Otimização de uma Carteira de Investimentos com Evolução Diferencial

Bruno Borges de Souza

Palavras-chave

CVAR

Evolução Diferencial

Markowitz

Otimização de Portfólio

VAR

As decisões financeiras na prática não são tomadas em ambiente de total certeza com relação a seus resultados. Por essas decisões estarem fundamentalmente voltadas para o futuro, é imprescindível que se introduza a variável incerteza (risco) com um dos mais significativos aspectos do estudo das finanças corporativas. O objetivo deste trabalho é a otimização de um portfólio utilizando como medida de risco o VAR condicional. Para isso, será utilizado uma técnica evolucionária conhecida como Evolução Diferencial, que otimiza um problema imitando a seleção natural, melhorando iterativamente a solução candidata. Os resultados obtidos utilizando o algoritmo de Evolução Diferencial se assemelham aos resultados utilizando as técnicas clássicas de programação linear.

INTRODUÇÃO

Até a década de 50, os modelos para risco e retorno eram em grande parte subjetivos e variavam de investidor para investidor. A partir do desenvolvimento da moderna teoria do portfólio por pesquisadores como Markowitz, Sharp, Black, Scholes entres outros, foi desenvolvido um arcabouço teórico quantitativo com base na teoria econômica neoclássica. O desenvolvimento da teoria do portfólio em 1952 por Markowitz, revolucionou o estudo das finanças, fato que lhe rendeu o prêmio Nobel de economia em 1990.

O principal aspecto da teoria do portfólio é que o risco individual de um ativo é diferente de seu risco na carteira, tornando a diversificação capaz de minimizar o risco não-sistemático dos ativos em conjunto. Utilizando métodos de programação não-linear, é possível escolher a proporção ideal de cada ativo no portfólio, otimizando a relação retorno/risco da carteira de títulos.

Kato (2004), Alexander (2008) e Roman e Mitra (2009) citam duas categorias para mensuração do risco financeiro. A primeira é caracterizada pelas medidas que consideram a dispersão em relação a um retorno-alvo e podem somente assumir valores positivos; nesse caso, a categoria é dividida em dois grupos: medidas simétricas e assimétricas. A segunda categoria é formada pelas medidas de risco baseadas em quantis.

Na primeira categoria, o grupo das medidas simétricas é representado pela variância (ou desvio padrão) e pelo desvio médio absoluto. Como retorno-alvo utiliza-se o valor esperado. Nesse caso, os riscos são considerados como dispersões acima ou abaixo das expectativas. No segundo grupo, o risco é mensurado somente em relação aos desvios abaixo do retorno-alvo (*downside risk*), podendo ser o valor esperado ou algum *benchmarking*. Os principais representantes do grupo são a semivariância e o conjunto de medidas baseadas no momento parcial inferior da distribuição de probabilidade dos retornos (*Lower Partial Moment* [LPM]).

A segunda categoria é formada pelas medidas de risco baseadas em quantis (percentis). Nesse campo, são classificadas as medidas Value-at-Risk (VaR) e Conditional Value-at-Risk (CVaR). A dimensão é caracterizada pelo foco no chamado *tail risk measures*, ou seja, o objetivo é estimar o risco a partir da área da cauda da distribuição de probabilidade dos retornos do ativo, dado certo nível de confiança (1%, 5% ou 10%).

DESENVOLVIMENTO

Medidas de Risco

Dado algum instrumento de referência, existe uma maneira natural de definir uma medida de risco descrevendo quão próxima ou distante uma posição é da aceitação pelo regulador (Artzner et al. [3]).

Definição 1 Medida de Risco *Seja Ω o conjunto de estados da natureza e assuma que é finito. Seja também* ***X*** *o conjunto de todos os riscos, ou seja, o conjunto de todas as funções de valor real x ∈* ***X****, que representam o patrimônio líquido final de um instrumento, ou de uma carteira de cada elemento de Ω. Uma medida de risco ρ (x) é uma aplicação de* ***X*** *para* $\mathbb{R}$.

Quando positivo, o número *ρ(x)* atribuído pela medida $\rho: \boldsymbol{X} \rightarrow \mathbb{R}$ ao risco x será interpretado como a quantidade de capital que um agente deve adicionar à posição de risco x para torná-lo uma posição aceitável. Pelo contrário, se *ρ (x)<0*, a quantia em dinheiro −*ρ (x)* pode ser retirada da posição já aceitável e investida de uma forma mais lucrativa.

Esta noção permite escolher um membro adequado do conjunto **X** de variáveis aleatórias admissíveis. Se preferimos valores menores de x ∈ **X**, usamos o seguinte problema de otimização para fazer a seleção:

$$\min_{X \in \mathbf{X}} \rho(\mathrm{x}) \tag{1}$$

A noção de coerência foi introduzida por Artzner et al. [3] e atualmente, é um conceito fundamental relacionado à aceitabilidade de uma medida de risco. A literatura introduz um número de propriedades que são usadas para determinar um medida de risco. O seguinte descreve as propriedades mais importantes para a medida de risco $\rho: \mathbf{X} \rightarrow \mathbb{R}$:

(P1) Invariância à Translação: $\rho(X+c)=\rho(X)+c$ para qualquer X ∈ **X** e $c \in \mathbb{R}$;

(P2) Subaditividade: $\rho(X_1+X_2) \leq \rho(X_1)+\rho(X_2)$ para todo $X_1, X_2 \in \mathbf{X}$;

(P3) Monotonicidade: Se $X_1, X_2 \in \mathbf{X}$ e $X_1 \leq X_2$, então $\rho(X_1) \leq p(X_2)$;

(P4) Homogeneidade Positiva: $\rho(\lambda X)=\lambda \rho(X)$ para todo X ∈ **X** e $\lambda \geq 0$.

Estudos anteriores propuseram várias medidas de risco que foram usadas por pesquisadores. As expectativas (ρ_E) e as medidas de risco de pior caso ρ_w são exemplos de duas medidas de riscos elementares coerentes:

$$\rho_E(X):=\mathrm{E}(X),\ \rho_w(X):=ess \sup(X) \tag{2}$$

Volatilidade*, VAR e *VAR condicional

A volatilidade de uma carteira é uma medida de risoc introduzida por Markowitz (1952) que explica o quanto o portfólio tende a se mover. O método mais popular de calcular a volatilidade é abordando-a como o desvio padrão médio do retorno R da carteira, ou seja,

$$\sigma=\sqrt{\frac{1}{N-1} \sum_{i=1}^{N}(R_i-\bar{R})^2},\ \mathrm{R}=(\mathrm{R}_1, \ldots, R_n)$$

O Valor condicional em risco (CVaR) examina as perdas que excedem o limite do Valor em Risco (VaR). O VaR e o CVaR estão intimamente relacionados e, ao minimizar o CVaR, também levará a uma redução do VaR da carteira. Para determinar o CVaR do portfólio, apresentamos a abordagem de R. T. Rockafellar e S. Uryasev, 2000 [11] para derivar uma expressão de CVaR que pode ser minimizada.

Definição 2 Value-at-Risk (VAR): *Seja X uma variável aleatória representando a perda de uma determinada carteira. Dado um parâmetro 0 <α<1, o α-VaR de X é:*

$$\tilde{F}_\beta(x, \alpha)=\alpha+\frac{1}{n(1-\beta)} \sum_{k=1}^{n}[f(x, y)-\alpha]^{+} \tag{3}$$

em que $\Psi(x, \alpha)$ *é a função de distribuição acumulada par a perda associada a x.*

Dada a Definição 2.2, o VaR pode ter várias interpretações equivalentes:

• VaRα (X) é a perda mínima que não será excedida com probabilidade α;

• VaRα (X) é o quantil-α da distribuição de X.

• VaRα (X) é a menor perda nos casos (1 - α) × 100% piores.

• VaRα (X) é a perda mais alta nos melhores casos de α × 100%.

O CVaR é uma medida de risco estimada pela média das perdas desde o pior resultado até o percentil (Valor em Risco –VaR) selecionado (geralmente 1%, 5% ou 10%), o que possibilita conhecer a informação a respeito da extensão das perdas (Alexander, 2008). Para o caso de distribuição contínua, o CVaR é apresentado por meio da equação 4:

$$CVaR_{\alpha} = -\alpha^{-1}\int_{-\infty}^{x_{\alpha}} x \cdot f(x) \cdot dx \qquad (4)$$

em que $f(x) \cdot dx$ é a densidade de probabilidade de obter um retorno com valor $x \in \mathbb{R}^n$; x_{α} é o VaR; α é o percentil da distribuição de x; e f(x) é uma função de perda associada a x. Para a estimação do CVaR, podem ser utilizadas metodologias como as simulações por Monte-Carlo, análise de séries históricas e realização de aproximações da distribuição de probabilidade na função de perda por meio da distribuição normal ou da t de Student.

A representação do CVaR na Equação 4 sofreu uma modificação proposta por [11]; a mudança pode ser visualizada na expressão dada pela Equação 5:

$$\phi_{\beta}(x) = (1-\beta)^{-1} \int_{f(x,y) \geq \alpha_{\beta}(x)} f(x,y) \cdot p(y) \cdot dy \qquad (5)$$

em que p(y) é a função densidade de probabilidade de variáveis de mercado; β é o nível de probabilidade a ser escolhido; e f(x,y) é uma função de perda associada à carteira x e a variáveis de mercado y. Rockafellar e Uryasev (2000) propuseram a combinação das Expressões [4] e [5] em termos de uma função F_{β}, como é demonstrado na Equação 6.

$$F_{\beta}(x,\alpha) = \alpha + (1-\beta)^{-1} \int_{y \in \mathbb{R}^m} [f(x,y) - \alpha]^{+} \cdot p(y) \cdot dy$$

(6)

em que $[c]^{+} = \max\{0, c\}$.

A expressão dada pela Equação [6] é aplicada para distribuições contínuas de probabilidade. No caso discreto, na fórmula há algumas modificações (Equação [7]):

$$\tilde{F}_{\beta}(x,\alpha) = \alpha + \frac{1}{n(1-\beta)} \sum_{k=1}^{n} [f(x,y) - \alpha]^{+} \qquad (7)$$

A partir da definição do CVaR para o caso discreto, na Equação [7] é definida a forma para a otimização de uma carteira de ações utilizando o CVaR como medida de risco.

$$\begin{aligned} &Min \ \alpha + \frac{1}{n(1-\beta)} \sum_{k=1}^{n} [-w_i R_i - \alpha]^{+} \\ &sa \quad \sum_{k=1}^{n} w_i = 1 \\ &\qquad 0 \leq w_i \leq 1 \end{aligned} \qquad (8)$$

em que n é o tamanho da amostra.

O uso de técnicas de programação clássica para resolver problemas relacionados à otimização de portfólio ainda são bastante utilizadas, principalmente em problemas onde a função objetivo é convexa, como o modelo de Markowtiz [10], por exemplo. Mas para funções objetivos muito complexas, nem sempre é possível garantir que tais métodos serão eficientes. O algoritmo de Evolução Diferencial introduzido por Storm e Price [13], provou ser bastante poderoso para problemas desta natureza, como o VaR e CVaR, pois tenta otimizar um problema melhorando iterativamente uma solução candidata sem fazer nenhuma suposição sobre o problema que está sendo otimizado.

Evolução Diferencial

O algoritmo evolutivo *Differential Evolution*, em português Evolução Diferencial, abreviado por DE, é um método heurístico, que não usa derivadas, e visa solucionar problemas de otimização contínua. O DE tem se apresentado como um simples, mas poderoso algoritmo de otimização numérica para busca da solução ótima global, sendo aplicado com sucesso na solução de vários problemas de otimização complexa.

De acordo com Cheng e Hwang [4], as principais características do algoritmo DE são:

- É um algoritmo de busca estocástica, originado dos mecanismos de seleção natural;
- O algoritmo é simples e de fácil entendimento, com poucos parâmetros de controle para conduzir a otimização;
- É eficaz para solucionar problemas de otimização com função objetivo descontínua, pois não

necessita de informações sobre derivadas da mesma;

- É mais robusto quanto a ótimo locais, pois busca a solução ótima global manipulando uma população de soluções, ou seja, utiliza diferentes regiões no espaço de busca;
- É eficaz mesmo trabalhando com uma população pequena;
- Permite as variáveis serem otimizadas como números reais, sem processamento extra;

Para otimizar um portfólio utilizando o algoritmo DE melhoramos iterativamente uma solução candidata de uma população gerada de vetores, $\boldsymbol{v}1,...,\boldsymbol{v}i$, i = 1,..., onde cada vetor contém n elementos e representa as variáveis objetivas, ou seja, os pesos do portfólio. A DE visa otimizar o trade-off entre risco e retorno, em vez de minimizar o risco de um determinado retorno específico. O objetivo é assim definido como

$$Max\ (1\text{-}\lambda)\mathbf{w}'\mathbf{R} - \lambda risk \tag{9}$$

onde $\lambda \in \mathbb{R}+$ é o nível de aversão ao risco do investidor e $\boldsymbol{risk}$ pode ser qualquer medida de risco.

A ideia básica do DE é produzir uma nova solução para cada vetor atual, onde a nova solução é uma combinação de quatro soluções atuais na população. Funciona da seguinte maneira: Primeiro, selecionamos um vetor de destino, v_0, da população atual. Em seguida, selecione aleatoriamente três vetores diferentes, usamos um deles como um vetor base e adicionamos a diferença ponderada dos outros dois para construir uma nova solução. Assim, formulamos o novo vetor como

$$v_m = v_1 + F(v_2 - v_3) \tag{10}$$

onde F: [0,1] é um fator de mutação que controla a taxa na qual a população evolui. Finalmente, realizamos uma solução cruzada com o pai e o vetor mutado. Cada elemento no novo vetor de trilha será determinado por uma taxa de *crossover* definida pelo usuário, CR: [0,1], e um número gerado pseudo-aleatório, ϵ. O cruzamento controla a fração de valores de parâmetros copiados do vetor mutante, de modo que

$$v_{n_j} = \begin{cases} v_{0j}, se\ z_j < CR \\ v_{mj}, \text{se } z_j \geq CR \end{cases} \tag{11}$$

Se o número gerado for menor que a razão de cruzamento, o vetor de trilha herdará o elemento j do vetor de destino. Da mesma forma, se o número gerado for maior ou igual à razão de cruzamento, o vetor de trilha herdará o elemento j do vetor mutado.

Manipulando as restrições

Para lidar com as descontinuidades do espaço de busca devido às restrições nos pesos, é introduzida uma função de reparo que garante que o vetor de trilha permaneça sempre dentro do conjunto viável de soluções:

$$\sum_{I=1}^{N} \mathbf{v}_{ni} = 1 \tag{12}$$

Neste caso, a restrição (12) é satisfeita normalizando os comprimentos:

$$v_n = \frac{v_n}{\sum_{j=1}^{N} v_j} \tag{13}$$

Da mesma forma, podemos impor uma função de penalidade para lidar com restrições de cardinalidade, de modo que o portfólio consista apenas em H elementos, ou seja,

$$\left\{ i \in \mathbb{Z}^+ \mid \mathbf{v}_n(i) \neq 0 \right\} \leq H \tag{14}$$

Assim, durante a amostragem da população, cada indivíduo construído deve ser reparado se a carteira representativa não satisfizer as restrições do problema. Se o número dos ativos selecionados for menor ou maior que H, então um operador de reparo seleciona ou exclui um ativo usando uma heurística que prioriza os ativos. Em [5] e [9], é dado um método para encontrar o valor de prioridade:

$$\Upsilon_i = 1 + (1-\lambda) \times \mu_i, \quad i = 1, \ldots, \mathcal{N},$$
$$\Lambda_i = 1 + \lambda \times (\sum_{j=1}^{\mathcal{N}} \sigma_{ij}/\mathcal{N}) \quad i = 1, \ldots, \mathcal{N},$$
$$\Theta = -1 \times min(0, \Upsilon_1, \ldots, \Upsilon_{\mathcal{N}}),$$
$$\mho = -1 \times min(0, \Lambda_1, \ldots, \Lambda_{\mathcal{N}},$$
$$a_i = \frac{\Upsilon_i + \Theta}{\Lambda_i + \mho}, \quad i = 1, \ldots, \mathcal{N}.$$

Uma vez que o vetor de trilha esteja bem protegido, ele é comparado com o vetor alvo e aquele com o maior valor objetivo irá sobreviver na próxima geração. Um esquema do algoritmo pode ser visto na Figura 1

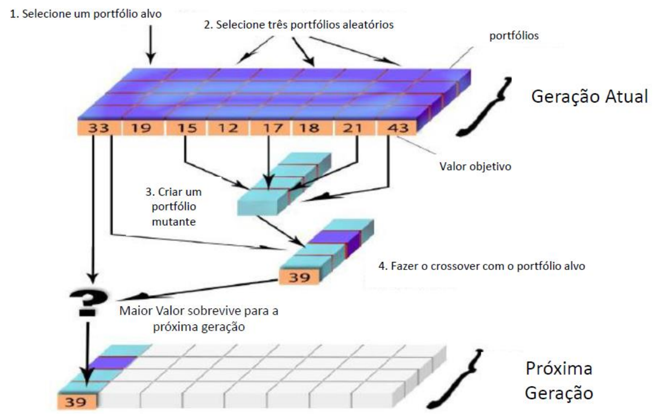


Figura 1 – Algoritmo para otimização com CVaR

A precisão do resultado do algoritmo depende dos parâmetros especificados pelo usuário e, ajustando esses parâmetros, o resultado pode ser significativamente melhorado. Portanto, é importante especificar corretamente essas quatro variáveis:

- N: o número de vetores da população;
- F: fator de mutação;
- CR: proporção de cruzamento;
- ε: Taxa de arredondamento dos pesos das carteiras;

O algoritmo termina quando, após cada geração, os elementos da população concordam com a média da população, aproximando os pesos da carteira usando uma razão ε. Uma vez que todas as soluções são consideradas “próximas o suficiente” para o ótimo global, o algoritmo para.

RESULTADOS

Apresentamos aqui a escolha de parâmetros para o algoritmo DE e comparamos seus resultados com outras técnicas de otimização. Consideramos uma carteira composta apenas por ações e assumimos que os retornos diários dos registros são normalmente distribuídos.

Com relação aos dados, foram utilizados 25 ações da BM&Fbovespa, usando cotações históricas diárias de 5 anos. Além disso, assumimos que há 252 dias de negociação por ano e, assim, escalamos os retornos proporcionalmente ao tempo. Para a programação dos algoritmos, foi utilizado o software Matlab. No caso do VaR condicional, o algoritmo é baseado em [11]. Para o algoritmo DE, sua implementação foi baseada no artigo [13] e utilizando como modelo-base o algoritmo dado em [7].

Volatilidade, VaR e CvaR com DE

Para mostrar que o DE pode lidar com qualquer medida de risco, usamos a abordagem de R. T Rockafellar e S. Uryasev [13] para minimizar o CVaR comparando as soluções do algoritmo DE com o de programação linear. Usamos aqui a mesma suposição do último caso e aproximamos os retornos. Em seguida, vamos otimizar o CVaR no nível de probabilidade de 95% usando os valores dos parâmetros N = 25, F = 0,5, CR = 0,7 e ε = 1%. Esses valores foram escolhidos pois são que oferecem melhor acurácia em um tempo admissível.

Para o caso da volatilidade, podemos ver na Figura 2, que o retorno esperado ideal sobre o investimento varia de 16% a 35%, e a volatilidade de 13% a 22%, ou seja, a tendência de movimento da carteira. Além disso, ao analisar os retornos e riscos dos ativos individualmente, vemos que a diversificação representa um risco menor e, ao mesmo tempo, produz um retorno maior, onde cada portfólio ideal fica ao longo da fronteira eficiente.

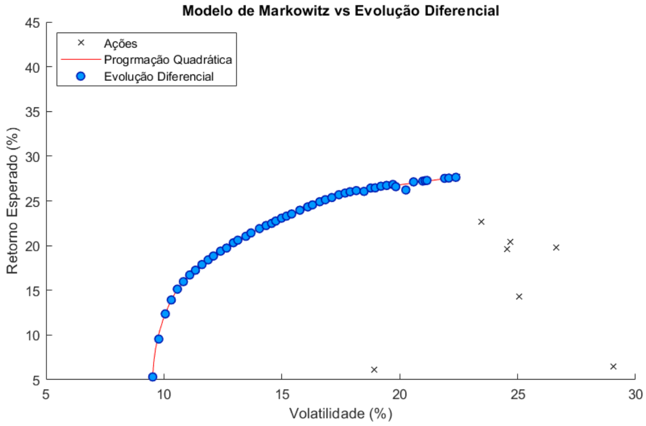


Figura 2 – Fronteira Eficiente para o Modelo de Markowtiz

Vendo que as soluções estão ao longo da fronteira eficiente, podemos confirmar que o algoritmo DE aproximou bem o Modelo de Markowitz, comparando a distribuição dos pesos da carteira, nas Figuras 4 e 5.

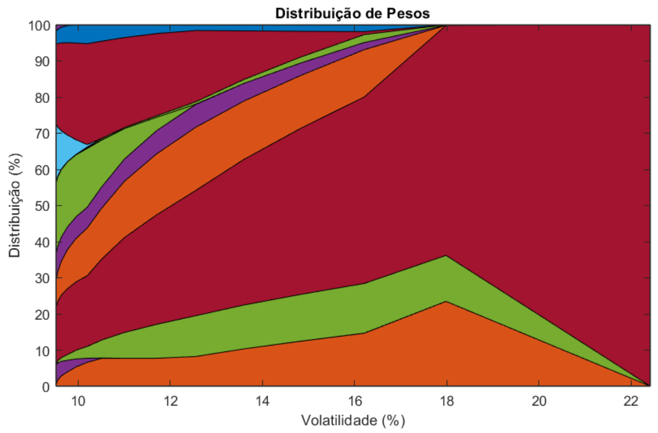


Figura 3 – Distribuição para o Modelo de Markowitz usando Programação Quadrática

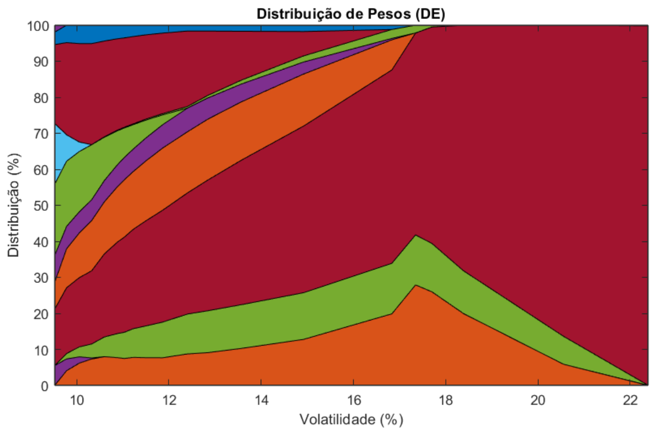
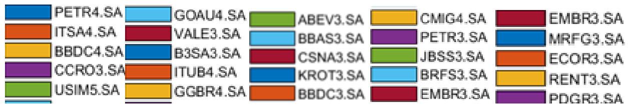


Figura 4 – Distribuição para o Modelo de Markowitz usando Evolução Diferencial



O $CVaR_{95\%}$ varia de 30% a 100%, que é a proporção média de um capital de investidores que será perdido para o nível de probabilidade de 95%. Essa abordagem

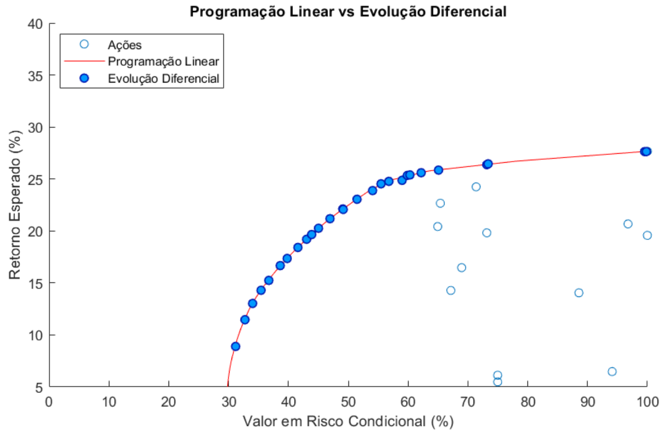


Figura 5 – Fronteira Eficiente para o CVAR

também levou a uma redução no $VaR_{95\%}$, uma vez que é um fator envolvido no processo de minimização do $CVaR_{95\%}$. Para o VaR, na figura 6 vemos $VaR_{95\%}$ varia de 22% a 70%. De qualquer forma, vemos que o DE está na fronteira eficiente e podemos confirmar que as soluções são as mesmas do LP, investigando novamente a distribuição do peso do portfólio.

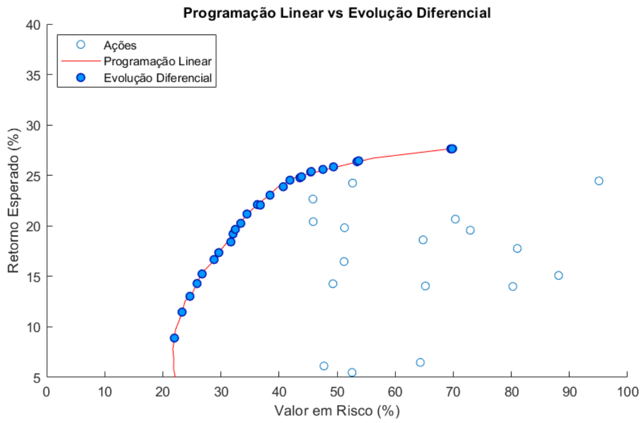


Figura 6 – Fronteira Eficiente para o VAR

A principal desvantagem do DE, é que o algoritmo não é tão eficiente quanto o LP para o Valor em Risco Condicional, já que levou mais tempo para convergir. O tempo médio de duração da otimização CVaR pelo algoritmo DE levou aproximadamente 172 segundos, com F = 0,5, CR = 0,7 e ε =

1%, enquanto que o tempo de duração do LP foi de apenas 10 segundos. No caso do modelo quadrático de Markowitz, o tempo de processamento para o algoritmo utilizando DE foi de 13 segundos, enquanto que utilizando técnicas clássicas de otimização quadrática, o tempo foi de 2 segundos.

Para efeitos de comparação, nas Figuras 7 e 8, vemos a composição dos portfólios utilizando o método CVaR-LP e CVaR-DE. Percebe-se que o segundo algoritmo conseguiu aproximar bem a distribuição de pesos do modelo com os parâmetros dados.

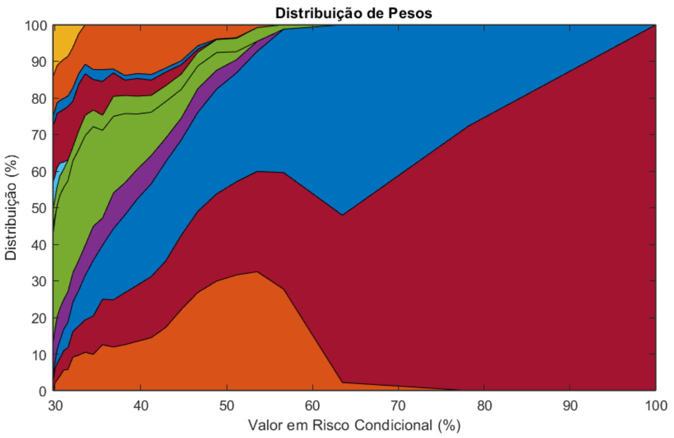


Figura 7 – Distribuição para o CVAR usando Programação Linear

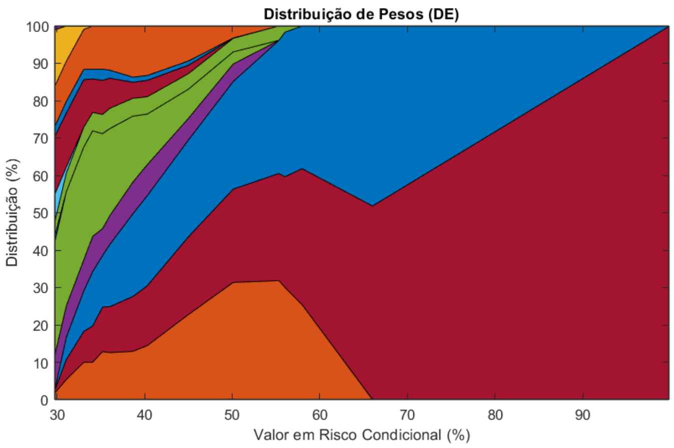
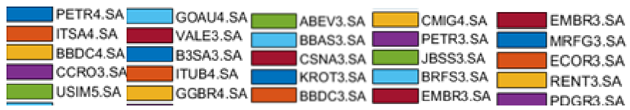


Figura 8 – Distribuição para o CVAR usando Evolução Diferencial



CONCLUSÂO

Verificou-se que para as três medidas de risco estudadas, o algoritmo de Evolução Diferencial conseguiu aproximar de forma satisfatória a distribuição de pesos e a fronteira eficiente da carteira comparando com as técnicas clássicas de otimização. Desta forma, mostramos que a Evolução Diferencial é um algoritmo de busca robusto que pode ser aplicado ao problema do portfólio ótimo. Com relação ao tempo médio de execução do algoritmo, outras estratégias no processo de mutação podem ser úteis no processo de obtenção de um tempo mais rápido. O que foi usado neste artigo é o dado por Storn e Price [11], mas desde a introdução da Evolução Diferencial, outras estratégias têm sido desenvolvidas, o que poderia melhorar a duração do tempo.

BIBLIOGRAFIA

[1] Ahmadi-javid, Amir. **Entropic value-at-risk: A new coherent risk measure**. Journal of Optimization Theory and Applications, v. 155, n. 3, p. 1105-1123, 2012.

[2] ALEXANDER, C. (2008). **Market risk analysis volume IV: Value at risk models**. London: John Wiley & Sons.

[3] Artzner, P., Delbaen, F., Eber, J., & Heath, D. (1999, July). **Coherent measures of risk. Mathematical Finance**, 9(3), 203-228. DOI: 10.1111/1467-9965.00068

[4] CHENG, S. L.; HWANG, C. **Optimal approximation of linear systems by a differential evolution algorithm**, IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, v. 31, n. 6, pp. 698-707, 2001

[5] Cura, T. (2009). **Particle Swarm Optimization Approach to Portfolio Optimization**.. Nonlinear Analysis: Real World Applications, 10(4):2396 – 2406.

[6] DE ARAÚJO, Alcides Carlos; DE ÁVILA MONTINI, Alessandra. **Análise de métricas de risco na otimização de portfolios de ações. Revista de Administração**, v. 50, n. 2, p. 208-228, 2015.

[7] Pcheco, André (2015). **O algoritmo evolutivo Differential Evolution**. Disponível em : http://www.computacaointeligente.com.br/algoritmos/o-algoritmo-evolutivo-differential-evolution/

[8] Kato, F. H. (2004). **Análise de carteiras em tempo discreto**. Dissertação de Mestrado, Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo, SP, Brasil.

[9] Lwin, Khin Thein (2015) **Evolutionary approaches for portfolio optimization**. PhD thesis, University of Nottingham.

[10] MARKOWITZ, Harry. Portfolio selection. **The journal of finance**, v. 7, n. 1, p. 77-91, 1952.

[11] Rockafellar, R. T., & Uryasev, S. (2000). **Optimization of conditional value-at-risk. Journal of Risk**, 2(3), 21-41.

[12] Roman, D., & Mitra, G. (2009, April). **Portfolio selection models: a review and new directions**. Wilmott Journal, 1(2), 69-85. DOI: 10.1002/wilj.4

[13] STORN, Rainer; PRICE, Kenneth. **Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces**. Journal of global optimization, v. 11, n. 4, p. 341-359, 1997.