



Algoritmos Genéticos e Evolução Diferencial na Otimização de Carteiras de Investimento: Uma Revisão Teórica

Genetic Algorithms and Differential Evolution in Investment Portfolio Optimization: A Theoretical Review

Tiago Jazynski

Mestre em Economia pela Universidade Estadual de Ponta Grossa (UEPG). Ponta Grossa, Paraná, Brasil.

Clóvis Airon de Quadros

Mestre em Direito Econômico e Social pela Pontifícia Universidade Católica do Paraná. Ponta Grossa, Paraná, Brasil.

Resumo: Este estudo apresenta uma revisão teórica sobre os algoritmos genéticos (AG) e sua aplicação à otimização de carteiras de investimento, com ênfase na evolução diferencial (ED) como técnica de otimização numérica contínua. O estudo tem como objetivo sistematizar os fundamentos dos algoritmos evolutivos e analisar evidências empíricas nacionais sobre sua utilização na composição e otimização de portfólios de ações, considerando sua aplicação em problemas complexos e não lineares. Metodologicamente, caracteriza-se como pesquisa bibliográfica, de abordagem qualitativa e de natureza exploratória-descritiva. A revisão contempla as características dos algoritmos genéticos e da evolução diferencial, bem como sua aplicação em comparação com métodos tradicionais, particularmente o modelo de média-variância de Markowitz. As evidências identificadas na literatura nacional indicam resultados favoráveis à utilização de técnicas evolutivas na otimização de carteiras, com desempenho superior ou equivalente aos métodos tradicionais em diferentes contextos. Conclui-se que os algoritmos evolutivos constituem uma abordagem metodológica relevante para a otimização de portfólios, especialmente em situações caracterizadas por elevada complexidade do espaço de busca e limitações dos métodos convencionais.

Palavras-chave: algoritmo genético; evolução diferencial; otimização de carteiras; computação evolucionária; Markowitz.

Abstract: This study presents a theoretical review of genetic algorithms (GAs) and their application to investment portfolio optimization, with emphasis on differential evolution (DE) as a continuous numerical optimization technique. The study aims to systematize the foundations of evolutionary algorithms and analyze Brazilian empirical evidence on their use in the construction and optimization of equity portfolios, considering their application to complex and nonlinear problems. Methodologically, the study is characterized as a bibliographic research with a qualitative approach and an exploratory-descriptive nature. The review addresses the characteristics of genetic algorithms and differential evolution, as well as their application in comparison with traditional methods, particularly the Markowitz mean-variance model. The evidence identified in the Brazilian literature indicates favorable results for the use of evolutionary techniques in portfolio optimization, with performance superior or equivalent to traditional methods in different contexts. It is concluded that evolutionary algorithms constitute a relevant methodological approach to portfolio optimization, particularly in situations characterized by a highly complex search space and limitations of conventional methods.

Keywords: genetic algorithm; differential evolution; portfolio optimization; evolutionary computation; Markowitz.

INTRODUÇÃO

O crescimento dos mercados financeiros e a ampliação das alternativas de investimento têm demandado modelos matemáticos mais robustos e estruturados, capazes de subsidiar, de forma objetiva, a prospecção, a avaliação e a seleção de ativos (Mota, 2013). Nesse contexto, desde a contribuição seminal de Markowitz (1952), a moderna teoria de carteiras, também conhecida como modelo de média-variância, consolidou-se como uma das principais referências para a composição de portfólios. Seu pressuposto central consiste na determinação dos pesos dos ativos de modo a otimizar a relação entre o retorno esperado e o risco, este último mensurado pela variância dos retornos.

Embora amplamente difundido, o modelo de Markowitz (1952) apresenta limitações teóricas e operacionais importantes. Entre elas, destacam-se a pressuposição de normalidade na distribuição dos retornos, a elevada sensibilidade a erros na estimação das médias e das matrizes de covariância e a possibilidade de que os métodos tradicionais de otimização conduzam a soluções associadas a ótimos locais, em vez de identificarem o ótimo global (Oliveira; Maciel, 2020). Chan, Karceski e Lakonishok (1999) já haviam ressaltado a significativa imprecisão das estimativas de variâncias e covariâncias dos retornos, parâmetros fundamentais para os modelos tradicionais de otimização, como uma importante fonte de fragilidade desses procedimentos. Além disso, Lee e Eid Júnior (2018) observaram que grande parte dos gestores de investimentos no Brasil não utiliza, na prática, as recomendações acadêmicas relacionadas à otimização quantitativa de portfólios, preferindo adotar regras simplificadas, como a definição de limites máximos de concentração por ativo.

Diante dessas limitações, os algoritmos genéticos (AGs) passaram a ser considerados uma alternativa metodológica para a otimização de carteiras. Desenvolvidos com base em princípios da evolução biológica e da seleção natural, esses algoritmos foram sistematizados por Holland (1975) e estruturam a busca por soluções por meio de procedimentos como seleção, cruzamento e mutação. Diferentemente do modelo de Markowitz, os AGs não dependem necessariamente de pressupostos como a normalidade dos retornos, a continuidade ou a diferenciabilidade da função-objetivo. Essa flexibilidade permite sua aplicação a problemas complexos e não lineares, além de favorecer uma exploração mais ampla do espaço de soluções e reduzir a possibilidade de aprisionamento em máximos ou mínimos locais. Assim, os algoritmos genéticos configuram-se como uma abordagem promissora para a identificação de carteiras eficientes em contextos nos quais os métodos tradicionais podem apresentar limitações (Oliveira; Maciel, 2020).

O problema de pesquisa que orienta este estudo pode ser assim formulado: quais são os fundamentos teóricos, a formulação matemática e as evidências

empíricas nacionais que sustentam o uso de algoritmos genéticos e, em particular, da evolução diferencial, como técnica de otimização de carteiras de investimento? A justificativa do estudo reside na relevância de se sistematizar, de forma robusta e autocontida, o instrumental teórico dos algoritmos evolutivos aplicados a finanças, tema que, apesar de crescente interesse acadêmico, ainda carece de textos de revisão que integrem fundamentos, exemplos numéricos e evidência empírica nacional.

O objetivo geral deste estudo é sistematizar, por meio de revisão teórica, os fundamentos conceituais, os operadores e a formulação matemática dos algoritmos genéticos, com destaque para a evolução diferencial, enquanto técnica de otimização de carteiras de investimento, bem como revisar as evidências empíricas nacionais sobre sua aplicação. Como objetivos específicos, busca-se: (i) apresentar os fundamentos conceituais dos algoritmos genéticos e seus operadores de seleção, cruzamento e mutação; (ii) demonstrar, por meio de exemplo numérico, o funcionamento do algoritmo genético binário na otimização de uma função simples e de uma função multimodal; (iii) apresentar a formulação matemática da evolução diferencial e sua aplicação ao problema de otimização de carteiras; e (iv) revisar a evidência empírica nacional sobre o desempenho de carteiras otimizadas por algoritmos genéticos.

Metodologicamente, o estudo caracteriza-se como uma pesquisa bibliográfica, de abordagem qualitativa e natureza exploratória-descritiva, cujos procedimentos são detalhados na seção 3. O texto está organizado em seis seções, além desta introdução: a seção 2 apresenta o referencial teórico, subdividido em fundamentos dos algoritmos genéticos, operadores genéticos, exemplo numérico de otimização binária, algoritmos genéticos em funções multimodais, evolução diferencial e aplicação à otimização de carteiras; a seção 3 descreve os procedimentos metodológicos; a seção 4 revisa a evidência empírica nacional; a seção 5 discute os achados de forma integrada; e a seção 6 apresenta as considerações finais.

REFERENCIAL TEÓRICO

Fundamentos dos Algoritmos Genéticos

A computação evolucionária constitui um subcampo da inteligência artificial voltado ao desenvolvimento de métodos de otimização global. Seus algoritmos fundamentam-se em estratégias estocásticas e meta-heurísticas, nas quais soluções candidatas são geradas, avaliadas e aprimoradas de forma iterativa. Em geral, o processo tem início com uma população aleatória de soluções, que é submetida a mecanismos inspirados na seleção natural. Ao longo das gerações, os indivíduos mais bem avaliados tendem a permanecer e a se reproduzir, conduzindo à evolução gradual da população e à maximização da função de aptidão do problema (Sarmas, 2020).

Segundo Sarmas (2020), os algoritmos evolutivos operam sobre uma população de indivíduos, cada um deles representando uma possível solução dentro do espaço de busca. A partir de uma população inicial, o processo evolutivo direciona as soluções para regiões potencialmente mais promissoras por meio da aplicação combinada de seleção, recombinação e mutação. A seleção favorece os indivíduos que apresentam melhor desempenho, aumentando sua probabilidade de contribuir para as gerações seguintes. A recombinação, por sua vez, possibilita a combinação de características provenientes de diferentes indivíduos, transferindo informações dos progenitores para seus descendentes. Já a mutação introduz alterações aleatórias nas soluções, ampliando a diversidade da população e contribuindo para a exploração de novas regiões do espaço de busca.

Os algoritmos genéticos (AG) são os algoritmos evolutivos mais conhecidos na literatura, tendo seus primeiros estudos iniciados por John Holland, em 1962, a partir do conceito de processos adaptativos de um sistema que recebe entradas sensoriais do ambiente por meio de detectores binários (Sarmas, 2020). Segundo Goldberg (1989), os objetivos centrais da pesquisa original de Holland e seus colaboradores da Universidade de Michigan eram dois: (i) abstrair e explicar rigorosamente os processos adaptativos dos sistemas naturais; e (ii) projetar sistemas artificiais de software capazes de reter os mecanismos relevantes desses sistemas naturais.

Ainda segundo Goldberg (1989), os algoritmos genéticos são algoritmos de busca baseados no processo da mecânica da seleção natural e da genética natural, que combinam a sobrevivência do mais apto entre estruturas do tipo string com uma troca de informações estruturada, ainda que aleatória. Em cada nova geração, um novo conjunto de soluções é criado a partir das mais aptas da geração anterior; embora envolvam componentes aleatórios, os AG não configuram um passeio aleatório simples, pois exploram com eficiência informações históricas para investigar novos pontos de busca com desempenho esperado ainda melhor.

Dias (2008) define o algoritmo genético como um método baseado na teoria da evolução natural de Darwin, segundo a qual os organismos de uma população que melhor se adaptam ao meio conseguem se reproduzir e possuem mais chances de sobrevivência, enquanto os menos adaptados são eliminados; a combinação de boas características entre os indivíduos produz descendentes cada vez mais bem adaptados, que geram resultados cada vez melhores. Seita (2014) complementa que se trata de uma técnica heurística baseada na biologia evolutiva, cuja lógica de sobrevivência do mais apto permite a convergência iterativa do problema de otimização para um conjunto de soluções ótimas.

Em termos de linguagem de otimização, cada indivíduo da população corresponde a uma solução codificada, formada por uma série de genes chamada de cromossomo, e cada cromossomo representa uma possível solução para o problema. O algoritmo genético, portanto, pode ser utilizado como ferramenta para buscar, em um problema de otimização, o ótimo global dos pesos de um portfólio por meio de programação, exigindo em primeiro lugar a definição de uma função de adequação (fitness) que represente a referência do modelo (Dias, 2008).

Tayler (1995) observa que praticamente qualquer problema de otimização de funções matemáticas pode ter sua solução encontrada por meio de algoritmo genético, ressaltando que, para funções simples com uma única solução ótima, não há vantagem relevante em utilizá-lo; contudo, para funções mais complexas e com múltiplas variáveis, especialmente aquelas que envolvem interações não lineares, o AG revela-se ferramenta poderosa para localizar pontos de máximo ou mínimo globais. Essa vantagem decorre do fato de que os algoritmos genéticos não são fundamentalmente limitados por suposições restritivas sobre o espaço de busca, como a continuidade da função ou a existência de derivadas (Goldberg, 1989).

Não existe, na literatura de computação evolucionária, uma definição única e rigorosamente consensual para o conceito de algoritmo genético que permita distingui-lo de forma absoluta dos demais métodos evolutivos. Entretanto, os algoritmos geralmente classificados nessa categoria compartilham um conjunto de características fundamentais: trabalham com populações de cromossomos, utilizam uma função de aptidão para orientar a seleção dos indivíduos, aplicam operadores de cruzamento para gerar novos descendentes e recorrem à mutação para introduzir variações na descendência (Mitchell, 1998).

Operadores Genéticos: Seleção, Cruzamento e Mutação

Para Mitchell (1998), as formas mais simples de algoritmo genético apresentam três operadores centrais: seleção, cruzamento (crossover de ponto único) e mutação. O operador de seleção escolhe cromossomos da população para reprodução, de modo que, quanto mais apto for o cromossomo, maior a frequência com que é selecionado para se reproduzir. A probabilidade de seleção de um cromossomo i é dada pela razão entre sua aptidão e a soma das aptidões de toda a população:

$$p_i = f(x_i) / \sum_k f(x_k), \quad k = 1, \dots, N$$

O operador de crossover escolhe aleatoriamente uma posição (locus) e troca as subsequências anteriores e posteriores a esse ponto entre dois cromossomos, criando dois descendentes; esse processo imita, de forma aproximada, a recombinação biológica entre dois organismos de cromossomo único (Mitchell, 1998). Já o operador de mutação transforma aleatoriamente um ou mais bits (genes) de um cromossomo, podendo ocorrer em cada posição com uma probabilidade geralmente muito pequena, por exemplo, 0,001, o suficiente para assegurar a diversidade de cromossomos na população, sem ser tão elevada a ponto de descaracterizar o processo de seleção (Mitchell, 1998).

Exemplo Numérico de Otimização por Algoritmo Genético Binário

A lógica de funcionamento do algoritmo genético pode ser ilustrada por um exemplo simples de maximização da função $f(x) = x^2$, sujeita à restrição $0 \leq x \leq 31$. O primeiro passo consiste em definir uma população inicial escolhida aleatoriamente, codificada em números binários, conforme a Tabela 1.

Tabela 1 - População inicial de cromossomos.

Cromossomo	x	f(x)	Probabilidade de seleção
A1 = 1 1 0 0 1	25	625	54,5%
A2 = 0 1 1 1 1	15	225	19,6%
A3 = 0 1 1 1 0	14	196	17,1%
A4 = 0 1 0 1 0	10	100	8,7%

Fonte: elaborada pelo autor a partir da dissertação de origem.

Cada cromossomo representa o número binário correspondente ao valor de x, por exemplo, o número binário de 25 corresponde a 1 1 0 0 1. A probabilidade de seleção de cada indivíduo, calculada pela equação de pi apresentada na seção 2.2, favorece os cromossomos A1 e A2, os de maior aptidão. Selecionados esses dois indivíduos como pais, aplica-se o operador de crossover:

Pais: 1 1 0 0 1 e 0 1 1 1 1 → Filhos: 1 1 0 1 1 e 0 1 1 0 1

Quando o crossover é aplicado, os pais trocam suas “caudas” a partir do ponto de corte sorteado, gerando dois filhos; caso o operador não fosse aplicado, os filhos seriam cópias exatas dos pais. Em seguida, aplica-se a mutação, que inverte os bits dos filhos: considerando o filho 0 1 1 0 1, ao sofrer mutação em sua segunda posição, ele se torna 0 0 1 0 1. A partir dessa primeira geração de filhos, obtêm-se os resultados apresentados na Tabela 2.

Tabela 2 - Primeira geração de filhos.

Cromossomo (1ª geração)	x	f(x)	Probabilidade de seleção
1 1 0 1 1	27	729	29,1%
1 1 0 0 1	25	625	24,9%
1 1 0 0 1	25	625	24,9%
1 0 1 1 1	23	529	21,1%

Fonte: elaborada pelo autor a partir da dissertação de origem.

Procedendo de forma recursiva, repetindo seleção, crossover e mutação a cada geração, o algoritmo produz gerações sucessivas de cromossomos com aptidão crescente, conforme sintetizado na Tabela 3.

Tabela 3 - Segunda a quinta gerações de filhos.

Geração	Cromossomos	x	f(x)
2ª geração	11011/11000/10111/10101	27 / 24 / 23 / 21	729 / 576 / 529 / 441
3ª geração	11011/10111/01111/00111	27 / 23 / 15 / 7	729 / 529 / 225 / 49
4ª geração	11111/11011/10111/10111	31 / 27 / 23 / 23	961 / 729 / 529 / 529
5ª geração	11111/11111/11111/10111	31 / 31 / 31 / 23	961 / 961 / 961 / 529

Fonte: elaborada pelo autor a partir da dissertação de origem.

Na quinta geração, os resultados convergem para x = 31, representado pelo cromossomo 1 1 1 1 1, exatamente o ponto de máximo da função-objetivo dentro da restrição $0 \leq x \leq 31$. O exemplo ilustra, de forma didática, como o algoritmo genético

converge para o ótimo global por meio da repetição dos operadores de seleção, crossover e mutação, sem necessidade de cálculo de derivadas.

Algoritmos Genéticos em Funções Multimodais

Em problemas cuja função-objetivo apresenta estrutura simples, os métodos tradicionais de otimização baseados no cálculo diferencial são, em geral, capazes de identificar pontos de máximo ou mínimo. Entretanto, quando a função possui maior complexidade e apresenta múltiplos máximos e mínimos locais, os algoritmos genéticos podem oferecer maior flexibilidade, pois realizam uma busca mais ampla no espaço de soluções e reduzem a probabilidade de convergência prematura para um ótimo local. Nesse sentido, a função seno, apresentada a seguir, constitui um exemplo de função para a qual a aplicação de um algoritmo genético pode ser particularmente adequada:

$$f(x) = x \cdot \sin(10\pi x) + 1$$

Essa função apresenta diversos pontos de máximo e mínimo locais dentro do domínio entre -1,00 e 2,00, mas um único ponto de máximo global. Em funções desse tipo, descontínuas, ruidosas ou com múltiplos ótimos locais, os métodos tradicionais de otimização tendem a ficar presos no primeiro ótimo local encontrado, ao passo que o algoritmo genético, por conduzir simultaneamente a busca a partir de múltiplos pontos da população, tem maior probabilidade de encontrar o ótimo global (Holland, 1992).

Segundo Eltaieb e Mahmood (2018), os algoritmos evolutivos constituem abordagens modernas e relevantes para a resolução de problemas de otimização complexos. Sua aplicação é particularmente adequada em situações caracterizadas por espaços de busca amplos, funções-objetivo dinâmicas ou instáveis e restrições que dificultam a identificação de soluções viáveis.

Evolução Diferencial: Fundamentos e Formulação Matemática

A evolução diferencial (ED) é um algoritmo estocástico baseado em população, utilizado para resolver problemas de otimização numérica contínua. Desenvolvida por Kenneth Price em 1994, a ED tornou-se um poderoso instrumento de otimização global, capaz de convergir para um ponto ótimo real sem demandar grande volume de recursos computacionais; seu desempenho foi validado no domínio evolucionário pela *IEEE Conference on Evolutionary Computation* (CEC), em 1996 (Eltaieb; Mahmood, 2018).

Diferentemente de outros algoritmos evolutivos, que geralmente geram novos indivíduos a partir da combinação de soluções por meio de operadores de cruzamento, a evolução diferencial (ED) utiliza um mecanismo específico de recombinação baseado em diferenças vetoriais entre indivíduos selecionados da população. A mutação diferencial constitui o principal elemento distintivo da ED, pois orienta a geração de novos candidatos a partir das relações entre soluções existentes, ampliando a capacidade de exploração do espaço de busca e favorecendo a identificação de descendentes promissores.

O desempenho da evolução diferencial depende, essencialmente, de dois aspectos: da estratégia de mutação adotada e dos parâmetros de controle utilizados no processo evolutivo. Entre esses parâmetros, destacam-se o tamanho da população (NP), a taxa de cruzamento (Cr) e o fator de escala (F), responsáveis por influenciar, respectivamente, a diversidade populacional, a probabilidade de recombinação e a intensidade das variações aplicadas aos indivíduos (Eltaeib; Mahmood, 2018).

Formalmente, busca-se um ótimo X^* representado pelo vetor X_i , $i = 1, \dots, D$, $X \in \mathbb{R}^D$, dentro das restrições de fronteira $L \leq X \leq H$. A população inicial, representada por $P_{i,j} \in \mathbb{R}[D \times NP]$, é inicializada aleatoriamente dentro dos limites superior (H) e inferior (L), uma vez que não há informação prévia sobre o ótimo:

$$P_{i,j} = L + (H - L) \cdot \text{rand}_{i,j}[0,1], \quad i = 1, 2, \dots; \quad j = 1, 2, \dots, NP$$

A evolução ocorre em três processos: mutação, cruzamento e seleção. Na estratégia tradicional, selecionam-se aleatoriamente três vetores distintos da população, v_1 , v_2 e v_3 (com $v_1 \neq v_2 \neq v_3$), a partir dos quais se constrói, por recombinação, o vetor de mutação v_y :

$$v_y = v_3 + F \cdot (v_1 - v_2)$$

Em que F é o fator de escala da mutação, parâmetro de controle predefinido. Segundo Eltaeib e Mahmood (2018), a mutação é o principal componente da ED, existindo diferentes estratégias de mutação, cada uma identificada por uma notação baseada na estratégia clássica, por exemplo, DE/rand/1/bin, em que “DE” representa evolução diferencial, “rand” indica que os vetores são selecionados aleatoriamente, o número 1 indica a quantidade de pares de diferença utilizados (no caso, o par $v_1 - v_2$) e o termo final indica o tipo de crossover, exponencial (“exp”) ou binomial (“bin”).

Para complementar a etapa de mutação, a ED aplica um crossover uniforme para construir vetores de teste (vtrail), combinando parâmetros copiados de dois vetores distintos: um vetor-alvo v_x , selecionado aleatoriamente da população e diferente de v_y , e o próprio vetor mutante v_y . O cruzamento binomial é gerado da seguinte forma:

$$\text{vtrail}[i,j] = v_y[i,j] \text{ se } \text{rand}(0,1) < Cr; \quad v_x[i,j] \text{ caso contrário}$$

Em que a probabilidade de cruzamento $Cr \in (0,1)$ é uma taxa predefinida que especifica a fração do parâmetro transferido do indivíduo mutante, controlando quais fontes participam de um determinado parâmetro. A taxa de cruzamento é comparada a valores aleatórios uniformes formados a partir de $(0,1)$: se o valor aleatório for menor ou igual a Cr , o parâmetro de teste é copiado do vetor mutante v_y ; caso contrário, é herdado do vetor-alvo v_x (Eltaeib; Mahmood, 2018).

Na etapa final, de seleção, o vetor de teste (vtrail) compete com o vetor-alvo (v_x): se o vetor de teste apresentar aptidão igual ou superior à do vetor-alvo, ele substitui esse vetor na geração seguinte; caso contrário, o vetor-alvo permanece inalterado na população. Esse processo pode ser expresso, para um problema de minimização, por:

$$V_x[i,j] = \text{vtrail} \text{ se } f(\text{vtrail}) < f(v_x); \quad v_x \text{ caso contrário}$$

Em que $f(x)$ é a função-objetivo. As diferentes fases da ED têm como propósito evitar que a população se deteriore ao longo do tempo, de modo que ela permaneça estável ou melhore continuamente a cada iteração, um refinamento contínuo alcançado justamente pelo vetor de teste, fator crucial no funcionamento da ED (Eltaeib; Mahmood, 2018).

Apesar de seu grande potencial, a literatura aponta desafios relevantes ao desempenho da ED, como a estagnação, quando a população não melhora ao longo das iterações e o algoritmo não consegue encontrar um novo domínio de busca, frequentemente em razão de parâmetros de controle que se tornam ineficientes para um problema específico, a convergência prematura e a sensibilidade aos parâmetros de controle (Eltaeib; Mahmood, 2018).

Aplicação de Algoritmos Evolutivos à Otimização de Carteiras de Investimento

Na aplicação da evolução diferencial à otimização de carteiras, o problema assume uma formulação particular: diferentemente do modelo de Markowitz, que resulta em um conjunto de soluções eficientes de Pareto (a fronteira eficiente), a ED aqui empregada busca uma solução única, ao maximizar a proporção de períodos em que a carteira supera um índice de mercado. Sendo m o índice de mercado, rim o retorno do índice de mercado no período i e rij o retorno do título j no período i , o retorno da carteira p no período i é dado por:

$$rip = \sum_{j=1..m} w_j \cdot rij$$

Em que w_j é a proporção do título j na carteira p . Diz-se que a carteira construída supera o índice de mercado no período i se a seguinte inequação for satisfeita:

$$rip \geq rim$$

Segundo Sarmas (2020), nesse problema o algoritmo genético toma como entrada os dados históricos para T períodos e busca maximizar a porcentagem de casos em que a carteira construída supera o índice de mercado, o que pode ser expresso como:

$$\max(w_i) \sum_{i=1..T} b_i / T$$

$$\text{sujeito a: } \sum_{i=1..N} w_i = 1 \text{ e } w_i \geq 0, \forall i = 1, 2, \dots, N$$

Em que b_i é uma variável binária que assume o valor 1 se $rip \geq rim$ no período i , e 0 caso contrário. A evolução diferencial, ao ser um método estocástico baseado em população usado para encontrar o mínimo (ou máximo) global de uma função multivariada, opera modificando cada solução candidata a cada iteração, misturando-a com outras soluções candidatas para criar uma candidata de teste, conforme descrito na seção 2.5 (Sarmas, 2020).

Essa formulação evidencia a principal vantagem prática dos algoritmos evolutivos sobre o modelo de Markowitz no contexto da otimização de carteiras: ao não exigir a estimação de uma matriz de covariâncias entre todos os pares de ativos, cujo número de parâmetros cresce rapidamente com o número de ativos,

e cuja estimativa está sujeita a erros substanciais (Chan; Karceski; Lakonishok, 1999), a ED contorna diretamente uma das principais fontes de fragilidade prática do modelo tradicional, ao custo de perder a interpretação em termos de fronteira eficiente de risco-retorno.

PROCEDIMENTOS METODOLÓGICOS

Quanto à sua natureza, esta pesquisa caracteriza-se como básica, de finalidade teórica; quanto à abordagem, como qualitativa; e quanto aos objetivos, como exploratória-descritiva. O estudo tem por finalidade sistematizar, integrar e analisar criticamente conceitos, formulações matemáticas e evidências empíricas previamente desenvolvidos pelo autor em pesquisa de mestrado sobre a temática, não envolvendo a produção ou coleta de dados primários adicionais (Gil, 2008). Quanto aos procedimentos técnicos, caracteriza-se como pesquisa bibliográfica, fundamentada exclusivamente no referencial teórico e no material empírico levantados, analisados e sistematizados na dissertação de mestrado que deu origem ao presente estudo, constituindo esta a base documental e analítica para o desenvolvimento do estudo (Jazynski, 2022).

O levantamento privilegiou três eixos temáticos: (i) os fundamentos conceituais e os operadores dos algoritmos genéticos, com base, sobretudo, em Holland (1975, 1992), Goldberg (1989), Mitchell (1998), Dias (2008), Seita (2014) e Sarmas (2020); (ii) a formulação matemática da evolução diferencial, com base em Eltaieb e Mahmood (2018); e (iii) a evidência empírica nacional sobre a aplicação de algoritmos genéticos à otimização de carteiras, com base em Parreiras (2003), Mota (2013), Cartacho e Souza (2002) e Oliveira e Maciel (2020), entre outros estudos nacionais revisados na dissertação de origem.

O tratamento do material foi desenvolvido em três etapas. Na primeira, realizaram-se o fichamento e a sistematização dos conceitos, operadores e formulações matemáticas relacionados aos algoritmos genéticos e à evolução diferencial. Na segunda, o conteúdo foi organizado em uma sequência lógica e progressiva, partindo dos fundamentos gerais, passando pela descrição dos operadores genéticos e pela apresentação de exemplos numéricos de caráter didático, até alcançar a formulação específica da evolução diferencial aplicada à otimização de carteiras. Na terceira etapa, procedeu-se à revisão cruzada das evidências empíricas identificadas na literatura nacional, buscando estabelecer relações entre os fundamentos teóricos apresentados e as aplicações práticas documentadas em estudos brasileiros sobre otimização de carteiras por meio de algoritmos evolutivos.

Como limitação metodológica, destaca-se que o presente estudo não incorpora novas fontes bibliográficas além daquelas originalmente reunidas e analisadas na dissertação de mestrado que lhe deu origem. Essa delimitação decorre da opção metodológica de manter a análise circunscrita ao material previamente selecionado e sistematizado, concentrando-se na sua reorganização, aprofundamento e

articulação teórica. Dessa forma, o estudo não se propõe a constituir uma revisão sistemática e exaustiva da literatura internacional mais recente sobre algoritmos genéticos e evolução diferencial, mas uma sistematização e aprofundamento de conhecimentos e evidências anteriormente levantados e analisados.

REVISÃO EMPÍRICA: EVIDÊNCIAS SOBRE O USO DE ALGORITMOS GENÉTICOS NA OTIMIZAÇÃO DE CARTEIRAS

A literatura nacional oferece um conjunto relevante de estudos aplicados sobre o uso de algoritmos genéticos na otimização de carteiras de ações. Oliveira e Maciel (2020) avaliaram o desempenho de carteiras otimizadas por algoritmo genético no mercado acionário brasileiro, comparando os resultados a benchmarks como o índice Ibovespa e o CDI, com base em diferentes metodologias de construção de carteira, variância mínima, máxima razão de Sharpe e diversificação ingênua. Os autores constataram que a carteira de variância mínima de Markowitz superou tanto a carteira de máximo Sharpe quanto o Ibovespa em indicadores como retorno cumulativo, índice de Modigliani e volatilidade anualizada; mais relevante para este estudo, concluíram que as carteiras otimizadas por AG apresentaram, no geral, desempenho superior às otimizadas pelos métodos tradicionais, tratando-se de alternativa viável e capaz de gerar ganhos econômicos ao investidor, justamente por não se limitar a máximos e mínimos locais.

Mota (2013) destaca que, embora o modelo de média-variância de Markowitz permaneça o mais conhecido e amplamente utilizado tanto no mercado quanto na academia, seu principal desafio de aplicabilidade está na dificuldade de estimação dos parâmetros do modelo, uma vez que essa estimação se fundamenta em séries históricas de retornos que podem retratar apenas uma estimativa grosseira do comportamento futuro dos ativos; além disso, as suposições do modelo quanto ao formato da distribuição dos retornos, ainda que simplifiquem a análise, também podem introduzir distorções relevantes nos resultados.

Parreiras (2003), por sua vez, destaca que o algoritmo genético é eficiente para a previsão de preços no mercado de ações, inclusive em contextos nos quais os dados apresentam alta taxa de ruído, uma vantagem relevante frente aos métodos tradicionais de otimização, sensíveis a esse tipo de imperfeição nos dados. O autor acrescenta que o uso do AG para composição de carteiras oferece benefícios adicionais, dada sua capacidade de analisar com rapidez e eficiência o espaço de soluções, o que se traduz em ganhos relevantes de rentabilidade.

Já Cartacho e Souza (2002), em estudo apresentado ao Encontro Nacional dos Programas de Pós-Graduação em Administração, propuseram um modelo composto por algoritmos genéticos e redes neurais para o processo de seleção de carteiras, explorando a complementaridade entre a capacidade de otimização do AG e a capacidade de previsão das redes neurais artificiais, evidência adicional de que os algoritmos genéticos vêm sendo incorporados à literatura brasileira de finanças não apenas de forma isolada, mas também combinados a outras técnicas de inteligência computacional.

Em conjunto, esses estudos convergem para uma conclusão comum: a de que os algoritmos genéticos representam alternativa metodologicamente robusta à otimização tradicional de carteiras, com desempenho empírico favorável no contexto brasileiro, ainda que, como destaca Mota (2013), o modelo de Markowitz continue sendo o mais difundido na prática de mercado, uma lacuna entre a evidência acadêmica e a adoção efetiva pelos gestores de investimento que já havia sido apontada por Lee e Eid Júnior (2018).

DISCUSSÃO

A leitura conjunta dos fundamentos teóricos e da evidência empírica revisada permite identificar três pontos centrais de discussão. O primeiro refere-se à vantagem estrutural dos algoritmos genéticos frente aos métodos tradicionais de otimização: por não dependerem de pressupostos restritivos sobre a continuidade da função-objetivo ou a existência de derivadas (Goldberg, 1989), os AG conseguem explorar todo o espaço de soluções em busca do ótimo global, o que os torna particularmente adequados a problemas com múltiplos ótimos locais, como ilustrado pelo exemplo da função seno na seção 2.4, situação frequente em problemas reais de otimização de carteiras, marcados por funções descontínuas e ruidosas.

O segundo ponto diz respeito à formulação específica da evolução diferencial aplicada à otimização de carteiras, que, ao maximizar diretamente a proporção de períodos em que a carteira supera o índice de mercado, em vez de minimizar a variância a partir de uma matriz de covariâncias estimada, contorna uma das principais fontes de fragilidade prática do modelo de Markowitz apontada por Chan, Karceski e Lakonishok (1999): a imprecisão das estimativas de variância e covariância. Essa característica sugere que a ED pode ser particularmente vantajosa em contextos de amostras curtas ou de maior instabilidade nas relações entre os ativos, situação na qual a matriz de covariância amostral tende a apresentar maior erro de estimação.

O terceiro aspecto diz respeito à convergência das evidências empíricas nacionais analisadas por Oliveira e Maciel (2020), Parreiras (2003) e Cartacho e Souza (2002), quanto ao desempenho favorável de carteiras otimizadas por algoritmos genéticos em comparação com métodos tradicionais. Os resultados indicam vantagens tanto em termos de retorno quanto de robustez diante de dados ruidosos. Esse conjunto de evidências sugere que a capacidade teórica dos algoritmos genéticos de explorar o espaço de soluções e reduzir a probabilidade de convergência para ótimos locais pode se refletir em ganhos econômicos mensuráveis no mercado acionário brasileiro. Entretanto, conforme observado por Mota (2013), o modelo de Markowitz ainda permanece como uma referência predominante na prática de mercado. Essa diferença entre os resultados da pesquisa acadêmica e as práticas de gestão de portfólios, também identificada por Lee e Eid Júnior (2018), evidencia uma oportunidade relevante para aproximar o conhecimento científico da tomada de decisões no âmbito dos investimentos.

CONSIDERAÇÕES FINAIS

Este estudo teve como objetivo sistematizar, por meio de uma revisão teórica, os fundamentos conceituais, os operadores e a formulação matemática dos algoritmos genéticos, com ênfase na evolução diferencial, enquanto técnica aplicada à otimização de carteiras de investimento. Também se buscou revisar as evidências empíricas nacionais acerca da utilização desses métodos. Os objetivos específicos foram alcançados. Inicialmente, foram apresentados os fundamentos conceituais dos algoritmos genéticos e descritos os operadores de seleção, cruzamento e mutação. Em seguida, o funcionamento de um algoritmo genético binário foi demonstrado, por meio de exemplos numéricos, tanto em uma função simples quanto em uma função multimodal. Também foi apresentada a formulação matemática da evolução diferencial, acompanhada de sua aplicação ao problema de otimização de carteiras. Por fim, foram analisadas as evidências empíricas nacionais relacionadas ao desempenho de carteiras otimizadas por algoritmos genéticos.

A principal contribuição deste estudo consiste na sistematização do instrumental teórico relacionado à aplicação dos algoritmos genéticos à otimização de carteiras. O estudo reúne, em um único texto, os fundamentos conceituais desses métodos, a demonstração didática de seu funcionamento, a formulação matemática da evolução diferencial e as evidências empíricas produzidas no contexto nacional. Essa organização é relevante porque, na literatura brasileira de finanças, tais conteúdos frequentemente se encontram distribuídos em diferentes trabalhos. Dessa forma, o estudo pode servir como material de referência para pesquisadores interessados na aplicação de algoritmos evolutivos a problemas de otimização financeira, bem como para investidores que desejem compreender os fundamentos, as possibilidades e as limitações dessas técnicas antes de utilizá-las.

Entre as limitações da pesquisa, destaca-se a opção metodológica de não incorporar estudos internacionais adicionais além daqueles já contemplados na dissertação de mestrado que originou este estudo. Essa delimitação restringe o alcance da revisão empírica, concentrando-a em estudos nacionais e em um conjunto previamente definido de referências teóricas. Embora essa escolha tenha contribuído para preservar o foco do trabalho, ela também limita a possibilidade de comparação com aplicações desenvolvidas em outros mercados e contextos institucionais.

Para pesquisas futuras, recomenda-se a ampliação da revisão bibliográfica por meio da inclusão de estudos internacionais recentes sobre diferentes variantes da evolução diferencial e de outros algoritmos evolutivos, como algoritmos genéticos multiobjetivo, otimização por enxame de partículas e otimização por colônia de formigas. Sugere-se, igualmente, a realização de estudos empíricos que comparem diretamente o desempenho da evolução diferencial e do algoritmo genético binário clássico na otimização de carteiras compostas por ativos do mercado brasileiro. Essas investigações poderão contribuir para avaliar, de maneira mais sistemática, a

eficiência, a estabilidade e a aplicabilidade dessas técnicas em diferentes condições de mercado e sob distintas restrições de composição de portfólio.

REFERÊNCIAS

- CARTACHO, Marcelo Soares; SOUZA, Antônio Artur de. **A utilização de um modelo composto por algoritmos genéticos e redes neurais no processo de seleção de carteiras**. In: Encontro Nacional Dos Programas De Pós-Graduação Em Administração, 2002, Salvador. Anais [...]. Salvador: ANPAD, 2002.
- CHAN, Louis K. C.; KARCESKI, Jason; LAKONISHOK, Josef. On portfolio optimization: forecasting covariances and choosing the risk model. **The Review of Financial Studies**, v. 12, n. 5, p. 937-974, 1999.
- Días, Carlos Henrique. **Um novo algoritmo genético para a otimização de carteiras de investimento com restrições de cardinalidade**. 2008. Dissertação (Mestrado) - Campinas, 2008.
- ELTAEIB, Tarik; MAHMOOD, Ausif. **Differential evolution: a survey and analysis**. Applied Sciences, v. 8, n. 10, p. 1945, 2018.
- GIL, Antonio Carlos. **Métodos e técnicas de pesquisa social**. 6. ed. São Paulo: Atlas, 2008.
- GOLDBERG, David E. **Genetic algorithms in search, optimization, and machine learning**. Reading: Addison-Wesley, 1989.
- HOLLAND, John H. **Adaptation in natural and artificial systems**. Ann Arbor: University of Michigan Press, 1975.
- JAZYNSKI, Tiago. **Seleção e otimização de carteiras de ações do mercado brasileiro: momentum, Markowitz e algoritmo genético**. 2022. 59 f. Dissertação (Mestrado em Economia) — Universidade Estadual de Ponta Grossa, Ponta Grossa, 2022.
- LEE, Stefan Colza; EID JÚNIOR, William. **Portfolio construction and risk management: theory versus practice**. RAUSP Management Journal, v. 53, n. 3, p. 345-365, 2018.
- MARKOWITZ, Harry. Portfolio selection. **The Journal of Finance**, v. 7, n. 1, p. 77-91, 1952.
- MITCHELL, Melanie. **An introduction to genetic algorithms**. Cambridge: MIT Press, 1998.
- MOTA, Lucas da Silveira. **Otimização de carteiras de investimento através da aplicação de algoritmos genéticos multiobjetivo**. 2013. Dissertação (Mestrado) - Universidade Federal de Juiz de Fora, Juiz de Fora, 2013.

OLIVEIRA, Guilherme Isaac Giorgete; MACIEL, Leandro dos Santos. **Algoritmos genéticos na otimização de carteiras de ações no Brasil**. [S.l.: s.n.], 2020.

PARREIRAS, Luiz Paulo Rodrigues de Freitas. **Modelo genético-neural de gestão de carteiras de ações**. São Paulo, 2003.

SARMAS, Elissaio. **Multicriteria portfolio construction with Python**. Cham: Springer, 2020.

SEITA, Marcelo R. **Simulação multiagente em mercados financeiros artificiais utilizando algoritmos genéticos**. 2014. Dissertação - 2014.

TAYLER, Paul. **Modeling artificial stock markets using genetic algorithms**. In: GOONATILAKE, Suran; TREELEAVEN, Philip (ed.). *Intelligent systems for finance and business*. Chichester: John Wiley & Sons, 1995. p. 271- 288.