



Instituto de
MATEMÁTICA
E ESTATÍSTICA

UFRGS



UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

INSTITUTO DE MATEMÁTICA E ESTATÍSTICA

DEPARTAMENTO DE ESTATÍSTICA

**ESTIMAÇÃO NÃO PARAMÉTRICA DE MATRIZ DE COVARIÂNCIAS
CONDICIONAL: UMA APLICAÇÃO NA OTIMIZAÇÃO DE PORTFÓLIOS EM
FUNDOS MULTIMERCADOS**

TIAGO LUIS JANTSCH

Porto Alegre
2017

TIAGO LUÍS JANTSCH

**ESTIMAÇÃO NÃO PARAMÉTRICA DE MATRIZ DE COVARIÂNCIAS
CONDICIONAL: UMA APLICAÇÃO NA OTIMIZAÇÃO DE PORTFÓLIOS EM
FUNDOS MULTIMERCADOS**

Artigo submetido como requisito
parcial para a obtenção do Grau de
Bacharel em Estatística

Orientador: Prof. Dr. Hudson da Silva
Torrent

Porto Alegre

2017

Instituto de Matemática e Estatística

Departamento de Estatística

**ESTIMAÇÃO NÃO PARAMÉTRICA DE MATRIZ DE COVARIÂNCIAS
CONDICIONAL: UMA APLICAÇÃO NA OTIMIZAÇÃO DE PORTFÓLIOS EM
FUNDOS MULTIMERCADOS**

Tiago Luís Jantsch

Banca examinadora:

Prof. Dr. Eduardo de Oliveira Horta

Universidade Federal do Rio Grande do Sul

Demore o tempo que for para decidir o que você quer da vida, e depois que decidir não recue ante nenhum pretexto, porque o mundo tentará te dissuadir.

Nietzsche, Friedrich

Agradecimentos

Este trabalho representa o ápice de anos de estudo em que pude expandir meus horizontes e contribuir de alguma forma com o universo acadêmico. Primeiramente, quero agradecer à Universidade Federal do Rio Grande do Sul, em especial ao Instituto de Matemática e Estatística pelos profissionais e pela infraestrutura de qualidade disponibilizada gratuitamente; ao meu orientador Prof. Dr. Hudson Torrent, pela sugestão de pesquisa, e pelo apoio essencial à realização deste trabalho; aos meus familiares que mesmo distantes e sem compreender exatamente o que desenvolvi ao longo deste período, sempre acreditaram e confiaram no meu potencial. Aos amigos pelas revisões e discussões durante a elaboração deste trabalho; não caberá citar nomes, pois todos tiveram sua contribuição no meu desenvolvimento como profissional de estatística.

Resumo

Este trabalho fundamenta-se na teoria moderna do portfólio proposta por Markowitz (1952), que trata da diversificação de investimentos na relação entre a minimização do risco e na maximização dos retornos atrelados a uma estratégia de otimização. Este estudo, visando diminuir os erros de estimação do modelo de média-variância clássico, propõe uma reformulação na estimação da matriz de covariâncias baseado na estimação das variâncias e covariâncias condicionais a partir de estimadores não paramétricos *Kernel*. Sua aplicação foi feita com base nos retornos diários de 359 fundos multimercados no período de 02 de julho de 2012 e 31 de dezembro de 2016. Por ser um mercado com uma alta volatilidade, fez-se necessário um rebalanceamento da carteira, a fim de corrigir as flutuações dos ativos. Por fim, essas métricas modeladas foram comparadas pelo retorno e índice Sharpe e também, com os benchmarks IBOVESA e CDI, no mesmo período. Conclui-se que no âmbito dos retornos a estratégia não paramétrica apresentou 5,36% de retorno superior ao modelo de média-variância, entretanto o seu índice Sharpe foi inferior ao modelo tradicional.

Palavras-chave: Markowitz, Estimativa de Densidade por Kernel, Otimização de carteiras, Média-Variância

Abstract

This paper is based on the modern portfolio theory proposed by Markowitz (1952), which deals with the diversification of investments in the relationship between risk minimization and the maximization of returns tied to an optimization strategy. This study, aimed at reducing estimation errors of the classical mean-variance model, proposes a reformulation in the estimation of the covariance matrix based on the estimation of variances and conditional covariates from non-parametric Kernel estimators. Its application was made based on the daily returns of 359 multimarket funds in the period of July 2, 2012 and December 31, 2016. As a market with high volatility, it was necessary to rebalance the portfolio in order to correct Fluctuations of assets. Finally, these model metrics were compared by the return and Sharpe index and also, with the IBOVESPA and CDI benchmarks, in the same period. It was concluded that the non-parametric strategy presented a return return of 5.36% higher than the mean-variance model, although its Sharpe index was lower than the traditional model.

Keywords: Markowitz, Kernel Density Estimation, Portfolio Optimization, mean-variance

Sumário

1. Introdução.....	9
2. Modelagem.....	10
2.1 Excesso dos Retornos.....	11
2.2 Modelo de Média-Variância.....	11
2.3 Índice de Sharpe.....	14
2.4 Estimação das Covariâncias.....	15
3. Aplicação Empírica.....	21
3.1 Coleta e Tratamento dos Dados.....	21
3.3 Resultados.....	23
4. Conclusões.....	28
Referências.....	29
Anexo I – Tabelas.....	32
Anexo II – Código fonte do modelo de média-variância clássico.....	34
Anexo III – Código fonte do modelo de média-variância NP.....	37

1. Introdução

Estratégias de investimentos despertam o interesse de pessoas que aplicam recursos financeiros. Em termos de otimização, muito se tem discutido sobre a diversificação de uma carteira de investimentos para a geração de portfólios com uma melhor relação entre retorno e risco.

Popularizado como Teoria Moderna do Portfólio (**TMP**), o modelo de média-variância, introduzido por Markowitz (1952), apresenta uma estrutura de otimização de um portfólio de ativos com base nos retornos históricos e no *trade-off* entre retorno e risco. A concepção básica da **TMP** é a não preferir por ativos singularmente e sim por diversificar. Em termos práticos, Markowitz (1952), descreve que quando dois ativos pouco relacionados compõem uma carteira, o risco é menor do que a média ponderada dos riscos singulares. Sendo em alguns casos um risco inferior que o do ativo de menor risco com um retorno superior do que este ativo.

É importante considerar que a construção de portfólios ótimos é de fundamental importância para a aplicação de recursos no mercado financeiro. Assim, é necessária a adoção de modelos e métodos eficientes estatisticamente. Neste âmbito, diversos investidores demandam tempo e dinheiro a fim de encontrar um mecanismo de seleção de portfólio ótimo.

Apesar de ser uma estratégia consolidada e amplamente discutida no ramo acadêmico, o modelo de Markowitz sofre diversas críticas. Segundo Best & Grauer (1991), o erro de estimação advinda das médias amostrais sobre retornos históricos dos ativos pode afetar diretamente a qualidade destes estimadores na composição ótima do portfólio. A estimativa da matriz de covariância é um delicado desafio estatístico que requer métodos sofisticados (LEDOIT e WOLF, 2004). Como consequência, sua *performance* é diretamente afetada. Diversos autores propuseram mudanças no modelo tradicional por meio de ajustes nas pequenas variações das estimativas, a uma relação de rotatividades nas ponderações dos ativos que compõem a carteira final (BROADIE, 1993; CORNUEJÓLIS e TUTUNCU, 2007).

Neste contexto, de modo a contribuir com uma estratégia de otimização de um portfólio, este artigo apresenta uma proposta visando corrigir problemas na estimação da matriz de covariância. A proposta se baseia, embora não seja igual, àquela apresentada em Salah *et. al.* (2016), que demonstra uma nova alternativa para estimação da matriz de covariâncias de Markowitz (1952) através de uma estimação não paramétrica via *Kernel*.

No âmbito da sua aplicação, o objetivo proposto, foi em cotas de fundos de investimentos multimercados¹, em virtude da importância e do crescimento desta modalidade no mercado financeiro brasileiro (CALDEIRA *et al.*, 2014). Finalmente, o método tradicional e o método proposto foram comparados pelo retorno e índice Sharpe e também, com os *benchmarks* IBOVSPA² e CDI³, no mesmo período.

2. Modelagem

Este artigo propõe a otimização de uma carteira por meio da diversificação nos investimentos, com a intenção de maximizar os retornos correndo baixos riscos. Na literatura, diversas técnicas de análise são apresentadas, por exemplo, média-variância, mínima-variância, *Bayes-Stein*, entre outras. Este estudo utilizou a otimização via média-variância proposta por Markowitz (1952) com uma nova alternativa para estimação da matriz de covariâncias. A seguir, serão apresentados os conceitos básicos e as técnicas utilizadas neste estudo para a montagem da carteira.

¹São fundos que têm liberdade para operar diferentes ativos, papéis de renda fixa, ações de empresas, moedas, derivativos e investimento no exterior.

²Índice oriundo de uma carteira teórica de ativos elaborada pela BM&FBovespa.

³Certificado de Depósito Interbancário - taxa referencial para avaliar a rentabilidade das aplicações.

2.1 Excesso dos Retornos

Devido à grande instabilidade do mercado brasileiro, faz-se justificável, neste estudo, utilizar a técnica de excesso dos retornos, com a finalidade de definir uma carteira que avalie o impacto entre os investimentos de baixo risco. Foi utilizado, neste estudo, o excesso dos retornos atrelados à taxa CDI, sendo a medida do excesso de retorno definida como:

$$R_a - R_f$$

Onde R_a é o retorno do investimento do ativo e R_f é a taxa de retorno do ativo livre de risco.

2.2 Modelo de Média-Variância

O modelo tradicional de média-variância de Markowitz (1952), que será denotado por **MV**, é o ponto de partida para a construção de portfólios ótimos. Em seu formato mais básico, o modelo determina a proporção ideal de capital no m -ésimo ativo contido no portfólio, que será denotado por w_m . No entanto, são necessárias algumas restrições, como a não aceitação da venda a descoberto, ou seja, $w_m \geq 0$; e também que o portfólio, em sua totalidade, deva ser investido, isto é,

$$\sum_{m=1}^N w_m = 1$$

Onde N representa o total de ativos contidos na carteira. A equação que apresenta o retorno esperado da carteira é dada por:

$$E = \sum_{m=1}^N w_m \mu_m \quad (1.1)$$

Onde μ_m é o retorno esperado de cada ativo. O retorno esperado de uma carteira, equação (1.1), é simplesmente uma média ponderada dos retornos individuais.

Na construção de um portfólio têm-se m ativos candidatos, cujos retornos serão denotados por r_{jt} , isto é, r_{jt} denota o retorno do patrimônio j na data t , sendo $t = 1, \dots, N$. Definido isso, a estimativa da matriz de variância-covariância dos retornos ($r_1 \dots r_m$), que será denotada por \mathbf{M} é apresentada na equação 1.2:

$$\mathbf{M} = \frac{1}{N} \sum_{t=1}^N \begin{bmatrix} (r_{1t} - \bar{r}_1)^2 & (r_{1t} - \bar{r}_1) \cdot (r_{2t} - \bar{r}_2) & (r_{1t} - \bar{r}_1) \cdot (r_{mt} - \bar{r}_m) \\ (r_{2t} - \bar{r}_2) \cdot (r_{1t} - \bar{r}_1) & (r_{2t} - \bar{r}_2)^2 & (r_{2t} - \bar{r}_2) \cdot (r_{mt} - \bar{r}_m) \\ \vdots & \vdots & \vdots \\ (r_{mt} - \bar{r}_m) \cdot (r_{1t} - \bar{r}_1) & (r_{mt} - \bar{r}_m) \cdot (r_{2t} - \bar{r}_2) & (r_{mt} - \bar{r}_m)^2 \end{bmatrix} \quad (1.2)$$

Definidas as condições iniciais do modelo de média-variância de Markowitz (1952), o seu modelo de otimização que visa minimizar uma função objetivo está sujeita a:

$$\min_w w^T \mathbf{M} w \quad (1.3)$$

Sujeita à,

$$w^T \boldsymbol{\mu} = E$$

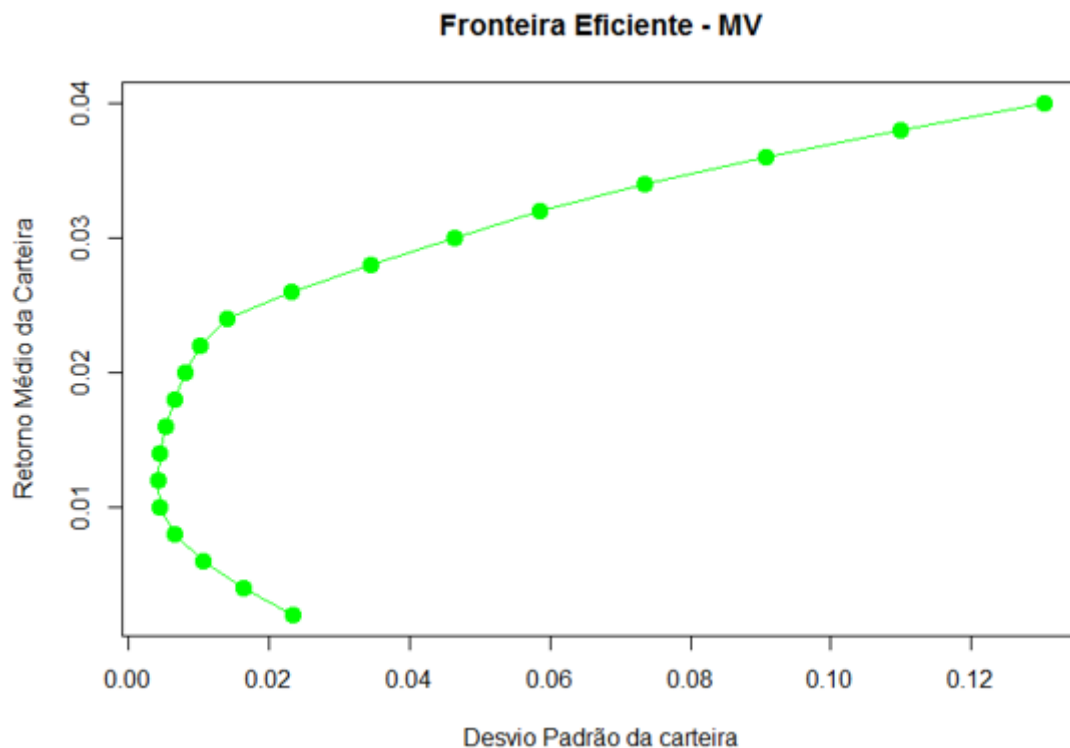
$$w^T \mathbf{1} = 1$$

$$w_m \geq 0$$

Onde, $w^T = (w_1, \dots, w_m)$ é o vetor de pesos do portfólio e, $\mu^T = (\bar{r}_1, \dots, \bar{r}_m)$ são os retornos médios empíricos.

Resolvendo o problema por diversas vezes sobre as restrições propostas e, com um valor esperado diferente em cada problema, adquire-se, associado a cada nível de E , a composição do portfólio de menor risco. Analisando graficamente a relação do risco versus retorno da otimização do problema (modelo de programação quadrática), é possível traçar uma curva entre os portfólios dos ativos, definida como **fronteira eficiente**. Com esta fronteira eficiente o investidor pode decidir qual nível de risco deseja assumir associado a um retorno desejável, como pode ser visto na Figura 1.

Figura 1: Fronteira de Markowitz



Fonte: Elaborada pelo autor.

2.3 Índice de Sharpe

Na literatura são apresentadas diversas técnicas que mensuram performances de portfólios. O Índice Sharpe é uma das técnicas que mensuram a eficiência dos investimentos composto pelo retorno esperado, pelo ativo livre de risco e pela variação da carteira (desvio padrão). O índice Sharpe (Sharpe, 1966), representado por **IS**, é apresentado pela seguinte equação:

$$IS = \frac{(E(R_a) - R_f)}{\sigma_a} \quad (1.4)$$

Onde, $E(R_a)$ é o retorno esperado da estratégia a ; R_f é o retorno do ativo livre de risco (*risk-free*) e σ_a é o risco do ativo (desvio padrão) dos retornos da carteira da estratégia a . O **IS** mensura o prêmio oferecido por um ativo para cada unidade adicional de risco assumido (ASSAF NETO, 1999, p. 294).

Quanto maior for o valor do **IS**, mais eficiente é o desempenho da estratégia. A fim de testar se há igualdade no desempenho das estratégias, será realizado um teste estatístico entre os diferentes **IS**, com o intuito de verificar se há diferença significativa entre as duas medidas. Um teste de hipóteses sugerido por Jobson e Korbie (1981) trata de avaliar a diferença entre os **IS** das carteiras otimizadas. Para duas carteiras de índices (i, j) foi testada a seguinte hipótese:

$$H_0: Sh_i - Sh_j = 0 \quad (1.5)$$

$$H_1: Sh_i - Sh_j \neq 0$$

Para a estatística de índices de Sharpe, a distribuição sob a hipótese nula é normal com média dada por:

$$Sh_{ij} = S_j \bar{r}_i - S_i \bar{r}_j.$$

Sendo \bar{r} a média da série de índices e $S_{i,j}$ o desvio padrão amostral da série de índices (i, j) ; e variância dada por:

$$\theta = \frac{1}{T} \left[2\sigma_i^2 \sigma_j^2 - 2\sigma_i \sigma_j \sigma_{ij} + \frac{1}{2} \mu_i^2 \sigma_j^2 + \frac{1}{2} \mu_j^2 \sigma_i^2 - \frac{\mu_i \mu_j}{2\sigma_i \sigma_j} (\sigma_{ij}^2 + \sigma_i^2 \sigma_j^2) \right]$$

Onde T é o número de observações; $\sigma_{i,j}^2$ é a variância amostral da série de índices (i, j) ; σ_{ij} é a covariância amostral entre os índices; $\sigma_{i,j}$ é o desvio padrão amostral da série de índices e $\mu_{i,j}$ é a média da série de índices (i, j)

A estatística é dada por: $z_{Sh_{ij}} = \frac{Sh_{ij}}{\sqrt{\theta}}$

2.4 Estimação das Covariâncias

Para implementar as estratégias quantitativas de otimização em portfólios no modelo tradicional de média-variância, é necessário estimar médias e covariâncias amostrais dos retornos e, após, aplicá-las em uma solução analítica para a otimização. Sabendo disto, quanto maior o erro de estimação, maior será o erro na composição do portfólio (BEST & GRAUER, 1991).

A abordagem apresentada em Salah *et. al.* (2016) visa apresentar uma alternativa válida para o processo de estimação da matriz de covariâncias do modelo de média-variância. Este portfólio é obtido minimizando o chamado

risco condicional, sendo a ideia principal antecipar os valores dos retornos dos ativos na data $t + 1$ (conhecer o passado) e incorporar essa informação no modelo de otimização. Para constituir a matriz de covariâncias proposta por Salah *et. al.* (2016), inicia-se apresentando um modelo de regressão local (equação 1.6).

Tradicionalmente, em regressão, se deseja investigar a associação entre uma variável resposta (y) e uma variável explicativa (x). No cenário clássico, como em Mínimos Quadrados Ordinários (MQO), supõem-se que essa relação seja dada de forma linear. Quando a sua forma funcional for de natureza desconhecida, uma alternativa à utilização de regressão não paramétrica que utiliza de estimadores locais baseados em uma função *kernel*. O método univariado da estimação pode ser observado ao considerar o modelo de regressão dado:

$$Y_i = m(X_i) + \varepsilon_i \quad (1.6)$$

Onde ε é o termo de erro tal que $E(\varepsilon_i|X_i = x) = 0$, $Var(\varepsilon_i|X_i = x) = \sigma^2(x)$, Y representa a variável resposta, e X a covariável. A regressão local fará a estimativa da função $m(\cdot)$ usando dados na vizinhança de cada ponto de interesse x . Para o cálculo desta estimativa são necessárias duas escolhas: a densidade da função *Kernel* e o parâmetro h que determina o grau de suavidade das estimativas. Na literatura, são apresentadas diversas maneiras de estimar $m(\cdot)$. Será adotada a estimação do tipo Nadaraya-Watson (Nadaraya, 1964 e Watson, 1964) que estima uma função localmente considerando a média ponderada das observações que estão na vizinhança do ponto x . O estimador de Nadaraya-Watson para $m(x)$:

$$\hat{m}_{nw}(x) = \frac{\sum_{t=1}^T Y_t \cdot K_H(X_t - x)}{\sum_{t=1}^T K_H(X_t - x)} \quad (1.7)$$

Sendo $Y = (Y_1, \dots, Y_m)$ um vetor aleatório e a função *Kernel* uma função de densidade de probabilidade simétrica ao redor de zero, ou seja, satisfaz as seguintes suposições:

$$\int K(u)dx = 1$$

$$\int uK(u)dx = 0$$

$$K(u) = K(-u)$$

Utilizando da estimação não paramétrica via *Kernel*, onde o interesse está em estimar a matriz de covariâncias condicional, considere uma série estacionária m -multivariada $\{(r_{1t}, \dots, r_{mt}), t = 1, \dots, N\}$, onde se quer prever os valores dos retornos dos ativos na data $t + 1$. Primeiramente se faz necessário estimar a esperança condicional para cada j , onde $j = 1, \dots, m$ e um dado x .

$$\hat{m}_j(x) = \frac{\sum_{t=2}^N r_{jt} \cdot k\left(\frac{r_{jt-1}-x}{h_j}\right)}{\sum_{t=2}^N k\left(\frac{r_{jt-1}-x}{h_j}\right)} \quad (1.8)$$

Com $r_{jt} = m_j(r_{jt-1}) + \varepsilon_{jt}$ para $j = 1, \dots, m$ e $t = 2, \dots, N$, isto é, um auto regressivo com 1 defasagem, onde o regressor é o valor do retorno da carteira j no dia $t - 1$ e o regressando é o valor do retorno da carteira j no dia t . Sendo a componente residual dada por: $\hat{\varepsilon}_{jt} = r_{jt} - \hat{m}_j(r_{jt-1})$. A diagonal principal da matriz de covariâncias condicional é dada pela estimação das variâncias condicionais, definida por:

$$\sigma^2(x) = E((Y - m(x))^2 | X = x) \quad (1.9)$$

Onde o seu estimador *Kernel* univariado será dado por:

$$\hat{\sigma}^2(x) = \frac{\sum_{t=2}^T (\hat{\varepsilon}_{jt}^2) \cdot k\left(\frac{r_{jt-1}-x}{h_j}\right)}{\sum_{t=2}^T k\left(\frac{r_{jt-1}-x}{h_j}\right)} \quad (2.0)$$

A estimação tem o interesse em fazer x igual ao último valor observado para a amostra dos retornos, pois a ideia é estimar o dia $t+1$.

Para a estimação da covariância condicional o método de estimação através de regressão local pode ser facilmente estendido para duas covariáveis onde o modelo de regressão a ser estimado assume a forma:

$$Y_i = m(X_{1i}, X_{2i}) + \varepsilon_i \quad (2.1)$$

Sendo x_1 e x_2 duas variáveis regressoras e ε é o termo de erro tal que:

$$E(\varepsilon_i | X_{1i} = x_1, X_{2i} = x_2) = 0$$

$$Var(\varepsilon_i | X_{1i} = x_1, X_{2i} = x_2) = \sigma^2(x_1, x_2)$$

No caso bivariado a estimação local pode ser definida através do produto entre duas funções *kernel* univariadas para cada covariável $K_{h_1}(X_{1t} - x_1)K_{h_2}(X_{2t} - x_2)$. Assim, o estimador de $m(x_1, x_2)$ será dado por $\hat{m}(x_1, x_2)$, como no caso univariado. Geralmente, a utilização deste método é

o suficiente, caso tenha interesse em incluir mais covariáveis a estimação pode apresentar o problema da *curse of dimensionality*, ver Ziegelmann (2002).

A covariância condicional entre Y_1 e Y_2 dado $X = x$ é definida como:

$$\sigma_{12}(x) = E((Y_1 - m_1(x))(Y_2 - m_2(x))|X = x) \quad (2.2)$$

Onde o seu estimador *Kernel* é dado por:

$$\hat{\sigma}_{js}(x_j, x_s) = \frac{\sum_{t=2}^T (\hat{\varepsilon}_{jt}) \cdot (\hat{\varepsilon}_{st}) \cdot k\left(\frac{r_{jt-1} - x_j}{h_j}\right) \cdot k\left(\frac{r_{st-1} - x_s}{h_s}\right)}{\sum_{t=2}^T k\left(\frac{r_{jt-1} - x_j}{h_j}\right) \cdot k\left(\frac{r_{st-1} - x_s}{h_s}\right)} \quad (2.3)$$

O interesse na equação 2.3 é em considerar x_j e x_s iguais aos últimos valores disponíveis para os ativos j e s em cada janela, empregado em um produto de componentes univariadas, onde para cada par $(j, s) \forall j, s = 1, \dots, m$ e $j \neq s$.

A escolha da função de densidade do *Kernel* é apresentada na literatura por diversas funções, uniforme, triangular, quadrático (*biweight*), cúbico (*triweight*), *epanechnikov*, *gaussiano*. Na prática o *epanechnikov* e o *gaussiano* são os mais usados. O *Kernel epanechnikov* (EPANECHNIKOV, 1969) é apresentado como sendo o ótimo, entretanto, no estudo de Fan *et. al.* (1996), avalia os *kernels* subótimos para a estimação de densidade, e estes têm a mesma eficiência assintótica, sendo assim, este estudo utilizará do *gaussiano* por ser mais difundido na literatura.

A escolha da matriz de largura de banda h (*Bandwidth*) é o fator mais importante na precisão do estimador de *Nadaraya-watson*, uma vez que controla a orientação e a quantidade de alisamento induzida. A escolha por um h muito pequeno, os pontos mais afastados acabam por influenciar pouco na

estimativa gerando um modelo com baixa capacidade de generalização. Para uma largura grande, a suavização acaba não demonstrando a relação local entre a variável explicada e as explicativas e acaba se aproximando da estimação paramétrica.

Hall e Marron (1987) apresentam um método de escolha do parâmetro de alisamento baseado na minimização do erro quadrático médio integrado. Este conceito é aperfeiçoado por Sheather (1991) que traz uma reformulação no termo não estocástico, previamente omitido, o qual reduz o viés na estimativa sem inflar a variância.

Outras técnicas de estimação para determinar o h do *kernel* podem ser encontradas em Molanes-Lopez(2008). No presente estudo, foi tomado como base o conceito tratado por Scott (1992) que se utiliza do erro quadrático médio assintótico (AMISE)⁴.

Neste artigo não são demonstradas algebricamente as técnicas utilizadas até chegar à equação (2.4). Caso o leitor tenha interesse, as manipulações algébricas detalhadas podem ser encontradas em Scott (1992, p. 165).

$$h_j = \left(\frac{4\hat{\sigma}_j^5}{3T_j} \right)^{\frac{1}{5}} \approx 1.06\hat{\sigma}_j T_j^{-1/5} \quad (2.4)$$

Sendo T o tamanho da série histórica no ativo j e $\hat{\sigma}$ o desvio padrão da série no ativo j .

Por fim, na construção do novo portfólio, que será denotada por **NP**, usando a estimação não paramétrica via *kernel* no modelo de média-variância de Markowitz, a estrutura ficará da seguinte forma:

$$\min_w w_c^T \mathbf{M}_c w_c \quad (2.5)$$

⁴ O resultado presente na equação (2.4) é referente à estimação de densidade. Utiliza-se esse resultado aqui por simplicidade. Critérios mais apropriados para estimação de médias, variâncias e covariâncias condicionais serão abordados em trabalhos futuros.

Sujeita à,

$$w_c^T \mu_c = E$$

$$w_c^T \mathbf{1} = 1$$

$$w_{c,m} \geq 0$$

Onde, $\mu_c = (\bar{r}_{c,1}, \dots, \bar{r}_{c,m})^T$, E é o valor esperado dos retornos do portfólio requerido pelo investidor, \mathbf{M}_c é a matriz de covariâncias condicional cujos elementos agora são $\hat{\sigma}_{c,ij}(r_{1N}, \dots, r_{mN})$.

Resolvendo o problema inúmeras vezes, agora sobre as novas condicionalidades propostas, obtêm-se uma nova fronteira eficiente.

Na próxima seção, será apresentada a aplicação da teoria abordada.

3. Aplicação Empírica

Nesta etapa, foi elaborado um estudo com a aplicação dos conceitos da *Modern Portfolio Theory* para a montagem da carteira nos retornos dos fundos multimercados. Complementando, o interesse nessa pesquisa é comparar o desempenho das diferentes estratégias usadas, perante o CDI e IBOVESPA. Para a execução das análises foi utilizado o *software* R Studio versão 0.99.467.

3.1 Coleta e Tratamento dos Dados

Nesta etapa, foi elaborado um estudo com a base de dados final composta por 359 fundos de investimentos multimercados extraídas da plataforma da Economatica⁵, no período de 02 de julho de 2012 a 31 de dezembro de 2016. Os dados coletados perfazem um total de 1.113

⁵ Plataforma que disponibiliza dados do mercado financeiro.

observações compreendidas neste intervalo. Também foram coletados os retornos do índice IBOVESPA e do CDI para o mesmo período, a fim de utilizar como *benchmark* entre as carteiras. Os critérios de seleção dos fundos foram os seguintes:

- a) ser classificado na Comissão de Valores Mobiliários (CVM) como fundo multimercado;
- b) estar ativo em 31 de dezembro de 2016;
- c) apresentar Patrimônio Líquido superior a R\$ 30 milhões em 31 de dezembro de 2016.

Após esta classificação, restaram 674 fundos multimercados. Em decorrência da restrição de alguns fundos, foram excluídos os que tiveram ausência de informações na série histórica e dependência linear, restando assim os 359 fundos de investimentos multimercados os quais compuseram a amostra final.

Os fundos multimercados presentes na amostra são descritos na Tabela 1 pela sua classificação na ANBIMA⁶:

Tabela 1: Frequência das Características dos Fundos.

Classificação	Valor Absoluto	Percentual (%)
Trading	1	0,3
Balanceados	2	0,6
L/S - Direcional	2	0,6
L/S - Neutro	2	0,6
Estrat. Específica	3	0,8
Dinâmico	10	2,8
Macro	27	7,5
Juros e Moedas	40	11,1
Livre	101	28,1
Invest. no Exterior	171	47,6
Total	359	100

Fonte: Elaborada pelo autor com dados da ANBIMA.

⁶ Associação Brasileira das Entidades dos Mercados Financeiro e de Capitais.

3.2 Otimização

Para o procedimento de otimização dos portfólios, foi utilizado o conceito de *rolling-sample*. Neste estudo, foram definidos, D como dias de retornos, sendo $D = 1113$; e J para o tamanho de uma janela, onde $J = 500$ foi utilizado para a estimação dos parâmetros. Para cada dia d , iniciando em $d = J + 5$ adotou-se os 500 dias anteriores para estimar os parâmetros de cada método. O *rolling-sample* final será uma série de $D - J$ janelas de retornos diários gerados para cada estratégia. Por ser um mercado com uma alta volatilidade, fez-se necessário um rebalanceamento da carteira há cada 5 dias, a fim de corrigir as flutuações dos ativos.

Outro procedimento que se fez necessário, foi o incremento de um fator a banda de alisamento nas janelas, nas quais a matriz de covariâncias condicional não é positiva definida. Isso se equivale a todos os autovalores da \mathbf{M}_c serem positivos. Devido à capacidade computacional do *software R Studio* em alguns campos (i, j) da matriz de covariâncias condicional foi atribuído NA^7 , nestes casos foi feita uma substituição singular utilizando a estimação via covariância amostral.

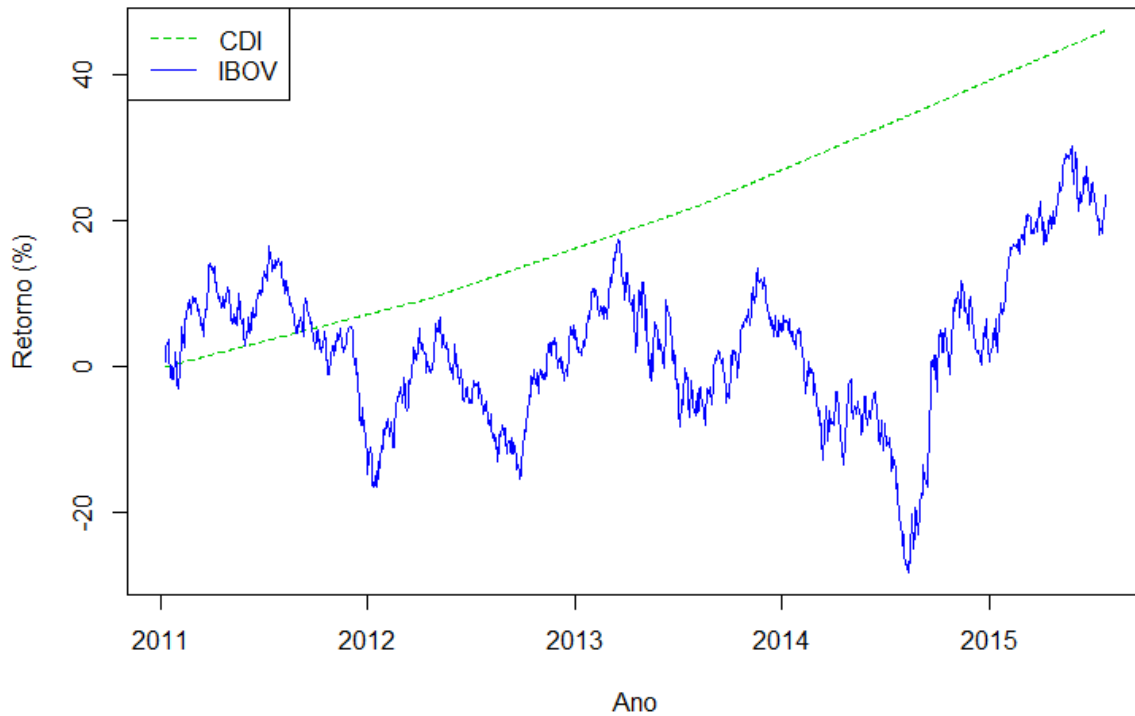
3.3 Resultados

O objetivo principal deste estudo foi comparar as otimizações dos portfólios formados e verificar suas características em termos empíricos. Sabe-se, a priori, que o mercado brasileiro é considerado por muitos investidores como possuidor de uma alta volatilidade. Um detalhamento desta oscilação é apresentado na Figura 2 das séries históricas do CDI e IBOVESPA. Neste período, investimentos atrelados ao CDI tiveram um retorno de 29,33% com um desvio padrão de 0,004%. Já investimentos atrelados ao índice IBOVESPA

⁷ *Not Available*. Não é um número, no entanto, geralmente é usado para valores ausentes, ou indeterminações numéricas.

trouxeram um retorno de 16,79% e um desvio padrão de 1,646%, apresentados na Tabela 2.

Figura 2: Retornos Acumulados CDI e Ibovespa



Fonte: Elaborada pelo autor.

A Tabela 2 apresenta os resultados das estatísticas descritivas dos excessos de retorno de cada estratégia. A estratégia que expôs maior retorno médio foi o modelo **NP** com 0,060% de excesso de retorno ao dia. Em compensação, este modelo apresentou um desvio padrão (volatilidade) superior ao **MV**, sendo um indicativo de maior risco associado.

Tabela 2: Estatísticas Descritivas.

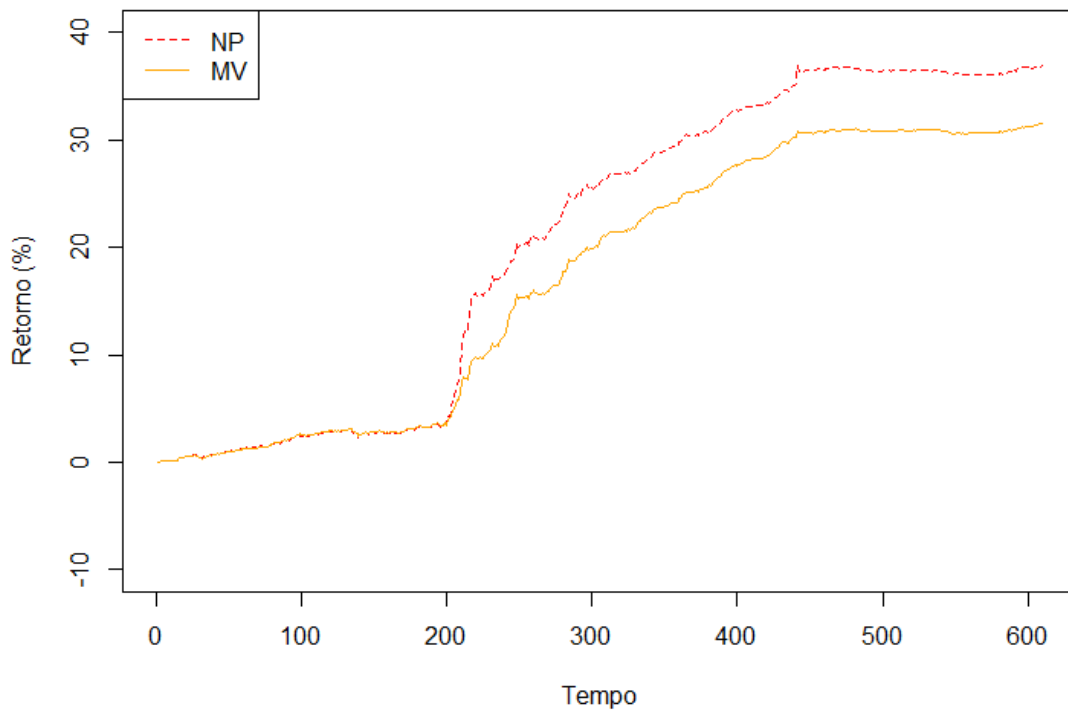
Estratégias	MV	NP	CDI	IBOVESPA
Média	0,052	0,060	0,047	0,027
Mediana	0,022	0,020	0,050	0,000
Desvio Padrão	0,141	0,200	0,004	1,646
Mínimo	-0,463	-0,618	0,040	-4,870
Máximo	1,196	2,152	0,050	6,600

Fonte: Elaborada pelo autor.

Outro fator importante que vale ser destacado é o fato dos excessos de retornos nos dois modelos serem positivos, o que significa que em ao menos empiricamente restringindo-se ao dados ora analisados o CDI, investimento de baixo risco, é menos eficaz que as estratégias propostas.

A Figura 3 apresenta as séries dos retornos acumulados de cada estratégia. As duas estratégias nas 200 observações iniciais apresentaram um desempenho muito similar. Após, o modelo **NP** teve um desempenho superior ao modelo de média-variância e manteve esta superioridade durante toda a série analisada. A rentabilidade acumulada pela estratégia **NP** foi de 36,88% contra 31,52% do método **MV**, apresentando um *trade-off* de 5,36% entre as duas estratégias.

Figura 3: Retornos Acumulados MV e NP



Fonte: Elaborada pelo autor.

A Tabela 3 contém o índice Sharpe para cada estratégia proposta. Os índices foram obtidos através da função “Sharpe Ratio” do pacote *PerformanceAnalytics* do R.

Tabela 3: Sharpe Ratio.

Estratégias	MV	NP
Sharpe Ratio	0,366	0,301

Fonte: Elaborada pelo autor.

Cabe ressaltar que quanto maior o valor do índice Sharpe, melhor é o desempenho da estratégia. Sendo assim, a estratégia **MV** possui maior valor para o **IS**, indicando que a mesma tem o melhor desempenho em relação ao risco versus retorno. Com o intuito de verificar se há diferença significativa entre as duas medidas foi realizado o teste de hipóteses proposto por Jobson e

Korbie (1981). Sob a hipótese nula de nenhuma diferença no desempenho dos índices, o valor da estatística de teste é $z_{sin} = 2.708533$. Estima-se com um nível de significância de 1% que as médias de desempenho dos índices diferem entre si.

Caso o leitor tenha interesse nas séries completas do excesso dos retornos e dos **IS**, foram disponibilizadas, em anexo, as tabelas 4 e 5 que podem auxiliar na interpretação.

4. Conclusões

Este trabalho propôs uma avaliação empírica de diferentes métodos de seleção de portfólios, utilizando os resultados do modelo de média-variância adaptando a estrutura de covariância ao método de estimação não paramétrico via *Kernel*, denominado **NP** e comparou seus resultados.

Dando sequência à estratégia proposta por Salah *et. al.* (2016), a qual aconselha aplicar em outras bases de dados, a aplicação realizada em fundos multimercado brasileiros apresentou resultados falta uma palavra com a nova técnica de estimação.

Analisando quanto ao excesso de retorno acumulado, os resultados foram satisfatórios, obtendo um ganho quanto à técnica tradicional. Em relação ao índice Sharpe, o modelo tradicional apresentou melhor *performance*.

Sugere-se que trabalhos futuros abordem a questão, da estimação de uma banda de alisamento “ótima” para o parâmetro suavizador no modelo **NP**, a qual poderá apresentar outros resultados. A aplicação desta técnica sobre outras bases de dados, como a Bolsa de Valores, fundos quantitativos⁸, entre outros, e também uma alteração na frequência dos rebalanceamentos podem levar a remates diferentes dos aqui encontrados.

Por fim, cabe ressaltar que rentabilidade passada não garante rentabilidade futura e também que grandes investidores – entenda-se aqui pessoas que possuem ganhos expressivos no mercado financeiro – estimam suas amostras com base em técnicas fundamentalistas minimizando ainda mais o seu risco.

⁸ São fundos geridos por “máquinas pensantes”, orientados por modelos matemáticos e estatísticos.

Referências

ANBIMA. **Associação Brasileira das Entidades dos Mercados Financeiro e de Capitais**. Disponível em: http://www.anbima.com.br/pt_br/index.htm. Acesso em: 23 abr. 2017.

ASSAF NETO, Alexandre. **Mercado Financeiro**. 3ª ed. São Paulo: Atlas, 1999.

BEST, M. J., GRAUER, R. R. *Positively weighted minimum-variance portfolios and the structure of asset expected returns*. **Journal of Financial and Quantitative Analysis**, 27, 513-537, 1992.

BROADIE, M. *Computing efficient frontiers using estimated parameters*. **Annals of Operations Research**, 45: 21-58. 1993.

CALDEIRA, João Frois; MOURA, Guilherme Valle; SANTOS, André Alves Portela e TESSARI, Cristina. Seleção de carteiras com modelos fatoriais heterocedásticos: aplicação para fundos de fundos multimercados. **RAM, Rev. Adm. Mackenzie** [online]. 2014, vol.15, n.2, pp.127-161.

CORNUEJÓLIS, G.; TÜTÜNCÜ, R. **Optimization methods in finance**. Cambridge University Press, Cambridge, UK. 2007.

ECONOMATICA. Disponível em: <https://economica.com/>. Acesso em: 05 mai. 2017.

EPANECHNIKOV, V. A. *Non-parametric estimation of multivariate probability density*. **Theory of Probability and Its Applications**, 14, 153-158, 1969.

FAN, J., Gasser, T., Gijbels, I., Brockmann, M. and Engel, J. *Local polynomial fitting: optimal kernel and asymptotic minimax efficiency*. **The annals of the Institute of Statistical Mathematics**, 49, 79-99, 1996.

HALL, P. e Marron, J.S. *Estimation of integrated squared density derivatives*. **Statist. Probab. Lett.**, 6, 109-115, 1987.

JOBSON, J. D., KORKIE, Bob M. *Performance Hypothesis Testing with the Sharpe and Treynor Measures. **The Journal of Finance***, Vol. 36, No. 4 (Sep., 1981), p. 889-908.

LEDOIT, O., WOLF, M. Honey, *I shrunk the sample covariance matrix. **Journal of Portfolio Management***, 30, 110-119, 2004.

MARKOWITZ, H.M. *Portfolio section. **The Journal of Finance***, Vol. 7, No. 1. (Mar., 1952), pp. 77-91.7:77-91.

MOLANES-LÓPEZ, E. M.; CAO, R. *Plug-in bandwidth selector for the kernel relative density estimator. **Annals of the Institute of Statistical Mathematics***, [S.I.], v.60, n.2, p.273–300, 2008.

NADARAYA, E. A. *On Estimating Regression. **Theory of Probability and its Applications***, v. 9, p. 141-142,1964.

SALAH, Hanene. GANNOUN, Ali.DE PERERRI, Christian.RIBATER, Mathieu. ***Conditional Mean-Variance and Mean-Semivariance models in portfolio optimization***. 2016. <hal-01404752>.

SHARPE, W. (1966). *Mutual fund performance. **Journal of Business***. v. 39: 119-138.

SHEATHER, S. J. and JONES, M. C. (1991) *A reliable data-based bandwidth selection method for kernel density estimation. **Journal of the Royal Statistical Society** series B*, 53, 683–690.

SCOTT, D. W. (1992) *Multivariate Density Estimation: **Theory, Practice, and Visualization***. Wiley.

WATSON, G. S. ***Smooth Regression Analysis, Sankhya***, Series A, v. 26, p. 359-372, 1964.

ZIEGELMANN, F. A. *Estimation of Volatility Functions: Nonparametrics and Semi-parametric Methods. Ph.D. Dissertations. University of Kent at Canterrbury*, UK, 2002.

Anexo I – Tabelas

Tabela 4: Excesso dos Retornos

Janela	MV	NP	Janela	MV	NP	Janela	MV	NP
1	0,200	0,207	42	2,546	4,404	83	0,053	0,038
2	-0,040	-0,052	43	0,994	3,332	84	0,183	0,164
3	0,224	0,235	44	1,245	1,868	85	0,545	0,476
4	0,191	0,205	45	-0,085	-0,233	86	0,686	0,599
5	0,066	0,068	46	0,821	0,863	87	0,069	0,150
6	-0,244	-0,257	47	0,334	0,620	88	0,545	0,644
7	0,152	0,172	48	1,003	0,572	89	0,338	1,213
8	0,138	0,147	49	2,291	1,139	90	0,014	0,029
9	0,186	0,190	50	1,175	1,521	91	0,136	0,124
10	0,117	0,134	51	0,218	0,210	92	-0,031	-0,040
11	0,102	0,116	52	0,491	0,658	93	0,175	0,190
12	0,162	0,156	53	-0,379	-0,321	94	-0,032	-0,045
13	0,017	0,014	54	0,296	0,480	95	0,070	0,048
14	0,154	0,127	55	0,500	0,891	96	-0,026	-0,096
15	0,073	0,043	56	1,238	1,460	97	-0,034	-0,132
16	0,255	0,208	57	1,073	1,138	98	0,003	0,005
17	0,064	0,062	58	0,379	0,209	99	-0,115	-0,180
18	0,303	0,279	59	0,609	0,467	100	0,048	0,099
19	0,370	0,374	60	0,149	0,112	101	0,025	0,024
20	0,065	0,052	61	0,514	0,351	102	0,012	0,033
21	-0,051	-0,043	62	0,776	0,586	103	0,005	-0,015
22	0,155	0,157	63	0,292	0,351	104	-0,069	-0,045
23	0,161	0,149	64	-0,004	0,017	105	0,163	0,094
24	0,147	0,143	65	0,179	0,143	106	-0,014	-0,019
25	-0,032	-0,024	66	0,238	0,168	107	0,035	-0,072
26	0,067	0,077	67	0,803	0,706	108	-0,099	-0,060
27	-0,202	-0,217	68	0,540	0,603	109	-0,132	-0,126
28	-0,186	-0,191	69	0,486	0,409	110	-0,172	-0,064
29	0,158	0,151	70	0,067	0,148	111	0,032	-0,102
30	0,080	0,090	71	0,378	0,365	112	0,025	0,019
31	0,029	0,028	72	0,381	0,420	113	0,069	0,027
32	-0,104	-0,087	73	0,617	0,680	114	-0,059	-0,057
33	0,024	0,027	74	0,015	-0,044	115	0,004	-0,015
34	0,091	0,090	75	0,268	0,264	116	0,036	0,007
35	0,184	0,192	76	0,200	0,101	117	0,147	0,295
36	0,230	0,216	77	0,588	0,511	118	0,187	0,291
37	0,025	0,016	78	0,740	0,668	119	0,084	0,023
38	0,173	0,155	79	0,433	0,526	120	0,111	0,052
39	-0,116	-0,126	80	0,291	0,254	121	0,118	0,011
40	0,131	0,142	81	0,313	0,246	122	0,193	0,158
41	1,491	2,707	82	0,179	0,149			

Tabela 5: Indíce Sharpe

Janela	MV	NP	Janela	MV	NP	Janela	MV	NP
1	1,184	1,160	42	1,238	1,207	83	0,166	0,150
2	-0,147	-0,188	43	0,422	1,047	84	0,441	0,492
3	0,589	0,575	44	0,725	0,655	85	1,325	1,162
4	0,770	0,836	45	-0,145	-0,324	86	2,185	1,953
5	0,285	0,293	46	2,545	1,672	87	0,169	0,409
6	-0,634	-0,643	47	0,204	0,214	88	1,452	1,384
7	0,371	0,383	48	1,306	1,964	89	0,392	0,382
8	0,522	0,545	49	1,581	1,600	90	0,097	0,336
9	0,875	0,860	50	0,491	0,536	91	0,598	1,750
10	1,641	1,652	51	0,824	0,537	92	-0,431	-0,633
11	0,506	0,523	52	0,280	0,269	93	0,805	1,179
12	2,118	1,282	53	-0,721	-0,564	94	-0,215	-0,275
13	0,080	0,050	54	0,605	0,605	95	0,948	0,482
14	0,701	0,633	55	0,925	1,109	96	-0,259	-1,180
15	0,267	0,160	56	1,567	1,669	97	-0,103	-0,354
16	0,903	0,777	57	0,719	0,589	98	0,009	0,020
17	0,204	0,225	58	0,539	0,345	99	-0,350	-1,101
18	0,719	0,692	59	0,601	0,306	100	0,930	0,890
19	1,207	1,206	60	0,135	0,090	101	0,459	0,118
20	0,192	0,153	61	0,581	0,368	102	0,113	0,263
21	-0,126	-0,115	62	1,089	0,845	103	0,038	-0,142
22	0,471	0,519	63	0,534	0,497	104	-0,641	-0,520
23	1,023	1,032	64	-0,045	0,146	105	0,933	0,841
24	0,920	0,855	65	0,491	0,621	106	-0,081	-0,165
25	-0,173	-0,148	66	0,467	0,370	107	0,185	-0,555
26	0,307	0,330	67	1,579	1,503	108	-0,440	-0,394
27	-0,261	-0,268	68	1,066	1,388	109	-0,999	-1,064
28	-0,217	-0,218	69	1,061	0,885	110	-0,713	-0,335
29	0,491	0,408	70	0,259	0,593	111	0,267	-0,305
30	0,190	0,227	71	2,259	1,968	112	0,245	0,171
31	0,114	0,121	72	0,653	0,750	113	0,723	0,363
32	-0,368	-0,326	73	1,255	1,690	114	-0,922	-0,872
33	0,120	0,136	74	0,058	-0,141	115	0,024	-0,117
34	0,370	0,392	75	0,401	0,306	116	0,217	0,066
35	0,462	0,456	76	0,545	0,243	117	0,455	0,599
36	0,777	0,747	77	1,431	1,306	118	0,486	0,365
37	0,071	0,050	78	1,657	1,509	119	0,380	0,033
38	0,481	0,477	79	1,917	2,981	120	0,398	0,117
39	-0,181	-0,201	80	0,914	0,621	121	0,350	0,025
40	0,160	0,178	81	0,685	0,445	122	1,273	1,209
41	1,763	1,701	82	0,544	0,481			

Anexo II – Código fonte do modelo de média-variância clássico

```
require(zoo)
require(xts) #exigido pelo PerformanceAnalytics
require (PerformanceAnalytics) #SharpeRatio
require (tseries)
require (timeDate)
require (timeSeries)
require(pvclust)
require(sparcl)
require(dendextend)
require(quadprog)
require(corpcor)
require(forecast)
#Importar os dados dos Fundos Multimercados
BancoFM <- read.csv2("Banco_FM_Final.csv", header = F, dec = ",")
Indices <- read.csv2("CDI_IBOV.csv", header = F, dec = ",")
#Excesso de retorno
BancoFM <- BancoFM2-Indices$V1
retns = round(BancoFM, 2)
# dependencia linear entre os ativos 2 e 3 e entre 168 e 169
retns = retns[,-c(2,168)]
retns.indice = data.frame(round(Indices$V1, 2))
#tamanho da janela para calcular md
wl=500
#lagv - length número de periodos até o rebalanceamento.
lwnd=5
#nwindow num de janelas - (total de observacoes)
nwnd= floor((dim(retns)[1]-wl)/lwnd)
# ponto inicial para usar a amostra até o final
n.in0=dim(retns)[1] - nwnd*lwnd - wl + 1
n.in1 = wl + n.in0 - 1
```

```

SR.Optim = vector() #Usado para guardar o índice de Sharpe
nwjanelas = vector() #Usado para guardar o tamanho da janela
m.optim = NULL
#Guardar em uma matriz os pesos usados nas janelas
Pesos = matrix(0,ncol=nwnd,nrow=ncol(retns))
for (n0 in 1:nwnd){
  #ativos 1 sempre tem o tamanho de wl (padrao eh 500 observações)
  ativos1=retns[n.in0:n.in1,]
  #tamanho da janela posterior para análise (5 observações posteriores)
  ativos2=retns[(n.in1+1):(n.in1+lwnd),]
  #otimização - mínimo de retorno esperado
  ativos3=retns.indice[n.in0:n.in1,]
  #atualização
  #ignora a informacao do passado e monta o grupo com os dados de
  uma janela
  n.in0=n.in0+lwnd
  n.in1=lwnd+n.in1
  #Otimização com o excesso de retorno
  x=data.matrix(ativos1)
  port_optim=portfolio.optim(x, pm=mean(ativos3), riskless=FALSE,
shorts=FALSE, rf=0.0, reslow=rep(1e-10,ncol(x)), reshigh=NULL,
covmat=cov(x))
  #Sharpe Ratio Portfolios
  R=data.matrix(ativos2)
  Rts=as.ts(R)
  X=zoo(Rts,seq(as.Date('2012-07-02'), by = 'days', length = lwnd))
  SR.Optim[n0]=SharpeRatio(X, Rf=0, p=0.95, FUN="StdDev",
weights=port_optim$pw, annualize=F, geometric=F)
  #Retorno da otimização
  m.optim=c(m.optim,apply(ativos2,1,mean.p,p=port_optim$pw))

  #Quantidade de pesos nas janelas onde os pesos >= margem

```

```
margin=1.00000e-07
nwjanelas[n0]=length(port_optim$pw[which(port_optim$pw >= margin)])
#Guardar os pesos
Pesos[,n0]=data.matrix(port_optim$pw)
print(n0)
}
```

Anexo III – Código fonte do modelo de média-variância NP

```
m.nw = function(x,y,h,k=dnorm){  
  
  n = length(x)  
  
  result = rep(0,n)  
  
  for(i in 1:n){  
  
    k1 = k((x-x[i])/h)  
  
    result[i] = sum(k1*y)/sum(k1)  
  
  }  
  
  result  
  
}  
  
m.nw.x0 = function(x,x0,y,h,k=dnorm){  
  
  k1 = k((x-x0)/h)  
  
  sum(k1*y)/sum(k1)  
  
}  
  
m.nw.x0.2d = function(x1,x2,x01,x02,y1,y2,h1,h2,k=dnorm){  
  
  k1 = k((x1-x01)/h1)  
  
  k2 = k((x2-x02)/h2)  
  
  sum(k1*k2*y1*y2)/sum(k1*k2)  
  
}  
  
reg.or.and <- function(colunas, ativos1_kernel, T1, yd, yind){  
  
  for(i in 1:colunas){  
  
    yd[,i] = ativos1_kernel[2:T1,i]  
  
    yind[,i] = ativos1_kernel[1:(T1-1),i]  
  
  }  
  
}
```

```

retns = round(BancoFM, 2)

# dependencia linear entre os ativos 2 e 3 e entre 168 e 169

retns = retns[,-c(2,168)]

retns.indice = data.frame(round(Indices$V1, 2))

wl=500 #tamanho da janela para calcular md

lwnd=5 #lagv - length número de periodos até o rebalanceamento.

#nwindow num de janelas - (total de observacoes)

nwnd= floor((dim(retns)[1]-wl)/lwnd)

n.in0=dim(retns)[1] - nwnd*lwnd - wl + 1 # ponto inicial para usar a amostra ate
o final

n.in1 = wl + n.in0 - 1

#Usado para guardar o indice de Sharpe

SR.Optim_kernel = Fatores = vector()

#Usado para guardar o n dos pesos usados >= margem

nwjanelas_kernel = vector()m.optim_kernel = NULL

contador.NaN=0

#Guardar em uma matriz os pesos de cada janela

Pesos_Kernel = matrix(0,ncol=nwnd,nrow=ncol(retns))

for (n0 in 1:nwnd){

#ativos 1 sempre tem o tamanho de n.in1 (padrao eh 500 observacoes) para
calcular a md

ativos1_kernel=retns[n.in0:n.in1,]

#tamanho da janela posterior para analise (5 observações posteriores)

ativos2_kernel=retns[(n.in1+1):(n.in1+lwnd),]

#otimizacao - minimo de retorno esperado

```

```

ativos3=retns.indice[n.in0:n.in1,]

T1 = nrow(ativos1_kernel)

colunas <- ncol(ativos1_kernel) #numero de ativos

linhas <- nrow(ativos1_kernel)

##Inicializando

yd <-matrix(rep(0,(linhas-1)),ncol=1,nrow=linhas-1)

yind <-matrix(rep(0,(linhas-1)),ncol=1,nrow=linhas-1)

#regressor e regressando para cada janela

yd = ativos1_kernel[-1,]

yind = ativos1_kernel[-T1,]

reg.or.and(colunas,ativos1_kernel,T1, yd, yind)

## bandwidth ##

h <-matrix(rep(0,(colunas)),ncol=colunas,nrow=1)

fator=1

for(i in 1:colunas){

aux = ativos1_kernel[,i]

h[1,i] = fator*1.06*sd(aux)*length(aux)^(-0.2)

}

## 1 passo ##

## Regressao Univariada para cada Ativo

r <-matrix(0,ncol=colunas,nrow=T1-1)

reg.np <-matrix(0,ncol=colunas,nrow=T1-1)

for(i in 1:colunas){

reg.np[,i] = m.nw(yind[,i],yd[,i],h[1,i])

```

```

r[,i] = yd[,i] - reg.np[,i]
}

fator = 1

cc = -1

while(cc < 0){ ## enquanto houver um autovalor nao positivo executar a rotina

## 2 passo ##

## Estimacao Univariada da variancia de cada ativo no último periodo de tempo

## 3 passo ##

## Estimcao das covariancias entros dois ativos

cov.T <- matrix(0,ncol=colunas,nrow=colunas)

for(i in 1:colunas){
for(j in i:colunas){

if(i == j){

## 2 passo ##

a=r[,i]^2

cov.T[i,i] = m.nw.x0(x0=yd[nrow(yd),i],x=yind[,i], y=a, h=fator*h[1,i])

} else {

## 3 passo ##

cov.T[i,j] = cov.T[j,i] = m.nw.x0.2d(x01=yd[nrow(yd),i], x1=yind[,i], y1=r[,i],
h1=fator*h[1,i], x02=yd[nrow(yd),j],x2 = yind[,j],y2=r[,j], h2=fator*h[1,j])

}

}

}

}

#Se tiver NaN nas covariancias retorna True

bool.True.False<-any(is.na(cov.T))

```



```

if(bool.True.False==T){
#Guardar a posicao de cada objeto
matriz.ij.NaN <- which(is.nan(cov.T),arr.ind = T)
#Guardar em x para calcular cov tradicional
x=data.matrix(ativos1_kernel)
matriz.ij.Sem.NaN <- cov(x)
#Guarda tamanho e quantidade de NaN
tam.matriz <-dim(matriz.ij.NaN)[1]
contador.NaN <- contador.NaN + tam.matriz
#Substituir em todos os campos com NaN pela Cov tradicional
for(p in 1:tam.matriz){
i <- matriz.ij.NaN[p,1]
j <- matriz.ij.NaN[p,2]
cov.T[i,j] <- matriz.ij.Sem.NaN[i,j]
}
}
auto.val = eigen(cov.T) #atualiza os autovalores
cc = min(auto.val$values)
print(fator)
fator=1.3*fator
}
Fatores[n0] = fator
#atualizacao
#ignora a informacao do passado e monta o grupo com os dados de uma
janela

```

```

n.in0=n.in0+lwnd

n.in1=lwnd+n.in1

#Otimizacao com o excesso de retorno

x=data.matrix(ativos1_kernel)

port_optim_kernel=portfolio.optim(x, pm = mean(ativos3), riskless = FALSE,
shorts = FALSE, rf = 0.0 , reslow = rep(1e-10,ncol(x)), reshigh = NULL, covmat
= cov.T)

#Sharpe Ratio Portfolios

R=data.matrix(ativos2_kernel)

Rts=as.ts(R)

X=zoo(Rts,seq(as.Date('2012-07-02'), by = 'days', length = lwnd))

SR.Optim_kernel[n0]=SharpeRatio(X, Rf=0, p = 0.95, FUN = "StdDev",
weights=port_optim_kernel$pw, annualize = F, geometric=F)

#Retorno da otimizacao

m.optim_kernel=c(m.optim_kernel,apply(ativos2_kernel,1,mean.p,p=port_optim
_kernel$pw))

#Quantidade de pesos nas janelas onde os pesos >= margem

margin = 1.00000e-07

nwjanelas_kernel[n0] =
length(port_optim_kernel$pw[which(port_optim_kernel$pw >= margin)])

#Quantidade de pesos nas janelas onde os pesos >= margem

#Guardar os pesos

Pesos_Kernel[,n0] = data.matrix(port_optim_kernel$pw)

print(n0)

}

```