

Otimização robusta de portfólios usando conjuntos de incerteza não paramétricos sobre dados financeiros

Tatiane Cazarin da Silva¹, Gislaine Aparecida Pericaro²,
Eduardo H. M. Pena³, Ademir Alves Ribeiro⁴

¹Departamento Acadêmico de Matemática, Universidade Tecnológica Federal do Paraná
UTFPR - Campus Campo Mourão, PR - Brasil

²Colegiado de Matemática, Universidade Estadual do Paraná
Unespar - Campus Campo Mourão, PR - Brasil

³Departamento de Ciência da Computação, Universidade Tecnológica Federal do Paraná
UTFPR - Campus Campo Mourão, PR - Brasil

⁴Departamento de Matemática, Universidade Federal do Paraná
UFPR - Curitiba, PR - Brasil

tatianecazarin@utfpr.edu.br, gislaine.pericaro@ies.unespar.edu.br,
eduardopena@utfpr.edu.br, ademir.ribeiro@ufpr.br

Resumo. A otimização de portfólios financeiros depende da qualidade dos dados históricos e da representação adequada da incerteza nos parâmetros estimados. O modelo de Markowitz é amplamente usado como referência, mas apresenta sensibilidade a erros de estimação dos retornos esperados, comprometendo sua robustez em cenários de alta volatilidade. Este trabalho propõe um modelo robusto cujo conjunto de incerteza é construído de forma não paramétrica usando Support Vector Clustering (SVC). O modelo elimina a necessidade de hipóteses restritivas de simetria ou normalidade, e um parâmetro de regularização controla o suporte desse conjunto, funcionando simultaneamente como mecanismo de tratamento de outliers e de calibração do perfil de risco do investidor. O modelo proposto foi avaliado em contratos futuros de commodities internacionais durante um período de alta volatilidade, superando consistentemente o modelo clássico de Markowitz e a estratégia de pesos iguais, com menor volatilidade, perdas extremas reduzidas e índices Sharpe e Sortino positivos em todas as estratégias consideradas.

Abstract. Portfolio optimization depends on the quality of historical data and the adequate representation of uncertainty in estimated parameters. The Markowitz model is widely used as a benchmark, but is sensitive to estimation errors in expected returns, compromising its robustness in high-volatility scenarios. This work proposes a robust model whose uncertainty set is constructed non-parametrically using Support Vector Clustering (SVC). The model eliminates the need for restrictive assumptions of symmetry or normality, and a regularization parameter controls the support of this set, simultaneously serving as a

mechanism for outlier handling and investor risk profile calibration. The proposed model was evaluated on international commodity futures contracts during a high-volatility period, consistently outperforming classical Markowitz model and the equal-weight strategy, with lower volatility, reduced extreme losses, and positive Sharpe and Sortino ratios across all evaluated strategies.

1. Introdução

Modelos de otimização dependem diretamente da qualidade dos dados que os alimentam. Os parâmetros dos modelos não são observados de forma direta, mas estimados a partir de amostras históricas, o que os torna suscetíveis a ruído, viés de estimação e instabilidade temporal [Delage and Ye 2010]. A literatura em otimização sob incerteza oferece diferentes estratégias para lidar com esse cenário [Ben-Tal et al. 2009]. A programação estocástica pressupõe conhecimento prévio da distribuição de probabilidade dos parâmetros. A otimização distributivamente robusta relaxa essa hipótese, trabalhando com famílias de distribuições [Delage and Ye 2010]. A otimização robusta, por sua vez, substitui estimativas pontuais dos parâmetros por intervalos de variação e protege a solução contra o pior caso dentro desse conjunto de incerteza [Bertsimas and Thiele 2006]. Em todas essas abordagens, a escolha da representação da incerteza é determinante para a qualidade da solução obtida.

Essa questão é particularmente crítica no contexto de portfólios financeiros. O modelo de média variância de Markowitz [Markowitz 1952] estima seus parâmetros de entrada, como retornos esperados e covariâncias, a partir de séries temporais com propriedades estatísticas, incluindo assimetria, caudas pesadas e correlações que variam ao longo do tempo. Há evidência empírica de que o modelo é substancialmente mais sensível a erros nos retornos esperados do que a erros na matriz de covariância [Chopra and Ziemba 1993], o que torna a modelagem da incerteza nos retornos esperados o foco central de abordagens robustas. Ainda assim, muitos conjuntos de incerteza usados na literatura impõem uma simetria geométrica que raramente reflete a estrutura real dos dados financeiros, podendo produzir soluções excessivamente conservadoras ou insuficientemente resilientes [Bertsimas and Sim 2004].

Trabalhos relacionados. Uma alternativa mais aderente à realidade dos dados é construir o conjunto de incerteza diretamente a partir das observações disponíveis, sem impor sua geometria a priori. Nessa direção, abordagens baseadas em aprendizado de máquina permitem que a fronteira do conjunto de incerteza seja determinada pela estrutura dos dados em vez de por hipóteses geométricas pré-estabelecidas, capturando assimetrias e dependências que conjuntos em caixa ou elipsoidais não representam [Bertsimas et al. 2018, Shang et al. 2017, Zhang et al. 2022]. Em particular, o uso de funções de *kernel* permite representar a fronteira da distribuição de forma não paramétrica e computacionalmente tratável [Shang et al. 2017], dispensando hipóteses de normalidade ou simetria que raramente se verificam em séries financeiras. [Chenreddy et al. 2022] avançam nessa direção ao construir conjuntos de incerteza via agrupamento profundo de dados históricos, sem impor geometria paramétrica à região de proteção. Diferentemente do SVC, essa abordagem requer covariáveis externas para particionar o espaço de incerteza e não oferece um mecanismo direto para calibrar o conservadorismo segundo o perfil do investidor. Outra linha de trabalho aprimora os parâmetros de entrada dos modelos clássicos via previsão de retornos com redes neurais. [Soares and Canuto 2025] integram esse tipo de arquitetura

ao modelo do Índice de Sharpe para estimar retornos e correlações futuras de criptoativos e obtêm ganhos expressivos frente ao modelo sem previsão em mercados de alta volatilidade. No contexto brasileiro, [Bomfim and Nascimento 2024] propõem um pipeline para detecção de bolhas financeiras no mercado nacional e documentam episódios de instabilidade que reforçam a necessidade de modelos robustos a mudanças de regime.

Contribuições. Este trabalho propõe um modelo robusto de otimização de portfólio com conjunto de incerteza não paramétrico construído via *Support Vector Clustering* (SVC) [Shang et al. 2017], sem assumir simetria na distribuição dos retornos. O parâmetro de regularização $\psi \in (0, 1)$ permite ajustar o grau de conservadorismo do modelo conforme diferentes perfis de investidor. O modelo é formulado como um problema de programação quadrática e avaliado em treze contratos futuros de commodities internacionais entre 2019 e 2022, período marcado por elevada volatilidade e instabilidade nos mercados financeiros [Bomfim and Nascimento 2024]. Os resultados são comparados ao modelo clássico de Markowitz, à estratégia de pesos iguais e ao índice Bloomberg Commodity (BCOM), considerando métricas de desempenho, risco e diversificação. As principais contribuições são: (i) a utilização de um conjunto de incerteza não paramétrico baseado em SVC; (ii) a interpretação do parâmetro ψ como mecanismo de ajuste ao perfil de risco do investidor; e (iii) a avaliação empírica do modelo em um cenário de alta volatilidade de mercado.

2. Modelo matemático

A formulação matemática do modelo clássico de média-variância de Markowitz [Markowitz 1952] é dada por:

$$\begin{aligned} \max_w \quad & \mu^T w - \gamma w^T \Sigma w \\ \text{s. a.} \quad & w \in \mathcal{W}, \end{aligned} \tag{1}$$

em que $\mathcal{W} = \{w \in \mathbb{R}^n \mid w^T \mathbf{1} = 1 \text{ e } w \geq 0\}$, com $\mathbf{1} = [1, 1, \dots, 1]^T \in \mathbb{R}^n$, representa o conjunto de restrições para os pesos $w \in \mathbb{R}^n$. Os parâmetros $\mu \in \mathbb{R}^n$, $\Sigma \in \mathbb{R}^{n \times n}$ e $\gamma \in \mathbb{R}$ representam, respectivamente, o vetor de retornos esperados, a matriz de covariância dos retornos dos ativos e o coeficiente de aversão ao risco. Em outras palavras, o modelo busca alocar o capital disponível entre n ativos de forma a maximizar o retorno esperado da carteira, penalizado pela sua variância. O parâmetro γ regula o equilíbrio entre risco e retorno, que pode ser definido, por exemplo, de acordo com o perfil do investidor.

Considere que os retornos dos ativos pertencem a um conjunto de incertezas $\mathcal{U}$. A otimização robusta é uma abordagem que busca soluções viáveis e eficientes mesmo diante de incertezas nos parâmetros do problema, proporcionando resiliência às possíveis realizações desses cenários [Bertsimas and Sim 2004, Ben-Tal and Nemirovski 1998]. Esta metodologia tem sido amplamente aplicada na otimização de portfólios [Bertsimas and Pachamanova 2008, Danesh et al. 2024, Fabozzi et al. 2007, Sehgal and Jagadeesh 2023], contribuindo para a construção de alocações mais estáveis e adaptáveis a diferentes condições de mercado. A formulação robusta para o problema (1) é dada por um problema do tipo max-min, definido por:

$$\max_{w \in \mathcal{W}} \left[\left(\min_{\mu \in \mathcal{U}} \mu^T w \right) - \gamma w^T \Sigma w \right]. \tag{2}$$

Intuitivamente, o modelo escolhe a alocação w que maximize o retorno ajustado ao risco dentro do conjunto de incerteza $\mathcal{U}$. Usando o conjunto de incertezas definido em

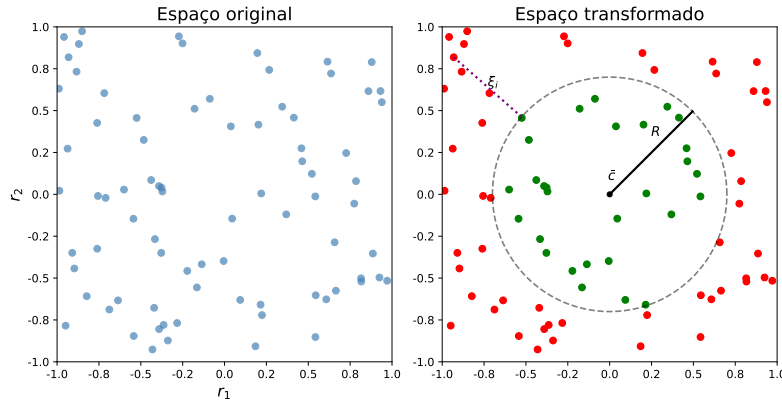


Figura 1. Ilustração do problema (3) para o caso bidimensional.

[Shang et al. 2017], incorporado ao problema (2), um novo problema é definido, denominado contraparte robusta, que utiliza Clusterização de Vetores de Suporte (*Support Vector Clustering* - SVC) [Ben-Hur et al. 2000] para construir um conjunto de incertezas flexível. Diferentemente das abordagens tradicionais, essa metodologia não pressupõe hipóteses de simetria, permitindo um melhor ajuste ao otimizar a relação risco-retorno considerando flutuações nos retornos esperados e riscos de volatilidade.

2.1. Formulação matemática usando SVC na construção do conjunto de incertezas

A técnica de SVC consiste em realizar um mapeamento da forma $r^k \in \mathbb{R}^n \rightarrow \phi(r^k) \in \mathbb{R}^p$. Nesse novo espaço, o objetivo é determinar uma hipersfera de volume mínimo que englobe os pontos mapeados. Para evitar raios excessivos, são introduzidas variáveis de folga, $\xi > 0$, que controlam a inclusão ou não de todos os pontos. Matematicamente, o problema de otimização pode ser formulado por:

$$\begin{aligned} \underset{R, \bar{c}, \xi}{\text{minimizar}} \quad & R^2 + \frac{1}{N\psi} \sum_{k=1}^N \xi_k \\ \text{sujeito a} \quad & \|\phi(r^k) - \bar{c}\|^2 \leq R^2 + \xi_k, \quad k = 1, \dots, N \\ & \xi_k \geq 0, \quad k = 1, \dots, N \end{aligned} \quad (3)$$

sendo $R \in \mathbb{R}$ e $\bar{c} \in \mathbb{R}^p$ o raio e o centro da hipersfera, respectivamente, N o número de observações e $\psi \in (0, 1]$ um parâmetro de regularização. Uma ilustração para o modelo, para o caso em que $d = 2$ e $\phi = I$, pode ser observado na Figura 1. Note que o parâmetro ψ regulariza a função objetivo. Valores maiores tendem a penalizar a função objetivo, resultando em menores raios, e por consequência, considerando menos pontos pertencentes à hipersfera. Essa característica pode então ser empregada na determinação de pontos *outliers*, à medida que desconsidera os pontos mais distantes do centro.

Usando o princípio da dualidade, o problema dual associado a (3) é dado por:

$$\begin{aligned} \underset{\alpha}{\text{maximizar}} \quad & \left(\sum_{k=1}^N \alpha_k \phi(r^k)^T \phi(r^k) - \sum_{k=1}^N \sum_{i=1}^N \alpha_k \alpha_i \phi(r^k)^T \phi(r^i) \right) \\ \text{sujeito a} \quad & 0 \leq \alpha_k \leq \frac{1}{N\psi} \quad k = 1, \dots, N. \end{aligned} \quad (4)$$

A solução desse problema fornece os multiplicadores de Lagrange α , a partir dos quais definem-se os conjuntos $SV = \{k \mid \alpha_k > 0\}$ e $BSV = \{k \mid 0 < \alpha_k < \frac{1}{N\psi}\}$. Quando $k \in BSV$, o ponto $\phi(r^k)$ encontra-se geometricamente sobre a fronteira da hipersfera no espaço transformado, sendo sua distância ao centro $\bar{c}$ igual ao raio R . Os pontos para os quais $\alpha_k = \frac{1}{N\psi}$ situam-se fora da hipersfera e são classificados como *outliers*. Por outro lado, quando $k \notin SV$, tem-se $\alpha_k = 0$, indicando que os pontos correspondentes estão localizados no interior da hipersfera. Assim, o conjunto de incertezas considerado é dado pelo conjunto de pontos pertencentes à hipersfera no espaço de transformado, isto é,

$$\mathcal{U}(\mathcal{R}) = \{r \in \mathbb{R}^n \mid \text{dist}(\phi(r), \bar{c})^2 \leq R^2\}. \quad (5)$$

Para evitar a determinação explícita da função ϕ , utiliza-se o denominado truque do *kernel*, técnica muito utilizada em aprendizado de máquina que consiste em trabalhar somente com o valor do produto interno, ou seja, $\phi(r)^T \phi(s) = K(r, s)$. Assim, ao adotar a função kernel $K(r, s) = \sum_{j=1}^n \ell_j - \|Q(r - s)\|_1$ com $Q = \Sigma^{-\frac{1}{2}}$ e $\ell_j > \max_{1 \leq k \leq N} q_j^T r^k - \min_{1 \leq k \leq N} q_j^T r^k$, proposta por [Shang et al. 2017], o conjunto de incertezas $\mathcal{U}(\mathcal{R})$ pode ser reescrito por:

$$\mathcal{U}(\mathcal{R}) = \left\{ r \in \mathbb{R}^n \mid \sum_{i \in SV} \alpha_i \mathbf{1}^T v^i \leq \theta, -v^i \leq Q(r - r^i) \leq v^i, i \in SV \right\}, \quad (6)$$

sendo $\theta = \min \left\{ \sum_{i \in SV} \alpha_i \|Q(r^k - r^i)\|_1, k \in BSV \right\}$ e $v^i \in \mathbb{R}^n$ variáveis auxiliares que representam o valor em módulo de cada elemento do vetor $Q(r - r^i)$, o que mantém as restrições lineares, superando o uso da kernel gaussiana. A partir da formulação interna do problema robusto definido em (2) para o conjunto de incertezas (6), o dual do problema interno pode ser então escrito por:

$$\begin{aligned} & \underset{\eta, \lambda, \rho}{\text{maximizar}} && -\eta\theta + \sum_{i \in SV} (\lambda_i - \rho_i)^T Q r^i \\ & \text{sujeito a} && w = \sum_{i \in SV} Q(\lambda_i - \rho_i) \\ & && \eta \alpha_i \mathbf{1} - \lambda_i = \rho_i, \forall i \in SV \\ & && \eta \geq 0, \lambda_i, \rho_i \geq 0, \forall i \in SV. \end{aligned} \quad (7)$$

E o modelo proposto passa a ser:

$$\begin{aligned} & \underset{w, \eta, \lambda, \rho}{\text{maximizar}} && -\gamma w^T \Sigma w - \eta\theta + \sum_{i \in SV} (\lambda_i - \rho_i)^T Q r^i \\ & \text{sujeito a} && w^T \mathbf{1} = 1 \\ & && w = \sum_{i \in SV} Q(\lambda_i - \rho_i) \\ & && \eta \alpha_i \mathbf{1} - \lambda_i = \rho_i, \forall i \in SV \\ & && w \geq 0, \eta \geq 0, \lambda_i, \rho_i \geq 0, \forall i \in SV, \end{aligned} \quad (8)$$

mantendo as características de um problema de programação quadrática, assim como o modelo clássico de Markowitz.

3. Metodologia

O modelo formulado na seção anterior requer como entrada os retornos esperados dos ativos para a construção do conjunto de incerteza com o SVC e a estimação da matriz de covariância Σ . A seguir, descrevemos os dados utilizados, o protocolo de treinamento e teste, e as métricas empregadas para avaliação das carteiras geradas.

Os dados utilizados correspondem aos preços de fechamento das principais commodities negociadas no mercado internacional, representadas pelos seguintes contratos futuros: soja (ZS=F), milho (ZC=F), café (KC=F), açúcar (SB=F), algodão (CT=F), trigo (ZW=F), boi gordo (LE=F), suíno (HE=F), minério de ferro (TIO=F), petróleo WTI (CL=F), petróleo Brent (BZ=F), alumínio (ALI=F) e ouro (GC=F), no período de 2019 a 2022. Os códigos entre parênteses correspondem aos *tickers* do provedor *Yahoo Finance*, acessados com a biblioteca *yfinance*. O sufixo =F indica séries de contratos futuros. Neste estudo, utilizamos as séries *contínuas* (encadeadas por rolagem do contrato mais próximo), conforme disponibilizadas pelo provedor.

Como *benchmark* (Bench), empregou-se o índice Bloomberg Commodity Index (BCOM), que fornece exposição diversificada a commodities físicas por meio de contratos futuros. As cotações também foram obtidas por meio da biblioteca *yfinance*. Os retornos esperados foram estimados por meio dos retornos diários históricos, que foram calculados a partir da variação percentual entre os preços de fechamento consecutivos. Por fim, foram selecionados apenas os ativos com séries históricas completas ao longo de todo o período considerado.

O período analisado, 01/01/2019 a 01/01/2022, foi escolhido por abranger a pandemia de COVID-19, caracterizada por elevada volatilidade e aumento significativo da incerteza nos mercados financeiros. Estudos documentam que a pandemia elevou tanto o número quanto a amplitude das sobrereações de preços em futuros de commodities [Borgards et al. 2021], além de intensificar a co-dependência entre diferentes classes de ativos e ampliar a volatilidade de forma heterogênea entre categorias de commodities [Umar et al. 2021]. Esse contexto é apropriado para a avaliação de metodologias voltadas ao controle de risco e à robustez dos modelos de otimização de portfólio, uma vez que a ocorrência frequente de eventos extremos justifica o uso de técnicas para identificação e tratamento de observações discrepantes, permitindo reduzir a influência de movimentos atípicos sobre as estimativas estatísticas e a alocação de ativos.

Para construção e validação do modelo, os dados foram divididos em períodos de treinamento e teste, permitindo a avaliação do modelo sob diferentes condições de mercado. Foram empregadas as estratégias de rebalanceamento e *Buy and hold* (Compra e manutenção), usando a abordagem de horizonte rolante. Tal abordagem atualiza periodicamente a carteira usando apenas os dados mais recentes de um intervalo temporal fixo, descartando observações antigas e incorporando novas, o que permite adaptação dinâmica às condições de mercado e avaliação fora da amostra. Para ambas as estratégias, considerou-se o período fixo de treinamento de 252 dias úteis (aproximadamente um ano). Para avaliar as estratégias de otimização de portfólio, foram definidos dois parâmetros:

- Coeficiente de Aversão ao Risco (γ): Regula o equilíbrio entre retorno esperado

e risco (variância da carteira). Foi adotado $\gamma = 1$ em todos os experimentos, visando manter a consistência das comparações realizadas.

- Parâmetro de Regularização (ψ): Determina o tamanho do conjunto de incerteza, onde $\psi \in (0, 1]$. Para a análise dos métodos baseados em SVC, foram considerados três valores distintos, que podem ser relacionados a diferentes perfis de investidores: conservador, moderado e arrojado, correspondentes a $\psi = 0.2$, 0.5 e 0.8 , respectivamente. Esses casos são denotados como SVC_{ψ} .

Para a comparação da carteira, foram consideradas diferentes abordagens: (i) modelo clássico de Markowitz (MK), (ii) alocação de pesos iguais (*Equal-Weights* - EW) e (iii) modelo robusto baseado em SVC com diferentes valores do parâmetro de regularização. Para avaliação da carteira, foram empregadas métricas de risco (volatilidade, perda máxima, Valor em Risco - VaR, Valor Condicional em Risco - CVaR), desempenho (índices: alpha, beta, sharpe e sortino, ômega e retorno médio) e composição/diversificação (índice Herfindahl-Hirschman - HHI, *turnover* e o número de ativos).

O modelo proposto foi implementado em Python e resolvido com o solver *Gurobi Optimizer* (via *gurobipy*) [Gurobi Optimization, LLC 2026]. Os problemas de otimização apresentaram baixo custo computacional e tempos de resolução muito reduzidos, motivo pelo qual não foram contabilizados nas análises.

4. Resultados

Esta seção apresenta e discute os resultados obtidos com o modelo proposto, organizados em três partes: métricas de risco e desempenho das carteiras, análise de diversificação e, por fim, um comparativo entre a abordagem SVC e versões do modelo MK com tratamento de *outliers*.

As métricas utilizadas para avaliar e comparar os modelos são amplamente adotadas na literatura de otimização de portfólios [Bertsimas and Pachamanova 2008]. As métricas de risco compreendem: volatilidade (desvio padrão dos retornos), máximo draw-down (maior queda acumulada no período), Valor em Risco (VaR) e Valor Condicional em Risco (CVaR) ao nível de 95%, que estimam, respectivamente, a perda máxima esperada e a perda média nos piores cenários. As métricas de desempenho incluem os índices Sharpe e Sortino, que medem o retorno ajustado ao risco total e ao risco de queda, respectivamente, o índice Omega, que relaciona ganhos e perdas em relação a um limiar, e o retorno médio diário. Por fim, a composição e diversificação das carteiras são avaliadas pelo número de ativos selecionados, pelo Índice Herfindahl-Hirschman (HHI), que mede a concentração da carteira, onde valores mais baixos indicam maior diversificação, e pelo *turnover*, que quantifica a rotatividade dos ativos entre rebalanceamentos e serve como indicador indireto dos custos de transação.

A Tabela 1 apresenta as medidas de risco e desempenho das carteiras construídas pelos modelos considerando diferentes frequências de investimentos (*Buy and hold* e rebalanceamentos - trimestral, semestral e anual). Os dados originais foram mantidos para servir como referência na análise, permitindo avaliar a efetividade das estratégias de tratamento de *outliers* consideradas neste estudo. Consequentemente, eventuais valores atípicos ou inconsistências presentes nos dados foram mantidos, refletindo as condições reais dos dados disponíveis aos investidores e justificando possíveis resultados fora dos intervalos usualmente observados.

Tabela 1. Medidas de risco e desempenho das carteiras

Rebalanceamento Trimestral								
Modelo	Risco				Desempenho			
	Volatilidade	Max Drawdown	VaR (95%)	CVaR (95%)	Sharpe	Sortino	Omega	Retorno médio
SVC _{0.2}	0.0207	-0.3963	-0.0219	-0.0502	0.0057	0.0057	0.0007	0.0001
SVC _{0.5}	0.0257	-0.6484	-0.0273	-0.0658	-0.0030	-0.0029	0.0044	-0.0001
SVC _{0.8}	0.0139	-0.2431	-0.0198	-0.0363	0.0453	0.0423	0.0000	0.0006
MK	0.0173	-0.3599	-0.0230	-0.0467	-0.0107	-0.0111	0.0003	-0.0002
EW	0.0158	-0.2726	-0.0170	-0.0373	-0.7353	-0.9408	0.0000	0.0003
Bench	0.1619	-0.2714	-0.0165	-0.0259	0.0448	0.0557	0.0000	0.0005
Rebalanceamento Semestral								
Modelo	Risco				Desempenho			
	Volatilidade	Max Drawdown	VaR (95%)	CVaR (95%)	Sharpe	Sortino	Omega	Retorno médio
SVC _{0.2}	0.0683	-0.3963	-0.0258	-0.1147	-0.0272	-0.0196	0.0023	-0.0019
SVC _{0.5}	0.1067	-0.6802	-0.0399	-0.1922	-0.0391	-0.0270	0.0131	-0.0042
SVC _{0.8}	0.0129	-0.2816	-0.0179	-0.0331	0.0429	0.0439	0.0000	0.0006
MK	0.0551	-0.4117	-0.0295	-0.1064	-0.0453	-0.0331	0.0000	-0.0025
EW	0.0158	-0.2726	-0.0170	-0.0373	-0.7353	-0.9408	0.0000	0.0003
Bench	0.1619	-0.2714	-0.0165	-0.0259	0.0448	0.0557	0.0000	0.0005
Rebalanceamento Anual								
Modelo	Risco				Desempenho			
	Volatilidade	Max Drawdown	VaR (95%)	CVaR (95%)	Sharpe	Sortino	Omega	Retorno médio
SVC _{0.2}	0.0685	-0.3963	-0.0276	-0.1193	-0.0340	-0.0245	0.0023	-0.0023
SVC _{0.5}	0.1070	-0.6802	-0.0483	-0.1912	-0.0435	-0.0315	0.0130	-0.0047
SVC _{0.8}	0.0093	-0.2333	-0.0147	-0.0243	0.0527	0.0638	0.0000	0.0005
MK	0.0559	-0.4986	-0.0323	-0.1131	-0.0442	-0.0324	0.0000	-0.0025
EW	0.0158	-0.2726	-0.0170	-0.0373	-0.7353	-0.9408	0.0000	0.0003
Bench	0.1619	-0.2714	-0.0165	-0.0259	0.0448	0.0557	0.0000	0.0005
Estratégia Buy and hold								
Modelo	Risco				Desempenho			
	Volatilidade	Max Drawdown	VaR (95%)	CVaR (95%)	Sharpe	Sortino	Omega	Retorno médio
SVC _{0.2}	0.0681	-0.3963	-0.0280	-0.1165	-0.0350	-0.0246	0.0022	-0.0024
SVC _{0.5}	0.1075	-0.6802	-0.0497	-0.1943	-0.0340	-0.0242	0.0132	-0.0037
SVC _{0.8}	0.0084	-0.2079	-0.0129	-0.0217	0.0742	0.0883	0.0000	0.0006
MK	0.0544	-0.3746	-0.0255	-0.0991	-0.0269	-0.0187	0.0000	-0.0015
EW	0.0158	-0.2726	-0.0170	-0.0373	-0.7353	-0.9408	0.0000	0.0003
Bench	0.1619	-0.2714	-0.0165	-0.0259	0.0448	0.0557	0.0000	0.0005

De modo geral, observa-se que o modelo SVC_{0.8} apresentou desempenho superior e mais consistente nas estratégias analisadas, apresentando maior controle de risco e menor exposição a perdas extremas. Esse comportamento pode ser observado por menores níveis de volatilidade, VaR, CVaR e máximo drawdown. Além disso, foi o único modelo que resultou em índices de Sharpe e Sortino positivos de forma recorrente, bem como retorno médio positivo. Por outro lado, os modelos SVC_{0.2} e SVC_{0.5} apresentaram desempenho inferior, com maiores níveis de volatilidade e drawdowns, e índices Sharpe e Sortino predominantemente negativos. Esse comportamento sugere que valores menores de ψ não são suficientes para filtrar adequadamente os ativos mais instáveis, resultando em carteiras mais expostas ao risco.

O modelo MK apresentou os piores resultados em termos de máximo drawdown, índice Sharpe e Sortino. Tais resultados indicam elevada instabilidade e baixa robustez do método frente às características dos dados utilizados, possivelmente em razão da sensibilidade do modelo à estimação da matriz de covariância e à presença de ativos altamente voláteis.

A carteira EW apresentou volatilidade relativamente baixa, porém sem ganhos

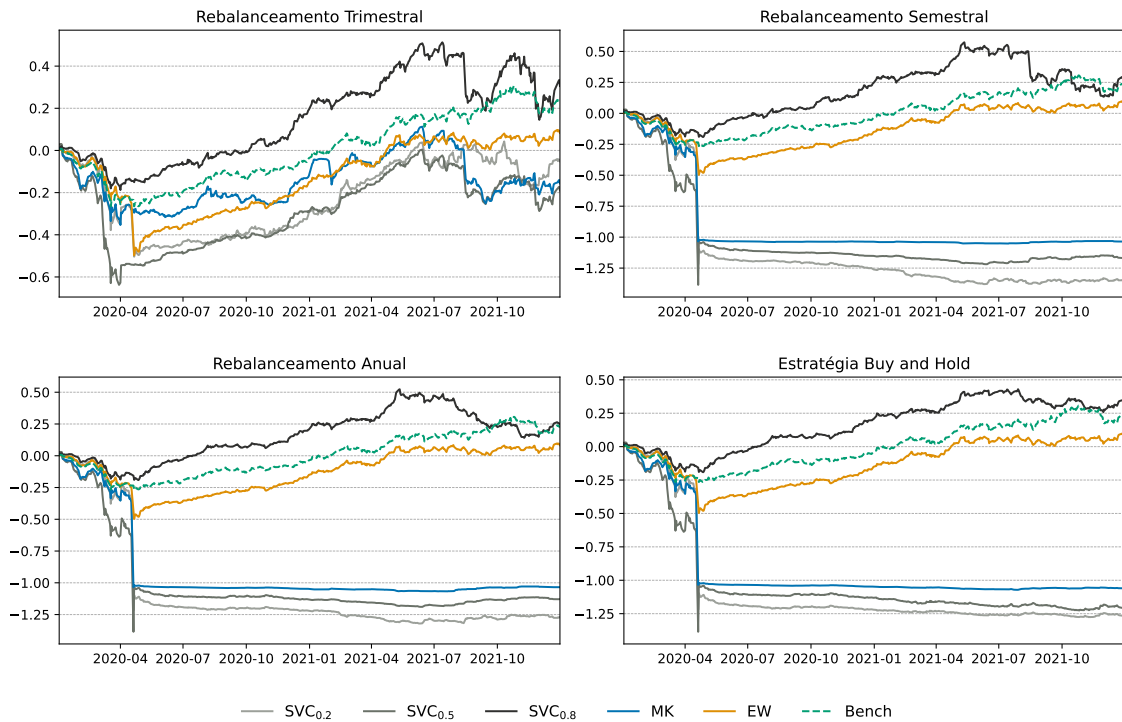


Figura 2. Evolução do retorno acumulado das carteiras ao longo do tempo.

expressivos em termos de retorno médio e com índices Sharpe negativos. Isso indica que, apesar de sua simplicidade e estabilidade, a estratégia não foi eficiente em gerar retorno ajustado ao risco superior aos demais modelos considerados.

O *benchmark* apresentou retorno médio, índices Sharpe e Sortino positivos, no entanto a volatilidade foi maior do que a observada para o modelo $SVC_{0.8}$. Esse resultado evidencia que a variação do modelo proposto é capaz de alcançar uma relação risco-retorno mais eficiente do que o mercado, ao reduzir substancialmente o risco sem comprometer o retorno médio.

Quanto à frequência de rebalanceamento, os resultados mantêm padrões qualitativos semelhantes entre as estratégias trimestral, semestral, anual e *Buy and hold*, sem alteração significativa na hierarquia dos modelos. Contudo, o impacto do rebalanceamento depende do perfil da carteira. Para a configuração mais conservadora $SVC_{0.2}$, o rebalanceamento trimestral reduz de forma relevante as medidas de risco extremo, especialmente o máximo Drawdown e o CVaR, indicando maior controle de perdas. Já para a configuração mais agressiva $SVC_{0.8}$, frequências menores tendem a melhorar o desempenho ajustado ao risco, com aumento do índice de Sharpe e do retorno médio, sem agravamento substancial das métricas de risco. Assim, os efeitos da frequência de rebalanceamento não são uniformes, variando conforme o nível de exposição ao risco da estratégia.

A Figura 2 apresenta uma evolução das carteiras ao longo do período analisado, exibindo o retorno acumulado para cada uma das estratégias. Nota-se uma estabilidade entre o desempenho dos métodos utilizados, independente da estratégia empregada. Isso se torna mais evidente nas estratégias *Buy and hold* e nos rebalanceamentos semestral e anual. Já no rebalanceamento trimestral, há uma maior proximidade entre os métodos

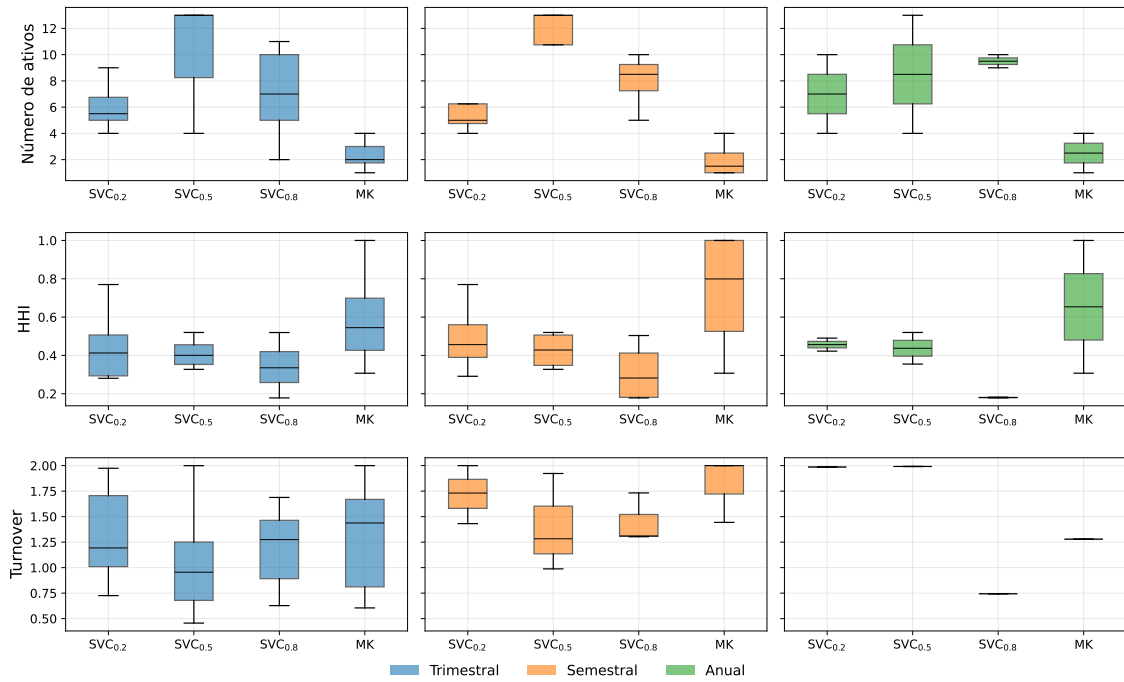


Figura 3. Número de Ativos, Índice Herfindahl-Hirschman (HHI) e Turnover: diversificação das carteiras segundo diferentes estratégias de rebalanceamentos.

utilizados, sendo o cenário no qual MK e as duas versões de SVC, com valores mais baixos de ψ ficaram mais próximas dos demais métodos.

Em síntese, os resultados indicam que a utilização do modelo SVC com valores mais elevados de ψ constitui uma estratégia mais eficiente para a construção de carteiras no cenário analisado, proporcionando menor exposição ao risco, menores perdas extremas e melhor desempenho ajustado ao risco quando comparada aos modelos tradicionais e ao *benchmark* de mercado. Para comparar o desempenho ajustado ao risco das carteiras, foi aplicado o teste pareado para o Índice de Sharpe proposto em [Ledoit and Wolf 2008]. Os resultados indicaram que as diferenças observadas entre os métodos não foram estatisticamente significativas ao nível de 5% de significância. Ainda assim, as análises descritivas mostraram que o modelo SVC apresentou métricas superiores às do modelo de Markowitz, com maiores valores de Índice de Sharpe e menor exposição ao risco. Além disso, observou-se que o SVC foi capaz de capturar retornos mais expressivos em determinados períodos do horizonte de análise. Em conjunto, esses resultados apontam para uma vantagem prática do método proposto, especialmente no que se refere ao controle de risco e ao desempenho ajustado ao risco.

Os índices referentes à composição e diversificação nas carteiras, quanto às estratégias de rebalanceamentos, podem ser observados na Figura 3. Pode-se observar que as estratégias baseadas no método SVC apresentam carteiras mais diversificadas quando comparadas ao modelo MK, conforme indicado por menores valores do HHI e maior número de ativos selecionados, respectivamente. Observa-se ainda que MK produz carteiras altamente concentradas, sobretudo no rebalanceamento anual. Em relação ao *turnover*, as estratégias SVC apresentam maior estabilidade ao longo do tempo, implicando em menores custos de transação. Esses resultados indicam que a abordagem SVC é mais adequada

para a construção de carteiras diversificadas e operacionalmente mais eficientes.

4.1. Comparativo para tratamento de *outliers*

O parâmetro de regularização ψ na estratégia baseada em SVC permite reduzir o conjunto de dados utilizado no treinamento do modelo em aproximadamente 20%, 50% e 80%. Para investigar se os ganhos observados no modelo $SVC_{0.8}$ decorrem efetivamente da abordagem de otimização robusta, ou se são simplesmente consequência dessa redução do conjunto de dados, conduzimos um experimento de controle.

Para isso, aplicamos ao modelo MK uma estratégia direta de remoção de *outliers*. Inicialmente, calculamos o vetor de retornos médios dos ativos. Em seguida, para cada observação do conjunto de dados, representada pelo vetor de retornos r^k , calculamos sua distância Euclidiana em relação a esse vetor médio. As observações com maiores distâncias foram então removidas segundo as proporções de 20%, 50% e 80%, originando, respectivamente, os modelos $MK_{0.2}$, $MK_{0.5}$ e $MK_{0.8}$.

A Figura 4 apresenta os valores dos retornos médios, volatilidade e Índice Sharpe para as estratégias de construção das carteiras, usando as 3 variações de eliminação de *outliers* aplicadas ao modelo MK, comparando-os aos resultados do modelo clássico MK, já apresentado anteriormente.

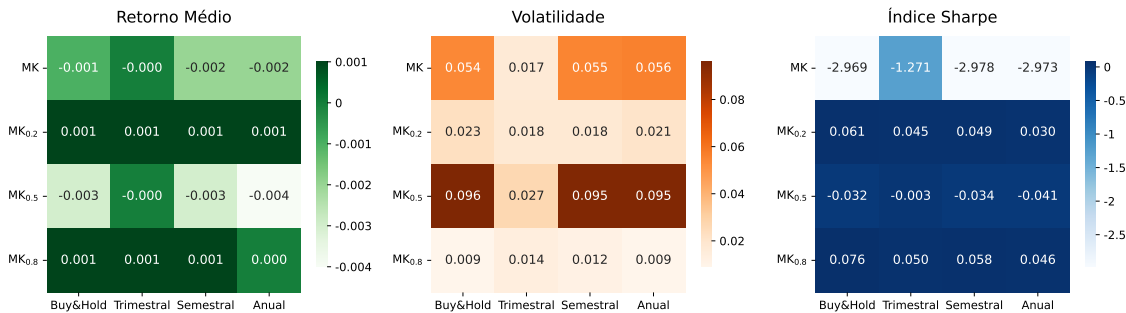


Figura 4. Comparativo de retorno médio, volatilidade e índice Sharpe entre o modelo MK clássico e suas variantes com remoção de *outliers* ($MK_{0.2}$, $MK_{0.5}$ e $MK_{0.8}$) e o modelo $SVC_{0.8}$.

Em relação ao retorno médio, observa-se que os modelos MK e $MK_{0.5}$ apresentam valores negativos em todas as frequências de rebalanceamento analisadas, indicando incapacidade de gerar ganhos consistentes ao longo do período estudado, enquanto o $MK_{0.2}$ apresenta retornos ligeiramente positivos, porém muito próximos de zero. Em contraste, o modelo $SVC_{0.8}$ apresenta retornos médios positivos em todas as estratégias consideradas, sugerindo desempenho superior em termos de geração de retorno.

No que se refere ao risco, medido pela volatilidade, os modelos baseados em Markowitz apresentam valores mais elevados quando comparados ao $SVC_{0.8}$. Em particular, os modelos MK e $MK_{0.5}$ exibem as maiores volatilidades, indicando maior instabilidade das carteiras geradas. O modelo $MK_{0.2}$ apresenta volatilidade intermediária, enquanto o $SVC_{0.8}$ se destaca por apresentar os menores níveis de volatilidade entre todos os modelos analisados, independentemente da frequência de rebalanceamento.

Resultados semelhantes são observados na análise do índice Sharpe. O modelo MK tradicional apresenta valores fortemente negativos em todas as estratégias, caracterizando uma relação risco-retorno desfavorável. O modelo $MK_{0.5}$ também apresenta índice Sharpe negativo, embora de menor magnitude, enquanto o $MK_{0.2}$ apresenta valores positivos. Ainda assim, o modelo $SVC_{0.8}$ apresenta os maiores valores de Sharpe em todas as frequências analisadas, indicando melhor desempenho em termos de retorno ajustado ao risco.

De forma geral, os resultados sugerem que o modelo $SVC_{0.8}$ apresenta desempenho superior às variantes do modelo de Markowitz consideradas, tanto em termos de redução do risco quanto de melhoria da relação risco-retorno. Enquanto os modelos MK permanecem sensíveis à estimação da matriz de covariância e à elevada volatilidade dos ativos, $SVC_{0.8}$ demonstra maior robustez, produzindo carteiras mais estáveis e com melhor relação risco-retorno em todas as frequências de rebalanceamento analisadas. Esses resultados sugerem que a incorporação do critério baseado em SVC com parâmetro ψ elevado constitui uma alternativa promissora à abordagem clássica de Markowitz no contexto dos dados analisados.

5. Conclusão

O modelo proposto mostrou que construir o conjunto de incerteza diretamente a partir dos dados, sem impor hipóteses de simetria sobre a distribuição dos retornos, produz carteiras mais estáveis e diversificadas no cenário analisado. O parâmetro de regularização permite ajustar o grau de conservadorismo da carteira ao perfil do investidor: valores mais altos resultaram em conjuntos de incerteza menores, com melhor controle de risco e desempenho superior ao índice de mercado e aos modelos tradicionais. O experimento com variantes do modelo de Markowitz com remoção de outliers reforça que os ganhos não se explicam apenas pela redução do conjunto de treinamento, mas pela formulação robusta em si. As carteiras geradas também apresentaram maior diversificação e menor rotatividade de ativos, o que reduz os custos de transação na prática.

Ressalta-se, contudo, que a escolha do parâmetro ψ depende das características do conjunto de dados analisado, não sendo possível generalizar os valores obtidos neste estudo para outros mercados ou períodos. Assim, como trabalhos futuros, pretende-se investigar métodos de calibração automática desse parâmetro, além de avaliar o modelo em diferentes classes de ativos e cenários de mercado com distintos níveis de volatilidade. E, também, investigar o impacto de diferentes valores para o coeficiente de aversão ao risco γ sobre o desempenho das estratégias.

Referências

- Ben-Hur, A., Horn, D., Siegelmann, H., and Vapnik, V. (2000). A support vector method for clustering. In *Advances in Neural Information Processing Systems*, pages 346–352.
- Ben-Tal, A., Ghaoui, L., and Nemirovski, A. (2009). *Robust Optimization*. Princeton, NJ.
- Ben-Tal, A. and Nemirovski, A. (1998). Robust convex optimization. *Mathematics of Operations Research*, 23:769–805.
- Bertsimas, D., Gupta, V., and Kallus, N. (2018). Data-driven robust optimization. *Mathematical Programming*, 167(2):235–292.

- Bertsimas, D. and Pachamanova, D. (2008). Robust multiperiod portfolio management in the presence of transaction costs. *Computers & Operations Research*, 35:3–17.
- Bertsimas, D. and Sim, M. (2004). The price of robustness. *Operations Research*, 52:35–53.
- Bertsimas, D. and Thiele, A. (2006). A robust optimization approach to inventory theory. *Operations Research*, 54:150–168.
- Bomfim, U. and Nascimento, F. (2024). Monitoramento do mercado de ativos brasileiro: uma proposta de pipeline de dados para detecção de bolhas financeiras. In *Anais Estendidos do XXXIX Simpósio Brasileiro de Bancos de Dados*, pages 58–64, Porto Alegre, RS, Brasil. SBC.
- Borgards, O., Czudaj, R. L., and Hoang, T. H. V. (2021). Price overreactions in the commodity futures market: An intraday analysis of the covid-19 pandemic impact. *Resources Policy*, 71:101966.
- Chenreddy, A. R., Bandi, N., and Delage, E. (2022). Data-driven conditional robust optimization. In Koyejo, S., Mohamed, S., Agarwal, A., Belgrave, D., Cho, K., and Oh, A., editors, *Advances in Neural Information Processing Systems*, volume 35, pages 9525–9537. Curran Associates, Inc.
- Chopra, V. and Ziemba, W. (1993). The effect of errors in means, variances, and covariances on optimal portfolio choice. *Journal of Portfolio Management - J PORTFOLIO MANAGE*, 19:6–11.
- Danesh, S., Mohammadi, E., Makui, A., and Jafari, M. (2024). Data-driven robust optimization based on position-regulated support vector clustering. *Journal of Computational Science*, 76:102210.
- Delage, E. and Ye, Y. (2010). Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations Research*, 58:595–612.
- Fabozzi, F., Kolm, P., Pachamanova, D., and Focardi, S. (2007). Robust portfolio optimization - recent trends and future directions. *The Journal of Portfolio Management*, 33:40–48.
- Gurobi Optimization, LLC (2026). Gurobi optimizer reference manual. <https://www.gurobi.com/documentation/>. Acesso em: 23 fev. 2026.
- Ledoit, O. and Wolf, M. (2008). Robust performance hypothesis testing with the sharpe ratio. *Journal of Empirical Finance*, 15.
- Markowitz, H. (1952). Portfolio Selection. *Journal of Finance*, 7(1):77–91.
- Sehgal, R. and Jagadesh, P. (2023). Data-driven robust portfolio optimization with semi mean absolute deviation via support vector clustering. *Expert Systems with Applications*, 224:120000.
- Shang, C., Huang, X., and You, F. (2017). Data-driven robust optimization based on kernel learning. *Computers & Chemical Engineering*, 106:464–479.
- Soares, I. R. and Canuto, S. (2025). Criando portfólios de alto desempenho: Otimização de portfólios de ativos de alta volatilidade através da previsão de retorno baseada em

- cnn+bilstm. In *Anais do XL Simpósio Brasileiro de Bancos de Dados*, pages 685–698, Porto Alegre, RS, Brasil. SBC.
- Umar, Z., Gubareva, M., and Teplova, T. (2021). The impact of covid-19 on commodity markets volatility: Analyzing time-frequency relations between commodity prices and coronavirus panic levels. *Resources Policy*, 73:102164.
- Zhang, C., Wang, Z., and Wang, X. (2022). Machine learning-based data-driven robust optimization approach under uncertainty. *Journal of Process Control*, 115:1–11.