



## **Análise Envoltória de Dados para a Gestão de Portfólios**

**Nathan Arnaldo Bido**

n204090@dac.unicamp.br

**Orientadora: Profa. Dra. Priscila Cristina Berbert Rampazzo**

Faculdade de Ciências Aplicadas

### **RESUMO**

Este trabalho teve como objetivo inicial aplicar a metodologia de Análise por Envoltória de Dados (DEA) na Gestão de Portfólios, por meio da avaliação da eficiência financeira de um conjunto de empresas listadas na Bolsa de Valores de São Paulo. Paralelamente a isto, um Algoritmo Genético (NSGA-II) foi utilizado para abordar o problema de otimização de portfólios com múltiplos objetivos. Desta forma, este projeto aplicou um modelo DEA como método para selecionar as DMUS mais eficientes e o algoritmo NSGA-II, para identificar o quanto deveria ser investido em cada um dos ativos selecionados na primeira etapa. A hipótese que motivou a pesquisa é que o processo de seleção (com DEA) pudesse contribuir para melhores resultados do processo de otimização (através do modelo de Markovitz com múltiplos objetivos). Os problemas foram modelados matematicamente e resolvidos por um *solver* disponível no módulo *ortools*, em linguagem de Programação Python.

**PALAVRAS CHAVE.** Análise Envoltória de Dados, Multiobjetivo, Portfólios.

### **1. Introdução**

Dentro do mercado de investimentos, a busca por retornos financeiros motiva diversos investidores a realizar aplicações expressivas em determinados ativos, assim como também há outros que procuram investir naqueles que representam menores riscos. No mercado financeiro, um investidor racional sempre busca minimizar os riscos e maximizar o retorno de seu investimento em um portfólio ideal [Sarker, 2013]. Apesar de parecer simples, selecionar uma carteira de investimento, o qual é um estudo de como investir um patrimônio [Junior et al., 2002], é algo complexo. Pode-se dizer que a seleção de carteira de investimento enquadra-se como um problema de otimização de portfólios. Portanto, para buscar soluções que representassem melhor a realidade, modelos de Otimização Multiobjetivo passaram a ser utilizados [Markovitz, 1952]. Sendo assim, este trabalho tem interesse no problema de otimização multiobjetivo, com o intuito da maximização de retornos e minimização de riscos. Para este caso, o Modelo de Média-Variância de Markowitz [Markovitz, 1952] foi utilizado. Assim, uma das formas para abordar o modelo não-linear e multiobjetivo de Markowitz é através de Algoritmos Genéticos Multiobjetivo. É através desta abordagem que obtemos a *fronteira não-dominada*, que representa o conjunto de soluções de um problema de otimização multiobjetivo.

Paralelamente a isto, um modelo de programação linear pode ser utilizado para auxiliar na seleção de investimentos, através da Análise Envoltória de Dados (DEA), o qual é uma técnica de mensuração de desempenho com o objetivo de determinar a eficiência relativa de unidades produtivas [Lopes e Carneiro, 2010]. O modelo DEA tem como principal foco avaliar as eficiências das DMUs (ativos). Do ponto de vista multicritério, a hipótese que se deseja analisar, é se a Análise Envoltória de Dados pode contribuir na pré-definição do conjunto de portfólios que será utilizado na etapa de otimização. Como temos múltiplos objetivos, o resultado é uma fronteira de soluções não-dominadas, o que possibilita o tomador de decisão escolher o retorno que deseja obter e o risco que está disposto a assumir.



## 2. Materiais e Métodos

As seções a seguir descrevem os procedimentos utilizados para a seleção de portfólios eficientes através dos Modelos DEA e o processo de otimização dos portfólios selecionados através do Modelo de Markovitz.

### 2.1. Definição de Inputs e Outputs

Para utilizar o DEA para análises das eficiências, é necessário definir os recursos disponíveis ou entradas (inputs) e o que foi produzido ou saídas (outputs). De acordo com [Dyson et al., 2001], para a escolha dos inputs e outputs, é necessário que os parâmetros apresentem baixa correlação, ou seja, que os parâmetros não sejam múltiplos um dos outros ou estejam relacionados diretamente. Para a seleção dos inputs e outputs buscou-se na literatura sobre indicadores financeiros, uma vez que trata-se de uma análise financeira, assim como protocolos e cuidados que devem ser seguidos; dentre o material estudado pode-se destacar [Dyson et al., 2001], [Fonseca, 2009], [Bodie et al., 2014], [Ross et al., 2013], [Damodaran, 2012], [Alexandrino et al., 2017]

De acordo com [Silva et al., 2014], alguns índices de mercado foram criados com o objetivo de medir a performance das empresas em diferentes áreas, servindo de base para comparação entre elas. Estes indicadores, também conhecidos como indicadores fundamentalistas, acabam por servir como apoio na prática de investimentos em ações em diferentes empresas. Dentre os indicadores fundamentalistas que são mais usualmente utilizados pelos investidores, foram escolhidos para a presente pesquisa, os seguintes: dividend yields, ROE, retornos, Lucro por Ação (LPA), a volatilidade e o coeficiente beta, constituindo-se então os *inputs* e *outputs*. O *coeficiente beta* e a *volatilidade* foram tratados como inputs, uma vez que estes retratam os riscos associados ao processo de investimento, e tem como objetivo por parte do investidor de minimizá-los. Por outro lado, o *ROE*, *LPA*, *Dividend Yields* e *Retornos* foram considerados como outputs, pois estas tornam-se mais atrativas conforme são maiores, ou seja, há um objetivo de maximizá-las por parte do investidor. Sobre a unidade de medição dos indicadores, os valores de LPA foram medidos em reais, enquanto que os demais parâmetros, com exceção do coeficiente beta, foram medidos em porcentagem.

### 2.2. Análise Envoltória de Dados: Construção dos Modelos DEA

Nesta etapa, foi realizado o estudo dos principais modelos que compõem a metodologia DEA: CCR Multiplicador, CCR Envelope, BCC Multiplicador e BCC Envelope, assim como suas particularidades. Os modelos foram implementados utilizando-se de ferramenta computacional. Após a implementação do modelo, o programa é executado por um *solver*. Os programas foram implementados em linguagem de programação *Python* em ambiente Google Colab, importando o pacote *ortools*. O OR-Tools [OR-Tools, 2019] é uma biblioteca de código aberto utilizada para escrever os modelos de otimização.

O modelo escolhido, que melhor se adaptou à problemática do trabalho, foi o modelo BCC [de Mello et al., 2005]. A escolha se deu por conta do BCC trabalhar com um retorno variável de escala. Em outras palavras, pode-se dizer que um aumento nos *inputs* pode não variar proporcionalmente nos *outputs*. O modelo matemático do BCC é apresentado a seguir. No modelo,  $x_{ik}$  : *input*  $i$  da  $k$ -ésima DMU;  $y_{jk}$  : *output*  $j$  da  $k$ -ésima DMU;  $v_i$  : pesos dos *inputs*;  $u_j$  : pesos dos *outputs*.



$$\text{Max Effo} = \sum_{j=1}^s u_j \cdot y_{jo} + u_* \quad (1)$$

$$\sum_{i=1}^r v_i \cdot x_{io} = 1 \quad (2)$$

$$\sum_{j=1}^s u_j \cdot y_{jk} - \sum_{i=1}^r v_i \cdot x_{ik} + u_* \leq 0 \quad \forall k \quad (3)$$

$$v_i, u_j \geq 0 \quad \forall i, \forall j, \quad u_* \in \mathbb{R} \quad (4)$$

### 2.3. Modelo de Otimização: Explicando Modelo de Markowitz

Sabe-se que, dentro de um cenário de investimentos, um investidor tem basicamente dois objetivos: maximizar seus lucros e minimizar os riscos [Markovitz, 1952]. Para isto o modelo de Otimização de Markowitz pode ser utilizado, a fim de analisar estes dois principais parâmetros. O modelo é apresentado a seguir, onde  $x_i$ ,  $i = 1, 2, \dots, n$ , é o percentual do capital  $C$  a ser investido no ativo  $i$ , sendo  $n$  o número de ativos;  $\sigma_{ij}$  é a covariância entre os ativos  $i$  e  $j$ ;  $\bar{r}_i$  é o retorno esperado (médio);  $f_1(x)$  : função de retorno e  $f_2(x)$  : função de risco.

$$\text{Maximizar} \quad f_1(x) = \sum_{i=1}^n \bar{r}_i x_i \quad (5)$$

$$\text{Minimizar} \quad f_2(x) = \sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} x_i x_j \quad (6)$$

$$\text{Sujeito a} \quad \sum_{i=1}^n x_i = 1 \quad (7)$$

$$x_i \geq 0 \quad (8)$$

### 2.4. Algoritmo Genético Multiobjetivo - NSGA-II

Para abordar o modelo de Markowitz, foi utilizado o Algoritmo Genético *Non-dominated Sorting Genetic Algorithm* (NSGA-II) [Deb et al., 2000]. O resultado desta etapa de otimização é uma fronteira com as soluções não-dominadas em relação aos objetivos de minimização de risco e maximização de retorno. Os pontos pertencentes à esta fronteira dominam quaisquer outras soluções e são indiferentes entre si, no sentido de que a melhora em um dos objetivos, ocasiona piora em outro (objetivos conflitantes). O objetivo da obtenção desta fronteira não-dominada é disponibilizar ao tomador de decisão alternativas, de acordo com teu perfil de investimento.

## 3. Resultados

As empresas escolhidas para o procedimento de gestão de portfólios, foram as empresas cotadas na IBRX50, índice o qual seleciona as 50 empresas com maior negociabilidade e representatividade dentro da Bolsa de Valores de São Paulo (IBOVESPA). Para a pesquisa foi levado em consideração o período de 1 de janeiro de 2019 até 31 de dezembro de 2019, pois, era necessário um conjunto de dados de pelo menos um ano e devido a isto, das 50 empresas na IBRX50, 49 foram consideradas. Por conta da pandemia o foco se deu nos dados do ano de 2019.

Durante o processo de implementação, foi feito uma pré-seleção dos ativos, retirando aqueles com a presença de outputs e inputs negativos, resultando num total de 36 empresas. Destas 36 empresas, o modelo DEA BCC considerou que 9 são eficientes (de acordo com os critérios que foram adotados e com o período dos dados analisados), como podem ser vistas na Figura 1. Nota-se uma concentração de ativos eficientes no setor financeiro. Tal concentração demonstra o bom



desempenho deste setor no ano de 2019, frente a queda de outros setores durante mesmo período [SEEB; Tauhata], sendo o foco maior para o subsetor de intermediários financeiros (bancos). Um outro modo de analisar a eficiência da carteira de investimentos propostas, é analisando o retorno da IBRX50 e da carteira selecionada com os 9 ativos definidos como eficientes. Para esta análise, foi feito o gráfico através do retorno das ações dos ativos durante o início dos meses de janeiro à dezembro, o qual pode ser visto na Figura 2. Verifica-se que a carteira selecionada apresenta pico de retorno maiores que a do IBRX50, e menores picos de perdas, fato este que prova sua maior eficiência. Em números, o retorno oferecido pela Carteira selecionada no ano, é de 57% enquanto que o do Índice Bovespa é de 25%, além de que, este aumento de retorno é seguido por um aumento de apenas 2% de volatilidade histórica, sendo o do IBRX50 18% e da Carteira selecionada de 20%.

Eficiência dos Ativos					
Ativos	Eficiência	Ativos	Eficiência	Ativos	Eficiência
ABEV3	1,46%	GGBR4	6,90%	PETR4	100%
B3SA3	17,33%	GOAU4	9,15%	RADL3	62,32%
BBAS3	100%	IRBR3	0%	RAIL3	13,06%
BBDC3	85,05%	ITSA4	96,45%	RENT3	20,86%
BBDC4	90,18%	ITUB4	100%	SBSP3	62,94%
BBSE3	100%	JBSS3	60,69%	VIVT4	100%
BRDT3	100%	LAME4	0%	WEGE3	44%
BRML3	0%	LREN3	1,55%	YDUQ3	61,87%
CCRO3	64,37%	MGLU3	45,50%	BPAC11	100%
C3NA3	20,33%	MRVE3	0%	COGN3	0%
ELET3	100%	MULT3	0%	GNDI3	100%
ELET6	17,02%	PETR3	0%	NTCO3	19,82%

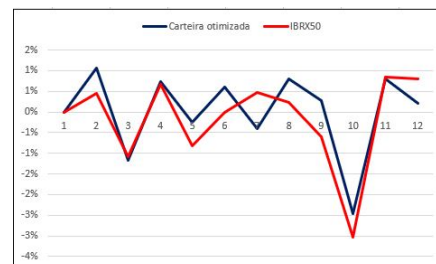


Figura 1: Eficiência dos 36 ativos pelo BCC.

Figura 2: Comparação da Carteira vs IBRX50.

As análises do modelo de otimização multiobjetivo foram divididas em dois casos de estudo: com 49 ativos iniciais e com os 9 ativos eficientes resultantes do modelo DEA. Neste caso, em ambas as situações, foram testados três cenários distintos para os parâmetros de tamanho da população e gerações, sendo os melhores resultados obtidos com um tamanho de população 300 e um número de gerações também de 300. Além disso, a partir destes testes, foi possível identificar uma diferença nas fronteiras não-dominadas dos dois casos de estudo. A comparação pode ser vista nas Figuras 3 e 4. Através do Algoritmo Genético, é possível identificar qual o montante que deve ser investido em cada um dos ativos, escolhendo uma solução da fronteira não-dominada obtida.

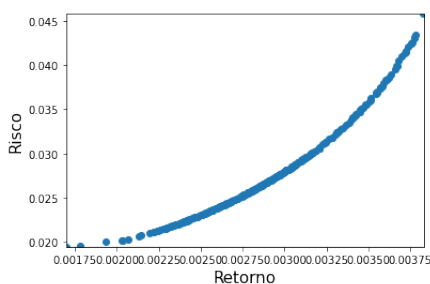


Figura 3: Fronteira não-dominada, 49 ativos

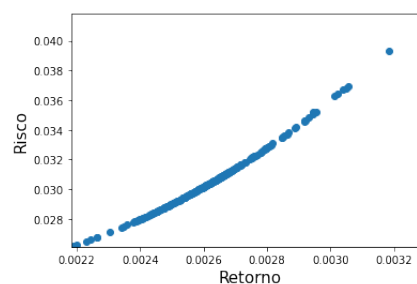


Figura 4: Fronteira não-dominada, eficientes

É possível identificar que a consideração dos 49 ativos da IBRX50 pode gerar soluções com maior retorno, mas também gera soluções com riscos maiores. Outro ponto que deve ser considerado, é que o investimento em uma quantidade menor de ativos é um processo mais realista, uma vez que uma quantidade grande de ativos torna o investimento em alguns ativos quase irrisório.

#### 4. Conclusão

Através da discussão realizada, pode-se concluir que a aplicação da Análise Envolvória de Dados é viável como ferramenta para auxiliar na seleção de uma carteira de investimentos diversi-



ficada e eficiente, uma vez que torna possível a construção de uma carteira com ativos de diferentes setores e com um retorno maior do que o mercado, assim como menor presença de quedas de rentabilidade. Nota-se que o resultado proveniente após filtragem do BCC é fundamental para redução do que chama-se de risco diversificável. Risco diversificável refere-se aquele tipo de risco que afeta um número pequeno de empresas, ou seja, se algo acontecer de errado apenas uma ou poucas empresas serão afetadas [Junior et al., 2002]. Outro ponto importante é que, ao utilizar-se do Algoritmo Genético, a fim de indentificar o montante a ser investido em cada ativo, leva em consideração diferentes cenários de risco e retorno, disponibilizando ao tomador de decisão diferentes soluções alternativas, de acordo com o perfil do investidor.

### Referências

- Alexandrino, F. Q. L., Santos, R. A. B., Meza, L. A., e Rego, R. B. (2017). Utilização da análise envoltória de dados na composição de carteira de investimento diversificada e eficiente. *Produção Online*, p. 8–11.
- Bodie, Z., Kane, A., e Marcus, A. J. (2014). *Fundamentos de Investimentos*. AMGH Editora Ltda, São Paulo - Brasil, 9<sup>a</sup>ed. edition.
- Damodaran, A. (2012). *Investment Valuation: Tools and techniques for Determining the Value to Any Asset*. John Wiley Sons, United States, 5<sup>a</sup>ed. edition.
- de Mello, J. C. C. B. S., Meza, L. A., Gomes, E. G., e Neto, L. B. (2005). Curso de análise de envoltória de dados. XXXVII Simpósio Brasileiro de Pesquisa Operacional.
- Deb, K., Agrawal, S., Pratab, A., e Meyarivan, T. (2000). A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II. In *Proceedings of VI Conference Parallel Problem Solving from Nature*, volume 1917, p. 849–858, Paris, France. Springer-Verlag.
- Dyson, R. G., Allen, R., Camanho, A. S., Podinovski, V. V., Sarrico, C. S., e Shale, E. A. (2001). Pitfalls and protocols in dea. *European Journal of Operational Research*, p. 245–259.
- Fonseca, J. W. F. (2009). *Análise e Decisão de Investimentos*. Parana, Brasil, 1 edition.
- Junior, C. G., Pamplona, E. O., e Montevechi, J. A. B. (2002). Seleção de carteiras através do modelo de markowitz para pequenos investidores (com uso de planilhas eletrônicas). *IX Simpósio de Engenharia de Produção*, p. 3–5.
- Lopes, A. L. M. e Carneiro, M. L. (2010). Markowitz na otimização de carteiras selecionadas por data envelopment analysis - dea. *Gestão e Sociedade*, p. 1–17.
- Markovitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1):77–91.
- OR-Tools, G. (2019). Route. schedule. plan. assign. pack. solve. or-tools is fast and portable software for combinatorial optimization. <https://developers.google.com/optimization/>.
- Ross, S. A., Westerfield, R. W., Jordan, B. D., e Lamb, R. (2013). *Fundamentos da Administração Financeira*. Escola de Administração da UFRGS, AMGH - Editora, São Paulo - Brasil, 9<sup>a</sup>ed. edition.
- Sarker, M. R. (2013). Markowitz portfolio model: Evidence from dhaka stock exchange in bangladesh. *Journal of Business and Management*, p. 1–6.
- SEEB. Bancos têm lucros de r\$ 118,7 bilhões em 2019. <https://bancariose.com.br/conteudo/22130>.
- Silva, P. M. S., Moreira, B. C. M., e Francisco, G. A. (2014). Linear programming applied to finance - building a great portfolio investment. *Gestão, Finanças e Contabilidade*, 4(3):1–4.
- Tauhata, S. Mercado de seguros cresceu entre 11% e 12,5% em 2019, aponta a cnseg. <https://valor.globo.com/financas/noticia/2020/01/09/mercado-de-seguros-cresceu-entre-11-pontos-percentuais-e-125-em-2019-aponta-a-cnseg.ghtml>.