

# Otimização de portfólio utilizando previsão e regularização: um estudo aplicado ao mercado de ativos brasileiro

Luiz Bezerra de Oliveira Lima Neto\*

<sup>1</sup>Universidade Federal do Rio Grande do Sul

**Resumo.** Este artigo tem como objetivo analisar a utilização de modelos de previsão e regularização em problemas de otimização de portfólio. Testes causalidade Granger indicam a existência significativa de relações lead-lag entre as séries de retorno de ativos no mercado acionário brasileiro. O modelo de Vetores Autoregressivos (VAR) é proposto para explorar as relações de dependência encontradas e conduzir previsões um passo a frente. Em função dos problemas inerentes à alta dimensionalidade do VAR, foram adotadas duas abordagens. A primeira utilizando o método de regularização Ridge Regression e a segunda um VAR Bayesiano, contornando os problemas de dimensionalidade e gerando estimadores mais estáveis. Outro modelo proposto é um GARCH multivariado ortogonal que é utilizado para previsão um passo a frente da matriz de covariância. O modelo VAR com a regularização Ridge Regression produziu os melhores resultados fora da amostra entre os modelos de previsão.

**Classificação JEL:** C01, C58, G11, G17.

**Palavras chave:** Métodos de previsão, Vetores Autoregressivos (VAR), GARCH Ortogonal (O-GARCH), Ridge Regression, VAR Bayesiano (BVAR), Alocação de ativos.

**Abstract.** The aim of this paper is to analyze the application of predictive models in portfolio optimization problems. Granger causality tests indicate the existence of significant lead-lag relations among Brazilian's stock returns. A vector-autoregressive (VAR) model is proposed to explore such dependency relations and to conduct one-step ahead forecasts. Due to the large number of parameters and thus the over-parametrization problems two approaches are proposed. The first is to use the Ridge Regression regularization method. Secondly a Bayesian VAR is proposed circumventing the dimensionality problems and also getting more stable parameters. A multivariate orthogonal GARCH is proposed to make one-step ahead forecasts of the covariance matrix. The VAR model using Ridge Regression yielded the best out-of-sample results among the forecasting models.

**JEL classification:** C01, C58, G11, G17.

**Keywords:** Forecasting methods, Vector-Autoregressive (VAR), Orthogonal GARCH (O-GARCH), Ridge Regression, Bayesian VAR, Asset allocation.

---

\*Mestrando em Economia Aplicada pela Universidade Federal do Rio Grande do Sul – UFRGS. Contato: netolui bezerra@gmail.com

## 1. Introdução

A previsibilidade do retorno de ativos é um dos temas mais debatidos e longevos na área de finanças, tendo como trabalho precursor a tese de doutoramento de Bachelier (1900). Até os dias atuais este continua sendo um assunto controverso em sua investigação empírica. Sob a ótica da teoria da eficiência de mercado, os ativos financeiros já refletem todas as informações de mercado, e por isso estratégias de investimento que utilizam previsão para seleção de ativos não capazes de gerar retornos anormais.

De fato, trabalhos como Fama e French (1970) documentam que apesar da evidência de alguma dependência diária nos preços e retornos de ações, estas porém são próximas de zero, o que as torna economicamente insignificantes. Também ficou evidenciado que a tentativa de utilizar essa dependência em estratégias de investimento acabaria em tantas transações que se tornariam inviáveis até com custos mínimos de transação. Uma corrente mais recente de trabalhos, como Lo e MacKinlay (1988), encontrou evidências empíricas que preços históricos de ações não seguem um passeio aleatório. Em um outro trabalho, Lo e MacKinlay (1990) evidenciaram relações de *lead-lag* semanal, onde as ações de empresas com maior capitalização lideraram o retorno de empresas menores. Jegadeesh e Titman (1993), documentam que estratégias que compram ações que performaram bem em um horizonte intermediário de tempo de 3 a 12 meses continuam performar melhor do que as ações perdedoras, indicando a ocorrência de “momento” no mercado de ações norte americano. O estudo desenvolvido por De Miguel et al (2014) trouxe consistente evidência empírica sobre possibilidade de se explorar a dependência serial no retorno de ativos para seleção de portfólio.

Apesar da extensa literatura existente sobre estratégias de investimento, ainda são raros os trabalhos que buscam estudar as relações dependência entre ativos e a sua aplicabilidade na previsão de retornos no mercado brasileiro. O trabalho de Müller et al (2015), por exemplo, encontrou sinais de dependência serial no retorno de ativos brasileiros. Ainda de acordo com Müller et al (2015) a constatação de dependência serial pode ser um sinal de previsibilidade no retorno de ações. Neste contexto este estudo visa identificar as relações de dependência entre as séries de retornos de ativos que são negociados na Bovespa. Neste artigo as relações de dependência são oriundas de duas fontes. A primeira é do ativo com ele mesmo considerando um grau de defasagem, ou seja, a capacidade do retorno de um ativo hoje ser útil na na previsão do seu retorno no período seguinte. A segunda é de um ativo com outro ativo com um grau de defasagem, ou seja, a capacidade do retorno um determinado ativo na previsão do retorno de outro ativo. Assim a primeira contribuição do artigo é testar as relações de causalidade-Granger no mercado de ações brasileiro. Grande parte da literatura foca nas relações de *lead-lag* entre o índice Bovespa e índices estrangeiros. Neste são testadas as relações de causalidade com relação as ações por *book-to-market* e também por setor. Também é feita a análise do índice Bovespa principais índices das bolsas americanas. Foi constatada a existência de causalidade-granger na maioria dos testes realizados.

Consecutivamente será proposta a abordagem de Vetores Autoregressivos (VAR) para a modelar as relações de dependência entre as séries de retorno. Este modelo é adequado ao objetivo pois é capaz de captar quaisquer relações lineares entre os diferentes retornos no modelo. Na literatura pode-se encontrar a utilização dos modelos VAR na

otimização estratégica de ativos<sup>1</sup>. Contudo a implementação de um modelo VAR com o intuito de se explorar a capacidade do retorno de um ativo de risco prever o retorno de outro ativo de risco foi feita de forma pioneira no artigo De Miguel et al (2014). A análise das séries de retorno do mercado de ações brasileiro mostra que um grau de defasagem é o suficiente para modelar as relações de dependência entre os retornos. Dado o elevado número de variáveis incorporadas ao modelo e o subsequente problema de dimensionalidade ocasionado pela estimação através do VAR, são utilizadas duas abordagens de forma a contornar essa situação. A primeira é a utilização da *Ridge Regression* com o intuito de se obter parâmetros mais estáveis. A outra abordagem utilizada é a estimação de um VAR Bayesiano. Os métodos Bayesianos foram impulsionados após o desenvolvimento de novas técnicas mais eficientes no tratamento das distribuições *posteriori*. Este método consiste na atribuição de restrições probabilísticas ao modelo, conseguindo fornecer previsões com maior grau de assertividade. Neste artigo, será utilizada a priori de Minnessota.

Na abordagem tradicional de alocação de portfólio proposta por Markowitz, a matriz de covariância tem papel central. Ao longo dos últimos anos diversos métodos foram propostos de forma melhorar a sua estimação. Conforme Ledoit e Wolf (2003) “ninguém deve utilizar a matriz de covariância amostral com o proposito de otimização de portfólio”. Neste artigo será utilizada a metodologia desenvolvida por Ledoit e Wolf para todas as abordagens de alocação de portfólio que fazem uso da matriz de covariância dos retornos. Outra questão relacionada a matriz de covariância emerge com o aumento da dimensionalidade. A estimação de matrizes de covariâncias com um numero elevado de parâmetros ainda é um desafio na área de finanças. Por tanto neste trabalho também será proposto um modelo GARCH ortogonal generalizado, que será utilizado para fazer a previsão um passo a frente na matriz de covariância.

Por fim, as estratégias que utilizam a previsão um passo a frente serão comparadas com os portfólios clássicos da literatura. Entre eles, o portfólio de média-variância, de mínima-variância global e a carteira “ingênua”(1/N). A principal medida de desempenho das estratégias será o índice de Sharpe. As carteiras também serão avaliadas com relação ao retorno e volatilidade anualizada. A análise será realizada primeiramente desconsiderando os custos de transação. Uma segunda análise será realizada incorporando 20 pontos percentuais de custos de transação.

O restante do artigo está dividido da seguinte forma. A segunda seção faz o estudo detalhado sobre as principais características dos modelos VAR propostos no artigo. A quarta seção aborda a modelagem da matriz de covariância. A quinta seção trata dos métodos de otimização de portfólio. A sexta seção trata das medidas de desempenho utilizadas para avaliação dos portfólios. A sétima seção discute os resultados empíricos, e a oitava seção discute as conclusões finais.

## **2. A modelagem das séries de retorno através de um Vetor Autoregressivo (VAR)**

Os modelos de Vetores Autoregressivos (VAR) são uma classe de séries temporais que tem ampla aplicação em macroeconomia e finanças. Geralmente modelos econômicos contam com um grande número de variáveis, sendo necessário um instrumento que seja

---

<sup>1</sup>Como referência ver: Campbell e Viceria(1999,2002), Campbel, Chan e Viceira (2003), Balduzzi e Lynch (1999), Barberis (2000)

flexível para sua estimação. O VAR é considerado uma boa opção para estas situações, pois tem a habilidade de se ajustar aos dados permitindo a estimação dos parâmetros do modelo. A grande popularidade dos modelos VAR teve início com a publicação dos trabalhos de Granger e Sims. Em seu trabalho seminal, Granger (1969b) propôs a ideia da causalidade de Granger. Este conceito é fundamental para previsão das séries que estão sendo modeladas. O trabalho de Sims (1980) propôs a utilização de modelos VAR em aplicações econômicas. Desde então os modelos VAR vêm sendo amplamente utilizados em diversos campos de aplicação.

Neste artigo, o modelo VAR é utilizado no contexto de múltiplos ativos, e assume que o retorno futuro esperado de cada ativo dependa linearmente do retorno de cada ativo no período atual e também do seu próprio retorno. Além de estudar as dependências lineares, deseja-se identificar a origem da preditividade dos dados, analisando assim as autocorrelações e correlações cruzadas destes ativos com a defasagem de um período.

Uma série temporal multivariada  $r_t$  é um processo VAR de ordem 1, se segue o modelo:

$$r_{t+1} = \phi_0 + \Phi r_t + a_t, \quad (1)$$

Onde,  $\phi_0$  é um vetor de dimensão  $k$ ,  $\Phi$  é uma matriz  $k \times k$ , e  $a_t$  é uma sequência serialmente não correlacionada de vetores aleatórios com média zero e matriz de covariância  $\Sigma$ . É exigido que a matriz  $\Sigma$  seja positiva-definida, caso contrário, a dimensão de  $r_t$  poderia ser diminuída. Como é possível inferir da equação (1), a dependência linear é caracterizada pela matriz  $\Phi$ . Para entendermos a dinâmica dos efeitos marginais sobre os retornos futuros vamos primeiramente considerar o caso bi-variado. O nosso modelo VAR(1) consiste de duas equações:

$$\begin{aligned} r_{1t+1} &= \phi_{10} + \Phi_{11}r_{1,t-1} + \Phi_{12}r_{2,t-1} + a_{1t} \\ r_{2t+1} &= \phi_{20} + \Phi_{21}r_{1,t-1} + \Phi_{22}r_{2,t-1} + a_{2t} \end{aligned}$$

Onde  $\Phi_{ij}$  é o (i,j) elemento de  $\Phi$  e representa efeito marginal de  $r_{j,t}$  sobre  $r_{i,t+1}$  condicional a  $r_t$ . Assumindo que  $r_t$  é um processo fracamente estacionário, com média finita  $E(r_t) = E(r_{t-1}) = \dots = \mu$ , matriz de covariância positiva definida  $\Gamma_0 = E[(r_{t-1} - \mu)(r_t - \mu)]$ , e matriz de covariância cruzada finita  $\Gamma_0 = E[(r_{t-1} - \mu)(r_t - \mu)]$ .

## 2.1. O efeito lead-lag no mercado de ativos brasileiro

Esta subseção tem como objetivo contextualizar o leitor com relação a dinâmica das variáveis que serão exploradas ao longo do artigo e abordar com maior profundidade a causalidade de Granger.

O teste de causalidade de Granger tem como intuito testar o sentido da casualidade entre as variáveis que compõem o modelo. Este conceito diz respeito a precedência temporal entre as variáveis. Suponha por exemplo as séries de retorno  $R_{1t}$  e  $R_{2t}$ , se a variável

$R_{1t}$  é capaz de ajudar a prever a variável  $R_{2t}$  então podemos afirmar que a variável  $R_{1t}$  causa Granger  $R_{2t}$ . Pode-se utilizar um teste  $F$  convencional da seguinte maneira:

1. Estime:

$$R_{1t} = \beta_{10} = \sum_{i=1}^p \beta_{11} R_{1,t-i} + \sum_{j=1}^p \beta_{12} R_{2,t-j} + e_{1t} \quad (2)$$

$$R_{2t} = \beta_{20} = \sum_{i=1}^p \beta_{21} R_{1,t-i} + \sum_{j=1}^p \beta_{22} R_{2,t-j} + e_{2t}$$

Onde  $\beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}$  são os parâmetros a serem estimados e  $e_{1t}$  e  $e_{2t}$ , são os erros aleatórios  $\sim N(0, \sigma^2)$ .

2. Teste se  $R_{1t}$  não causa Granger  $R_{2t}$  usando o teste de F, sob:

$$H_0 = \beta_{1,21} = \beta_{2,21} = \dots = \beta_{p,21} = 0 \quad X \quad \beta_{i,21} \neq 0 \quad (3)$$

O teste F é utilizado para determinar se há, e qual o sentido de causalidade envolvendo as séries do modelo. Os resultados sobre as relações causalidade encontradas entre as séries de retorno estudadas neste artigo serão descritas na seção de resultados empíricos.

## 2.2. A maldição da dimensionalidade

Os modelos VAR são um instrumento flexível para se trabalhar com um número razoável de variáveis. Porém a medida que estas aumentam acontece a sobre-parametrização do modelo, comprometendo os graus de liberdade. Este fenômeno é popularmente conhecido como a maldição da dimensionalidade e ocorre em função do crescimento quadrático do número de parâmetros em relação às variáveis presentes no modelo. Por exemplo estimar um VAR(2) com 5 séries temporais ( $p=5$ ), requer estimar  $dp^2 = 50$  parâmetros. Ainda, a estrutura de autocorrelação que caracteriza as séries comumente incluídas no VAR tende a gerar multicolinearidade e sobre-parametrização (Jaramillo,2009), tornando os estimadores de MQO instáveis. Ao se trabalhar com séries de retorno de ativos inevitavelmente necessita-se trabalhar com um número elevado de variáveis. Para se contornar os problemas de dimensionalidade intrínsecos ao VAR deve-se lançar mão de algum instrumento que gere estimadores robustos mesmo na presença de um grande número de parâmetros. Neste trabalho serão adotadas duas abordagens para lidar com este problema, a Ridge regression e a estimação de um VAR Bayesiano (BVAR).

## 2.3. Estimação dos parâmetros VAR

As observações de uma série temporal permitem considerar:

$$z_t = \phi_0 + \phi_1 z_{t-1} + \dots + \phi_p z_{t-p} + a_t \quad t = p + 1, \dots, T,$$

Onde a matriz de covariância  $a_t$  é  $\Sigma_a$  e onde esteja disponíveis  $T - p$  pontos para estimação. Por simplificação, pode-se escrever o modelo VAR(p) como:

$$z_t^T = x_t^T \beta + a_t^T, \quad (4)$$

Onde  $x_t = (1, z_{t-1}^T, \dots, z_{t-p}^T)^T$  é um vetor de dimensões  $(kp + 1)$  e  $\beta^T = [\phi_0, \phi_1, \dots, \phi_p]$  é uma matriz  $k \times (kp + 1)$ . Com este formato podemos rescrever os dados como:

$$Z = X\beta + A \quad (5)$$

Onde  $Z$  é uma matriz  $(T - p) \times k$ ,  $X$  é uma matriz  $(T - p) \times (kp + 1)$  e  $A$  é uma matriz  $(T - p)$ .

Pode-se perceber da equação acima que o sistema de vetores autoregressivos podem ser vistos como um modelo de regressão multivariado. Desta forma cada equação pode resolvida individualmente através de mínimos quadrados. É válido observar que solução de mínimos quadrados generalizados é idêntica a de mínimos quadrados e que sob as hipóteses de normalidade, ou seja,  $a_t$  segue uma distribuição normal, as estimativas por máxima verossimilhança são assintoticamente equivalentes as estimativas de mínimos quadrados. O aprofundamento desta questão encontra-se no apêndice A. Através destes resultados chegamos a conclusão de que:

$$\hat{\beta} = (X^T X)^{-1} (X^T Z) \quad (6)$$

#### 2.4. Estimação através da Rigde Regression

Com a finalidade de ilustrar o método da Ridge Regression, será apresentado o exemplo proposto por Exterkate et al (2012). Suponha que deseja-se realizar a estimação de um modelo econométrico através de MQO para posterior previsão. Neste caso após o período  $T$  deseja-se estimar um valor de  $y$  projetado em uma eventual data futura e este valor seja denotado por  $y_*$ . Considere que temos o vetor  $N \times 1$  de variáveis endógenas, e também que as observações históricas  $t = 1, 2, \dots, T$  estão disponíveis para todas as variáveis. Desta forma temos o vetor  $y$ ,  $T \times 1$  e a matriz  $X$  com dimensões  $T \times N$ . Assumindo a função linear de previsão  $\hat{y}_* = x_*^T \hat{\beta}$ , podemos obter  $\hat{\beta}$  através da minimização de  $\|y - X\beta\|^2$ , onde  $\|\cdot\|$  denota a norma  $L_2$ . Dado que  $X$  tem posto  $N$ ,  $\hat{\beta} = (X^T X)^{-1} X^T y$ , o que gera seguinte previsão:  $\hat{y}_* = x_*^T (X^T X)^{-1} X^T y$ .

O procedimento de MQO pressupõe que  $N \leq T$ , e na prática  $N \ll T$  é necessário para que não ocorra o problema da sobre-parametrização. Isto é, se  $N$  não for pequeno com relação a  $T$ , então será obtido um excelente ajuste na amostra (no limite que se  $N = T$ , o ajuste na amostra é perfeito), porém a previsão fora da amostra em muitos caso se mostra de baixa qualidade. Uma solução possível para este problema é a estimação através da *Ridge Regression*. Esta regressão visa balancear o ajustamento (*goodness-of-fit*) e a magnitude do coeficiente  $\beta$ . A solução da Ridge Regression é dada pela minimização da seguinte equação, onde  $\lambda$  é o fator que penaliza os  $\beta$ 's:

$$\text{minimizar } \sum_{i=1}^n (y - \mathbf{X}\beta)^T (y - \mathbf{X}\beta) \quad S.A : \sum_{j=1}^p \beta_j^2 \leq t \quad (7)$$

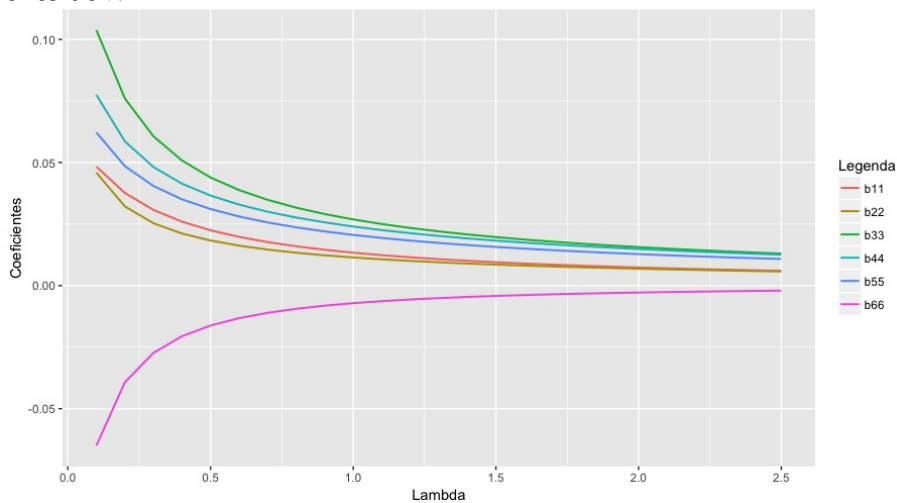
Tomando a primeira derivada e resolvendo para  $\beta$  temos:

$$\hat{\beta}_{\lambda}^{ridge} = (\mathbf{X}^T \mathbf{X} + \lambda \mathbf{I}_p)^{-1} \mathbf{X}^T \mathbf{y} \quad (8)$$

Pode-se ver que  $\lambda$  tem papel fundamental na regularização dos coeficientes. Segundo Hoerl e Kennard 1977, para a escolha de  $\lambda$  deve-se plotar os componentes  $\hat{\beta}_{\lambda}^{ridge}$  contra o fator  $\lambda$ . Deve-se escolher aquele parâmetro para o qual os coeficientes não estejam mudando rapidamente. Ou seja quando evidenciada a estabilização dos parâmetros.

Com o objetivo de fornecer a intuição a respeito deste método, a figura abaixo mostra a evolução no comportamento dos parâmetros estimados através do modelo VAR quando aumenta-se o fator  $\lambda$  :

**Figura 1. Comportamento dos parâmetros do setor industrial de acordo com o aumento de  $\lambda$**

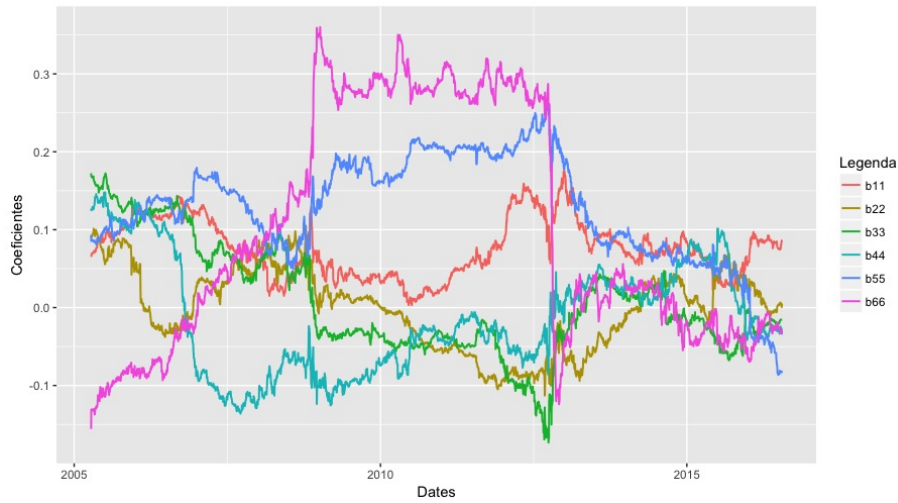


O que podemos observar da figura 1 é que os parâmetros estimados são muito sensíveis a mudanças no parâmetro  $\lambda$ . De tal forma que quando  $\lambda = 1$  todos os parâmetros encontram-se próximos a zero e em trajetória constante. Desta maneira pode-se ver claramente a manifestação de multicolinearidade nos parâmetros. Com relação as figuras 2 e 3, fica claro o encolhimento do intervalo na escala dos parâmetros. Evidencia-se também que a suavização das trajetórias dos parâmetros ao longo do tempo.

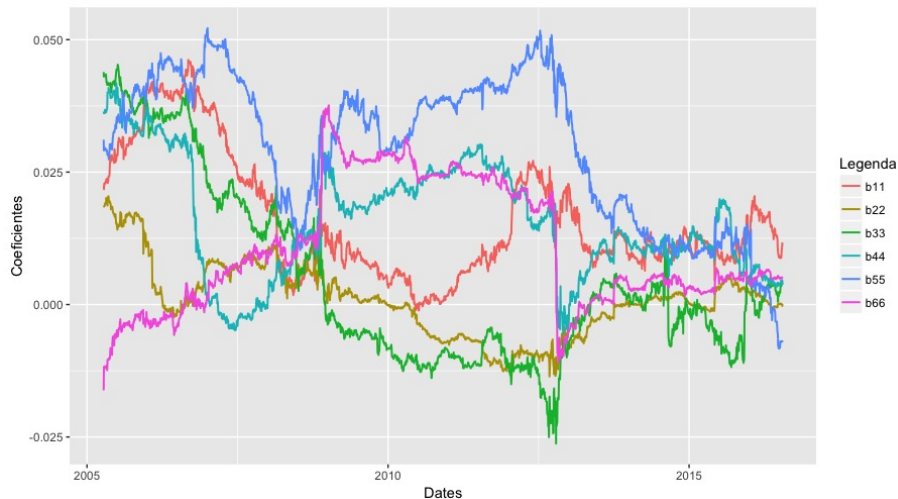
## 2.5. Estimação através do VAR Bayesiano

Considere  $\theta$  um parâmetro à ser estimado a partir de uma amostra de observações  $\mathbf{X}$ . Quando a inferência estatística é realizada através dos métodos clássicos, presume-se que  $\theta$  assume um determinado valor. Apesar de desconhecido, este valor de  $\theta$  é considerado fixo. Desta forma um estimador é construído para  $\theta$  a partir do conjunto de dados disponíveis. Um estimador conhecido na literatura é o valor de  $\theta$  que maximiza a função de

**Figura 2. Comportamento dos parâmetros autorregressivos do setor industrial de acordo a evolução no tempo considerando  $\lambda = 0$**



**Figura 3. Comportamento dos parâmetros autorregressivos do setor industrial de acordo a evolução no tempo considerando  $\lambda = 0.5$**



verossimilhança da amostra. Uma outra forma de se abordar este problema é utilizando a inferência Bayesiana, neste caso os parâmetros populacionais são tratados como variáveis aleatórias. Além do conjunto de dados disponíveis, o pesquisador possui um conjunto de *crenças* sobre os parâmetros populacionais e que são representados por uma função de distribuição *priori*, denotada aqui por  $P(\theta)$ . O produto desta *priori* com a função de máxima verossimilhança formam a distribuição conjunta, neste caso representados por  $X$  e  $\theta$ .

$$f(\theta|X) = \frac{f(\theta, X)}{f(X)} = \frac{f(X|\theta)P(\theta)}{f(X)} \quad (9)$$

Onde a distribuição marginal  $f(X)$  pode ser obtida por:



$$f(X) = \int f(X, \theta) d\theta = \int f(X|\theta) P(\theta) d\theta \quad (10)$$

A distribuição  $f(\theta|X)$  na equação (9) é chamada de distribuição posterior de  $\theta$ . Utiliza-se o teorema de Bayes para se obter:

$$f(\theta|X) \propto f(X, \theta) P(\theta) \quad (11)$$

Onde  $P(\theta)$  é a distribuição *a priori* e  $f(X|\theta)$  é a função de máxima verossimilhança. Nesta esteira, podemos conectar os conceitos Bayesianos com o modelo VAR desenvolvido aqui por meio da equação já apresentada anteriormente:

$$Z = X\beta + A \quad (12)$$

Onde a função de máxima verossimilhança pode ser derivada a partir densidade amostral.

### 2.5.1. A *priori* de Minnessota

Na prática a escolha da uma *priori* para um modelo Bayesiano é uma questão subjetiva, sendo possível encontrar na literatura uma grande quantidade de trabalhos com os mais diversos tipos de aplicação. Neste artigo será utilizada a *Priori* de Minnessota, que originalmente foi utilizada por Litterman (1980) na modelagem de séries macroeconômicas. Este método realiza a substituição de  $\Sigma_a \otimes C^{-1}$  por uma estimativa  $\mathbf{V}$ , que é definida como uma matriz diagonal:

$$\mathbf{V} = \begin{bmatrix} \sigma_{11} & 0 & 0 & \dots & 0 \\ 0 & \sigma_{22} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \sigma_{nn} \end{bmatrix}$$

Considerar este formato para a matriz de covariância dos resíduos permite uma grande simplificação computacional ao passo que é possível calcular cada equação do VAR individualmente e obter uma forma analítica para a sua distribuição posterior. Outra vantagem deste método é possibilitar uma grande flexibilidade da escolha da *priori*, ao contrario da abordagem inteiramente Bayesiana, onde a *priori* que gera resultados analíticos é uma necessariamente conjugada natural. A desvantagem deste método é substituir a matriz desconhecida  $\Sigma$  por uma estimativa aleatória, no lugar fazer a integração como é de praxe nos modelos Bayesianos. Para os coeficientes autoregressivos  $\phi_{l,ij}$  a *priori* da matriz de covariância é dada por:

$$Cov(\phi_{l,ij}) = \begin{pmatrix} \frac{\delta^2}{l} & se \ i = j \\ \frac{\delta\theta^2}{l} \times \frac{\sigma_{ii}}{\sigma_{jj}} & se \ i \neq j \end{pmatrix}$$

Onde  $\delta$  é um número real,  $0 < \theta < 1$ ,  $\sigma_{ii}$  é o  $i$ -ésimo elemento da diagonal na matriz  $\Sigma_a$  e  $l = 1, \dots, p$ . Através desta especificação nota-se a propriedade de encolhimento progressivo do coeficiente  $\phi$  conforme o aumento do número de defasagens. Outro fato importante é notar que a especificação de  $\theta$  entre 0 e 1 confere aos coeficientes autorregressivos maior poder preditivo encolhendo os coeficientes das outras variáveis e desta forma conferindo menor importância para estes coeficientes.

### 3. A modelagem da matriz de covariância através de modelos GARCH

A estimação da volatilidade do retorno de ativos financeiros tem papel fundamental na alocação de portfólios sob a abordagem de média-variância. Apesar de não ser diretamente observável, algumas características são comuns nas séries de retorno de ativos. A correlação positiva dos retornos ao quadrado revela que volatilidade do mercado financeiro se apresenta em *clusters*, em que períodos tranquilos são intercalados com períodos de grande volatilidade registrando grandes retornos (ou grandes perdas). Este fenômeno é conhecido como heterocedasticidade condicional autoregressiva. Uma classe de modelos econométricos que captam a estrutura citada acima que e tem extensa aplicação empírica são os modelos da família ARCH. Estes modelos foram inicialmente introduzidos por Engel (1982) e são considerados um *benchmark* na análise de volatilidade. Um modelo  $ARCH(p)$ , de ordem  $p \geq 1$  é definido como:

$$\begin{aligned} Y_t &= \sigma_t \epsilon_t \\ \sigma_t^2 &= \alpha_0 + \alpha_1 Y_{t-1}^2 + \dots + \alpha_p Y_{t-p}^2. \end{aligned} \quad (13)$$

$\sigma_t^2$  é a variância no tempo  $t$  condicional à  $I_t$ , isto é,  $\sigma_t^2 = Var(Y_t | I_{t-1})$ , onde  $I_{t-1} = (Y_{t-1}, Y_{t-2}, \dots)$  e  $\alpha_t \geq 0$ . Pode-se ver na equação acima que a variância condicional  $\sigma_t^2$  é modelada como uma função linear do quadrado das observações passadas de  $p$  do processo de  $Y_t$ .  $\{\epsilon_t\}$  é uma sequência i.i.d de variáveis aleatórias com média 0 e variância 1.

Apesar de simples o modelo  $ARCH$  geralmente requer um grande número de parâmetros para descrever o processo de volatilidade de forma adequada. Como solução para este problema Bollerslev (1986) propôs uma extensão, conhecida como  $ARCH$  generalizado ( $GARCH$ ), que adicionalmente contém as variâncias condicionais na equação da variância. Um modelo  $GARCH(p, q)$ , de ordem  $p \geq 1$  e  $q \geq 1$  é definido como:

$$\begin{aligned} Y_t &= \sigma_t \epsilon_t \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^p \alpha_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2, \end{aligned} \quad (14)$$

Onde  $\alpha_0 \geq 0$ ,  $a_i, \beta_j \geq 0$ ,  $i = 1, 2, \dots, p$ ;  $j = 1, 2, \dots, q$ . e  $\sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i) < 1$ .  $\sigma^2 = Var(Y_t|I_{t-1})$ , onde  $I_{t-1} = (Y_{t-1}, Y_{t-2}, \dots)$ . Novamente  $\{\epsilon_t\}$  é uma sequencia i.i.d de variáveis aleatórias com média 0 e variância 1.

### 3.1. GARCH multivariado

Na literatura voltada à finanças é possível encontrar diversos modelos multivariados para volatilidade. Contudo a modelagem multivariada de  $\Sigma_t$  encontra duas maiores dificuldades. A primeira é a maldição da dimensionalidade, já citada nos modelos VAR, é um problema recorrente destes modelos GARCH também. Como exemplo, para uma série temporal com número de dimensões  $k$ , a matriz de  $\Sigma_t$  consiste em  $k \frac{(k+1)}{2}$  diferentes elementos que variam no tempo. A segunda é assegurar que a matriz  $\Sigma_t$  seja positiva definida para todo  $t$ . Para superar o problema de dimensionalidade, algumas soluções são possíveis. Neste artigo será utilizado o GARCH ortogonal (O-GARCH).

## 4. Otimização de carteiras de investimento

Nesta seção serão apresentadas as principais estratégias de escolha de portfólio utilizadas neste artigo. Primeiramente serão expostas as estratégias que são mais familiares do público em geral: portfólio de média-variância, carteira ingênua ( $1/N$ ) e o portfólio de mínima variância. Em sequencia serão explicadas as que levam em consideração a dependência serial nas séries de retorno. São elas: o portfólio de média variância condicional e a carteira de média variância.

### 4.1. Carteira de Média-Variância

A análise de estratégias que fazem a alocação ótima de ativos é um dos mais importantes temas na área de finanças. O trabalho seminal de Markowitz (1952) revolucionou esta área derivando uma regra ótima para alocação da renda do investidor entre ativos de risco em um panorama estático em que o investidor apenas se importa com a média variância do retorno do portfólio. Esta abordagem ficou conhecida como carteira de Média-Variância. Apesar de seu caráter inovador, esta abordagem foi criticada por pesquisadores como Michaud (1989), Best e Grauer (1991) e Litterman (2003). Estudos empíricos mostraram que portfólios de média variância são suscetíveis a possuir ativos com pesos extremos e elevada sensibilidade com relação à mudanças nos dados de entrada do problema. Algebricamente, a alocação dos recursos nesta estratégia é equivalente a resolver o seguinte problema:

$$\begin{aligned} \min : \quad & \mathbf{w}^T \Sigma \mathbf{w} - \frac{1}{\gamma} E[R_{p,t+I}] \\ \text{s.a} : \quad & \mathbf{w}^T \boldsymbol{\mu} = \mu_c \\ & \mathbf{w}^T \mathbf{1} = 1, \quad w_i \geq 0 \forall i \end{aligned} \tag{15}$$

Onde  $w$  é o vetor de pesos ótimo é definido como  $\mathbf{w} = (w_1, \dots, w_n)$ . A terceira restrição diz que todos os recursos são investidos sem venda a descoberto. Isto é não

admite-se pesos negativos.  $R_{t+I}$  é o retorno aleatório dos fundos de investimento. O retorno da carteira entre os tempos  $t$  e  $t+I$  é dado por  $R_{p,t+I} = \sum_{i=1}^N w_i R_{i,t+I} = w^T R$ . Admita que  $R_t \sim N(\mu_t, \Sigma_t)$ , com  $\mu_t = (\mu_{1,t}, \mu_{2,t}, \dots, \mu_{N,t})$  e  $\Sigma_t = \sigma_{ij}$  a média e covariância respectivamente. O retorno do portfólio  $R_{p,t} = w_t^T R_t$  segue uma normal com média  $\mu_{p,t} = w_t^T \mu_t$  e variância  $\sigma_{p,t}^2 = w_t^T \Sigma_t w_t$ . Por fim,  $\gamma$  é grau relativo de aversão ao risco do investidor.

## 4.2. Carteira de Mínima-Variância

Esta estratégia é na verdade um caso particular da carteira de média-variância. Sob esta estratégia o investidor decide em incorrer no mínimo de risco possível, ou seja, o valor do grau de aversão ao risco é infinito ( $\gamma = \infty$ ). Pode-se verificar que a segunda parte da função objetivo ( $\frac{1}{\gamma} E[R_{p,t+I}]$ ) tende a zero e a estratégia não leva em consideração o retorno esperado. A popularidade desta estratégia decorre das dificuldades de se trabalhar com os retornos, comentadas na subseção sobre as carteira de média-variância. De acordo com Merton (1980), é mais difícil obter estimativas precisas da média dos retornos em comparação com a sua variância. Outras evidências empíricas corroboram esta posição. Entre as quais pode-se citar Jagannathan e Ma (2003) e de Miguel et al 2009, que mostraram que a carteira de variância mínima tem desempenho superior as carteiras que contam com as estimativas de retornos esperados.

### 4.2.1. Encolhimento da matriz de covariância

Conforme exposto nas últimas duas subseções, os portfólios de média e mínima variância são implementados através da utilização da matriz de covariância. Uma proposta natural nestes problemas seria a utilização da matriz de covariância amostral. Ocorre que esta estimativa é obtida utilizando informação histórica e assim a sua precisão é incerta e a performance da otimização fora da amostra fica comprometida. Outra fonte de problemas deve-se ao fato de que não são raras as vezes em que estas matrizes são singulares, inviabilizando a solução do problema. Mais ainda, quando esta matriz contém erros, a sua inversão gera o efeito denominado Michaud (1989) como a maximização do erro, ou seja, sua potencialização devido o processo de inversão. Nas próprias palavras de Ledoit e Wolf (2003), ninguém deve utilizar a matriz de covariância com o propósito de otimização de portfólio. Desta forma optou-se por utilizar o estimador de *encolhimento* da matriz de covariância  $\Sigma$  proposto por Ledoit e Wolf (2003):

$$\hat{\Sigma}_{Shrink} = \hat{\delta}^* F + (1 - \hat{\delta}^*) S \quad (16)$$

### 4.3. Carteira Ingênua (1/N)

A carteira ingênua considera um portfólio com pesos  $w_t = 1/N$  para cada um dos  $N$  ativos de risco. O ponto forte desta carteira é não requerer nenhum tipo de otimização ou processo gerador dos dados. Desta forma, os retornos esperados desta estratégia são proporcionais a todo o risco da carteira e não ao risco sistemático. De Miguel et al (2009) mostraram que o desempenho desta carteira foi sistematicamente melhor do que um grupo de 14 outras estratégias, incluindo o portfólio de Média-Variância.

#### 4.4. Carteira de média variância condicional baseada no modelo VAR

A carteira de média variância condicional explora a dependência serial nas séries de retorno substituindo as estimativas amostrais por estimativas obtidas através dos modelos VAR. Esta abordagem segue os moldes de De Miguel (2014), e são ótimos para um investidor míope, que acredita que o retorno de ações se comportam de acordo com o modelo linear VAR. As equações (19) são resolvidas considerando a média do retorno do período seguinte condicional com o retorno no período atual:  $\mu_v = a + Br_t$ , onde  $a$  e  $B$  são os estimadores do coeficiente VAR obtidos através da Ridge Regression. A matriz de covariância amostral é substituída neste caso pela matriz de covariância dos resíduos:

$$\Sigma_v = \sum_{i=t-\tau+1}^t (r_i - a - Br_{i-1})(r_i - a - Br_{i-1})^T \quad (17)$$

#### 4.5. Carteiras de arbitragem

Carteiras de arbitragem são comumente chamadas de carteiras de custo zero. Sua estratégia baseia-se em assumir posição comprada em determinados ativos e em vendida em outros de tal forma que a soma destas posições seja igual a zero.

#### 4.6. Carteira de arbitragem através de um modelo VAR

Analogamente as carteiras de arbitragem contrária e de momento, o peso é dos ativos é definido através dos resultados obtidos no instante  $t$ . Contudo, neste caso a informação incorporada ao modelo são os coeficientes estimados no modelo VAR para os retornos. Algebricamente:

$$\begin{aligned} w_{v,t+1} &= \frac{1}{n}(\hat{r}_{t+1} - \bar{r}_{vt}e) \\ w_{v,t+1} &= \frac{1}{n}(a + Br_t - \bar{r}_{vt}e) \end{aligned} \quad (18)$$

Sendo  $e$  um vetor  $n$ -dimensional de 1. Note que  $a + Br_t$  é a previsão do modelo VAR para o retorno no tempo  $t + 1$  condicional ao retorno no tempo  $t$ , e que os pesos  $w$  somam zero, e por isso este é um portfólio de arbitragem. Também é válido ressaltar que o portfólio  $w_{v,t+1}$  designa pesos positivos para os ativos os quais o retorno esperado condicionado ao VAR está acima do retorno igualmente ponderado do portfólio. Para melhor visualização, a equação acima em sua forma explícita é:

$$\begin{pmatrix} w_{1,t+1} \\ w_{2,t+1} \\ \vdots \\ w_{n,t+1} \end{pmatrix} = \begin{pmatrix} \hat{y}_{(1)t+1} - \bar{y} \\ \hat{y}_{(2)t+1} - \bar{y} \\ \vdots \\ \hat{y}_{(n)t+1} - \bar{y} \end{pmatrix}$$

## 5. Medidas de desempenho

Nesta seção são apresentadas as medidas de desempenho utilizadas no trabalho para a comparação das estratégias de investimento. Todos os portfólios estão sendo analisados fora da amostra.

### 5.1. Média, desvio padrão e Turnover

A média e o desvio padrão são obtidos através da série de retornos da implantação respectiva de cada uma das estratégias. Estas medidas são anualizadas de forma que seja facilitada a comparação destas medidas com outros estudos. Estas medidas são calculadas da seguinte maneira:

$$\bar{\mu}_p = \hat{\mu}_p * 252 \quad (19)$$

$$\bar{\sigma}_p = \hat{\sigma}_p^2 * \sqrt{252}$$

O turnover é calculado da seguinte maneira:

$$turnover = (T - 1)^{-1} \sum_{t=1}^{T-1} \sum_{i=1}^n |w_{i,t+1} - w_{i,t}| \quad (20)$$

Esta estatística pode ser interpretada como a porcentagem média diária da carteira que é reponderada. Os custos de transação são diretamente influenciados por esta media. Assim uma estratégia pode performar bem quando não forem considerados os custos de transação, porém se esta estratégia tiver um alto turnover, quando adicionados estes custos pode ser que ela se torne economicamente inviável. Desta forma, para se ter um cenário mais adequado a realidade é vital considerar os custo de transação.

### 5.2. Índice de Sharpe - *Sharpe ratio*

O índice de Sharpe (*Sharpe ratio*) é uma das medidas de desempenho mais importantes na área de finanças. Esta pondera o retorno de um portfólio com o seu risco. Apesar de amplamente utilizada, já foi alvo de severas criticas principalmente quando os retornos não são distribuídos normalmente. Andrew Lo (2002) encontrou que a correlação serial presentes em séries de retornos mensais podem gerar Índices de Sharpe exagerados. Por se tratar de uma estatística amplamente aceita, será utilizada neste artigo como principal parâmetro de comparação do desempenho dos portfólios. O índice de Sharpe diário pode ser calculado através de:

$$\hat{S}R_p = \frac{\hat{\mu}_p - r_f}{\hat{\sigma}_p} \quad (21)$$

$$\bar{S}R_p = \hat{S}R_p * \sqrt{252}$$

Neste artigo o índice de Sharpe será utilizado como principal medida para avaliação das estratégias de alocação de portfólio, utilizando-se como benchmark o portfólio de mínima variância global. Dada a característica das séries de retorno apresentarem caudas espessas e desta forma se tornarem inválidos os testes que assumem normalidade, será utilizado neste artigo o teste estatístico proposto Ledoit e Wolf (2008). Os parâmetros utilizados foram: um número de reamostragens de bootstrap = 5000 e bloco de tamanho = 5.

### 5.3. Divisão e nomenclatura dos portfólios

Os portfólios foram divididos em três grupos. Grupo A são as estratégias que utilizam apenas os momentos amostrais para otimização. Grupo B é formado pelas estratégias que exploram a dependência serial no retorno dos ativos para os modelos VAR e a previsão da matriz de covariâncias através do modelo GARCH ortogonal. O grupo C contempla os portfólios de arbitragem. Após a análise dos resultados segue a relação da nomenclatura dos portfólios e respectivamente uma breve descrição.

Por economia de espaço na exposição dos resultados, serão utilizadas as seguintes siglas: PMVG - Portfólio de mínima variância global incondicional, PMV - Portfólio de média variância incondicional, PI - Portfólio Ingênuo, PI - Portfólio ingênuo, R.VAR - Portfólio de média variância utilizando VAR com regularização Ridge Regression, B.VAR - Portfólio de média variância utilizando VAR Bayesiano, GR.VAR- Portfólio de média variância com VAR na previsão retornos e GO-GARCH previsão da matriz de covariância, GR - Portfólio de mínima variância global utilizando GO-GARCH, R.VAR - Portfólio de arbitragem utilizando VAR com regularização Ridge Regression.

## 6. Performance empírica

Nesta seção serão detalhados os dados utilizados no artigo e os principais resultados.

### 6.1. Base de dados

A base de dados é formada pelas séries de retorno de ações sobreviventes da bolsa de valores BOVESPA obtidas através do site ECONOMATICA, e séries formadas por portfólios ponderados pelo valor de mercado das empresas e agrupados por *book-to-market*, tamanho e por tipo de indústria. Estes foram coletados no site do NEFIN (Brazilian Center for Research in Financial Economics of the University of São Paulo). Para o caso das ações individuais, foram coletados retornos diários de janeiro de 2001 a dezembro de 2016. Após coleta das ações individuais foi realizado filtro excluindo aquelas ações que por algum motivo tiveram um número alto de retornos não computados, resultando em 45 ações sobreviventes. Com relação aos portfólios obtidos no site do NEFIN, foram utilizados os arquivos: 4 portfólios organizados por tamanho e *book-to market*, 7 portfólios organizados por indústria. No portfólio da indústria, foi retirada a série de retornos referente à indústria da construção, pois só começou a ser registrada a partir de 2007. Ainda referente à indústria, aos portfólios restantes foram atribuídas as seguintes legendas: 1= Produtos Básicos, 2 = Consumo, 3 = Energia, 4 = Finanças, 5 = Manufatura e 6 = Outros. Todas as séries têm periodicidade diária.

### 6.2. Resultados

Os testes estatísticos realizados para detecção de relações de lead-lag evidenciaram que de fato existe uma certa dinâmica causal entre as séries de retorno dos ativos da Bovespa.

O setor em que esta dinâmica ficou mais clara foi no de produtos básicos, onde o retorno de todos os outros setores causam o de produtos básicos. Em alguns casos essa dinâmica é bidirecional como no caso do setor de consumo, que tem seus retornos causados por produtos básicos, energia e manufatura e ao mesmo tempo causa os retornos do setores de produtos básicos e manufatura. Em outros casos a relação é unidirecional como no caso dos retornos do setor energia que não é causado por nenhum setor e causa produtos básicos e consumo. Estes resultados são considerados com o nível de significância de 10%.

Os principais resultados com relação as estratégias de investimento são apresentados nas tabelas 1 e 2. A única estratégia que obteve resultados melhores em índice de Sharpe em comparação ao benchmark foi o portfólio de média variância utilizando VAR com regularização Ridge Regression (R.VAR). Este portfólio foi estatisticamente melhor que o Benchmark considerando um nível de significância de 1% para as primeiras duas bases de dados, porém performou pior para a base de dados com 45 ações sobreviventes do Bovespa. Estes resultados estão destacados em negrito nas tabelas 1 e 2.

Um dos problemas observados em relação a utilização dos modelos de previsão foi a instabilidade dos pesos dos ativos. Esta instabilidade gera um alto turnover nestas carteiras e quando são introduzidos custos de transação estas obtêm índice de Sharpe baixo, ou em alguns casos assumem valores negativos (o que em tese não tem interpretação econômica). Entre os portfólios do grupo B e C, o R.VAR performou relativamente bem na presença dos custos de transação, isto deve-se ao fato de que a regularização da matriz de covariância estabiliza os coeficientes autorregressivos e assim reduz o turnover e os custos finais de transação.

## **7. Conclusão**

O objetivo deste artigo foi analisar se modelos de previsão podem ser usados no lugar dos momentos amostrais com o intuito gerar melhores resultados fora da amostra em problemas de seleção de portfólio. Embora nenhuma estratégia tenha sido sistematicamente melhor do que o benchmark, algumas ideias surgem como potenciais linhas de pesquisa. Primeiramente, a utilização de modelos não paramétricos. Estes modelos são capazes de captar também as relações não lineares entre as variáveis do modelo. Desta forma seria possível explorar relações de dependência que os tradicionais modelos VAR não são capazes de captar. Adicionalmente, outras abordagens podem ser adotadas na modelagem e previsão da matriz de covariância. Um modelo apropriado para tal objetivo é o E-GARCH, que considera os efeitos de assimetria comuns nas séries de retorno.

Por último vale ressaltar que este trabalho utiliza de forma pioneira o método de regularização Ridge Regression para problemas de seleção de ativos no mercado acionário Brasileiro. Os resultados indicam a superioridade deste modelo em relação aos que utilizam somente previsão, abrindo uma janela para posterior investigação deste e de outros métodos de regularização como por exemplo: LASSO (least absolute shrinkage and selection operator). A Ridge Regression também mostra-se como um eficiente mecanismo de controle de turn-over, mantendo bons resultados mesmo com a inclusão de custos de transação. Este resultado é particularmente importante do ponto de vista prático, haja visto que, decisões na área de alocação de portfólio são tomadas sob este tipo de fricção.



**Tabela 1. Performance dos portfólios - Índice de Sharpe anualizado**

	Portfólios A			Portfólios B			Portfólios C		
	<i>PMVG</i>	<i>PMV</i>	<i>PI</i>	<i>R.VAR</i>	<i>B.VAR</i>	<i>GR.VAR</i>	<i>GR</i>	<i>R.VAR</i>	
<b>S/ Custos</b>									
4P - S/BM	0.7801	0.6861	0.7079	<b>0.8065</b>	0.7122	0.7301	0.7757	0.6862	
6P - Ind	0.8919	0.7194	0.7244	<b>0.9668</b>	0.7162	0.6633	0.6696	0.7564	
Seleção Bovespa	1.1513	0.7941	0.8408	0.9343	—	0.7730	0.7958	0.9069	
<b>C/ custos (20 pts)</b>									
4P - S/BM	0.7555	0.550	0.6950	0.7027	-1.280	0.6321	0.7497	0.5993	
6P - Ind	0.869	0.6020	0.7027	<b>0.8867</b>	-1.460	0.5250	0.5433	0.6218	
Seleção Bovespa	1.069	0.5693	0.8003	0.6210	—	0.4881	0.6186	0.7817	

**Tabela 2. Performance dos portfólios - Retorno anualizado**

	Portfólios A			Portfólios B			Portfólios C		
	<i>PMVG</i>	<i>PMV</i>	<i>PI</i>	<i>R.VAR</i>	<i>B.VAR</i>	<i>GR.VAR</i>	<i>GR</i>	<i>R.VAR</i>	
<b>S/ Custos</b>									
4P - S/BM	0.1633	0.1713	0.1633	<b>0.1814</b>	0.1732	0.1884	0.1834	0.1613	
6P - Ind	0.1850	0.1926	0.1841	<b>0.2267</b>	0.1960	0.1812	0.1789	0.1945	
Seleção Bovespa	0.1882	0.2302	0.2063	0.1939	(-)	0.2546	0.1847	0.4664	
<b>C/ custos (20 pts)</b>									
4P - S/BM	0.1582	0.1372	0.1602	0.1581	-0.3147	0.1630	0.1772	0.1338	
6P - Ind	0.1805	0.1611	0.1797	<b>0.2079</b>	-0.401	0.1434	0.1457	0.1599	
Seleção Bovespa	0.1748	0.1650	0.1964	0.1289	(-)	0.1609	0.1436	0.4021	

## Referências

- [1] C. Alexander. Modelos de mercados. *Um guia para análise de informações financeiras*. trad. Jose Carlos de Souza Santos. São Paulo: Bolsa de Mercadorias & Futuros, 2005.
- [2] A. Bakke. How good is the out-of-sample performance of optimized portfolios: an empirical comparison of optimal versus naive diversification. Master's thesis, Universitet i Agder/University of Agder, 2014.
- [3] J. Caldeira, G. Moura, A. A. Santos, et al. Portfolio optimization using a parsimonious multivariate garch model: application to the brazilian stock market. *Economics Bulletin*, 32(3):1848–1857, 2012.
- [4] M. Carrasco and N. Noumon. Optimal portfolio selection using regularization. Technical report, Citeseer, 2011.
- [5] O. R. de Medeiros and G. R. de Oliveira. Testando a existência de efeitos lead-lag entre os mercados acionários norte-americano e brasileiro.
- [6] V. DeMiguel, L. Garlappi, and R. Uppal. Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *Review of Financial Studies*, 22(5):1915–1953, 2009.
- [7] V. DeMiguel, F. J. Nogales, and R. Uppal. Stock return serial dependence and out-of-sample portfolio performance. *Review of Financial Studies*, 27(4):1031–1073, 2014.
- [8] P. Exterkate, P. J. Groenen, C. Heij, and D. van Dijk. Nonlinear forecasting with many predictors using kernel ridge regression. *International Journal of Forecasting*, 32(3):736–753, 2016.
- [9] E. F. Fama and K. R. French. Permanent and temporary components of stock prices. *Journal of political Economy*, 96(2):246–273, 1988.
- [10] L. L. Godeiro, W. P. S. D. F. Souza, A. C. Anegues, J. F. Caldeira, et al. Seleção de carteiras com retornos serialmente correlacionados: uma aplicação de modelos autoregressivos. In *Anais do XLIII Encontro Nacional de Economia [Proceedings of the 43rd Brazilian Economics Meeting]*, number 128. ANPEC-Associação Nacional dos Centros de Pósgraduação em Economia [Brazilian Association of Graduate Programs in Economics], 2016.
- [11] N. Jegadeesh. Evidence of predictable behavior of security returns. *The Journal of Finance*, 45(3):881–898, 1990.
- [12] N. Jegadeesh and S. Titman. Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of finance*, 48(1):65–91, 1993.
- [13] O. Ledoit and M. Wolf. Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of empirical finance*, 10(5):603–621, 2003.
- [14] O. Ledoit and M. Wolf. Honey, i shrunk the sample covariance matrix. *The Journal of Portfolio Management*, 30(4):110–119, 2004.
- [15] A. W. Lo and A. C. MacKinlay. When are contrarian profits due to stock market overreaction? *Review of Financial studies*, 3(2):175–205, 1990.

- [16] A. Meucci. *Managing diversification*. 2010.
- [17] F. M. MÜLLER, M. B. RIGHI, and P. S. CERETTA. Análise da eficiência de mercado do ibovespa: Uma abordagem com o modelo autorregressivo quantílico. *Revista Base (Administração e Contabilidade) da UNISINOS*, 12(2):122–134, 2015.
- [18] L. M. Trovati. Análise bayesiana da dependência temporal de séries de ações no mercado brasileiro. 2015.
- [19] R. S. Tsay. *Analysis of financial time series*, volume 543. John Wiley & Sons, 2005.
- [20] R. S. Tsay. *Multivariate Time Series Analysis: with R and financial applications*. John Wiley & Sons, 2013.
- [21] W. N. van Wieringen. Lecture notes on ridge regression. *arXiv preprint arXiv:1509.09169*, 2015.
- [22] L. Yang. An application of principal component analysis to stock portfolio management. 2015.
- [23] F. A. Ziegelmann. *Estimation of Volatility Functions: Nonparametric and Semi-Parametric Methods*. PhD thesis, Tese de Doutorado, University of Kent at Canterbury (UK), 2002.