

08 a 11 de Outubro de 2018
Instituto Federal Fluminense
Búzios - RJ

INVESTIGANDO A EFICIÊNCIA DE MERCADOS POR MEIO DE UM MODELO DE PREVISÃO DE TENDÊNCIAS DE RETORNOS DE ATIVOS

Charlene Cássia de Resende¹ - charlenecassia@gmail.com

Arthur Rodrigo Bosco de Magalhães² - magalhaes@des.cefetmg.br

Adriano César Machado Pereira³ - adrianoc@dcc.ufmg.br

¹Centro Federal de Educação Tecnológica de Minas Gerais - Belo Horizonte, MG, Brasil

²Centro Federal de Educação Tecnológica de Minas Gerais - Belo Horizonte, MG, Brasil

³Universidade Federal de Minas Gerais - Belo Horizonte, MG, Brasil

Resumo. *Definimos um modelo de previsão de tendências de retornos de ativos financeiros, baseado em um sistema de equações diferenciais lineares acopladas discretizado. Neste, consideramos que a relação entre os desvios dos preços das ações em relação a um preço considerado justo traz informação sobre a dinâmica dos preços das ações. O modelo de previsão de tendências se assemelha a um modelo de regressão linear multivariado. A validação estatística como tal será realizada em etapa futura. Parâmetros do modelo são explorados com o objetivo de encontrar uma combinação que apresente taxas de acurácia relevantes. Testes de hipóteses serão aplicados aos resultados da previsão para investigar se o modelo tem uma taxa de acerto de tendência de preços diferente de um processo totalmente aleatório.*

Palavras-chave: *Equações Diferenciais, Mercado de Ações, Séries Financeiras, Hipótese de Mercado Eficiente, Tendências*

1. INTRODUÇÃO

A teoria econômica clássica defende que não é possível obter informação de dados futuros baseando-se em informações passadas que levem a ganhos demais (Samuelson 1965). Seus teóricos defendem que os mercados seriam, assim, eficientes, e que os preços caminham em direção a um ponto de equilíbrio no longo prazo. Esta discussão inclui a definição da hipótese de mercado eficiente (HME) (Malkiel e Fama 1970; Shiller 1980). Tal hipótese considera que não há possibilidades de arbitragem, que o mercado *descobre* o preço correto dos ativos, não sendo possível ganhos acima da média sem correr riscos acima da média. Considera, também, que os investidores são racionais e que toda informação relevante é disponível e absorvida rapidamente pelo mercado (Malkiel e Fama 1970). No entanto, um grande número de analistas do mercado, fundamentalistas ou técnicos, estão em busca de tais informações (Feng et al. 2012).

Investigações empíricas sobre a eficiência do mercado têm sido realizadas em muitos contextos, utilizando diversas ferramentas estatísticas (Cajueiro e Tabak 2004; Ntim et al. 2015; Zunino et al. 2007; Choudhry e Jayasekera 2014; Choudhry e Jayasekera 2012; Caginalp e DeSantis 2011;

Belaire-Franch e Opong 2010; Podobnik et al. 2006; Resende et al. 2017). Pesquisas mostram que o mercado não é tão eficiente em termos de absorção de informação. Nem toda informação relevante está disponível a todos investidores (Ross 1989; Healy e Palepu 2001; Brunnermeier 2005). No estudo realizado por (Cajueiro e Tabak 2004), os expoentes de Hurst das séries de países desenvolvidos estão bem próximos de 0,5. Já os valores obtidos para países emergentes se distanciam mais fortemente deste valor, indicando que mercados emergentes são menos eficientes que mercados desenvolvidos. A ineficiência dos mercados emergentes, dentre outros fatores, pode estar relacionada à baixa liquidez do mercado (Chung e Hrazdil 2010) e ao fluxo de informação (Ross 1989; Healy e Palepu 2001; Brunnermeier 2005). Também há pesquisas no campo de finanças comportamentais (Kahneman e Tversky 1979), onde o comportamento irracional dos agentes evidencia desvios da hipótese de mercado eficiente.

Motivados pela lacuna deixada pela hipótese de mercado eficiente, sugerida por modelos focados na previsão de dinâmicas de mercados (Atsalakis e K 2013; Atsalakis e Valavanis 2009), e por modelos empenhados na previsão do sinal de retornos de ativos (Chevapatrakul 2013; Pönkä 2014; Anatolyev e Gospodinov 2012; Nyberg 2011), propomos também um modelo de previsão. Este é baseado em um sistema de equações diferenciais lineares acopladas. Discretizando o sistema, obtemos um modelo que se assemelha a um modelo de regressão linear multivariado. O objetivo geral é apresentar um modelo de previsão de tendências de preços de todas as ações que compõem os índices DOW JONES e IBOVESPA.

2. METODOLOGIA

A definição do modelo é apresentada na seção 3.. O modelo de previsão foi definido baseando-se em um sistema de equações diferenciais lineares acopladas (Resende et al. 2017). Este sistema foi discretizado de forma a se aproximar de um modelo de regressão linear multivariado. A matriz de coeficientes do modelo, obtida até o tempo t , será determinada baseada em um estimador de mínimos quadrados, descrito por (Aguirre 2007), e será utilizada como instrumento para a previsão da tendência dos preços no tempo $t + 1$.

As séries utilizadas nas análises são de preços de fechamento das ações que participaram da composição da carteira teórica dos índices DOW JONES e IBOVESPA. A amostra é composta por dados com frequência de cotação de 5 minutos no primeiro e segundo quadrimestres de 2017. Em dados *intraday*, maiores correlações são esperadas (Allez e Bouchaud 2011). O índice IBOVESPA foi escolhido por ser um mercado de país emergente, onde a ineficiência tende a ser mais evidenciada, possibilitando, assim, maiores oportunidades de sucesso na previsão de tendências de retornos (Cajueiro e Tabak 2004; Zunino et al. 2007). O mercado Americano foi escolhido para contrastar com este cenário e também devido à sua importância.

Na subseção 3.1, será apresentado de forma detalhada o tratamento dos dados. Pois, no modelo analisado, consideramos que os desvios dos preços em relação a um preço considerado justo traz informações sobre a dinâmica dos preços das ações. Tais desvios são calculados utilizando-se a diferença logarítmica entre o preço bruto e o preço justo. Porém, calcular o preço justo de cada ação envolve diversas análises e variáveis que não estão no foco do trabalho. Por isto, utilizamos como aproximação para tal preço justo uma média móvel exponencial.

Nas subseções 4.1 e 4.2, variamos o tamanho da janela de pontos utilizada no ajuste e avaliamos os resultados em relação a manutenção dos parâmetros do modelo e sua capacidade de previsão de tendência de preços dos ativos.

3. MODELO

3.1 Tratamento dos dados

O preço bruto de cada uma das n ações no tempo t é representado por $x_k(t)$ ($k = 1, 2, \dots, n$). Utilizamos dados com *preços tratados*, que consistem nos desvios destes preços brutos $x_k(t)$ em relação a preços considerados justos. A forma que estabelecemos como aproximação para o cálculo foi uma média móvel exponencial $\bar{x}_k(t)$, calculada em um período precedente ao preço bruto analisado, definida como:

$$\bar{x}_k(t) = (x_k(t) - \bar{x}_k(t-1)) \times \frac{2}{(S+1)} + \bar{x}_k(t-1), \quad (1)$$

onde S é um parâmetro relacionado ao peso dado ao último valor $x_k(t)$ no cálculo e $\bar{x}_k(t-1)$ é a média móvel exponencial calculada no período anterior. Quanto maior o tamanho de S menor o peso dado ao último preço. Neste trabalho, utilizamos $S = 12$ para o cálculo das médias móveis. Com isto, a nova série de preços tratados será dada pela diferença entre o logaritmo dos preços brutos e o logaritmo dos preços considerados justos, definida como

$$X_k(t) = \ln x_k(t) - \ln \bar{x}_k(t). \quad (2)$$

Consideramos que os desvios dos preços brutos em relação aos preços justos é que são importantes para a representação das dinâmicas das séries. Desta forma, a partir deste ponto, em nossas análises utilizaremos as novas séries de *preços tratados*.

3.2 Ajuste linear do dados

O modelo de equações diferenciais lineares, que propusemos em (Resende et al. 2017), é dado na forma

$$\frac{d}{dt} \mathbf{X}_a(t) = \mathbf{B} \mathbf{X}_a(t), \quad (3)$$

onde

$$\mathbf{X}_a(t) = \begin{bmatrix} X_1(t-n+1) & \cdots & X_1(t-1) & X_1(t) \\ X_2(t-n+1) & \cdots & X_2(t-1) & X_2(t) \\ \vdots & \ddots & \vdots & \vdots \\ X_n(t-n+1) & \cdots & X_n(t-1) & X_n(t) \end{bmatrix}$$

e

$$\mathbf{B} = \begin{bmatrix} B_{1,1} & B_{1,2} & \cdots & B_{1,n} \\ B_{2,1} & B_{2,2} & \cdots & B_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ B_{n,1} & B_{n,2} & \cdots & B_{n,n} \end{bmatrix}.$$

A equação (3) equivale a n equações do tipo

$$\frac{d}{dt} \begin{bmatrix} X_1(\tau) \\ X_2(\tau) \\ \vdots \\ X_n(\tau) \end{bmatrix} = \mathbf{B} \begin{bmatrix} X_1(\tau) \\ X_2(\tau) \\ \vdots \\ X_n(\tau) \end{bmatrix}, \quad (4)$$

com $t - n + 1 \leq \tau \leq t$.

Neste modelo, consideramos que os desvios dos preços brutos em relação aos preços considerados justos são os elementos importantes para a caracterização das dinâmicas das séries.

Discretizando o modelo de equações diferenciais, representado pela equação (3), obtemos

$$\frac{\mathbf{X}_a(t) - \mathbf{X}_a(t-1)}{\Delta t} = \mathbf{B}\mathbf{X}_a(t), \quad (5)$$

equivalente a

$$\begin{aligned} \mathbf{X}_a(t-1) &= \mathbf{X}_a(t) - \Delta t \mathbf{B}\mathbf{X}_a(t), \\ &= (\mathbf{I} - \Delta t \mathbf{B})\mathbf{X}_a(t), \end{aligned} \quad (6)$$

onde \mathbf{I} é a matriz identidade $n \times n$. Assumindo que $\mathbf{A}_a = (\mathbf{I} - \Delta t \mathbf{B})^{-1}$ existe, podemos escrever

$$\mathbf{X}_a(t) = \mathbf{A}_a \mathbf{X}_a(t-1). \quad (7)$$

No modelo apresentado, se existir a inversa de $\mathbf{X}_a(t-1)$, podemos encontrar as matrizes \mathbf{A}_a perfeitamente ajustadas para a equação (7), isto é, o sistema definido desta maneira apresenta soluções que passam em todos os pontos no intervalo $[t-n; t]$. Agora, adaptaremos o modelo buscando matrizes \mathbf{A}_a que são aproximadamente ajustadas em intervalos maiores. Para a definição, suponha que a evolução dos preços tratados ainda possa ser expressa pela combinação linear das n ações analisadas, sendo

$$\mathbf{X}(t) = \mathbf{A}\mathbf{X}(t-1) + \epsilon, \quad (8)$$

onde

$$\mathbf{X}(t) = \begin{bmatrix} X_1(t-R+1) & \dots & X_1(t-1) & X_1(t) \\ X_2(t-R+1) & \dots & X_2(t-1) & X_2(t) \\ \vdots & \ddots & \vdots & \vdots \\ X_n(t-R+1) & \dots & X_n(t-1) & X_n(t) \end{bmatrix},$$

$$\mathbf{A} = \begin{bmatrix} A(1,1) & \dots & A(1,n-1) & A(1,n) \\ A(2,1) & \dots & A(2,n-1) & A(2,n) \\ \vdots & \ddots & \vdots & \vdots \\ A(n,1) & \dots & A(n,n-1) & A(n,n) \end{bmatrix},$$

$$\epsilon = \begin{bmatrix} \epsilon_1(t-R+1) & \dots & \epsilon_1(t-1) & \epsilon_1(t) \\ \epsilon_2(t-R+1) & \dots & \epsilon_2(t-1) & \epsilon_2(t) \\ \vdots & \ddots & \vdots & \vdots \\ \epsilon_n(t-R+1) & \dots & \epsilon_n(t-1) & \epsilon_n(t) \end{bmatrix},$$

ϵ é o erro associado ao ajuste, R é a quantidade de pontos utilizados no ajuste e n a quantidade de ações analisadas.

Observe que a equação (8) é diferente da equação (7), uma vez que existe um erro associado ao ajuste aproximado e também pelo fato de que $\mathbf{X}_a(t)(\mathbf{X}_a(t-1))$ é $n \times n$ e $\mathbf{X}(t)(\mathbf{X}(t-1))$ é da ordem $n \times R$ e, conseqüentemente, a matriz de coeficientes \mathbf{A} é diferente de \mathbf{B} .

Expandindo a equação (8), para $t - R + 1 \leq \tau \leq t$, o sistema fica na forma

$$\begin{cases} X_1(\tau) = A(1, 1)X_1(\tau - 1) + A(1, 2)X_2(\tau - 1) + \dots + A(1, n)X_n(\tau - 1) + \varepsilon_1(\tau) \\ X_2(\tau) = A(2, 1)X_1(\tau - 1) + A(2, 2)X_2(\tau - 1) + \dots + A(2, n)X_n(\tau - 1) + \varepsilon_2(\tau) \\ \vdots \\ X_n(\tau) = A(n, 1)X_1(\tau - 1) + A(n, 2)X_2(\tau - 1) + \dots + A(n, n)X_n(\tau - 1) + \varepsilon_n(\tau) \end{cases} \quad (9)$$

A equação (8) se assemelha a um *modelo de regressão linear multivariado*. Entretanto, a sua validação como tal modelo depende da capacidade do mesmo em representar a dinâmica que produziu os dados. Esta análise de validação será tratada em outros trabalhos.

Como conhecemos $\mathbf{X}(t)$ e $\mathbf{X}(t - 1)$, podemos determinar a matriz de coeficientes \mathbf{A} fazendo (Aguirre 2007)

$$\mathbf{A} = (\mathbf{X}(t)\mathbf{X}(t - 1)^T)(\mathbf{X}(t - 1)\mathbf{X}(t - 1)^T)^{-1}. \quad (10)$$

A matriz \mathbf{A} pode variar de acordo com a escolha de t , conforme a equação (10). Deste modo, usaremos a notação $\mathbf{A}(t)$ para denotar a matriz \mathbf{A} , calculada utilizando pontos do tempo $t - R$ até o tempo t .

Com a matriz $\mathbf{X}(t)$ e $\mathbf{X}(t - 1)$, podemos calcular a variação nos preços das ações, que é a diferença do preço da ação no período t em relação ao período anterior $t - 1$. Esta variação do preço $\Delta\mathbf{X}(t)$ é denominada *retorno* do ativo, definida por

$$\Delta\mathbf{X}(t) = \begin{bmatrix} \Delta X_1(t) \\ \Delta X_2(t) \\ \vdots \\ \Delta X_n(t) \end{bmatrix}, \quad (11)$$

sendo

$$\begin{aligned} \Delta X_k(t) &= X_k(t) - X_k(t - 1), \\ &= \log\left(\frac{x_k(t)}{\bar{x}_k(t)}\right) - \log\left(\frac{x_k(t - 1)}{\bar{x}_k(t - 1)}\right), \end{aligned} \quad (12)$$

onde usamos no passo final a equação (2).

3.3 Modelo de previsão

Para tentar prever o preço no tempo $t + 1$, vamos utilizar como parâmetro a matriz $\mathbf{A}(t)$, calculada utilizando-se $R + 1$ preços passados. Para tanto, é desejável que $\mathbf{A}(t)$ varie pouco no tempo, sendo $\mathbf{A}(t)$ próxima a $\mathbf{A}(t - 1)$. Com esta perspectiva, empregamos $\mathbf{A}(t)$ como estimativa para $\mathbf{A}(t + 1)$ no modelo de previsão, definido por

$$\hat{\mathbf{X}}(t + 1) = \mathbf{A}(t)\mathbf{X}_c(t), \quad (13)$$

onde $\mathbf{A}(t)$ foi obtida com os $R + 1$ preços até o tempo t ,

$$\hat{\mathbf{X}}(t + 1) = \begin{bmatrix} \hat{X}_1(t + 1) \\ \hat{X}_2(t + 1) \\ \vdots \\ \hat{X}_n(t + 1) \end{bmatrix}$$

é a matriz com preços previstos das ações no tempo $t + 1$ e

$$\mathbf{X}_c = \begin{bmatrix} X_1(t) \\ X_2(t) \\ \vdots \\ X_n(t) \end{bmatrix}$$

é uma matriz coluna com os valores dos preços tratados no tempo t .

Mesmo que haja pouca variação entre as matrizes $\mathbf{A}(t)$ e $\mathbf{A}(t + 1)$, é natural obtermos um erro associado à previsão.

O retorno previsto das ações é dado por

$$\hat{\Delta}\mathbf{X}(t + 1) = \begin{bmatrix} \hat{\Delta}_1 X(t + 1) \\ \hat{\Delta}_2 X(t + 1) \\ \vdots \\ \hat{\Delta}_n X(t + 1) \end{bmatrix}, \quad (14)$$

onde

$$\hat{\Delta}X_k(t + 1) = \hat{X}_k(t + 1) - X_k(t). \quad (15)$$

A partir deste retorno, definimos o vetor de tendência

$$\mathbf{T} = \begin{bmatrix} T_1(t) \\ T_2(t) \\ \vdots \\ T_n(t) \end{bmatrix}, \quad (16)$$

onde T são variáveis booleanas, assumindo os valores -1 e 1 , sendo 1 (-1) associado à tendência de alta (baixa) do preço. O bom resultado do modelo ocorrerá quando a tendência do preço previsto for igual a tendência observada do preço real. O desempenho do modelo será dado pela acurácia

$$a = \frac{N_c}{N}, \quad (17)$$

onde N_c é o número de previsões de tendências corretas e N é o número total de previsões executadas pelo modelo de previsão.

Nos resultados que apresentamos em (Resende et al. 2017), mostramos que o tamanho do retorno previsto traz informação sobre a tendência. Logo, dividimos o conjunto de retornos previstos U em 4 subconjuntos U_1 , U_2 , U_3 e U_4 , de forma que cada subconjunto contenha um quarto dos retornos previstos. Os maiores retornos (em módulo) estão em U_4 , seguidos pelos retornos em U_3 , U_2 e U_1 , nesta ordem. Os melhores resultados em (Resende et al. 2017) foram obtidos para o subconjunto U_4 . Portanto, neste trabalho, vamos analisar apenas as tendências previstas obtidas com retornos previstos em U_4 .

4. EXPLORAÇÃO E RESULTADOS DO MODELO

4.1 Distâncias entre as matrizes $\mathbf{A}(t)$ em tempos adjacentes

Para analisar a variação da matriz $\mathbf{A}(t)$ no tempo, definimos a distância

$$d(\mathbf{A}(t), \mathbf{A}(t-1)) = \sqrt{\sum_{i=1}^n \sum_{j=1}^n |A(t)_{i,j} - A(t-1)_{i,j}|^2}, \quad (18)$$

onde $A(t)_{i,j}$ ($A(t-1)_{i,j}$) é o elemento da matriz $\mathbf{A}(t)$ ($\mathbf{A}(t-1)$) na i^{th} linha e j^{th} coluna, entre as matrizes adjacentes $\mathbf{A}(t)$ e $\mathbf{A}(t-1)$ para vários valores de R e mostramos os resultados na Figura 1.

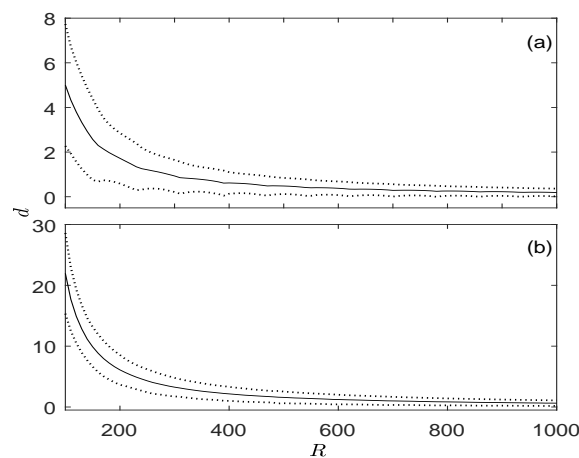


Figure 1- Distribuição das distâncias entre as matrizes \mathbf{A} em tempos adjacentes para valores de R variando de 100 a 1.000. Foram realizadas 5.000 execuções para as ações do IBOVESPA e 3.900 para as ações que compõem o DOW JONES. Utilizamos séries de preços com granularidade de 5 minutos, no período de 02/01/17 a 28/04/17 (primeiro quadrimestre de 2017). A linha contínua representa a média das distâncias e a linha pontilhada mostra a média de $d \pm$ os desvios dos índices analisados. (a) DOW JONES. (b) IBOVESPA.

Analisando os resultados da Figura 1, obtidos com a Eq. (10), percebemos que, à medida em que aumentamos o valor de R , as distâncias entre as matrizes calculadas em tempos adjacentes diminuíram. Este resultado nos motiva a explorar o efeito da variação do tamanho de R nos valores das acurácias a do modelo de previsão.

4.2 Variação do tamanho de R

O objetivo do modelo é prever a tendência do preço no tempo $t + 1$. Para isto, ele utiliza a matriz de coeficientes \mathbf{A} que minimiza a diferença quadrática entre a matriz de preços reais $\mathbf{X}(t + 1)$ e a matriz de preços previstos $\hat{\mathbf{X}}(t + 1)$. Para calcular esta matriz \mathbf{A} é preciso $R + 1$ pontos. Aqui, analisamos os resultados de acurácia do modelo na previsão de tendência de preços para diferentes valores de R . Para isto, definimos a acurácia média

$$\bar{a} = \frac{\sum_k^n a_k}{n} \quad (19)$$

e o desvio padrão

$$\sigma_a = \sqrt{\frac{\sum_k^n (a_k - \bar{a})^2}{n - 1}}, \quad (20)$$

onde a_k é a acurácia a para cada ação k e n é o número total de ações analisadas.

Nesta análise, uma acurácia média \bar{a} , envolvendo todas as k ações, é calculada para cada tamanho de R . São realizadas 5.000 execuções para as ações do IBOVESPA e 3.900 para as ações do DOW JONES. A diferença na quantidade de execuções entre DOW JONES e IBOVESPA é devida ao menor tempo diário de permanência do pregão de negociação do DOW JONES em relação ao IBOVESPA. A Figura 2 mostra a acurácia média \bar{a} e $\pm\sigma_a$ do modelo para todas as ações que compõem os índices DOW JONES e IBOVESPA. Os dados analisados são referentes ao primeiro quadrimestre de 2017. Utilizamos diferentes tamanhos de R variando de 100 a 1.000, com incremento de 10 pontos a cada nova execução do programa. O objetivo desta Figura é sinalizar o momento em que não há mais ganhos significativos na acurácia à medida que aumentamos o tamanho de R .

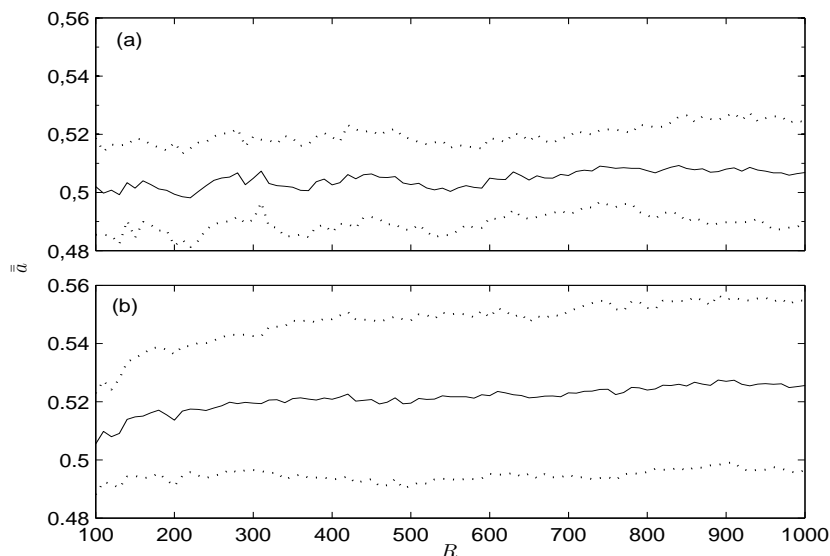


Figure 2- **Acurácia média para todas as n ações que compõem os índices DOW JONES e IBOVESPA.** Os dados analisados são referentes às séries de preços com granularidade de 5 minutos, no período de 02/01/17 a 28/04/17 (primeiro quadrimestre de 2017). Foram utilizados diferentes tamanhos de R . Com valor inicial de $R = 100$, com incremento de 10. Para cada valor de R , foram realizadas 5.000 execuções para a série do IBOVESPA e 3.900 para DOW JONES. A linha contínua representa a acurácia média \bar{a} obtida para todas as ações dos índices analisados e a linha pontilhada representa a acurácia média $\bar{a} \pm \sigma_a$ das acurácias obtidas. (a) DOW JONES e (b) IBOVESPA.

Podemos observar na Figura 2 que para valores maiores de R a curva da acurácia média \bar{a} tende a quase se estabilizar, porém ainda podemos perceber um leve crescimento. Para analisar se ainda há ganhos na acurácia para valores de $R > 500$, calculamos o coeficiente angular β e os coeficientes de determinação r^2 da reta de ajuste aos pontos com valores de R superiores a 500. Os coeficientes angulares β para todas as ações dos índices DOW JONES e IBOVESPA, calculados para o intervalo $500 \leq R \leq 1.000$, foram $\beta = 5,63 \times 10^{-5}$ e $\beta = 1,24 \times 10^{-5}$, respectivamente. Embora pequenos, não são irrelevantes: observe que numa variação de 0,02 na

acurácia (o que é significativo) entre $R = 500$ e $R = 1.000$, temos coeficiente angular 4×10^{-5} , ou seja, na ordem dos observados. Os coeficientes de determinação foram $r^2 = 0,2905$ para todas as ações do índice DOW JONES e $r^2 = 0,6295$ para todas as ações do IBOVESPA. Eles indicam que o ajuste linear não é satisfatório; Contudo, isto não é relevante aqui, uma vez que estamos interessados sobretudo no caráter crescente das curvas. Valores dos coeficientes angulares β , calculados para o intervalo $500 \leq R \leq 1.000$, pequenos, porém positivos, indicam crescimento da curva e nos motivam a investigar o comportamento do gráfico para valores de R maiores que 1.000.

Analisando a Figura 2, percebemos que à medida que aumentamos os valores de R , a curva da acurácia média tende a oscilar menos. Isso acontece devido ao incremento de 10 novos pontos ser cada vez menos significativos à medida em que aumentamos a quantidade de pontos na janela. O efeito é mais evidente ao compararmos os valores de R próximos a 10 com aqueles no final da série.

A Figura 2 nos sugere que a acurácia para as ações do índice DOW JONES tende a se estabilizar para valores de $R \geq 400$ pontos. Porém, a curva ainda mostra um leve crescimento dos resultados. Para as ações do índice IBOVESPA, é evidente que a curva de acurácia apresenta ganhos significativos à medida em que aumentamos o tamanho de R . Estas observações são evidenciadas pelos valores dos coeficientes angulares obtidos.

Os resultados apresentados na Figura 2 são as acurácias médias \bar{a} . Como estes valores se referem às médias das acurácias a de todas as ações para cada valor de R , não podemos visualizar o comportamento *individual* das ações para os diferentes tamanhos de R . Com isto, para mostrar a dispersão das acurácias das ações para cada valor de R , construímos os boxplots dos resultados (Montgomery 2008). A dispersão, localização e simetria das acurácias podem ser representadas pela Figura 3.

Como esperado, a Figura 3 mostra que as ações que compõem o índice IBOVESPA apresentam mais *outliers* com altas taxas de acurácias quando comparadas às ações do índice DOW JONES.

5. CONCLUSÃO

Observamos que a variação do tamanho da janela utilizada no ajuste reflete nos resultados de acurácias do modelo. Os ganhos de a tendem a aumentarem com valores maiores de R . Diante dos resultados, nosso objetivo foi mostrar que o modelo de previsão de tendência definido apresenta uma taxa de acurácia superior a um processo totalmente aleatório. Para validar estes resultados, realizamos análise estatística baseada em testes de hipóteses. Onde a hipótese a ser testada é que a acurácia do modelo é estatisticamente diferente de 0,5. Para os tamanhos de amostras analisadas, qualquer taxa de acurácia superior a 53% obtêm um p -valor menor que o nível de significância estabelecido de 0,05 para o teste. Este resultado indica que podemos rejeitar a hipótese nula de que o modelo é um processo totalmente aleatório. Portanto, encontramos desvios da hipótese de mercado eficiente nas séries analisadas por meio das acurácias obtidas. Como esperado, as maiores acurácias foram obtidas para as ações que compõem o índice IBOVESPA.

Agradecimentos

Agradecemos o suporte dado pelo CEFET/MG e FAPEMIG.

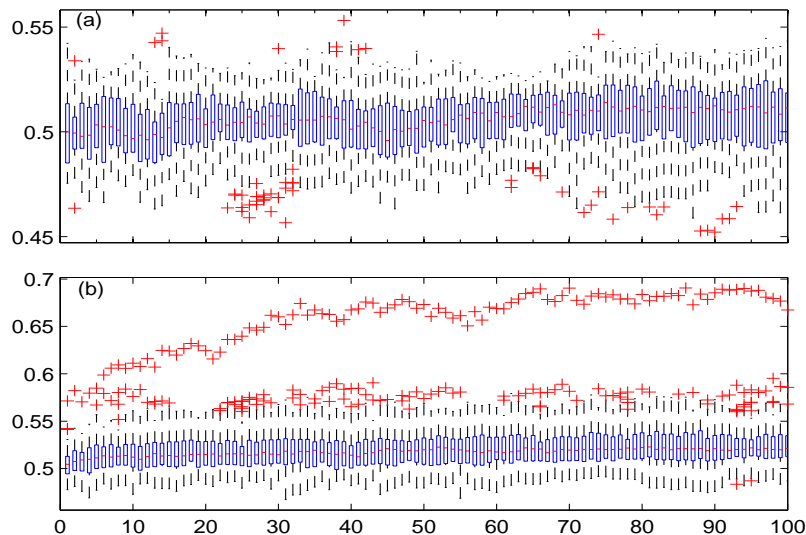


Figure 3- **Boxplot da Acurácia para todas as ações que compõem os índices DOW JONES e IBOVESPA.** Os dados utilizados foram preços de fechamento das ações que compõem os índices DOW JONES e IBOVESPA com granularidade de 5 minutos, no período de 02/01/17 a 28/04/17 (primeiro quadrimestre de 2017). Foram analisados 100 diferentes tamanhos de R . Para cada valor de R , foram realizadas 3.900 execuções para as ações do índice DOW JONES e 5.000 execuções para as ações que compõem a carteira teórica do IBOVESPA. (a) DOW JONES. (b) IBOVESPA.

REFERENCES

- [Aguirre 2007]AGUIRRE, L. *Introdução à Identificação de Sistemas – Técnicas Lineares e Não-Lineares Aplicadas a Sistemas Reais*. Editora UFMG, 2007. ISBN 9788570415844. Disponível em: <<https://books.google.com.br/books?id=f9IwE7Ph0fYC>>.
- [Allez e Bouchaud 2011]ALLEZ, R.; BOUCHAUD, J.-P. Individual and collective stock dynamics: intra-day seasonalities. *New Journal of Physics*, IOP Publishing, v. 13, n. 2, p. 025010, 2011.
- [Anatolyev e Gospodinov 2012]ANATOLYEV, S.; GOSPODINOV, N. Modeling financial return dynamics via decomposition. *Journal of Business & Economic Statistics*, Taylor & Francis, 2012.
- [Atsalakis e K 2013]ATSALAKIS, G.; K, V. Surveying stock market forecasting techniques - part i: Conventional methods. *Computation optimization in economics and finance research compendium*, p. 49–104, 01 2013.
- [Atsalakis e Valavanis 2009]ATSALAKIS, G. S.; VALAVANIS, K. P. Surveying stock market forecasting techniques—part ii: Soft computing methods. *Expert Systems with Applications*, Elsevier, v. 36, n. 3, p. 5932–5941, 2009.
- [Belaire-Franch e Opong 2010]BELAIRE-FRANCH, J.; OPONG, K. K. Testing for random walk in euro exchange rates using the subsampling approach. *Applied Economics Letters*, Taylor & Francis, v. 17, n. 12, p. 1145–1151, 2010.
- [Brunnermeier 2005]BRUNNERMEIER, M. K. Information leakage and market efficiency. *Review of Financial Studies*, Soc Financial Studies, v. 18, n. 2, p. 417–457, 2005.
- [Caginalp e DeSantis 2011]CAGINALP, G.; DESANTIS, M. A paradigm for quantitative behavioral finance. *American Behavioral Scientist*, Sage Publications, v. 55, n. 8, p. 1014–1034, 2011.

- [Cajueiro e Tabak 2004]CAJUEIRO, D. O.; TABAK, B. M. The hurst exponent over time: testing the assertion that emerging markets are becoming more efficient. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 336, n. 3, p. 521–537, 2004.
- [Chevapatrakul 2013]CHEVAPATRAKUL, T. Return sign forecasts based on conditional risk: Evidence from the uk stock market index. *Journal of Banking & Finance*, Elsevier, v. 37, n. 7, p. 2342–2353, 2013.
- [Choudhry e Jayasekera 2012]CHOUDHRY, T.; JAYASEKERA, R. Comparison of efficiency characteristics between the banking sectors of us and uk during the global financial crisis of 2007–2011. *International Review of Financial Analysis*, Elsevier, v. 25, p. 106–116, 2012.
- [Choudhry e Jayasekera 2014]CHOUDHRY, T.; JAYASEKERA, R. Market efficiency during the global financial crisis: Empirical evidence from european banks. *Journal of international Money and Finance*, Elsevier, v. 49, p. 299–318, 2014.
- [Chung e Hrazdil 2010]CHUNG, D.; HRAZDIL, K. Liquidity and market efficiency: A large sample study. *Journal of Banking & Finance*, Elsevier, v. 34, n. 10, p. 2346–2357, 2010.
- [Feng et al. 2012]FENG, L. et al. Linking agent-based models and stochastic models of financial markets. *Proceedings of the National Academy of Sciences*, National Acad Sciences, v. 109, n. 22, p. 8388–8393, 2012.
- [Healy e Palepu 2001]HEALY, P. M.; PALEPU, K. G. Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. *Journal of accounting and economics*, Elsevier, v. 31, n. 1, p. 405–440, 2001.
- [Kahneman e Tversky 1979]KAHNEMAN, D.; TVERSKY, A. Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the econometric society*, JSTOR, v. 47, n. 2, p. 263–291, 1979.
- [Malkiel e Fama 1970]MALKIEL, B. G.; FAMA, E. F. Efficient capital markets: A review of theory and empirical work. *The journal of Finance*, Wiley Online Library, v. 25, n. 2, p. 383–417, 1970.
- [Montgomery 2008]MONTGOMERY, D. C. *Design and analysis of experiments*. [S.l.]: John Wiley & Sons, 2008.
- [Ntim et al. 2015]NTIM, C. G. et al. On the efficiency of the global gold markets. *International Review of Financial Analysis*, Elsevier, v. 41, p. 218–236, 2015.
- [Nyberg 2011]NYBERG, H. Forecasting the direction of the us stock market with dynamic binary probit models. *International Journal of Forecasting*, Elsevier, v. 27, n. 2, p. 561–578, 2011.
- [Podobnik et al. 2006]PODOBNIK, B. et al. Fractionally integrated process for transition economics. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 362, n. 2, p. 465–470, 2006.
- [Pönkä 2014]PÖNKÄ, H. Predicting the direction of us stock markets using industry returns. *Empirical Economics*, Springer, p. 1–30, 2014.
- [Resende et al. 2017]RESENDE, C. C. de et al. Investigating market efficiency through a forecasting model based on differential equations. *Physica A: Statistical Mechanics and its Applications*, v. 474, n. C, p. 199–212, 2017.
- [Ross 1989]ROSS, S. A. Information and volatility: The no-arbitrage martingale approach to timing and resolution irrelevancy. *The Journal of Finance*, Wiley Online Library, v. 44, n. 1, p. 1–17, 1989.
- [Samuelson 1965]SAMUELSON, P. A. *Proof that properly anticipated prices fluctuate randomly*. [S.l.: s.n.], 1965. 41–45 p.

[Shiller 1980]SHILLER, R. J. *Do stock prices move too much to be justified by subsequent changes in dividends?* [S.l.]: National Bureau of Economic Research Cambridge, Mass., USA, 1980. 421–436 p.

[Zunino et al. 2007]ZUNINO, L. et al. Inefficiency in latin-american market indices. *The European Physical Journal B*, Springer, v. 60, n. 1, p. 111–121, 2007.

APPENDIX A

INVESTIGATING MARKET EFFICIENCY THROUGH A MODEL OF FORECASTING RETURN TRENDS OF ASSETS

Abstract. *We define a predictive model of financial asset returns, based on a discretized coupled linear differential equations system. In this, we consider that the relationship between stock price deviations from a price considered fair brings information about the dynamics of stock prices. The trend prediction model resembles a multivariate linear regression model. Statistical validation as such will be carried out in the future. Parameters of the model are explored in order to find a combination that presents relevant accuracy rates. Hypothesis tests will be applied to the forecast results to investigate whether the model has a price trend hit rate other than a totally random process.*

Keywords: *Differential Equations, Stock Market, Financial Series, Efficient Market Hypothesis, Trends*