre os ativos que compõem o portfólio (equações 3 e 4).

$$E = \sum_{i=1}^N X_i \mu_i \quad (3)$$

$$V = \sum_{i=1}^N \sum_{j=1}^N \sigma_{ij} X_i X_j \quad (4)$$

Onde X_i é o peso do ativo i na carteira, μ_i a média dos retornos do ativo i , N o número de ativos na carteira e σ_{ij} a covariância entre os retornos dos ativos i e j . Assim, para encontrar as carteiras eficientes deve-se buscar os valores de $X_1 \dots X_N$ que maximizam E e minimizam V , sujeito às restrições expressas pelas equações 5 e 6.

$$\sum_{i=1}^N X_i = 1 \quad (5)$$

$$X_i \geq 0 \quad (6)$$

3. Metodologia

Esta seção apresenta a metodologia empregada neste trabalho, bem como as principais tecnologias utilizadas para processamento e visualização dos dados, simulação e otimização.

Neste trabalho, foi realizada uma pesquisa aplicada com uma abordagem quantitativa através da realização de experimentos em ambiente simulado computacionalmente.

Foi utilizada a linguagem de programação Python para a implementação dos algoritmos, simulações, análises e visualizações de dados. Dentre as principais bibliotecas utilizadas estão: Pandas, Plotly Express, Datashader, Scikit-learn (PEDREGOSA, 2011) e Pymoo (BLANK, 2020). Para construção do conjunto de dados utilizado nas simulações foram realizadas algumas filtrações e transformações no conjunto de dados original que possuía um total de 13.326.639 linhas. Para processar essa grande quantidade de dados, foi utilizado o Google BigQuery. O resultado foi exportado para um arquivo CSV que foi usado em todas as simulações. Mais detalhes sobre a fonte dos dados e as transformações aplicadas neles são descritas na seção 4 deste artigo.

Todos os códigos, parâmetros e dados utilizados nas simulações assim como os resultados obtidos podem ser encontrados no GitHub¹.

4. Conjunto de Dados

Nesta seção é descrito o processo de construção do conjunto de dados utilizado nas simulações, dos dados brutos ao *dataset* final. A primeira parte fala sobre as fontes de dados utilizadas. A segunda descreve a análise realizada nesses dados e os problemas identificados. Na terceira parte, são detalhadas as ações realizadas para mitigar esses problemas, gerando assim o conjunto de dados definitivo, que é descrito na quarta e última parte desta seção.

4.1. Fonte dos Dados

Os dados originais foram obtidos do site da B3, a bolsa de valores brasileira. Foi feito o download das séries históricas dos anos de 1986 a 2022 no site da B3 na seção de séries históricas². Os dados são disponibilizados em formato TXT de difícil

¹ <https://github.com/aldo-fsm/b3-portfolio-optimization/tree/article-2023-09>.

²

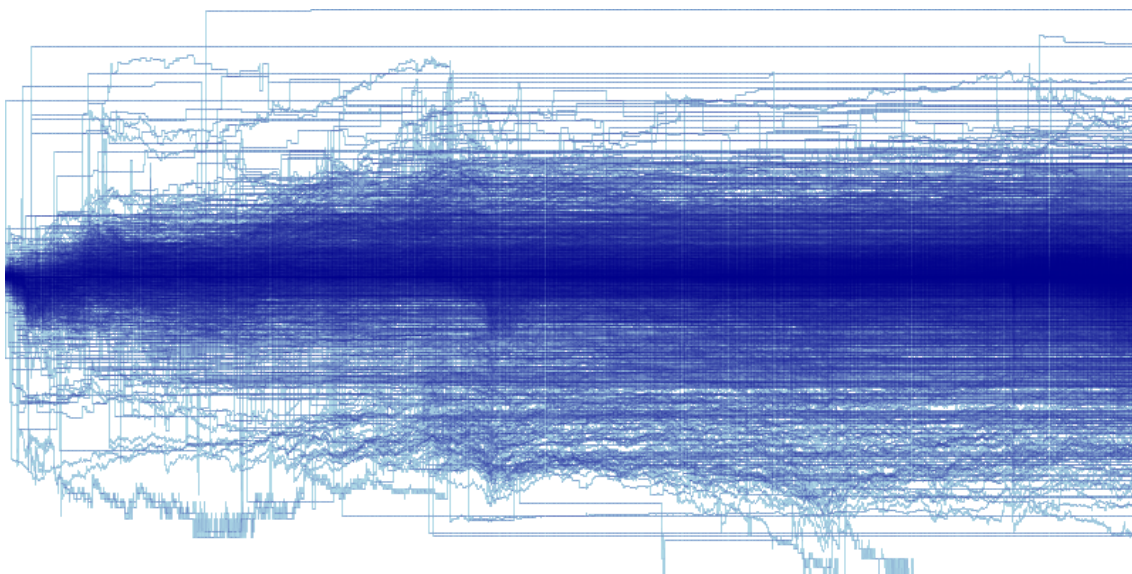
https://www.b3.com.br/pt_br/market-data-e-indices/servicos-de-dados/market-data/historico/mercado-a-vista/series-historicas/

manipulação, por isso foi utilizado o conversor, disponível no GitHub³, para converter os arquivos em Parquet. Além das séries históricas, também foram extraídos dados de eventos corporativos disponíveis no site da B3 na seção de empresas listadas⁴. Dados de eventos corporativos como desdobramentos e grupamentos são importantes para realizar o ajuste da série de retornos, uma vez que tais eventos geram variações bruscas nos preços resultando em distorções no cálculo dos retornos.

4.2. Análise Exploratória

Os dados convertidos em Parquet foram então salvos em tabelas no BigQuery para realização de análises e transformações. Utilizando a biblioteca Python Datashader foi possível plotar todas as séries temporais dos log-retornos acumulados referentes aos 4188 códigos de negociação distintos do mercado a vista no período de 16/03/1998 à 29/12/2022.

Figura 1 – Log-retornos acumulados no período de 16/03/1998 à 29/12/2022



Fonte: Elaborado pelo autor

Pode-se observar na figura 1, diversas linhas verticais e horizontais. As verticais são geradas por variações bruscas no preço do papel, indicando possíveis inconsistências na base de dados como, por exemplo, desdobramentos ou grupamentos não listados na base de dados. Já as linhas horizontais, representam papéis que não variaram o preço ou

³ <https://github.com/aldo-fsm/b3-datasets-parsing>

⁴ https://www.b3.com.br/pt_br/produtos-e-servicos/negociacao/renda-variavel/empresas-listadas.htm

que não possuem dados de negociação disponíveis no período, seja por interrupção da negociação do papel no período, entrada ou saída dele na bolsa, ou até mesmo mudança do código de negociação. Esses problemas identificados poderiam gerar resultados incorretos nas simulações, por isso aplicamos uma série de restrições ao conjunto de dados tornando-o mais próximo da realidade.

4.3. Restrições Aplicadas

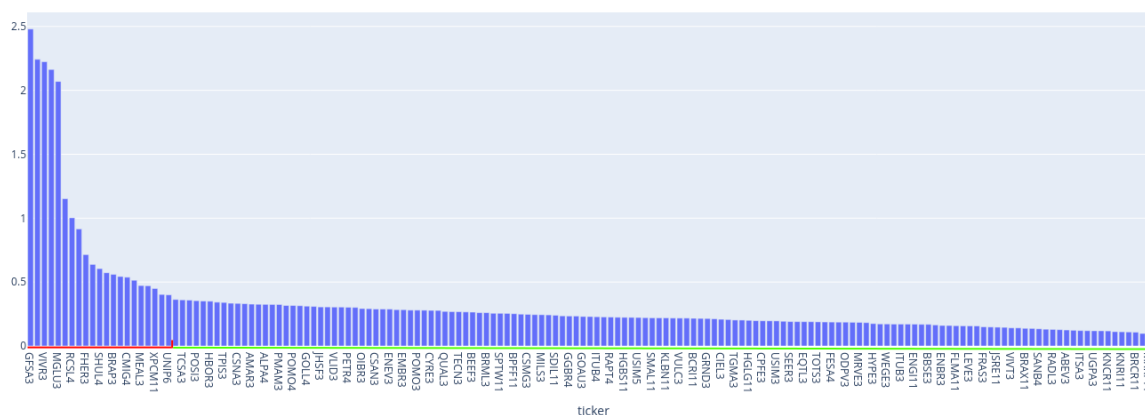
Para simplificar a análise e assegurar uma melhor qualidade dos dados, foram impostas algumas restrições para a construção do conjunto de dados final, a ser utilizado nas simulações. Apenas o período de 01/01/2017 à 31/12/2022 foi considerado, totalizando em 1485 datas distintas (dias úteis em que houve negociação na bolsa). Em relação aos papéis, foram considerados apenas os negociados no mercado à vista com código de negociação terminando com 3, 4, 5, 6, 7, 8, 11, 32, 33, 34 ou 35 e com o fator de cotação igual a 1.

Também foi imposta uma restrição na aplicação do ajuste do retorno devido a eventos corporativos: o ajuste foi aplicado apenas onde o valor absoluto do retorno é reduzido após o ajuste. Esse ajuste condicional foi feito para contornar a ausência de informações na base de eventos corporativos. Um exemplo disso é o que ocorreu com o ITUB4, que teve um desdobramento (1 para 100) e grupamento (100 para 1) ambos no dia 31/10/2011 (conforme indicado na página de relações com investidores no site do Itaú⁵), apesar de constar apenas o grupamento no site da B3. Nesse cenário, aplicar o ajuste faria com que o retorno do dia 31/10/2011 ficasse em torno de -99%, fato que não ocorreu.

Mesmo após os ajustes e restrições impostas, ainda haviam picos de retornos que poderiam indicar inconsistências. Para eliminar esses valores destoantes, foram retirados os códigos de negociação que tiveram o valor absoluto do log-retorno em um dia superior a 0,3633 (43,81%). A definição do ponto de corte foi feita a partir da análise do gráfico de barras dos valores absolutos dos log-retornos (figura 2).

⁵ <https://www.itaubr.com.br/relacoes-com-investidores/itau-unibanco/nossas-acoes/historico-de-eventos>

Figura 2 – Gráfico de barras usado para definir o ponto de corte



Fonte: Elaborado pelo autor

Para simplificar a simulação do ajuste da posição acionária devido a eventos corporativos, foram eliminados também os papéis que tiveram grupamentos no período. Após aplicadas todas as restrições, restaram apenas 137 códigos de negociação distintos no conjunto de dados.

4.4. Conjunto de Dados Definitivo

As restrições estabelecidas foram então incorporadas na consulta SQL para extração do dataset definitivo a partir das tabelas do BigQuery resultando em uma planilha CSV com 203.445 linhas e 6 colunas. A tabela 1 mostra o nome das colunas e uma breve descrição destas.

Tabela 1 – Colunas do dataset definitivo

Coluna	Descrição
date	Data de referência
ticker	Código de negociação do papel
price	Preço de abertura do papel no pregão na data de referência
adjust_factor	Fator de ajuste usado no cálculo do retorno
return	log-retorno em relação ao preço de abertura do dia seguinte
total	log-retorno acumulado

Fonte: Elaborado pelo autor

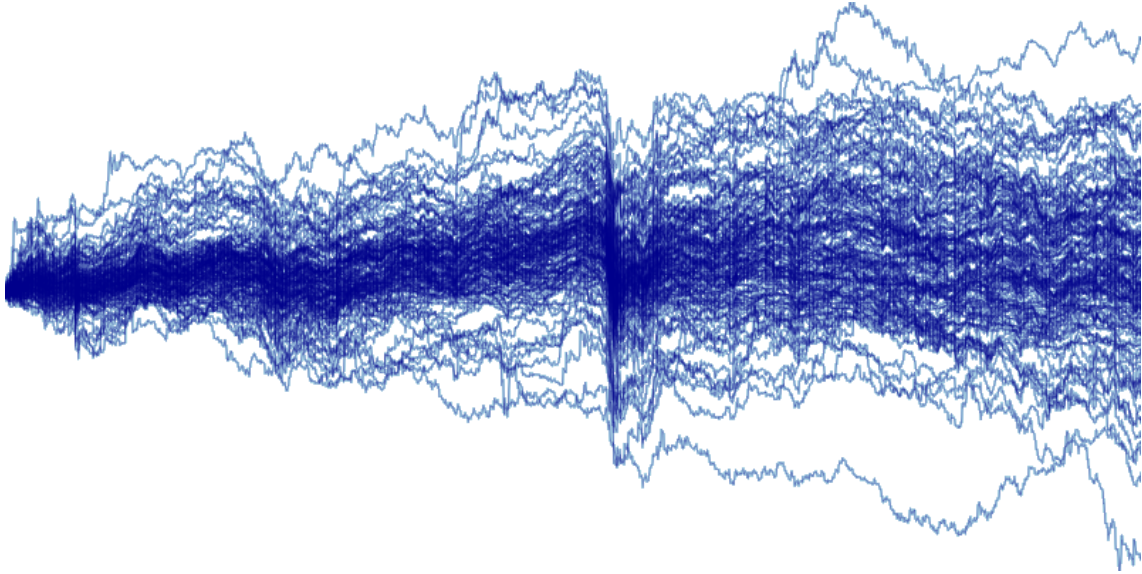
As colunas return e total foram calculadas, respectivamente, através das equações 7 e 8, onde a é o fator de ajuste (“adjust_factor”).

$$r_t = \ln\left(\frac{p_{t+1}}{a \cdot p_t}\right) \quad (7)$$

$$R_t = \sum_{i=1}^t r_t \quad (8)$$

A figura 3 mostra o retorno acumulado (coluna “total”) de todos os 137 códigos de negociação distintos presentes no conjunto de dados em relação ao tempo (coluna “date”). Em comparação com a figura 1, pode-se observar que as linhas verticais e horizontais não estão mais presentes, demonstrando a eficácia das restrições aplicadas, conforme descrito na seção 4.3.

Figura 3 – log-retornos acumulados no conjunto de dados definitivo



Fonte: Elaborado pelo autor

5. Experimentos

Nesta seção é descrito em detalhes o arranjo experimental utilizado neste trabalho, assim como o algoritmo utilizado para simular a aplicação das estratégias na bolsa de valores brasileira, suas parametrizações e os algoritmos de otimização e estimação de matriz de covariância utilizados. A seção é dividida em 4 partes:

primeiramente é descrito o algoritmo de simulação, onde os experimentos foram realizados. A segunda parte detalha o processo de otimização de portfólio utilizado. Na terceira parte são listadas as parametrizações de cada experimento. Por fim, na quarta parte são apresentados os resultados obtidos.

5.1. Simulador

Para a realização de experimentos da aplicação de otimização de portfólio na bolsa de valores brasileira, foi construído um simulador utilizando a linguagem de programação Python. Cada simulação pode ser configurada por uma série de parâmetros, conforme indicados e descritos na tabela 2.

Tabela 2 – Parâmetros de simulação

Parâmetro	Descrição
dataset	Conjunto de dados
start_date	Data inicial
initial_amount	Capital inicial
window_size	Tamanho da janela temporal usado nas estimações de médias e matrizes de covariância
rebalancing_interval	Intervalo em dias para rebalanceamento da carteira
relocation_interval	Intervalo em dias para realocação da carteira
initial_target_allocation	Alocação alvo inicial da carteira
covariance_matrix_estimator	Estimador de matriz de covariância

Fonte: Elaborado pelo autor

No início da simulação são inicializadas algumas variáveis de estado que serão modificadas e registradas numa lista de histórico no decorrer da simulação. As variáveis de estado e os respectivos valores iniciais são listados na tabela 3.

Tabela 3 – Variáveis de estado da simulação

Variável	Valor inicial
Data atual	Data anterior a especificada pelo parâmetro start_date
Montante total	Parâmetro initial_amount
Alocação alvo	Parâmetro initial_target_allocation
Alocação efetiva	0% para todos os ativos
Percentual desalocado	100%
Posição acionária	0 para todos os ativos
Valor em caixa	Parâmetro initial_amount

Fonte: Elaborado pelo autor

O simulador itera em ordem crescente as datas distintas presentes no dataset a partir da data inicial. Para cada iteração são executados as seguintes 6 etapas:

1. O montante total é calculado baseado nos preços de abertura na data de referência, as posições acionárias e o valor em caixa;
2. Caso seja a primeira iteração ou já se tenha passado o número de dias especificado no parâmetro relocation_interval desde a última relocação, é executado o processo de otimização de portfólio e o resultado deste é usado para atualizar a alocação alvo, isto é, o percentual desejado para cada ativo;
3. Caso seja a primeira iteração ou já se tenha passado o número de dias especificado no parâmetro rebalancing_interval desde o último rebalanceamento, é executado o processo rebalanceamento, onde as posições acionárias são modificadas para que o percentual aplicado em cada ativo seja o mais próximo possível da alocação alvo;
4. São calculados a alocação efetiva, percentual não alocado e o valor em caixa a partir dos preços, novas posições acionárias e montante total;
5. São registrados a data, montante total, alocação alvo, alocação efetiva, percentual não alocado, posição acionária e valor em caixa no histórico;
6. Finalmente, a posição acionária é ajustada usando o fator de ajuste (adjust_factor).

5.2. Otimização do Portfólio

Uma etapa importante dessa simulação é a aplicação da otimização de portfólio para a escolha automática do percentual de cada ativo na carteira. Para isso, foi aplicada a teoria de seleção de portfólio proposta por Markowitz (1952), com variações no conjunto de restrições do problema de otimização, visando maior proximidade com um cenário real de investimento na bolsa de valores. Nem toda alocação é possível de ser obtida na prática, uma vez que as ações são negociadas em quantidades inteiras do lote mínimo. O lote mínimo pode variar de acordo com o tipo de ativo, sendo os lotes de 1 e de 100 unidades os mais comuns. Para modelar essa característica da bolsa de valores dentro do problema de otimização, aplicamos a chamada restrição Round Lot, que exige que a posição acionária de um certo ativo deve ser de valores múltiplos do lote mínimo deste ativo (JIN, 2016). Pela ausência de dados sobre o lote mínimo dos ativos no conjunto de dados utilizado, foi considerado nas simulações um lote mínimo de 1 unidade para todos os ativos. A restrição de Round Lot conflita com a restrição imposta pela equação 5, visto que nem sempre é possível dividir o capital em múltiplos exatos do lote mínimo para todos os ativos. Por isso, flexibilizamos aquela restrição, permitindo que a soma das alocações seja menor que 1, colocando o restante do capital como valor em caixa. Como as simulações realizadas consideram o montante em dinheiro disponível para investimento, é preciso outra restrição para garantir que a quantia a ser investida não seja maior que o capital disponível. O conjunto de restrições aplicadas ao problema de otimização está expresso pelas equações 9, 10, 11 e 12, onde X_i é o percentual do ativo i na carteira, M o montante, p_i o preço do ativo i e N o número de ativos na carteira.

$$X_i = \frac{p_i}{M} \left\lfloor \frac{X_i \cdot M}{p_i} \right\rfloor \quad (9)$$

$$\sum_{i=1}^N X_i \leq 1 \quad (10)$$

$$X_i \geq 0 \quad (11)$$

$$X_i \cdot M \geq p_i \quad (12)$$

Para encontrar a alocação ótima dos ativos na carteira, foi utilizado o algoritmo de otimização multiobjetivo SMS-EMOA (BEUME, 2007), implementado na biblioteca Pymoo (BLANK, 2020) utilizando como objetivos a minimização de V (equação 4) e a maximização de E (equação 13). A equação original (3) foi substituída pela 13 devido a dois fatores: primeiro pela utilização do log-retorno, exigindo a conversão para retorno linear (através da exponencial) e segundo pelo fato da soma dos percentuais poder ser menor que 1, gerando um novo termo na equação.

$$E = \ln\left(\left(\sum_{i=1}^N X_i \cdot e^{\mu_i}\right) + 1 - \sum_{i=1}^N X_i\right) \quad (13)$$

Por ser uma otimização multiobjetivo, a solução gerada é um conjunto de soluções plausíveis. Dentre as carteiras eficientes encontradas pelo algoritmo, foi escolhida a que possui a maior razão E / V. Para garantir que as soluções geradas pelo algoritmo estarão de acordo com as restrições impostas, foi utilizado um mecanismo de “reparo”, onde as soluções geradas pelo algoritmo são modificadas para que as restrições sejam atendidas.

Para estimar a matriz de covariância, foram utilizados dois métodos diferentes. O primeiro deles é a matriz de covariância amostral, onde calcula-se a covariância dos log-retornos entre cada par de ativos em uma janela temporal. O segundo deles, foi o método proposto por Ledoit (2004) utilizando a implementação disponível na biblioteca Scikit-learn (PEDREGOSA, 2011).

5.3. Parametrizações das Simulações

Foram realizadas 23 simulações com diferentes parâmetros, sendo uma delas um *baseline* para comparar com o desempenho das estratégias de investimento utilizando a otimização de portfólio. Tal *baseline* consistiu em simular o investimento de 100% do capital no EFT BOVA11, que é um fundo de índice que replica o índice IBovespa. As demais 22 restantes simularam a aplicação da otimização de portfólio para definição da carteira de investimentos com diferentes parâmetros, como a janela temporal para média e covariância, frequência de realocação e rebalanceamento da carteira e estimador de matriz de covariância. Todas as simulações partiram da mesma data (11/01/2018) e com o mesmo capital inicial (R\$5000,00), para facilitar a comparação dos resultados. Quanto à janela temporal, foram utilizadas duas variações: anual (252 dias úteis) e trimestral (63

dias úteis). Já em relação a frequência de realocação e rebalanceamento, foram testadas 6 variações: diária (intervalo 1), semanal (intervalo 5), mensal (intervalo 21), trimestral (intervalo 63), anual (intervalo 252) e a execução da alocação apenas uma vez, no início da simulação (intervalo 10.000, ou seja, um intervalo superior ao número de dias úteis na base de dados). A cada simulação foi atribuído um ID único para posterior identificação na análise comparativa dos resultados. A simulação de baseline é identificada pelo ID “bova11”. Na tabela 4 são apresentados os valores dos parâmetros para cada simulação.

Tabela 4 – Valores dos parâmetros utilizados em cada simulação

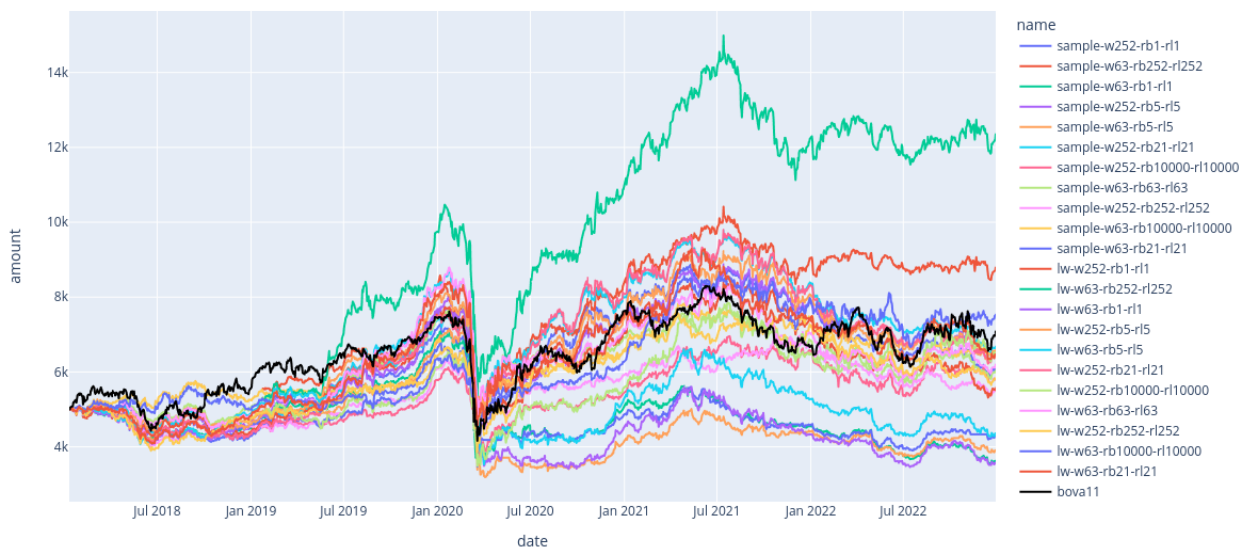
ID	start_date	initial_amount	window_size	rebalancing_interval	relocation_interval	covariance_matrix_estimator
bova11	2018-01-11	5000	252	1		sample
sample-w252-rb1-rl1	2018-01-11	5000	252	1	1	sample
sample-w63-rb252-rl252	2018-01-11	5000	63	252	252	sample
sample-w63-rb1-rl1	2018-01-11	5000	63	1	1	sample
sample-w252-rb5-rl5	2018-01-11	5000	252	5	5	sample
sample-w63-rb5-rl5	2018-01-11	5000	63	5	5	sample
sample-w252-rb21-rl21	2018-01-11	5000	252	21	21	sample
sample-w252-rb10000-rl10000	2018-01-11	5000	252	10000	10000	sample
sample-w63-rb63-rl63	2018-01-11	5000	63	63	63	sample
sample-w252-rb252-rl252	2018-01-11	5000	252	252	252	sample
sample-w63-rb10000-rl10000	2018-01-11	5000	63	10000	10000	sample
sample-w63-rb21-rl21	2018-01-11	5000	63	21	21	sample
lw-w252-rb1-rl1	2018-01-11	5000	252	1	1	LedoitWolf
lw-w63-rb252-rl252	2018-01-11	5000	63	252	252	LedoitWolf
lw-w63-rb1-rl1	2018-01-11	5000	63	1	1	LedoitWolf
lw-w252-rb5-rl5	2018-01-11	5000	252	5	5	LedoitWolf
lw-w63-rb5-rl5	2018-01-11	5000	63	5	5	LedoitWolf
lw-w252-rb21-rl21	2018-01-11	5000	252	21	21	LedoitWolf
lw-w252-rb10000-rl10000	2018-01-11	5000	252	10000	10000	LedoitWolf
lw-w63-rb63-rl63	2018-01-11	5000	63	63	63	LedoitWolf
lw-w252-rb252-rl252	2018-01-11	5000	252	252	252	LedoitWolf
lw-w63-rb10000-rl10000	2018-01-11	5000	63	10000	10000	LedoitWolf
lw-w63-rb21-rl21	2018-01-11	5000	63	21	21	LedoitWolf

Fonte: Elaborado pelo autor

5.4. Resultados

As 23 configurações da tabela 4 foram executadas e seus resultados salvos em planilhas para posterior análise. Na figura 4 podemos visualizar a evolução do patrimônio ao longo do tempo nas 23 simulações.

Figura 4 – Evolução do patrimônio nas 23 simulações



Fonte: Elaborado pelo autor

Analisando os gráficos apresentados na figura 4, pode-se observar que algumas configurações da otimização de portfólio geraram resultados bem superiores ao benchmark (bova11). Entretanto, houveram configurações que, no período analisado, performaram de forma bem inferior ao benchmark. Por outro lado, a volatilidade, calculada como o desvio padrão dos retornos diários, foi menor que a do benchmark em todas as configurações, conforme apresentado na coluna “Risco” da tabela 5, demonstrando a eficácia das estratégias baseadas na teoria de Markowitz (1952) na geração de portfólios diversificados e menos voláteis.

Tabela 5 – Métricas de performance obtidas nas simulações

ID	Capital Inicial	Montante	Retorno	Risco	Retorno/Risco
lw-w63-rb252-rl252	R\$5,000.00	R\$12,351.94	147.04%	1.56%	94.26
sample-w63-rb252-rl252	R\$5,000.00	R\$8,789.59	75.79%	1.13%	67.07
lw-w63-rb10000-rl10000	R\$5,000.00	R\$7,531.24	50.62%	1.12%	45.2
sample-w63-rb10000-rl10000	R\$5,000.00	R\$6,573.74	31.47%	0.90%	34.97
sample-w252-rb21-rl21	R\$5,000.00	R\$6,676.57	33.53%	1.09%	30.76
sample-w252-rb5-rl5	R\$5,000.00	R\$6,509.19	30.18%	1.10%	27.44
lw-w252-rb21-rl21	R\$5,000.00	R\$6,508.66	30.17%	1.17%	25.79
lw-w252-rb5-rl5	R\$5,000.00	R\$6,586.61	31.73%	1.27%	24.98
bova11	R\$5,000.00	R\$7,083.62	41.67%	1.68%	24.8
lw-w252-rb10000-rl10000	R\$5,000.00	R\$6,514.43	30.29%	1.43%	21.18
sample-w252-rb1-rl1	R\$5,000.00	R\$6,097.40	21.95%	1.07%	20.51
lw-w63-rb21-rl21	R\$5,000.00	R\$6,467.39	29.35%	1.47%	19.97
sample-w63-rb63-rl63	R\$5,000.00	R\$6,033.14	20.66%	1.16%	17.81
lw-w63-rb63-rl63	R\$5,000.00	R\$6,197.69	23.95%	1.38%	17.36
sample-w252-rb10000-rl10000	R\$5,000.00	R\$5,798.34	15.97%	1.09%	14.65
lw-w252-rb252-rl252	R\$5,000.00	R\$5,829.38	16.59%	1.28%	12.96
sample-w252-rb252-rl252	R\$5,000.00	R\$5,510.82	10.22%	1.01%	10.12
lw-w252-rb1-rl1	R\$5,000.00	R\$5,571.43	11.43%	1.22%	9.37
lw-w63-rb5-rl5	R\$5,000.00	R\$4,347.75	-13.05%	1.54%	-8.47
sample-w63-rb21-rl21	R\$5,000.00	R\$4,299.03	-14.02%	1.14%	-12.3
sample-w63-rb5-rl5	R\$5,000.00	R\$3,892.18	-22.16%	1.42%	-15.61
lw-w63-rb1-rl1	R\$5,000.00	R\$3,561.84	-28.76%	1.45%	-19.83
sample-w63-rb1-rl1	R\$5,000.00	R\$3,607.76	-27.84%	1.31%	-21.25

Fonte: Elaborado pelo autor

A configuração que gerou a maior razão retorno/risco foi a com janela de 63 dias e frequência de realocação/rebalanceamento anual usando o estimador de matriz de covariância proposto por Ledoit (2004), obtendo um retorno acumulado de 147,04% e desvio padrão de 1,56%. já o baseline bova11, valorizou apenas 41,67% no período e teve uma maior volatilidade, apresentando um desvio padrão de 1,68%.

6. Considerações Finais

Neste estudo, investigamos estratégias de alocação de ativos no mercado financeiro brasileiro, aplicando a teoria moderna do portfólio com otimização multiobjetivo em ambiente simulado computacionalmente. Nesta seção, serão

apresentados os principais resultados e conclusões obtidos, as limitações do estudo, bem como sugestões para trabalhos futuros.

6.1. Principais Resultados e Conclusões

Um dos resultados mais notáveis deste estudo foi que algumas configurações da otimização de portfólio superaram consistentemente o benchmark (BOVA11) em termos de retorno acumulado. Isso indica que a aplicação da teoria moderna do portfólio pode ser uma abordagem promissora para a construção de carteiras de investimento no contexto brasileiro. Além disso, observamos que todas as configurações da otimização de portfólio apresentaram um traço comum: a volatilidade (risco) das carteiras geradas foi menor do que a do benchmark. Isso enfatiza a capacidade das estratégias baseadas na teoria do portfólio de Markowitz em reduzir a exposição ao risco de mercado, um aspecto fundamental para investidores que buscam um equilíbrio entre risco e retorno.

6.2. Limitações

Apesar dos resultados promissores, é importante destacar que o sucesso dessas estratégias não é garantido e depende fortemente dos parâmetros escolhidos, incluindo a janela temporal, a frequência de realocação e os métodos de estimação da matriz de covariância. Além disso, as limitações e/ou simplificações aplicadas no processo de simulação podem afetar a validade destes resultados, uma vez que, num cenário real, tais limitações/simplificações não se aplicam. Dentre essas restrições, podemos destacar:

- O intervalo de tempo: Todas as simulações foram realizadas no período de 01/01/2017 à 31/12/2022;
- O conjunto de ativos: Apenas 137 papéis negociados na bolsa foram considerados nas simulações;
- Taxas e impostos: Não foram considerados nas simulações taxas como corretagem, taxa da B3, nem imposto de renda;
- Proventos: As simulações não levaram em consideração o pagamento de proventos como dividendos, juros sobre capital próprio, entre outros;

- Construção do conjunto de dados: As restrições aplicadas (descritas na seção 4.3) ao conjunto de dados para garantir maior qualidade dos dados, podem gerar viés aos experimentos.

6.3. Trabalhos Futuros

Para trabalhos futuros, podemos destacar a ampliação do escopo da simulação, tornando-a mais próxima da realidade, levando em consideração as taxas envolvidas nas operações na bolsa e os proventos pagos, bem como a ampliação do conjunto de ativos e do intervalo de tempo simulado. Além disso, pode-se testar outros algoritmos de otimização e de estimação de matriz de covariância. A inclusão de modelos preditivos em substituição ao retorno médio na estimação do retorno esperado, também parece ser uma oportunidade de aprimoramento dos resultados.

REFERÊNCIAS

- BEUME, Nicola; NAUJOKS, Boris; EMMERICH, Michael. SMS-EMOA: Multiobjective selection based on dominated hypervolume. **European Journal of Operational Research**, v. 181, p. 1653–1669, set, 2007.
- BLANK, Julian; DEB, Kalyanmoy. pymoo: Multi-Objective Optimization in Python. **IEEE Access**, v. 8, p. 89497–89509, abr, 2020.
- JIN, Yan; QU, Rong; ATKIN, Jason. Constrained Portfolio Optimisation: The State-of-the-Art Markowitz Models. **International Conference on Operations Research and Enterprise Systems**, v. 5, p. 388–395, 2016.
- LEDOIT, Olivier; WOLF, Michael. A well-conditioned estimator for large-dimensional covariance matrices. **Journal of Multivariate Analysis**, v. 88, p. 365–411, 2004.
- MARKOWITZ, Harry. Portfolio Selection. **Journal of Finance**, v. 7, p. 77–91, mar, 1952.
- O QUE É RISCO nos investimentos?. **Dicionário Financeiro**, [2017-]. Disponível em: <https://www.dicionariofinanceiro.com/risco-nos-investimentos>. Acesso em: 24/08/2023.
- PEDREGOSA, Fabian et al. Scikit-learn: Machine Learning in Python. **JMLR**, v. 12, p. 2825–2830, out, 2011.