OTIMIZAÇÃO DE PORTIFÓLIO COM TRADEOFF LUCRO x RISCO

Lucas Lukasavicus Silva
Jonathan Martinez Tsen
Universidade Federal de São Paulo (UNIFESP) e Instituto Tecnológico de Aeronáutica (ITA)
UNIFESP: Av. Cesare Mansueto Giulio Lattes, 1201 - Jardim Santa Inês I, S. J. Campos - SP
ITA: Pç. Marechal Eduardo Gomes, 50 - Vila das Acácias, S. J. Campos - SP
lukasavicus@gmail.com / dudujonathantsen@gmail.com

Prof. Dr. Edson França Senne
Universidade Federal de São Paulo (UNIFESP)
Av. Cesare Mansueto Giulio Lattes, 1201 - Jardim Santa Inês I, S. J. Campos - SP
edson.senne@unesp.br

RESUMO

Otimização de portfólio é um dos problemas mais clássicos na área de finanças, visto que todo investidor quer minimizar o risco da sua carteira de investimentos. Para resolver esse problema diversos estudos foram propostos, sendo que um dos mais famosos foi a Teoria Moderna de Portifólio de Harry Markowitz de 1952, nela é proposto pela primeira vez um estudo que considera informações históricas quantitativas para avaliar uma medida de risco associada a uma carteira e uma proposta para diminuir esse risco através da diversificação. 70 anos depois da sua proposta original essa teoria continua a influenciar diversas pesquisas como essa inclusive. Nesse artigo propomos uma extensão do modelo linear inspirado por Markowitz para incluir parâmetros que promovam maior adequação aos investidores modernos. Para isso coletamos dados e realizamos um pré-processamento adequado e avaliamos um modelo linear exato e uma metaheurística e discutimos seus resultados.

PALAVRAS CHAVE. Otimização de Portifólio. Tradeoff Lucro x Risco. AMPL. Metaheurística.

ABSTRACT

Portfolio optimization is one of the most classic problems in the finance area since every investor wants to minimize the risk of his investment portfolio. To solve this problem, several studies have been proposed, one of the most famous being the Modern Portfolio Theory by Harry Markowitz from 1952 in which a portfolio risk measure and a proposal to reduce that risk through diversification was proposed. 70 years after its original proposal, this theory continues to influence several researches such as this one. In this article we propose an extension of the linear model inspired by Markowitz to include parameters that promote greater suitability for modern investors. For this we collect data and perform an adequate pre-processing and evaluate an exact linear model and a metaheuristic and discuss their results.

KEYWORDS. Portfolio Optimization. Profit x Risk Tradeoff. AMPL. Metaheuristic.

1. Introdução

O objetivo de todo investidor é o lucro. Se houvesse algum tipo de investimento que garantidamente teria um retorno lucrativo em qualquer cenário de forma constante e consistente, todo investidor aplicaria seu patrimônio nesse investimento. Porém, como em todo mercado moderno existem riscos que decorrem de fatores intrínsecos e extrínsecos à cada investimento que fazem com a relação entre investimento e lucro não seja constante (nem em valor, nem em sinal, implicando em casos extremos, mas não incomuns, em perdas substanciais). Fatores como erros de cálculo de precificação, imprecisão na previsão de um cenário econômico futuro, ou mesmo situações mais anômalas como pandemias e guerras territoriais inserem riscos à mercados financeiros modernos. Com base nessa nova variável, o risco, o objetivo do investidor consciente, será maximizar seu lucro minimizando o risco do seu investimento. Esses dois objetivos são conflitantes por natureza. Para explicar esse fenômeno podemos recorrer a um exemplo:

Seja o ativo A um ativo que pode trazer um lucro de 10% sobre o valor investido no máximo ou 10% de prejuízo na pior das hipóteses. Já um segundo ativo B pode trazer 5% de lucro ou 2% de prejuízo, em cenários similares ao anterior. Se os eventos que influenciam nos lucros/prejuízos de ambos os ativos tivessem uma distribuição de probabilidade uniforme, podemos ver que na média, o lucro/prejuízo do ativo A se concentra em 0, enquanto o lucro/prejuízo do ativo B se concentra em 1.5% de lucro, sendo assim, considerando essa situação específica, seria mais interessante realizar um investimento no ativo B. Porém não sabemos a priori a distribuição de probabilidade dos eventos que influenciam os ativos de forma que uma hipótese de investimento válida pode ser construída com base no ativo A, supondo que em um horizonte de tempo futuro, cenários favoráveis são mais prováveis do que cenários desfavoráveis.

Esse exemplo ilustra bem a necessidade de incluir estratégias que visem diminuir os riscos inerentes dos investimentos, sendo que existem diversos riscos diferentes, como risco de mercado, risco de liquidez, risco operacional, etc. Não há como mitigar completamente todos os riscos, porém existem métodos de medir e controlá-los através de algumas estratégias. Dentre as estratégias de controle de risco, podemos citar a diversificação. Nessa estratégia a teoria de investimento determina que uma carteira com mais ativos, mais diversos está na média sujeita a uma variação do valor e do lucro menor do que uma carteira com poucos ativos, porém não são todos ativos que apresentam um desempenho (relação entre o valor investido e o retorno financeiro obtido dentro de um espaço de tempo) positiva ou atraente. Sendo assim, uma questão que surge naturalmente é: Dentro de um conjunto finito de ativos, qual é o melhor subconjunto de ativos que um determinado investidor deve comprar para maximizar seu lucro e minimizar seu risco sujeito a restrições como um dado orçamento de investimento.

Delimitando o escopo desse trabalho, passamos a considerar um tipo específico de investimento: os fundos de investimento imobiliários. Fundos de Investimentos Imobiliários ou FIIs podem ser formalmente definidos como a comunhão de recursos captados por meio do sistema de distribuição de valores mobiliários e destinados à aplicação em empreendimentos imobiliários (Instrução CVM 472, de 31 de outubro de 2008). Esses investimentos são interessantes por apresentar um baixo risco e rentabilidade moderada quando comparados com outros investimentos ditos de renda variável.

2. Revisão Bibliográfica

Após A análise de cenários econômicos e a proposta de métodos de otimização de portfólio datam desde a criação dos sistemas financeiros modernos, sendo que um dos métodos mais disruptivos e amplamente adotado como base para metodologias modernas é a Teoria Moderna de Portfólio de Harry Markowitz (Markowitz, 1952). Na sua proposta original, um investidor alocaria recursos, parcelas de um capital, em um determinado portfólio por um período que, iria produzir uma taxa de retorno aleatória, sendo que, a essa taxa aleatória no final do período apurado, o investidor teria um capital maior ou menor dependendo da média, ponderada dos rendimentos de cada ativo que compõem seu portfólio. Nesse sentido, Markowitz relaciona ideias fundamentais usadas até os dias de hoje, como, risco em função da volatilidade dos ativos e a taxa de retorno de uma carteira sendo dada em função do risco dela. Essas ideias formam a base da teoria que consegue formular objetivamente um modelo em que o objetivo é maximizar o lucro médio obtido (ou esperado) e minimizar o seu risco. Markowitz ainda modelou na sua proposta original uma medida de risco como sendo a variância dos valores dos ativos que compõem a carteira. Em termos objetivos podemos ver o modelo de Markowitz como:

O modelo proposto por Markowitz, no entanto, propõe que a medida de cálculo do risco se baseie na variância que é o quadrado do desvio-padrão e assim, esse modelo introduz uma variável quadrática. Apesar dos avanços na computação de problemas não-lineares, esse modelo ainda é bastante complexo e algumas vezes inviável de resolver com os *solvers* atuais, sendo assim, diversas propostas para linearizá-lo foram feitas (Sharpe, 1971, Speranza, 1993, Yitzhaki, 1982). Essas estratégias propõem diferentes métricas de riscos, diferentes funções-objetivos, adição de restrições reais, inteiras e binárias e o uso de modelos exatos e relativos, um resumo conciso da avaliação de cada um desses métodos e estratégias pode ser visto em (Mansini et. al., 2013). Dentre as propostas a que optamos por implementar principalmente pela facilidade de codificação é o modelo proposto por Konno e Yamazaki (Konno e Yamazaki, 1991), nesse modelo, a medida de risco se baseia no desvio-padrão. Konno e Yamazaki ainda provaram que a otimização de um problema linear cuja medida de risco é o desvio-padrão leva a otimização do modelo cuja medida é a variância, porém os resultados sempre serão piores ou iguais ao modelo original de Markowitz.

O uso de métodos computacionais também não é novidade, em (Karakalidis e Sifaleras, 2016), os autores avaliam diversos *solvers* comerciais, bem como a resolução que esses propõem para diferentes problemas de otimização de portfólios.

3. Descrição dos dados e metodologia de processamento

Segundo (B3, 2020 e Fiorini, 2012) fundos de investimento imobiliários podem ser entendidos como um grupo de investimentos que serão destinados para ativos (físicos ou não) relacionados ao mercado imobiliário (galpões logísticos, shoppings, prédios, corporativos, hospitais e cemitérios e títulos de dívidas imobiliárias (LCI)) (Assaf Neto, 2014). Uma cota de um fundo imobiliário, portanto é uma garantia que o investidor possui parte daquele grupo de investimentos tendo determinados direitos como votos em assembleias, mas principalmente direitos sobre os lucros advindos do pagamento de rendimentos sobre os ativos investidos. Isso pode ser exemplificado da seguinte maneira: Quando um investidor compra uma cota de um FII, ele compra uma fração de um determinado imóvel. O objetivo do investidor nesse caso, é tanto ganhar lucro pela valorização do imóvel no mercado imobiliário, quanto também através do pagamento do aluguel do imóvel por um ou mais locatórios que o imóvel pode vir a ter.

O valor e o rendimento de um FII mudam conforme o tempo, dessa forma, saber se um fundo irá valorizar, ou irá pagar bons rendimentos ao longo de um intervalo de tempo é necessário para todo investidor que deseja aumentar seu capital. Assim, é imprescindível a aplicação de métodos que procurem analisar tanto os fundamentos de um fundo imobiliário, como também possíveis cenários e projeções de lucro/prejuízo dos ativos econômicos. Isso implica em termos que coletar dados relativos ao histórico da valorização/desvalorização dos fundos, como também outros indicadores importantes (considerados fundamentais para diversas casas de análise de investimentos) como liquidez, número de cotas emitidas, valor patrimonial, valor unitário da cota, etc. Para futuras referências, nesse trabalho, iremos chamar o primeiro conjunto de dados históricos e o segundo conjunto de dados estáticos.

Esse trabalho, portanto, se limita a manipular dados desse investimento em renda variável, ainda que seja possível transportar as ideias debatidas nesse trabalho para outras modalidades de investimentos, como, por exemplo ações.

Os dados que usamos para o nosso trabalho foram os dados dos Fundos Imobiliários listados na B3 na data de 15/06/2022. Esse conjunto de dados conta com 407 ativos, sendo que para cada ativo coletamos 24 KPIs (key-performance indicators). Esse conjunto de dados foram os dados estáticos que consideramos como entrada para nossos modelos. Ainda coletamos para esses ativos o valor e o rendimento percentual ao longo dos últimos 12 meses, sendo que esse segundo conjunto de dados formam os dados históricos que consideramos como entrada para os modelos. No total, portanto, coletamos informações históricas e estáticas de 407 fundos imobiliários para trabalhar com nossos modelos.

Para trabalhar com esses dados foi necessário o desenvolvimento de rotinas de pré-processamento nas quais realizamos os seguintes passos:

- Coleta dos dados: Coletamos os dados de dois sites diferentes: https://www.fundsexplorer.com.br/ obtivemos os dados que chamamos estáticos, esses dados incluem informações sobre o setor do fundo imobiliário, preço atual, liquidez diária, dividendo, dividendo percentual, entre outras informações. Os dados de preço médio mensal e rendimento médio percentual mensal por fundo, os dados históricos, coletamos do site https://fiis.com.br/;
- Limpeza e filtragem dos dados: Uma vez com todos os dados em mãos, cruzamos os dois conjuntos a partir do nome do fundo imobiliário (que deve ser único sempre) para completar os dados, sendo que quando não achávamos um correspondente em uma das bases descartamos esses dados, ou seja, se houvesse um fundo A na primeira base, porém não na segunda base, descartamos esse dado do fundo A;
- Engenharia de features: Para melhor representar determinadas informações criamos alguns atributos a partir de atributos existentes nas bases obtidas pelo passo anterior. Alguns exemplos de dados produzidos assim foram a variável de “setor” do fundo, a qual codificamos usando o método hot-one encode. Além disso, fizemos alguns cálculos que ajudaram os modelos, como, por exemplo determinar o lucro máximo e mínimo, diversidade máxima e mínima da carteira e etc. Esses cálculos foram particularmente importantes para criar a função-objetivo do nosso modelo exato, o qual usamos essas informações para normalizar no intervalo [0, 1] cada uma dessas medidas;
- Exportação dos dados: Por fim, visto que coletamos, cruzamos os dados, limpamos e criamos atributos, o último passo do nosso script de pré-processamento foi a exportação desses dados para um formato que pudesse ser interpretado pela linguagem AMPL e pelo Python, para os modelos, exato e meta-heurística, respectivamente;

Exemplo das bases obtidas no final do processamento:

Códigodo fundo	Setor	Preço Atual	Liquidez Diária	Dividendo	DividendYield
KINP11	Residencial	13.00	2476.0	0.370736	2.896379
IDFI11	Híbrido	82.40	195.0	1.809889	2.196467
HGLG11	Logística	164.42	29445.0	3.300000	2.007055
JPPA11	Títulos e Val. Mob.	106.10	7675.0	2.040000	1.922714
RZAK11	Títulos e Val. Mob.	96.61	20283.0	1.800000	1.837672
NEWL11	Híbrido	93.00	12482.0	1.700000	1.827957
MBRF11	Outros	515.22	59.0	10.000000	1.818182
PORD11	Títulos e Val. Mob.	97.47	8205.0	1.720000	1.764646
FLCR11	Títulos e Val. Mob.	99.99	1094.0	1.750000	1.750000
VGIP11	Outros	98.39	39023.0	1.750000	1.745114

Figura 1: Exemplo de dados estáticos obtidos ao final do processamento.

asset	ABCP11	AFHI11	AFOF11	AIEC11	ALZR11	APTO11	ARCT11	ARRI11	BARI11	BBFI11B
yearmonth										
202106	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
202107	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
202108	73.50	0.00	0.00	81.40	0.00	0.00	0.00	101.93	106.50	2444.99
202109	69.25	95.00	91.45	81.00	115.99	0.00	101.75	97.45	105.33	2380.00
202110	66.99	95.61	79.97	74.00	116.25	0.00	101.79	96.45	103.40	2265.00
202111	61.34	98.25	79.53	68.81	102.98	0.00	101.90	94.22	101.70	1977.00
202112	69.99	98.20	82.44	73.50	109.95	10.59	105.00	97.98	103.85	2322.00
202201	74.30	101.18	86.00	77.70	114.49	10.53	108.00	101.89	105.15	2275.00
202202	77.04	102.25	90.00	77.21	116.45	10.19	109.50	100.48	103.16	2140.00
202203	72.90	102.15	93.15	81.03	110.15	10.20	109.86	99.00	102.70	2028.05
202204	73.57	103.13	94.01	80.20	115.52	10.40	106.31	98.80	102.45	2060.00
202205	72.50	100.65	89.00	80.30	116.35	10.10	106.00	96.20	100.60	1936.00
202206	0.00	102.44	90.89	0.00	115.50	0.00	109.00	0.00	0.00	0.00

Figura 2: Exemplo de dados históricos obtidos ao final do processamento (o índice dessa base é formado pela concatenação do ano e mês daquela informação para cada fundo imobiliário).

4. Modelos de solução do Problemas de Otimização

Diversos trabalhos foram propostos para solução do problema de otimização de portfólio que variam em função de características do problema como custos variáveis (custo de alocação, custo de transação na compra/venda de ativos), custos fixos (taxas de administração de corretoras/bancos), tamanho de lotes de investimentos (se é possível comprar cotas de forma unitária, ou somente um conjunto fechado de cotas, como um cento (100) de ações), tipos de variáveis (inteiras, binárias e/ou reais) entre outros aspectos. Em (Mansini et. al., 2013) os autores classificam os diversos modelos levantados pela revisão bibliográfica em modelos relativos, absolutos, exatos e heurísticas. Os termos relativos e absolutos dizem respeite o ao tipo da variável de decisão, sendo que quando é possível usar somente variáveis reais, esse modelo é chamado relativo, quando são usadas somente variáveis inteiras ou binárias esse modelo é dito absoluto.

A diferença entre modelos exatos e heurísticas diz respeito ao modo como uma solução factível é encontrada. Um modelo é dito exato quando esse implementa um método de busca em um espaço de busca, definido como um espaço n-dimensional (onde n é o número de variáveis consideradas no modelo) e delimitado por um conjunto de restrições. O modelo exato através de algum método de otimização então procura o mínimo ou máximo de uma função-objetivo definida, sendo que essa busca termina quando é encontrada a melhor solução dentre todas as possíveis soluções factíveis do problema. Esse modelo matemático possui diversos algoritmos de busca para diferentes problemas, sendo que os algoritmos de otimização mais conhecidos são o método simplex usado principalmente para problemas de otimização com variáveis do domínio dos reais, e os métodos de branch-and-bound e branch-and-cut para problemas com variáveis mistas (variáveis inteiras, binárias e reais). Esses métodos possuem um mecanismo eficiente para "caminhar" entre o espaço de soluções factíveis e sempre melhorar a solução encontrada. Outras características relevantes desses métodos é que o critério de parada de forma geral, é quando não existe possibilidade de encontrar uma solução melhor do que a atual, sendo que essa é considerada a solução ótima do problema (solução ótima global) e esses métodos podem ser utilizados para resolver qualquer problema linear desde que o problema seja modelado para atender determinados requisitos. Um ponto que é importante ser observado que motiva também a procura por outros métodos de resolução de problemas de otimização é que ao realizar o procedimento de busca do ótimo global da solução, em geral, modelos exatos podem ter que testar um número exponencial de soluções viáveis, isso acontece, em geral com problemas de natureza combinatória e nesses casos, a busca pela melhor solução pode ser exaustiva computacionalmente demorando diversas horas (em casos extremos até mesmos anos, sendo que um dos exemplos mais famosos é o problema do caixeiro viajante). Portanto, para esses cenários é necessário procurar modelos que sejam "mais rápidos", no sentido de procurar modelos que investiguem um número menor de soluções e encontrarem uma boa solução para o problema. Naturalmente que ao investigar um espaço de soluções menor, pode não ser possível encontrar a melhor solução para o problema de otimização, sendo que modelos que possuem essa característica são considerados modelos aproximados, ou seja, modelos que estão próximos da melhor solução global do problema. Esse balanço entre qualidade da solução e velocidade na busca pela solução foi um dos motivadores para o estudo de modelos conhecidos como meta-heurística.

De forma geral, meta-heurísticas podem ser explicadas como métodos que implementam procedimentos de buscas locais, ou seja, ao invés de considerarem todo o espaço de busca, esse modelo procura pela melhor solução dentre um espaço menor, mais limitado do que o espaço original (Holland, 1975 e Luke, 2009). Como a definição desse espaço de busca local depende de algumas características do problema e do algoritmo empregado, é comum que essemodelo encontre uma dificuldade conhecida como "problema do ótimo local", em que a melhor solução para aquele espaço é encontrada em um tempo factível, porém, existe outra região do espaço de soluções com outro ótimo local melhor do que o encontrado até o momento. Para aumentar a robustez desses modelos e tentar mitigar ou diminuir os efeitos de problemas como o do ótimo local, meta-heurísticas mais comuns empregam estratégias que permitem "transpor" algumas regiões do espaço de soluções em busca de outra solução melhor do que a solução ótima encontrada até o momento. Em particular, uma das estratégias mais comuns para endereçar esse problema foram os procedimentos que implementaram o conceito de vizinhança, que estabelece uma relação entre soluções para ser possível essa "movimentação" do algoritmo para espaços de busca ainda não explorados (Glover e Kochenberger, 2003). Vale ressaltar que a ideia de modelos de meta-heurística é buscar por uma solução boa, renunciando à garantia de otimalidade global que temos em modelos exatos, em contrapartida, a termos essa solução, em muitos problemas, muito mais rápido do que em modelos exatos.

4.1. Modelo Exato

O modelo exato implementado então se baseia na ideia do modelo original de Konno e Yamazaki, porém introduzimos novas variáveis e parâmetros para que o modelo fosse mais versátil com a condição e perfil de cada indivíduo que utilizasse esse modelo. Para parametrizar o modelo consideramos as métricas mais comuns na construção de portfólios de investimentos resilientes, como diversificação, risco e liquidez. Cada uma dessas métricas foi utilizada na função-objetivo do modelo de programação linear com pesos que ajudam a função a refletir ainda mais o racional do investidor que irá usar esse modelo. Dessa forma, nosso modelo pode ser expresso a partir da seguinte formulação:

F conjunto finito de opções de Fundos Imobiliários.

T um intervalo de tempo contínuo na qual foram avaliados os preços e rendimentos de fundos imobiliários

$V_{f,t}$ o valor do fundo imobiliário f em um instante t

$R_{f,t}$ o valor do rendimento pago por fundo f em um instante de tempo t

X_f variável de decisão que determina quanto do orçamento é aportada na compra de uma cota do fundo imobiliário f

P_f o preço estático do fundo imobiliário.

Re_f o rendimento estático do fundo imobiliário.

$$\max(((1 - avr) * pos) - (avr * neg))$$

Figura 3: Definição das variáveis do modelo exato

Sendo que *pos* é uma variável auxiliar denominada *parte positiva*, enquanto *neg* é a *parte negativa*.
Ambas variáveis podem ser definidas como:

$$pos = (lucro_{normalizado} * lucro_{peso} + liquidez_{normalizado} * liquidez_{peso} + diversidade_{normalizado} * diversidade_{peso}) / (lucro_{peso} + liquidez_{peso} + diversidade_{peso})$$

$$neg = risco_{normalizado}$$

Figura 4: Definição das variáveis intermediárias do modelo exato

Nesses modelos ainda temos o parâmetro *avr* de *aversão ao risco*. Esse parâmetro é uma medida definida no intervalo $[0,1] \in$ R, que pode ser interpretado como sendo a aversão do investidor a uma possível perda de capital em que 0 determina que o investidor não possui aversão ao risco, ou seja, tem um perfil de investimento bastante agressivo e 1 diametralmente representa o investidor ultra conservador, que irá tentar minimizar ao máximo o risco da sua carteira.

Restrito a:

R1 (restrição de orçamento): $\sum(X_f * P_f) <= budget;$
R2 (restrição de concentração): $X_f * P_f <= budget * perc_{concentracao}\ \forall f \in F$
R3 (restrição de concentração inter-setores): $AS_s <= budget * perc_{concentracaosetor}\ \forall s \in S$

Figura 5: Definição das principais restrições do modelo exato

Explicando ainda algumas variáveis intermediárias do modelo, temos:

$$lucro = \sum(X_f * Re_f)$$
$$liquidez = \sum(X_f)$$
$$diversidade = \sum(Xb_f)$$

Onde Xb é definido como:
$Xb_f = 0$ se $X_f = 0$ e 1 caso contrário.

Figura 6: Definição dos principais parâmetros auxiliares do modelo exato

O risco para esse problema foi definido seguindo a formulação de Konno e Yamazaki:

$$\begin{array}{ll} \text{minimize} & w(x)E\left[\left|\sum_{j=1}^{n} R_j x_j - E\left[\sum_{j=1}^{n} R_j x_j\right]\right|\right] \\ \text{subject to} & \sum_{j=1}^{n} E[R_j]x_j \geq \rho M_0, \\ & \sum_{j=1}^{n} x_j = M_0, \\ & 0 \leq x_j \leq u_j, \qquad j = 1, \ldots, n. \end{array}$$

Figura 7: Definição do cálculo do risco proposto por Konno e Yamazaki (Konno e Yamazaki, 1991)

Esse modelo foi implementado na linguagem AMPL (Fourer et. al., 2003) e é constituído de 3 arquivos:

- O arquivo do modelo – Em que descrevemos o modelo, seus parâmetros, suas variáveis, função-objetivo e restrições;
- O arquivo da execução – Nesse arquivo descrevemos quais arquivos de dados serão lidos como entrada para o modelo, qual é o *solver* (o algoritmo que será usado para resolver o problema descrito no modelo), quais variáveis queremos inspecionar durante e após a execução do modelo entre outras funcionalidades;
- O arquivo de dados – Nesse arquivo descrevemos todos os dados em uma estrutura de dados adequada, como uma matriz, ou vetor. Esse arquivo será lido pelo modelo como entrada para o processamento.

Os arquivos referentes a esse modelo e ao modelo de metaheurística desse artigo se encontram no repositório online:

https://github.com/Lukasavicus/OtimizacaoDePortifolio-IPO

4.1. Modelo Metaheurística

Como modelo de meta-heurística utilizamos o algoritmo bioinspirados conhecido como algoritmo genético. Algoritmos Genéticos, são um subconjunto de modelos de algoritmos evolutivos que implementam uma estratégia de exploração do espaço de busca estocástica, seguindo padrões observados nas áreas da biologia e genética. Os AGs fazem um mapeamento entre conceitos próprios dos modelos e dessas áreas, por exemplo, relacionando indivíduo a uma solução de um problema de otimização, um conjunto de soluções como uma população, uma determinada característica de uma solução (como a presença de um ativo em uma carteira) como um fenótipo, entre outras relações.

Para o problema de otimização de carteiras, temos então a seguinte representação via algoritmo genético:

Tabela 1: Mapeamento entre os conceitos de algoritmos genéticos com o problema de otimização de portifólio

Conceito do Modelo Genético	Adaptação para o problema de otimização
Indivíduo	Uma carteira, ou seja, um conjunto de ativos e a alocação de capital em cada um deles.
População	Um conjunto de carteiras
Seleção de pais	Critério para verificar quais pais podem formar a nova geração
Tipo de mutação	Maneira de mudar os genes de um indivíduo para gerar maior diversidade
Cruzamento	Maneira pela qual os genes de pais são selecionados para compor os genes de um indivíduo filho
Quantidade de pais mantidos entre as gerações	Número que determina a quantidade de pais que iremos selecionar da geração anterior para mantermos, usando um critério de elitismo e reinserção

Como nesse trabalho, decidimos comparar o uso da meta-heurística algoritmo genético contra o modelo exato com objetivo de entendermos se conseguiríamos obter um modelo que fosse mais rápido em termos de processamento implementamos um modelo baseado nessa meta-heurística utilizando o pacote PyGAD (Gad, 2021) na linguagem python 3.7.2. Esse pacote apresenta uma abstração para a implementação de algoritmos genéticos extremamente útil, de modo que tivemos que definir somente os hiperparâmetros (como números de gerações, quantidade de pais mantida entre as gerações, tipo de seleção dentre outros) e a função de aptidão, que foi uma adaptação da função original que usamos no modelo exato.

Como o modelo que implementamos possuía diversos hiperparâmetros, cada um deles com um conjunto de possíveis valores, antes da aplicação do modelo, passamos pela etapa de tunning dos hiperparâmetros, nessa etapa são feitos ajustes que determinam o melhor conjunto de valores para os parâmetros para a resolução de um determinado problema por um modelo de otimização paramétrico (Faceli et. al., 2021). Existem algumas técnicas que podem ser empregadas para esse ajuste, dentre as mais conhecidas destacamos a busca em grade, que consiste em criar um conjunto de combinações dos valores dos parâmetros testá-los com o problema que pretende-se resolver e a busca simples, nessa última cada valor é hiperparâmetro é ajustado individualmente, sendo que os demais permanecem com um valor default, quando o melhor valor para um parâmetro é decidido passamos a utilizar esse valor para as próximas análises. Como na busca em grade teríamos que testar aproximadamente 19.000 possibilidades, decidimos utilizar a busca simples, passando então a testar somente 27 possibilidades. Para diminuir o efeito de possíveis más escolhas para a inicialização das variáveis aleatórias usadas no modelo (como as variáveis aleatórias que determinam uma população inicial) testamos cada escolha de parâmetros um número n (nesse trabalho n = 4) de vezes e tiramos a média de cada experimento.

A tabela a seguir mostra quais valores foram testados para cada hiperparâmetro, bem como o conjunto de valores que mostrou o melhor resultado:

Tabela 2: Valores dos hiperparâmetros testados e o melhor valor para cada um

Hiperparâmetro	Valores testados	Melhor valor
num_parents_mating	3, 5, 7, 10	3
sol_per_pop	10, 15, 20	20
keep_parents	1, 2, 3	2
parent_selection_type	'sss', 'rank', 'sus', 'rws'	‘rank’
crossover_type	'single_point', 'uniform'	‘uniform’
mutation_type	'random', 'swap', 'adaptive'	‘random’
mutation_percent_genes	10, 20, 30	10
num_generations	1000, 5000, 10000, 20000, 100000	100000

5. Avaliação e comparação de resultados

Para avaliar os modelos exato e de meta-heurística, consideramos um investidor hipotético que forneceria um conjunto de parâmetros de entrada para o modelo na expectativa de obter a melhor carteira que os respeitassem, são os parâmetros:

- Orçamento: 10.000,00 reais;
- Aversão ao risco: 0.8 -isso indica um perfil mais conservador do que arrojado;
- Restrição de concentração: 30% - esse número indica que o maior valor que poderia ser alocado para um único investimento seria de 30% do orçamento inicial;
- Restrição de concentração entre setores: 30% - esse número também indica uma restrição, porém agora entre os setores dos FIIs. Esse valor significa que no máximo 30% do orçamento pode ser alocado para um único setor.
- Restrição de P/VP: Entre 0.8 e 1.4 – Essa restrição indica para o modelo quais FIIs podem ser considerados, nesse caso, somente aqueles que tiverem um P/VP (Patrimônio / Valor Patrimonial) dentro do intervalo. Esse indicador é importante para avaliar se um determinado fundo está sendo considerado muito descontado (quando essa razão é menor do que 1) ou supervalorizado (quando o número é maior do que 1).
- Restrição de Liquidez: > 0 – Alguns fundos continuam a serem listados mesmo quando não aceitam novas negociações, para eliminá-los filtramos os fundos com liquidez menor ou igual a 0.

Considerando esses parâmetros, ambos modelos buscaram construir o melhor portfólio, sendo que a solução obtida pelo modelo exato foi a seguinte:

Dos 245 ativos que o modelo considerou para a construção do portfólio (ativos que respeitavam as restrições impostas pelo usuário), o modelo escolheu 97 ativos, alocando um capital total de R$ 9.998,98 comprando um número total de 634 cotas e obtendo um lucro através dos rendimentos dessa carteira de R$ 143,94 ao mês (se as mesmas condições passadas fossem mantidas), ou seja, com um rendimento médio de 1,43% do valor investido ao mês. Para efeitos de comparação, segundo (Kastner, 2022) no ano de 2021, o valor médio de uma carteira de FIIs renderia aproximadamente 0,87% ao mês, olhando ainda para outras carteiras recomendadas como (Itau, 2022 e BTG, 2022) ainda vemos que essas carteiras em média rendem 0,9% a 1,05% ao mês.

O método implementado através da meta-heurística algoritmo genético, não produziu uma solução satisfatória, ou seja, a meta-heurística não produziu uma solução que respeitasse as restrições impostas na nossa modelagem. Ainda assim, observamos que meta-heurística ainda que produzisse uma solução satisfatória ainda levaria muito mais tempo para encontrá-la, visto que o tempo médio para solução do problema através da meta-heurística era de 4840 segundos, contra menos de 2 segundos da execução do método exato.

6. Conclusões e próximos passos

Concluímos que apesar de ser um problema complexo, o método exato produz uma solução ótima em um excelente tempo de execução. Nosso método ainda proporciona maior customização do que métodos anteriores, pois permite que o investidor entre com informações particulares como orçamento para investimento, nível de aversão ao risco e limites para concentração de ativos e demais parâmetros como importância da liquidez, lucro e diversidade, isso faz com que o nosso método seja extremamente versátil e útil. Ainda podemos explorar outros aspectos desse problema, como utilizar outras métricas quantitativas para filtrar resultados indesejados, como, por exemplo, permitir excluir da carteira um determinado setor que o investidor acredita estar ameaçado pela economia, além de ponderar os valores históricos para atribuir um maior peso àqueles valores mais próximos, para tentar reduzir o impacto de eventos anômalos do passado (como por exemplo, outliers que podem ter aparecido em decorrência da pandemia do COVID19). Outra ideia interessante é trabalhar com modelos mais complexos que não se limitem a modelagens lineares ou terem multiobjetivos.

Referências

Markowitz, H.M., (1952). Portfolio selection. Journal of Finance 7, 77–91.

Sharpe, W.F., (1971). A linear programming approximation for the general portfolio analysis problem. Journal of Financial and Quantitative Analysis 6, 1263–1275.

Konno, H., Yamazaki, H., (1991). Mean-absolute deviation portfolio optimization model and its application to Tokyo stock market. Management Science 37, 519–531.

Yitzhaki, S., (1982). Stochastic dominance, mean variance, and Gini's mean difference. American Economic Revue 72, 178–185.

Speranza, M.G., (1993). Linear programming models for portfolio optimization. Finance 14, 107–123

B3. Fundos de Investimentos. B3. 2020. `http://www.b3.com.br/pt_br/produtos-e-servicos/negociacao/renda-variavel/fundos-deinvestimento-imobiliario-fii.htm`. Acessado: 2022-06-30.

Mansini, R., Ogryczak, W., Grazia Speranza, M., Twenty Years of Linear Programming Based Portfolio Optimization, European Journal of Operational Research (2013), doi: http://dx.doi.org/10.1016/j.ejor.2013.08.035

Glover, F. e Kochenberger, G. A. (2003). Handbook of Metaheuristics. Kluwer Academic Publishers, Boston.

Fiorini, Renato Maestre . Determinantes da rentabilidade dos fundos de investimento imobiliários no brasil. São Paulo, 2012. Dissertação (Economia) - Fundação Getúlio Vargas. https://bibliotecadigital.fgv.br/dspace/bitstream/handle/10438/10349/Dissertacao_Renato_Fio rini_20121228.pdf. Acessado: 2022-06-30.

Assaf Neto, Alexandre. Mercado Financeiro. 12. ed. São Paulo: Atlas, 2014.

Luke, S. (2009). Essentials of Metaheuristics. http://cs.gmu.edu/~sean/book/metaheuristics/

Holland, J. H. (1975). Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control and artificial intelligence. Estados Unidos da America: The MIT Press, 1975.

A. F. Gad. Genetic algorithm python: Building genetic algorithm in python. https://github.com/ahmedfgad/GeneticAlgorithmPython/tree/05a069abf43146e7f8eb37f37c539523bf62ac9a, 2021.

Kastner, Tássia (2022) Web page. https://vocesa.abril.com.br/mercado-financeiro/a-ruina-dos-fundos-imobiliarios-e-como-reerguer-a-sua-carteira. Acessado: 2022-06-01.

Fourer, R., Gay, D. M., e Kernighan, B. W. (2003). AMPL: A Modeling Language for Mathematical Programming. Acessado: 2022-07-05.

Faceli, K, Lorena, A.C., Gama, J. e de Carvalho, A. C. P. L. F. (2021) Inteligência Artificial: Uma abordagem de Aprendizado de Máquina, ed. LTC

Itau (2022). Carteira recomendada com Imóveis. https://ww69.itau.com.br/fileserver/relatorios/carteira_recomendada_FII_junho_2022.pdf. Acessado: 2022-07-10.

BTG (2022). Carteira recomendada de fundos imobiliários. https://digital-research-prd.s3-sa-east-1.amazonaws.com/file/2022-07-04T100951.901_2022-07-01T104044.672_BTGFII%20JUL22.pdf. Acessado: 2022-07-10.

Karakalidis, A., Sifaleras, A. (2017). Solving Portfolio Optimization Problems Using AMPL. In: Grigoroudis, E., Doumpos, M. (eds) Operational Research in Business and Economics. Springer Proceedings in Business and Economics. Springer, Cham. https://doi.org/10.1007/978-3-319-33003-7_8