

INSTITUTO DE MATEMÁTICA PURA E APLICADA - IMPA
MESTRADO PROFISSIONAL EM FINANÇAS

VICTOR DIONISIO DE LIMA

**UTILIZAÇÃO DO MODELO OCULTO DE MARKOV (HMM)
NUM SISTEMA DE NEGOCIAÇÃO DE AÇÕES
BRASILEIRAS**

DISSERTAÇÃO

RIO DE JANEIRO
2018

VICTOR DIONISIO DE LIMA

**UTILIZAÇÃO DO MODELO OCULTO DE MARKOV (HMM)
NUM SISTEMA DE NEGOCIAÇÃO DE AÇÕES
BRASILEIRAS**

Dissertação apresentado ao Mestrado Profissional em
Finanças do Instituto de Matemática Pura e Aplicada
- IMPA, como requisito parcial para a obtenção do título
de Mestre.

Orientador: Roberto Imbuzeiro Felinto de Oliveira
IMPA

RIO DE JANEIRO
2018

AGRADECIMENTOS

Primeiramente, agradeço a Deus por ter me dado a perseverança e fé necessárias para superar as adversidades que enfrentei ao longo desta jornada. Agradeço também aos meus pais, Jorge e Maria Júlia, à minha irmã Verônica, e à minha avó Izabel, por terem sido uma fonte de apoio e exemplo durante toda minha vida. Faço também um agradecimento especial ao meu orientador, professor Roberto Imbuzeiro de Oliveira, por todo o tempo e paciência dedicados durante a orientação deste trabalho. E por fim, agradeço aos professores, monitores, funcionários da Coordenação de Ensino e colegas de turma pelo convívio, suporte e o debate de ideias, que são essenciais para o desenvolvimento acadêmico.

RESUMO

LIMA, Victor Dionisio. Utilização do Modelo Oculto de Markov (HMM) num Sistema de Negociação de Ações Brasileiras. 2018. 74 f. Dissertação – Mestrado Profissional em Finanças, Instituto de Matemática Pura e Aplicada - IMPA. Rio de Janeiro, 2018.

O objetivo deste trabalho é analisar estratégias simples de negociação de ações no mercado brasileiro baseadas nas previsões de preços feitas pelo Modelo Oculto de Markov (HMM). Utilizaremos o Critério de Informação de Akaike (AIC) e o Critério de Informação Bayesiano (BIC) para escolher o número ideal de estados do modelo HMM. O sistema de negociação de ações projetado terá como elementos fundamentais o modelo HMM e a Janela Temporal Retrospectiva. As previsões serão feitas para os preços diários de fechamento das ações utilizando dados de observações múltiplas. Tendo como parâmetro a estratégia *buy and hold*, concluiremos que o sistema de negociação apresentado neste trabalho apresenta bom desempenho num cenário de desvalorização das ações e uma performance irregular quando as ações estão num período de alta. Sendo assim, recomendaremos que o HMM precisa ser mais estudado em múltiplas situações para conhecermos melhor as forças e fraquezas deste modelo num contexto de negociação de ações no mercado financeiro.

Palavras-chave: Modelo Oculto de Markov (HMM). Séries Temporais Financeiras. Sistemas de Negociação de Ações.

ABSTRACT

LIMA, Victor Dionisio. Application of the Hidden Markov Model (HMM) in a Stock Trading System of Brazilian's Shares. 2018. 74 f. Dissertação – Mestrado Profissional em Finanças, Instituto de Matemática Pura e Aplicada - IMPA. Rio de Janeiro, 2018.

The aim of this work is to analyze simple stock trading strategies in the Brazilian Stock Market based on the price forecasts made by the Hidden Markov Model (HMM). We will use the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) to choose the right number of HMM's hidden states. The projected stock trading system will have as fundamental elements HMM and the Temporal Retrospective Window. Predictions will be made for daily stock closing prices using multiple observation data. Taking as a parameter the *buy and hold* strategy, we will conclude that the trading system presented in this work performs well in a scenario of bear market and an irregular performance when faces a bull market. Therefore, we will recommend that HMM needs to be studied in multiple situations in order to better understand the strengths and weaknesses of this model in the context of financial market trading.

Keywords: Hidden Markov Model (HMM). Financial Time Series. Stock Trading Systems.

LISTA DE FIGURAS

Figura 1 – Desempenho do modelo HMM-3/80d num período de desvalorização (baixa).	34
Figura 2 – Comparação entre HMM-2/20d e HMM-2/10d num período de valorização (alta).	35
Figura 3 – Desempenho do modelo HMM-2/20d num período de desvalorização (baixa).	36
Figura 4 – Desempenho do modelo HMM-3/10d num período de valorização (alta).	38
Figura 5 – Desempenho do modelo HMM-2/10d num período de desvalorização (baixa).	39
Figura 6 – Desempenho do modelo HMM-2/10d num período de valorização (alta).	40
Figura 7 – Valores de AIC para cada HMM num período de desvalorização (baixa).	42
Figura 8 – Valores de BIC para cada HMM num período de desvalorização (baixa).	43
Figura 9 – Comparação entre HMM-3/20d e HMM-4/10d num período de desvalorização (baixa).	43
Figura 10 – Desempenho do modelo HMM-2/10d num período de valorização (alta).	45
Figura 11 – Valores de AIC para cada HMM num período de desvalorização (baixa).	47
Figura 12 – Valores de BIC para cada HMM num período de desvalorização (baixa).	48
Figura 13 – Comparação entre HMM-3/20d e HMM-2/40d num período de desvalorização (baixa).	48
Figura 14 – Valores de AIC para cada HMM num período de valorização (alta).	50
Figura 15 – Valores de BIC para cada HMM num período de valorização (alta).	51
Figura 16 – Desempenho do modelo HMM-4/10d num período de valorização (alta).	51
Figura 17 – Valores de AIC para cada HMM num período de estagnação (neutro).	53
Figura 18 – Valores de BIC para cada HMM num período de estagnação (neutro).	54
Figura 19 – Comparação entre HMM-3/20d e HMM-2/20d num período de estagnação (neutro).	54

LISTA DE TABELAS

Tabela 1 – Desempenho do HMM com a ação ABEV3 num período de desvalorização.	33
Tabela 2 – Desempenho do HMM com a ação ABEV3 num período de valorização.	34
Tabela 3 – Desempenho do HMM com a ação ITUB4 num período de desvalorização.	36
Tabela 4 – Desempenho do HMM com a ação ITUB4 num período de valorização.	37
Tabela 5 – Desempenho do HMM com a ação PETR4 num período de desvalorização.	39
Tabela 6 – Desempenho do HMM com a ação PETR4 num período de valorização.	40
Tabela 7 – Desempenho do HMM com a ação VALE3 num período de desvalorização.	42
Tabela 8 – Desempenho do HMM com a ação VALE3 num período de valorização.	45
Tabela 9 – Desempenho do HMM com o índice Ibovespa num período de desvalorização.	47
Tabela 10 – Desempenho do HMM com o índice Ibovespa num período de valorização.	50
Tabela 11 – Desempenho do HMM com o índice Ibovespa num período de estagnação.	53
Tabela 12 – Resumo - Desempenho dos modelos	56

SUMÁRIO

1 – INTRODUÇÃO	1
2 – PROCESSOS DE MARKOV E HMM	2
2.1 PROCESSOS DE MARKOV - CONCEITOS	2
2.2 DEFINIÇÃO DO MODELO OCULTO DE MARKOV	2
2.3 PROBLEMAS CANÔNICOS	4
2.4 SOLUÇÃO DO PROBLEMA DA AVALIAÇÃO	5
2.5 SOLUÇÃO DO PROBLEMA DA DECODIFICAÇÃO	7
2.5.1 ALGORITMO DE VITERBI	8
2.6 SOLUÇÃO DO PROBLEMA DO APRENDIZADO	10
3 – SÉRIES TEMPORAIS EM FINANÇAS	14
3.1 INTRODUÇÃO	14
3.2 FILTRAÇÃO	14
3.3 ESPERANÇA CONDICIONAL: DEFINIÇÃO E PROPRIEDADES	15
3.4 PROCESSO ADAPTADO	17
3.5 MARTINGAL	17
3.6 SISTEMAS DE NEGOCIAÇÃO POTENCIALMENTE LUCRATIVOS (SNPL)	20
4 – MODELAGEM DO MERCADO ATRAVÉS DO HMM	24
4.1 INTRODUÇÃO	24
4.2 CARACTERÍSTICAS DO HMM	24
4.3 FATOS ESTILIZADOS EM FINANÇAS E O HMM	25
4.4 HMM E O MERCADO ACIONÁRIO	26
5 – METODOLOGIA	28
5.1 INTRODUÇÃO	28
5.2 SELEÇÃO DOS MODELOS	28
5.3 PREVISÃO E NEGOCIAÇÃO DAS AÇÕES	29
5.3.1 PREVISÃO DO PREÇO DAS AÇÕES	29
6 – RESULTADOS	33
6.1 CASO 1 - AMBEV	33
6.2 CASO 2 - ITAÚ	36
6.3 CASO 3 - PETROBRÁS	39
6.4 CASO 4 - VALE	42

6.5 CASO 5 - IBOVESPA	47
7 – CONCLUSÃO	56
Referências	59
Apêndices	60
APÊNDICE A – CÓDIGO-FONTE EM R	61

1 INTRODUÇÃO

Os modelos quantitativos em finanças estão atraindo um crescente interesse devido às suas amplas aplicações (como modelagem de portfólio, gerenciamento de risco e criação de sistemas de negociação automatizados). Atualmente, existem inúmeros modelos, cada um com diferentes vantagens e desvantagens. Nesta dissertação, iremos destacar o Modelo Oculto de Markov (HMM), que embora venha sendo estudado desde a década de 1960 em áreas como o reconhecimento da fala e a biologia, ainda é relativamente pouco utilizado em finanças.

Incluindo esta introdução, este trabalho está dividido em 7 capítulos. No capítulo 2, faremos um resumo básico da teoria do HMM detalhando os problemas e soluções canônicas que deram origem ao modelo. No capítulo 3, faremos uma introdução sucinta sobre séries temporais em finanças utilizando conceitos importantes da Probabilidade (filtração, esperança condicional, processo adaptado e martingal). Ao final deste capítulo, apresentaremos o Teorema Fundamental da Gestão Ativa de [Moffitt \(2017\)](#), que estabelece quais características uma série temporal deve apresentar para que o investidor tenha potencialmente lucros consistentes num sistema de negociação. No capítulo 4, iremos falar sobre a utilidade do HMM na modelagem do mercado de ações. Nesta etapa iremos abordar os principais aspectos do artigo de [Rydén, Teräsvirta e Åsbrink \(1998\)](#), que mostra que o HMM reproduz a maioria dos fatos estilizados sobre as séries diárias de retornos financeiros. Por fim, veremos como estes resultados junto com o artigo de [Nguyen \(2017\)](#) nos motivaram a criar sistemas de negociação baseados no HMM para o mercado acionário brasileiro.

No capítulo 5, iremos desenvolver a metodologia necessária para se construir um sistema de negociação baseado no HMM capaz de prever e negociar o preço das ações. Também falaremos neste capítulo sobre o critério de informação de Akaike (AIC) e o critério de informação Bayesiano (BIC), que são essenciais para estabelecermos o número ideal de estados ocultos de determinado HMM. No capítulo 6, iremos apresentar os resultados obtidos com a metodologia apresentada no capítulo anterior. Escolheremos cinco ativos financeiros (ações das empresas Ambev, Itaú, Petrobrás, Vale e o índice Ibovespa) e mediremos a rentabilidade do sistema de negociação, cuja as variáveis são o número de estados do modelo HMM e o intervalo em dias da Janela Temporal Retrospectiva. Por fim, no capítulo 7 concluiremos que o sistema de negociação apresentado neste trabalho apresenta um bom desempenho num cenário de desvalorização das ações e uma performance irregular quando as ações estão num período de alta. Assim, finalizaremos este trabalho com a perspectiva de se continuar desenvolvendo o modelo HMM na análise do mercado acionário.

2 PROCESSOS DE MARKOV E HMM

2.1 PROCESSOS DE MARKOV - CONCEITOS

Seja $\{X_t : 0 \leq t \leq T\}$ um processo estocástico (ou seja, é uma sequência de variáveis aleatórias) que assume valores num espaço de estados finito $Q = \{q_1, q_2, \dots, q_N\}$. Além disso, qualquer estado futuro de $\{X_t\}$ depende apenas do estado presente. Com estas propriedades, podemos caracterizar $\{X_t\}$ como uma Cadeia de Markov da seguinte forma:

$\{X_t : 0 \leq t \leq T\}$, tal que

$$P(X_{t+1} = x_{t+1} | X_t = x_t, X_{t-1} = x_{t-1}, \dots, X_1 = x_1) = P(X_{t+1} = x_{t+1} | X_t = x_t), \forall t \in \{0, \dots, T\}$$

Assumimos comumente que a Cadeia de Markov é homogênea, isto é, a probabilidade condicional $P(X_{t+1} = q_j | X_t = q_i)$ é independente do tempo t . Nesse caso temos que $a_{ij} = P(X_{t+1} = q_j | X_t = q_i)$ é independente de t , logo podemos organizar os valores de a_{ij} numa matriz de transição de probabilidades. Uma Cadeia de Markov homogênea fica totalmente determinada pelos seguintes parâmetros: a matriz de transição $A = \{P(X_{t+1} = q_j | X_t = q_i)\}$ e o vetor de probabilidades iniciais $\Pi = \{\pi_j = P(X_0 = q_j)\}$.

2.2 DEFINIÇÃO DO MODELO OCULTO DE MARKOV

Os Modelos Ocultos de Markov (HMM), que tiveram como referência neste trabalho os livros de [Koski \(2001\)](#), [Ibe \(2013\)](#) e [Stamp \(2017\)](#), são definidos como modelos Markovianos onde a observação da evolução do sistema se dá de forma indireta, como função probabilística da transição entre os estados definidos num determinado espaço de estados. Por mais que conheçamos todos os parâmetros do modelo, continua oculta a evolução da Cadeia de Markov que governa esse processo. Em outras palavras, não se sabe qual o caminho ou sequência de passos exatos que levaram a uma determinada observação. Mais detalhadamente, o HMM é caracterizado pelas propriedades (I), (II) e (III) a seguir:

(I) Cadeia de Markov oculta: o HMM possui uma cadeia de Markov oculta $\{X_t\}$ assumindo valores num espaço de estado finito $Q = \{q_1, q_2, \dots, q_N\}$ com N estados. A probabilidade condicional

$$a_{ij} = P(X_t = q_j | X_{t-1} = q_i) \text{ tal que } t \geq 1; q_i, q_j \in Q$$

é considerada homogênea no decorrer do tempo (ou seja, o valor de a_{ij} independe do valor de t). Além disso, a matriz de transição A é definida como:

$$A = (a_{ij})_{N \times N} \text{ tal que } a_{ij} \geq 0 \text{ e } \sum_{q_j \in Q} a_{ij} = 1$$

No tempo $t = 0$, o estado X_0 é definido pela distribuição de probabilidade

$$\pi_j = P(X_0 = q_j) \text{ com } \Pi = (\pi_1, \pi_2, \dots, \pi_N).$$

(II) Processo aleatório observável: o HMM possui um processo aleatório $\{O_t\}$ assumindo valores num espaço de estado finito $R = \{r_1, r_2, \dots, r_M\}$ com M estados. Os processos $\{O_t\}$ e $\{X_t\}$ se relacionam pela seguinte distribuição de probabilidade condicional (válida e homogênea para qualquer valor de $t \geq 1$):

$$b_j(k) = P(O_t = r_k | X_t = q_j) \text{ tal que } t \geq 1; r_k \in R; q_j \in Q$$

Além disso, a matriz de probabilidade de emissão B é definida como:

$$B = (b_j(k))_{N \times M} \text{ tal que } b_j(k) \geq 0 \text{ e } \sum_{r_k \in R} b_j(k) = 1$$

(III) Independência condicional: Para qualquer sequência de estados $x_0 x_1 \dots x_{n-1}$, a probabilidade da sequência de observações $o_0 o_1 \dots o_{n-1}$ é obtida da seguinte forma:

$$\begin{aligned} P(O_0 = o_0, \dots, O_{n-1} = o_{n-1} | X_0 = x_0, \dots, X_{n-1} = x_{n-1}) &= b_{x_0}(o_0) \cdot b_{x_1}(o_1) \dots b_{x_{n-1}}(o_{n-1}) \\ &= \prod_{i=0}^{n-1} b_{x_i}(o_i) \end{aligned}$$

Ou seja, os símbolos emitidos (observados) são condicionalmente independentes dada uma sequência de estados.

Em virtude das propriedades (I), (II) e (III), nós podemos definir a probabilidade conjunta entre $x_0 x_1 \dots x_{n-1}$ e $o_0 o_1 \dots o_{n-1}$ da seguinte forma:

$$\begin{aligned} P(O_0 = o_0, \dots, O_{n-1} = o_{n-1}, X_0 = x_0, \dots, X_{n-1} = x_{n-1}) &= b_{x_0}(o_0) \cdot \pi_{x_0} \cdot \prod_{i=1}^{n-1} a_{x_{i-1}, x_i} \cdot \prod_{i=1}^{n-1} b_{x_i}(o_i) \\ &= b_{x_0}(o_0) \cdot \pi_{x_0} \cdot \prod_{i=1}^{n-1} a_{x_{i-1}, x_i} \cdot b_{x_i}(o_i) \quad (1) \end{aligned}$$

Então, obtemos a probabilidade conjunta de $o_0 o_1 \dots o_{n-1}$ somando-se todos os caminhos possíveis da sequência de estados $x_0 x_1 \dots x_{n-1}$ na expressão (1):

$$P(O_0 = o_0, \dots, O_{n-1} = o_{n-1}) = \sum_{x_0 \in Q} (\dots) \sum_{x_{n-1} \in Q} b_{x_0}(o_0) \cdot \pi_{x_0} \cdot \prod_{i=1}^{n-1} a_{x_{i-1}, x_i} \cdot b_{x_i}(o_i) \quad (2)$$

Portanto, as distribuições finitas de O_t são totalmente especificadas pela distribuição de probabilidade Π e pelas matrizes estocásticas A e B . Assim, podemos utilizar a notação compacta $\lambda = (A, B, \Pi)$ para caracterizar o modelo HMM.

2.3 PROBLEMAS CANÔNICOS

A modelagem de um determinado fenômeno é uma versão simplificada da realidade propriamente dita. Dessa forma, não há um modelo absoluto, existem apenas modelos mais ou menos adequados para um dado sistema. Tendo isso em mente, poderíamos tratar o processo de modelagem como sendo composto de duas etapas: a definição dos parâmetros do modelo e o ajuste do mesmo pela resolução de uma série de “problemas-controle”. No contexto de modelagens em HMM, há três problemas fundamentais (ou canônicos) a serem resolvidos antes de se fazer uso de um modelo. Esses problemas, responsáveis pelo ajuste da modelagem, são os seguintes:

Problema da Avaliação: Dado o modelo $\lambda = (A, B, \Pi)$ e uma sequência de observações $O = (O_0 = o_0, \dots, O_{n-1} = o_{n-1})$, determine $P(O; \lambda)$ de forma eficiente.

Problema da Decodificação: Dado o modelo $\lambda = (A, B, \Pi)$ e uma sequência de observações $O = (O_0 = o_0, \dots, O_{n-1} = o_{n-1})$, encontre a sequência de estados ótima $X^* = (X_0^* = x_0^*, X_1^* = x_1^*, \dots, X_{n-1}^* = x_{n-1}^*)$ do processo markoviano subjacente. Ou seja, dentre as diversas sequências de estados que poderiam ter gerado essa sequência de observáveis qual é a mais provável?

Problema do Aprendizado: Dado uma sequência de observações $O = (O_0 = o_0, \dots, O_{n-1} = o_{n-1})$ e o parâmetro N (número de estados do modelo), determine um modelo da forma $\lambda = (A, B, \Pi)$ que maximize a probabilidade de O (podemos interpretar esse problema como a construção de um modelo de aprendizagem que se adeque melhor aos dados observados).

2.4 SOLUÇÃO DO PROBLEMA DA AVALIAÇÃO

Se utilizarmos o método direto para realizar este cálculo (cujo resultado é a expressão (2) do item (2.2)), verificamos que a solução do problema envolve um total de $2nN^n$ operações. Esse procedimento é computacionalmente muito oneroso. Como exemplo, considere um sistema com $N = 5$ estados e uma sequência de $n = 100$ observações. Neste caso, a resolução do Problema da Avaliação envolve $2.100.5^{100} \approx 10^{72}$ operações. Contudo, existe um método mais eficiente para resolver o Problema. Esse algoritmo é conhecido como *forward-backward procedure*.

O algoritmo Forward: Considere a variável forward, definida como

$$\alpha_t(j) = P(O_0 = o_0, \dots, O_{t-1} = o_{t-1}, O_t = o_t, X_t = q_j; \lambda) \text{ tal que } 0 \leq t \leq (n-1) \text{ e } 1 \leq j \leq N.$$

Ou seja, $\alpha_t(j)$ é a probabilidade de estar no estado q_j no tempo t depois de ter observado a sequência $\{o_0, o_1, \dots, o_t\}$. Ao efetuarmos estes cálculos temos:

$$\begin{aligned} \alpha_t(j) &= P(O_0 = o_0, \dots, O_{t-1} = o_{t-1}, O_t = o_t, X_t = q_j; \lambda) \\ &= P(O_t = o_t | X_t = q_j; \lambda) \cdot P(O_0 = o_0, \dots, O_{t-1} = o_{t-1}, X_t = q_j; \lambda) \\ &= P(O_t = o_t | X_t = q_j; \lambda) \cdot \sum_{q_i \in Q} P(O_0 = o_0, \dots, O_{t-1} = o_{t-1}, X_{t-1} = q_i, X_t = q_j; \lambda) \\ &= P(O_t = o_t | X_t = q_j; \lambda) \cdot \sum_{q_i \in Q} P(X_t = q_j | X_{t-1} = q_i; \lambda) \cdot P(O_0 = o_0, \dots, O_{t-1} = o_{t-1}, X_{t-1} = q_i) \\ &= b_j(o_t) \cdot \sum_{q_i \in Q} a_{ij} \cdot \alpha_{t-1}(i) \end{aligned}$$

Assim, o algoritmo forward funciona da seguinte forma:

(Passo 1) - Inicialização: $\alpha_0(i) = b_i(o_0) \cdot \pi_i ; 1 \leq i \leq N$

(Passo 2) - Indução: $\alpha_{t+1}(j) = b_j(o_{t+1}) \cdot \sum_{q_i \in Q} a_{ij} \cdot \alpha_t(i) ; 1 \leq j \leq N ; 0 \leq t \leq (n-2)$

(Passo 3) - Atualização do algoritmo: Ajustamos $t = t + 1$. Se $t < (n-1)$, voltamos para o passo 2; em caso contrário, vamos para o passo 4.

(Passo 4) - Finalização: Sendo $O = (O_0 = o_0, \dots, O_{n-1} = o_{n-1})$, temos:

$$P(O; \lambda) = \sum_{q_i \in Q} \alpha_{n-1}(i) = \sum_{q_i \in Q} P(O, X_{n-1} = q_i; \lambda)$$

O algoritmo forward necessita de $N(N+1)(n-2) + N$ multiplicações e $N(N-1)(n-2)$ adições, dando ao método uma complexidade computacional da ordem de N^2n ao invés de $2nN^n$ (método direto).

O algoritmo Backward: O algoritmo Backward é uma outra forma de resolver o Problema da Avaliação. Inicialmente, começamos definindo a variável backward da seguinte maneira:

$$\beta_t(j) = P(O_{t+1} = o_{t+1}, O_{t+2} = o_{t+2}, \dots, O_{n-1} = o_{n-1} | X_t = q_j; \lambda) \text{ tal que } 0 \leq t \leq (n-1) \text{ e } q_j \in Q.$$

Ou seja, $\beta_t(j)$ é a probabilidade condicional da observação parcial $\{o_{t+1}, o_{t+2}, \dots, o_{n-1}\}$ dado que o HMM está no estado q_j no tempo t . Assim, temos:

$$\begin{aligned} \beta_t(j) &= P(O_{t+1} = o_{t+1}, O_{t+2} = o_{t+2}, \dots, O_{n-1} = o_{n-1} | X_t = q_j; \lambda) \\ &= \sum_{q_i \in Q} P(O_{t+1} = o_{t+1}, O_{t+2} = o_{t+2}, \dots, O_{n-1} = o_{n-1}, X_{t+1} = q_i | X_t = q_j; \lambda) \\ &= \sum_{q_i \in Q} \frac{P(O_{t+1} = o_{t+1}, O_{t+2} = o_{t+2}, \dots, O_{n-1} = o_{n-1}, X_{t+1} = q_i, X_t = q_j; \lambda)}{P(X_t = q_j; \lambda)} \\ &= \sum_{q_i \in Q} P(O_{t+1} = o_{t+1}, \dots, O_{n-1} = o_{n-1} | X_{t+1} = q_i, X_t = q_j; \lambda) \cdot P(X_{t+1} = q_i | X_t = q_j; \lambda) \\ &= \sum_{q_i \in Q} P(O_{t+1} = o_{t+1}, \dots, O_{n-1} = o_{n-1} | X_{t+1} = q_i; \lambda) \cdot P(X_{t+1} = q_i | X_t = q_j; \lambda) \\ &= \sum_{q_i \in Q} \frac{P(O_{t+1} = o_{t+1} | X_{t+1} = q_i; \lambda) \cdot P(O_{t+2} = o_{t+2}, \dots, O_{n-1} = o_{n-1}, X_{t+1} = q_i; \lambda) \cdot a_{ji}}{P(X_{t+1} = q_i; \lambda)} \\ &= \sum_{q_i \in Q} P(O_{t+1} = o_{t+1} | X_{t+1} = q_i; \lambda) \cdot P(O_{t+2} = o_{t+2}, \dots, O_{n-1} = o_{n-1} | X_{t+1} = q_i; \lambda) \cdot a_{ji} \\ &= \sum_{q_i \in Q} b_i(o_{t+1}) \cdot \beta_{t+1}(i) \cdot a_{ji} \end{aligned}$$

Assim, o algoritmo backward funciona da seguinte forma:

(Passo 1) - Inicialização: $\beta_{n-1}(i) = 1$; $1 \leq i \leq N$

(Passo 2) - Indução: $\beta_t(i) = \sum_{q_j \in Q} a_{ij} \cdot \beta_{t+1}(j) \cdot b_j(o_{t+1})$; $1 \leq i \leq N$; $0 \leq t \leq (n-2)$

(Passo 3) - Atualização do algoritmo: Ajustamos $t = t - 1$. Se $t \geq 0$, voltamos para o passo 2; em caso contrário, vamos para o passo 4.

(Passo 4) - Finalização: Sendo $O = (O_0 = o_0, \dots, O_{n-1} = o_{n-1})$, temos:

$$P(O; \lambda) = \sum_{q_i \in Q} \beta_0(i) \cdot \alpha_0(i)$$

O algoritmo forward-backward é obtido a partir da observação de que para qualquer t , $0 \leq t \leq (n-1)$, temos que $P(O; \lambda) = \sum_{q_i \in Q} \beta_t(i) \cdot \alpha_t(i)$.

2.5 SOLUÇÃO DO PROBLEMA DA DECODIFICAÇÃO

Nesta seção iremos resolver o problema da decodificação, que pretende encontrar a sequência de estados ótima $X^* = (X_0^* = x_0^*, X_1^* = x_1^*, \dots, X_{n-1}^* = x_{n-1}^*)$ associada a uma determinada sequência de observações $O = (O_0 = o_0, \dots, O_{n-1} = o_{n-1})$ com um parâmetro λ . Ou seja, X^* é a sequência de estados que possui a maior probabilidade de gerar a sequência de observações dada. Assim, encontraremos o X^* que maximiza a expressão $P(X|O; \lambda)$. Neste primeiro momento, iremos encontrar os estados mais prováveis em cada instante de tempo e não como uma sequência completa. Assim, para cada t ($0 \leq t \leq (n-1)$) definimos a variável $\gamma_t(i)$ da seguinte maneira:

$$\gamma_t(i) = P(X_t = q_i | O; \lambda) = \frac{P(X_t = q_i, O; \lambda)}{P(O; \lambda)}$$

$$\text{Assim, } \gamma_t(i) = \frac{P(X_t = q_i, O_0 = o_0, \dots, O_t = o_t, \dots, O_{n-1} = o_{n-1}; \lambda)}{P(O; \lambda)}$$

Pelas propriedades da probabilidade sabemos que

$$P(Z = z, V = v) = P(Z = z | V = v) \cdot P(V = v)$$

$$\text{Logo, } \gamma_t(i) = \frac{P(O_0 = o_0, \dots, O_t = o_t | O_{t+1} = o_{t+1}, \dots, O_{n-1} = o_{n-1}, X_t = q_i; \lambda) \cdot P(C)}{P(O; \lambda)}$$

onde $C = (O_{t+1} = o_{t+1}, \dots, O_{n-1} = o_{n-1}, X_t = q_i; \lambda)$

Pelas propriedades de Markov sabemos que

$$P(O_0 = o_0, \dots, O_t = o_t | C) = P(O_0 = o_0, \dots, O_t = o_t | X_t = q_i; \lambda)$$

$$\text{Logo, } \gamma_t(i) = \frac{P(O_0 = o_0, \dots, O_t = o_t | X_t = q_i; \lambda) \cdot P(C)}{P(O; \lambda)}$$

Sabemos que $P(C) = P(O_{t+1} = o_{t+1}, \dots, O_{n-1} = o_{n-1} | X_t = q_i; \lambda) \cdot P(X_t = q_i; \lambda)$

$$\text{Assim, } \gamma_t(i) = \frac{P(O_0 = o_0, \dots, O_t = o_t, X_t = q_i; \lambda) \cdot P(C)}{P(X_t = q_i; \lambda) P(O; \lambda)}$$

$$\Rightarrow \gamma_t(i) = \frac{\alpha_t(i) \cdot \beta_t(i)}{P(O; \lambda)}$$

pois $\alpha_t(i) = P(O_0 = o_0, \dots, O_t = o_t, X_t = q_i; \lambda)$ e

$$\beta_t(i) = P(O_{t+1} = o_{t+1}, \dots, O_{n-1} = o_{n-1} | X_t = q_i; \lambda)$$

Sabemos que $P(O; \lambda) = \sum_{q_i \in Q} \alpha_t(i) \cdot \beta_t(i)$.

$$\text{Assim, } \gamma_t(i) = \frac{\alpha_t(i) \cdot \beta_t(i)}{\sum_{q_i \in Q} \alpha_t(i) \cdot \beta_t(i)}$$

Notemos que $\sum_{i=1}^N \gamma_t(i) = 1$

Portanto, o estado mais provável de X no tempo t é $x_t^* = \arg \max_{1 \leq i \leq N} \gamma_t(i)$, onde $0 \leq t \leq (n-1)$.

Assim, este método gera a sequência de estados mais prováveis em cada instante dada a sequência de observações $O = (O_0 = o_0, \dots, O_{n-1} = o_{n-1})$. Observamos que este método é inapropriado para encontrar a sequência mais provável (que é diferente da sequência de estados mais prováveis que calculamos) pois não leva em consideração as probabilidades de transição entre os estados. Por exemplo, se tivermos uma sequência que inclua dois estados vizinhos q_i e q_j cuja probabilidade de transição seja $a_{ij} = P(X_t = q_j | X_{t-1} = q_i) = 0$ teremos como resultado uma sequência de estados inválida. Um método eficiente que evita essas sequências inválidas e produz como resultado a sequência mais provável é o algoritmo de Viterbi.

2.5.1 ALGORITMO DE VITERBI

O algoritmo de Viterbi maximiza simultaneamente a probabilidade conjunta $P(X, O)$ e a probabilidade condicional $P(X|O)$ devido a seguinte relação:

$$\arg \max_X P(X|O; \lambda) = \arg \max_X \frac{P(X, O; \lambda)}{P(O; \lambda)} = \arg \max_X P(X, O; \lambda)$$

Além disso, o algoritmo define a variável $\delta_t(i)$ da seguinte maneira:

$$\delta_t(i) = \max_{X_0, X_1, \dots, X_{t-1}} P(X_0 = x_0, X_1 = x_1, \dots, X_{t-1} = x_{t-1}, X_t = q_i, O; \lambda)$$

Ou seja, $\delta_t(i)$ é o percurso de maior probabilidade ao longo do caminho associado às

$(t + 1)$ observações $O = (O_0 = o_0, \dots, O_{t-1} = o_{t-1}, O_t = o_t)$ e que termina no estado q_i . Outra variável, $\psi_t(j)$, armazena o ponto que leva ao caminho mais provável associado às transições entre os estados. Ou seja,

$$\psi_t(j) = \arg \max_{1 \leq i \leq N} \{\delta_{t-1}(i) \cdot a_{ij}\}$$

Assim, o algoritmo de Viterbi funciona da seguinte forma:

(Passo 1) - Inicialização:

$$\delta_0(i) = \pi_i \cdot b_i(O_0), \text{ onde}$$

$$\pi_i = P(X_0 = q_i) \text{ e } b_i(O_0) = P(O_0 = o_0 | X_0 = q_i)$$

$$\psi_0(i) = 0, \quad 1 \leq i \leq N$$

(Passo 2) - Recursão:

$$\delta_t(j) = \max_{1 \leq i \leq N} \{\delta_{t-1}(i) \cdot a_{ij}\} \cdot b_j(O_t) \text{ e}$$

$$\psi_t(j) = \arg \max_{1 \leq i \leq N} \{\delta_{t-1}(i) \cdot a_{ij}\}, \text{ onde}$$

$$1 \leq j \leq N \text{ e } 1 \leq t \leq (n - 1)$$

(Passo 3) - Atualização:

Ajustamos $t = t + 1$. Se $t < (n - 1)$, voltamos para o passo 2; em caso contrário, vamos para o passo 4.

(Passo 4) - Finalização:

$$P^* = \max_{1 \leq i \leq N} \{\delta_{n-1}(i)\}$$

$$x_{n-1}^* = \arg \max_{1 \leq i \leq N} \{\delta_{n-1}(i)\}$$

(Passo 5) - Caminho Regressivo:

$$x_t^* = \psi_{t+1}(x_{t+1}^*), \quad t = (n - 2), (n - 3), \dots, 0$$

O passo 5 nos permite encontrar a sequência mais provável a partir dos sucessores armazenados no passo 2.

2.6 SOLUÇÃO DO PROBLEMA DO APRENDIZADO

O problema de aprendizagem aborda a forma como podemos ajustar os parâmetros do HMM para que o conjunto de observações seja representado da melhor maneira possível pelo modelo. Como estamos buscando a melhor forma de representar uma determinada observação, nós estamos de fato resolvendo um problema de otimização. Um critério de otimização muito utilizado é o critério da máxima verossimilhança, que procura encontrar os parâmetros do modelo que maximizam a probabilidade de uma determinada sequência de observações. Infelizmente, o Problema do Aprendizado não possui um método conhecido para obter a solução analítica de λ que maximiza $P(O; \lambda)$. Porém, podemos escolher os parâmetros do HMM de forma que encontremos o máximo local de $P(O; \lambda)$. Este método é uma solução iterativa e é chamado de algoritmo de Baum-Welch, que é um caso especial do método *Expectation-Maximization* (EM).

O algoritmo de Baum-Welch: O algoritmo começa configurando os parâmetros A , B e Π para alguns valores iniciais que podem ser escolhidos de algum conhecimento prévio ou de alguma distribuição uniforme. Em seguida, usando o modelo atual, todos os caminhos possíveis para cada conjunto de treinamento (dados observados) são considerados para se obter novas estimativas \bar{A} , \bar{B} e $\bar{\Pi}$. O procedimento é repetido até que haja mudanças insignificantes nos parâmetros do HMM.

Assim como o algoritmo forward-backward, o algoritmo de Baum-Welch utiliza as mesmas variáveis forward e backward ($\alpha_t(i)$ e $\beta_t(i)$) usadas no Problema da Avaliação. Temos então o seguinte:

$$\alpha_t(i) = P(O_0 = o_0, \dots, O_{t-1} = o_{t-1}, O_t = o_t, X_t = q_i; \lambda)$$

$$\beta_t(i) = P(O_{t+1} = o_{t+1}, O_{t+2} = o_{t+2}, \dots, O_{n-1} = o_{n-1} | X_t = q_i; \lambda)$$

Onde $1 \leq t \leq (n - 1)$ e $1 \leq i \leq N$. Lembramos que $\alpha_t(i)$ é a probabilidade de estar no estado q_i no tempo t depois de ter observado a sequência $\{o_0, o_1, \dots, o_t\}$ e $\beta_t(i)$ é a probabilidade condicional da observação parcial $\{o_{t+1}, o_{t+2}, \dots, o_{n-1}\}$ dado que o HMM está no estado q_i no tempo t . Além disso, lembramos que essas variáveis são computadas indutivamente da seguinte maneira:

$$\alpha_0(i) = b_i(o_0) \cdot \pi_i ; 1 \leq i \leq N$$

$$\alpha_{t+1}(j) = b_j(o_{t+1}) \cdot \sum_{q_i \in Q} a_{ij} \cdot \alpha_t(i) ; 1 \leq j \leq N ; 0 \leq t \leq (n-2)$$

$$\beta_{n-1}(i) = 1 ; 1 \leq i \leq N$$

$$\beta_t(i) = \sum_{q_j \in Q} a_{ij} \cdot \beta_{t+1}(j) \cdot b_j(o_{t+1}) ; 1 \leq j \leq N ; 0 \leq t \leq (n-2)$$

Como no algoritmo de Viterbi, definimos a variável $\gamma_t(i)$ da seguinte maneira:

$$\Rightarrow \gamma_t(i) = \frac{\alpha_t(i) \cdot \beta_t(i)}{P(O; \lambda)} = \frac{\alpha_t(i) \cdot \beta_t(i)}{\sum_{q_i \in Q} \alpha_t(i) \cdot \beta_t(i)}$$

Esta é a probabilidade de estar no estado q_i no tempo t dada uma sequência de observações $O = (O_0 = o_0, O_1 = o_1, \dots, O_{n-1} = o_{n-1})$ no modelo HMM. Também definimos a variável $\varepsilon_t(i, j)$ como a probabilidade de estar no estado q_i no tempo t e de estar no estado q_j no tempo $t+1$, dada uma sequência de observações $O = (O_0 = o_0, O_1 = o_1, \dots, O_{n-1} = o_{n-1})$ no modelo HMM. Assim, temos:

$$\begin{aligned} \varepsilon_t(i, j) &= P(X_t = q_i, X_{t+1} = q_j | O; \lambda) \\ &= \frac{P(X_t = q_i, X_{t+1} = q_j, O; \lambda)}{P(O; \lambda)} \\ &= \frac{\alpha_t(i) \cdot a_{ij} \cdot b_j(o_{t+1}) \cdot \beta_{t+1}(j)}{P(O; \lambda)} \\ &= \frac{\alpha_t(i) \cdot a_{ij} \cdot b_j(o_{t+1}) \cdot \beta_{t+1}(j)}{\sum_{q_i \in Q} \sum_{q_j \in Q} \alpha_t(i) \cdot a_{ij} \cdot b_j(o_{t+1}) \cdot \beta_{t+1}(j)} \end{aligned}$$

Notemos que $\gamma_t(i)$ e $\varepsilon_t(i, j)$ estão relacionados da seguinte forma:

$$\gamma_t(i) = \sum_{q_j \in Q} \varepsilon_t(i, j)$$

Além disso, se fizermos o somatório de $\gamma_t(i)$ durante o tempo de observação $(n-1)$, obteremos a estimativa do número de vezes que o estado q_i é “visitado” em todo esse período. Se quisermos saber o número de transições a partir de q_i , basta levarmos o somatório até o instante $(n-2)$. Analogamente, ao fazermos o somatório de $\varepsilon_t(i, j)$ até $(n-2)$ obtemos a estimativa do número de vezes que ocorreram transições entre os estados $X_{t-1} = q_i$ e $X_t = q_j$. Assim, temos:

$$\sum_{t=0}^{n-2} \gamma_t(i) = \text{número esperado de transições a partir de } q_i.$$

$$\sum_{t=0}^{n-1} \gamma_t(i) = \text{número esperado de vezes que o estado } q_i \text{ é “visitado”}.$$

$$\sum_{t=0}^{n-2} \varepsilon_t(i,j) = \text{número esperado de transições de } q_i \text{ para } q_j.$$

Usando estas definições, podemos estimar os parâmetros do modelo HMM da seguinte forma:

$$\bar{a}_{ij} = (\text{número esperado de transições de } q_i \text{ para } q_j) / (\text{número esperado de transições partindo de } q_i)$$

$$\bar{a}_{ij} = \frac{\sum_{t=0}^{n-2} \varepsilon_t(i,j)}{\sum_{t=0}^{n-2} \gamma_t(i)}$$

$$\bar{b}_j(k) = (\text{número esperado de transições de } q_j \text{ emitindo } r_k) / (\text{número esperado de estados } q_j)$$

$$\bar{b}_j(k) = \frac{\sum_{t=0, o_t=r_k}^{n-1} \gamma_t(j)}{\sum_{t=0}^{n-1} \gamma_t(j)}$$

Assim, o algoritmo de Baum-Welch funciona da seguinte forma:

(Passo 1): Obtemos a estimativa da distribuição do estado inicial com a distribuição de frequências com a qual cada estado é visitado no tempo $t = 0$, ou seja:

$$\bar{\Pi}_i = \gamma_0(i)$$

(Passo 2): Obtemos as estimativas de \bar{a}_{ij} e $\bar{b}_j(k)$ (conforme detalhado anteriormente).

(Passo 3): Seja o modelo atual com o parâmetro $\lambda = (A, B, \Pi)$, que é utilizado para calcular os valores de \bar{a}_{ij} e $\bar{b}_j(k)$. Assim, o modelo reestimado passa a ter como parâmetro $\bar{\lambda} = (\bar{A}, \bar{B}, \bar{\Pi})$. Faremos sucessivamente este processo de atualização do modelo (de λ para $\bar{\lambda}$) até alcançarmos a convergência do mesmo (ver passo 4).

(Passo 4): Se $P(O; \bar{\lambda}) - P(O; \lambda) < \delta$, então devemos parar o algoritmo (δ é um valor pré-definido).

O algoritmo EM (que é uma generalização do algoritmo de Baum-Welch) afirma que após cada iteração pode acontecer uma das duas coisas:

(i): $\bar{\lambda}$ é mais verossímil do que λ , ou seja, $P(O; \bar{\lambda}) > P(O; \lambda)$;

(ii): Atingimos um ponto estacionário da função de verossimilhança no qual $\bar{\lambda} = \lambda$.

Pode-se provar que a cada iteração $P(O; \bar{\lambda}) \geq P(O; \lambda)$ e que a sequência de modelos $\bar{\lambda}_i$ obtidos com o algoritmo de Baum-Welch converge para λ^* , um máximo local de $\arg \max_{\lambda} P(O|\lambda)$. Apesar do algoritmo de Baum-Welch convergir localmente (mas não globalmente) para um máximo, este método produz bons resultados na solução de problemas aplicados.

Em algumas situações, a modelagem com o HMM-contínuo é mais adequada para a análise do problema proposto. Como neste trabalho iremos modelar os retornos dos ativos financeiros através da distribuição Gaussiana, utilizaremos o HMM-contínuo redefinindo o processo aleatório observável $b_j(k)$ (detalhado no item 2.2) da seguinte forma:

$$b_j(k) = b_j(O_t) = \mathcal{N}(O_t, \mu_j, \sigma_j)$$

onde $b_j(O_t)$ é função de densidade de probabilidade de uma distribuição normal em O_t .

Assim, implementaremos o algoritmo de Baum-Welch de forma análoga ao que foi feito no HMM discreto. A estimação dos parâmetros no HMM-contínuo será feita da seguinte forma:

$$\bar{\mu}_j = \frac{\sum_{t=0}^{n-1} \gamma_t(j) \cdot O_t}{\sum_{t=0}^{n-1} \gamma_t(j)}$$

$$\bar{\sigma}_j = \frac{\sum_{t=0}^{n-1} \gamma_t(j) \cdot (O_t - \mu_j) \cdot (O_t - \mu_j)^T}{\sum_{t=0}^{n-1} \gamma_t(j)}$$

onde $(O_t - \mu_j)^T$ é a transposta de $(O_t - \mu_j)$ e $1 \leq j \leq N$.

3 SÉRIES TEMPORAIS EM FINANÇAS

3.1 INTRODUÇÃO

Série Temporal é qualquer conjunto de observações ordenado no tempo. Os exemplos de séries temporais são encontrados em diversas áreas do conhecimento. Temos diariamente as informações do mercado financeiro, onde os preços dos ativos constituem a informação mais relevante para os investidores. Já na área econômica, o PIB, a inflação e a demanda agregada são exemplos de dados importantes que evoluem ao longo do tempo. Uma característica fundamental das séries temporais é que os valores da variável em questão não são dados independentes de uma população (há uma dependência temporal entre os valores). Sendo assim, o objetivo do estudo das séries temporais é o de conhecer o mecanismo gerador dos dados e realizar previsões sobre o comportamento da variável que está sob análise. Logo, alguns conceitos da Probabilidade moderna (filtração, processo adaptado e martingal), que serão vistos no decorrer deste capítulo e tiveram como referência neste trabalho os livros de [Petters, Dong et al. \(2016\)](#) e [Aiube \(2013\)](#), são fundamentais para o alcance deste objetivo. Após a apresentação destas formulações será possível enunciar o Teorema Fundamental da Gestão Ativa de [Moffitt \(2017\)](#), que estabelece quais características uma série temporal deve apresentar para que o investidor tenha potencialmente lucros consistentes num sistema de negociação.

3.2 FILTRAÇÃO

Os processos estocásticos são ferramentas úteis na modelagem de eventos aleatórios que evoluem ao longo do tempo. Por exemplo, seja $X = \{X_t : t \geq 0\}$ um processo que representa o preço de uma ação no decorrer do tempo. Em circunstâncias normais, o preço de mercado de uma ação é o preço pelo qual um comprador e um vendedor concordam negociar este ativo. Assim, a mudança do preço de mercado de uma ação é causada pela lei da oferta e da demanda, que reflete as expectativas de lucratividade da empresa. Para inferir sobre essas possibilidades, é preciso mensurar como qualquer notícia sobre a empresa será interpretada por investidores e especuladores. Portanto, precisamos construir uma estrutura de informação que evolua ao longo do tempo para estudarmos o processo aleatório que governa o preço da ação. O conceito matemático que descreve a dinâmica deste fenômeno (a propagação das informações ao longo do tempo) é chamado de filtração.

Definição 3.2.1 Um espaço de probabilidade filtrado é uma quadrupla $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$, onde $(\Omega, \mathcal{F}, \mathbb{P})$ é o espaço de probabilidade e $\{\mathcal{F}_t\}$ é a filtração. Uma filtração é uma coleção não decrescente de σ -álgebras $\{\mathcal{F}_t \subseteq \mathcal{F}; t \geq 0\}$ tal que:

$\mathcal{F}_s \subseteq \mathcal{F}_t$, para $\forall s, t \geq 0$ com $s \leq t$.

Uma maneira intuitiva de descrever uma filtração é que ela funciona como um crivo que controla o fluxo de informações. Para o nosso propósito, basta sabermos o seguinte:

(i) \mathcal{F}_t representa a informação disponível no tempo t .

(ii) A estrutura da informação (formalizada na definição) assegura que a quantidade de informação cresce conforme o tempo evolui; além disso, nenhuma informação é perdida com o decorrer do tempo (ou seja, qualquer informação disponível no momento s ainda está disponível no tempo t desde que $t \geq s$).

Definição 3.2.2 A filtração $\{\mathcal{F}_t \subseteq \mathcal{F}; t \geq 0\}$ é chamada de filtração padrão do processo estocástico X_t se

$$\mathcal{F}_t = \sigma(X_s, 0 \leq s \leq t); t \geq 0;$$

onde $\sigma(X_s, 0 \leq s \leq t)$ é a σ -álgebra gerada pelas variáveis aleatórias X_s , onde $0 \leq s \leq t$.

Nós também dizemos que a filtração $\{\mathcal{F}_t : t \geq 0\}$ gerada por $\sigma(X_s : s \leq t) = \mathcal{F}_t$ é a filtração induzida por $\{X_t : t \geq 0\}$. Notemos que

$$\sigma(X_u, 0 \leq u \leq s) \subseteq \sigma(X_u, 0 \leq u \leq t), \text{ onde } 0 \leq s \leq t.$$

3.3 ESPERANÇA CONDICIONAL: DEFINIÇÃO E PROPRIEDADES

A Esperança Condicional é definida da seguinte maneira:

Definição 3.3.1 Sejam X e Y variáveis aleatórias no mesmo espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$. A esperança condicional de X dado que $Y = y$, se existir, é dada por

$$E(X|Y = y) = \int_{-\infty}^{+\infty} x dF_X(x|Y = y).$$

A esperança condicional é a esperança calculada com a probabilidade condicional. Sendo $E(X)$ finita, a existência da esperança condicionada à Y é garantida pelo Teorema de Radon-Nikodym. Se a variável Y for contínua, lembremos que o condicionamento em $Y = y$ deve ser interpretado como o limite de probabilidades calculadas na vizinhança de y . A expressão $E(X|Y = y)$ é uma função de y e, às vezes, é simplesmente denotada

por $E(X|y)$. Em Estatística ela recebe também o nome de *regressão de X em relação à Y* e corresponde à média de valores de X , restrita aos valores de Y iguais a y . Se nós não particularizarmos o valor de Y escrevemos $E(X|Y)$, que será denominada *esperança condicional de X dado Y* . Notemos que $E(X|Y)$ é uma função da variável aleatória Y e, portanto, é também uma variável aleatória no mesmo espaço de probabilidade.

A Esperança Condicional possui as seguintes propriedades:

Propriedade 1 (Mensurabilidade): Se X é \mathcal{F} -mensurável, então

$$E(X|\mathcal{F}) = X$$

Interpretação: X ser \mathcal{F} -mensurável significa simplesmente que \mathcal{F} contém informação suficiente de X . Isto quer dizer que a variável X é conhecida, dada a sigma-álgebra \mathcal{F} . Como a melhor estimativa para uma variável conhecida é a própria variável, temos que a expressão $E(X|\mathcal{F}) = X$ é válida.

Propriedade 2: Se X e Y são variáveis aleatórias e X é \mathcal{F} -mensurável, então

$$E(XY|\mathcal{F}) = XE(Y|\mathcal{F})$$

Interpretação: Sendo X \mathcal{F} -mensurável significa que a variável X é conhecida, dada a sigma-álgebra \mathcal{F} . Isto quer dizer que podemos considerar X como uma constante dado um conjunto de informações \mathcal{F} .

Propriedade 3: Se X é uma variável aleatória e \mathcal{F} é uma σ -álgebra, então

$$E(E(X|\mathcal{F})) = E(X)$$

Propriedade 4: Se $\mathcal{F} \subset \mathcal{G}$, isto é, \mathcal{F} é uma sub- σ -álgebra de \mathcal{G} , então

$$E(E(X|\mathcal{G})|\mathcal{F}) = E(X|\mathcal{F})$$

Propriedade 5 (linearidade): Sejam a_1, a_2 constantes e X_1, X_2 variáveis aleatórias, então

$$E(a_1X_1 + a_2X_2|\mathcal{F}) = a_1E(X_1|\mathcal{F}) + a_2E(X_2|\mathcal{F})$$

Propriedade 6 (independência): Se X é independente de \mathcal{F} , então

$$E(X|\mathcal{F}) = E(X)$$

Interpretação: Em algumas situações a variável aleatória X e a σ -álgebra \mathcal{F} são dadas mas X não é mensurável em \mathcal{F} . Assim, intuitivamente dizemos que \mathcal{F} não tem o contraste suficiente para ler todas as informações contidas em X (na situação oposta, quando X é mensurável por \mathcal{F} , temos o caso no qual \mathcal{F} tem o contraste suficiente para ler todas as informações contidas em X). Portanto, quando X é independente de \mathcal{F} estamos na situação em que \mathcal{F} não possui nenhuma informação de X .

3.4 PROCESSO ADAPTADO

Definição 3.4.1: Um processo estocástico $\{X_t : t \geq 0\}$ definido num espaço de probabilidade filtrado $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$ é considerado adaptado se X_t é \mathcal{F}_t -mensurável para cada $t \geq 0$.

Interpretação: Pela definição temos que

$$X_s = E(X_s|\mathcal{F}_t), \text{ para cada } t \geq s$$

Particularmente temos o seguinte:

$$X_s = E(X_s|\mathcal{F}_s)$$

Dizemos que X_s é considerado “conhecido” em qualquer tempo t ($t \geq s$) se $\{X_t : t \geq 0\}$ é adaptado à filtração $\{\mathcal{F}_t\}$ (ou seja, um processo estocástico $\{X_t : t \geq 0\}$ adaptado à filtração implica que o valor de X_t é quase certamente determinado pela filtração $\{\mathcal{F}_t\}$). Por outro lado, para cada $t \geq s$, X_t em geral não é mensurável no tempo s . A variável X_t é considerada “desconhecida” porque as probabilidades dos eventos descritos por X_t não podem ser calculados com base na informação disponível antes do tempo t .

A noção de adaptabilidade pode ser interpretada como a incapacidade de prever eventos futuros. Por essa razão, um processo adaptado também é chamado de não-antecipativo pois a propagação ou revelação progressiva da informação sob adaptabilidade não permite a antecipação de informações futuras.

3.5 MARTINGAL

Definição 3.5.1: Um processo estocástico $\{X_t : t \geq 0\}$ definido num espaço de probabilidade filtrado $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$ é um martingal em relação à filtração $\{\mathcal{F}_t\}$ se $\{X_t : t \geq 0\}$ é adaptado à $\{\mathcal{F}_t\}$ e satisfaz as seguintes condições:

$$\underline{\text{(i)}}: E(|X_t|) < \infty, \forall t \geq 0$$

$$\underline{\text{(ii)}}: E(X_t | \mathcal{F}_s) = X_s, \text{ para } \forall s < t, t \geq 0;$$

Ao analisarmos a condição (ii) concluímos que num martingal a melhor previsão para um evento futuro é o valor atual do processo. Além disso, temos que

$$E(E(X_t | \mathcal{F}_s)) = E(X_s) \implies E(X_t) = E(X_s)$$

$$\text{Logo, } E(X_t - X_s) = 0$$

Ou seja, o ganho futuro esperado é zero. Logo, os martingais modelam um jogo justo, aquele que não favorece nenhum dos participantes. Reescrevendo a definição de outra forma, sabemos que o processo estocástico martingal $\{X_t : t \geq 0\}$ se caracteriza por satisfazer à expressão $E(X_{t+1} - X_t | X_t, X_{t-1}, \dots, X_1) = 0$, portanto concluímos que um jogo é justo quando os ganhos incrementais em qualquer estágio é nulo ao se condicionar aos resultados obtidos historicamente. No contexto moderno das Finanças, a propriedade martingal é uma condição necessária para se ter um mercado eficiente, ou seja, um mercado no qual a informação sobre preços passados são incorporados completamente e instantaneamente no preço atual do ativo. A eficiência de um mercado implica que não é possível obter lucros excessivos em operações baseadas em tendências históricas dos preços (ou seja, a expectativa de mudança do preço futuro, condicionada ao histórico de preços, não pode ser nem positiva nem negativa e deve ser zero).

3.5.2 - SÉRIES DE DIFERENÇAS MARTINGAIS (SDM)

Sabemos que um martingal X_t (adaptado em relação à série temporal Z_t) produz uma série temporal com as seguintes características:

$$\underline{\text{(i)}} E(|X_t|) < \infty, \forall t \geq 0$$

$$\underline{\text{(ii)}} E(X_t | Z_{t-1}, Z_{t-2}, \dots) = X_{t-1}, \forall t \geq 0$$

Uma Série de Diferença Martingal (SDM) Y_t (adaptada em relação à série temporal Z_t) é uma série temporal com as seguintes características:

$$\underline{\text{(i)}} E(|Y_t|) < \infty, \forall t \geq 0$$

$$\underline{\text{(ii)}} \quad E(Y_t | Z_{t-1}, Z_{t-2}, \dots) = 0, \quad \forall t \geq 0$$

Uma série temporal de preços financeiros P_t (considerando-se a ausência de fluxos de caixa externos) é frequentemente modelada como um martingal. Já a série temporal dos retornos associados a estes preços ($R_t = (P_{t+1} - P_t)/P_t$ ou $R_t = \ln(P_{t+1}/P_t)$) é modelada como uma Série de Diferença Martingal (SDM).

O modelo SDM para retornos financeiros pode ser desenvolvido da seguinte maneira: seja X_t um martingal em relação a Z_t e $Y_t = X_t - X_{t-1}$. Então, Y_t é uma SDM em relação a Z_t tal que

$$\begin{aligned} E(Y_t | Z_{t-1}, Z_{t-2}, \dots) &= E(X_t - X_{t-1} | Z_{t-1}, Z_{t-2}, \dots) \\ &= E(X_t | Z_{t-1}, Z_{t-2}, \dots) - E(X_{t-1} | Z_{t-1}, Z_{t-2}, \dots) \\ &= E(X_t | Z_{t-1}, Z_{t-2}, \dots) - X_{t-1} \\ &= X_{t-1} - X_{t-1} \\ &= 0 \end{aligned}$$

Além disso, podemos observar que

$$X_{t-1} = E(X_t | Z_{t-1}, Z_{t-2}, \dots) = g(Z_{t-1}, Z_{t-2}, \dots) \quad (1)$$

e

$$X_{t-1} = E(X_t | X_{t-1}, X_{t-2}, \dots) \quad (2)$$

Ao observarmos estas expressões podemos fazer o seguinte questionamento: Por que definir um martingal em termos de Z_t já que a expressão (2) chega ao mesmo resultado utilizando apenas a variável X_t ? Em primeiro lugar, devemos observar que as expressões (1) e (2) não são equivalentes. A variável X_t é uma função de Z_t (pela definição (1)). Assim, X_t pode conter “menos informações” que Z_t (outras variáveis podem ser extraídas de Z_t ; por exemplo: variáveis independentes ou não correlacionadas com X_t). Além disso, a expressão (1) é fundamental para detectar os sistemas de negociação financeiros potencialmente lucrativos (SNPL) já que variáveis associadas somente a processos martingais não possuem esta característica (detalharemos este fato no próximo item).

3.6 SISTEMAS DE NEGOCIAÇÃO POTENCIALMENTE LUCRATIVOS (SNPL)

Um sistema de negociação da série temporal X_t (tendo como referência a série temporal Z_t) é definido como uma família de funções reais $f_t(Z_t, Z_{t-1}, \dots)$ que especificam o quanto se deve investir no tempo t para auferir ganhos no instante $t + 1$ (dado o histórico Z_t, Z_{t-1}, \dots da série temporal de referência). Assim, temos que:

$f_{u-1}(Z_{u-1}, Z_{u-2}, \dots)(X_u - X_{u-1})$: representa o ganho do sistema de negociação no período $[u - 1, u]$;

$G_{t,s}(f) = \sum_{u=t-s+1}^t f_{u-1}(Z_{u-1}, Z_{u-2}, \dots)(X_u - X_{u-1})$: representa o ganho acumulado do sistema de negociação no período $[t - s, t]$;

$\bar{G}_{t,s}(f) = \frac{G_{t,s}(f)}{s}$: representa o ganho médio do sistema de negociação no período $[t - s, t]$;

Salientamos que $G_{t,s}(f)$ e $\bar{G}_{t,s}(f)$ são variáveis aleatórias dependentes de X_t e Z_t . A inclusão da série temporal Z_t pode parecer inapropriada, mas é necessária para capturar situações multifatoriais que surgem no mercado financeiro. Na prática, Z_t é frequentemente uma variável vetorial que representa a informação conhecida no tempo t (por exemplo, Z_t pode representar os retornos, volumes e volatilidades de um ativo no instante t).

Definição 3.6.1: Um Sistema de Negociação Potencialmente Lucrativo (SNPL), definido originalmente por Moffitt (2017), é um sistema f_t para a série temporal financeira X_t (tendo como referência a série temporal Z_t , ou seja, X_t é adaptada à Z_t) tal que $G_{t,s}(f)$ possui um valor esperado não negativo para quaisquer $t, s > 0$ e possui valor esperado positivo para pelo menos um par (t, s) (com $t, s > 0$), ou seja

$$E(G_{t,s}(f)) \geq 0, \forall t, s > 0$$

e

$$E(G_{t_0, s_0}(f)) > 0, \text{ para algum par } (t_0, s_0) \text{ com } t_0, s_0 > 0$$

Na teoria financeira clássica não é admitida a existência dos SNPL's. Porém, um número considerável de investidores têm obtido na prática retornos financeiros elevados e regulares ao longo de décadas. Este fato nos fornece evidências de que os SNPL's real-

mente existem (o que nos motiva a pesquisar em quais situações o mercado favorece o aparecimento destes sistemas de negociação).

3.6.2 O TEOREMA FUNDAMENTAL DA GESTÃO ATIVA - CONCEITOS

O termo gestão ativa refere-se a uma estratégia de gerenciamento de portfólio que explicitamente tenta produzir retornos ajustados ao risco em excesso de um portfólio de referência. A gestão ativa de carteiras baseia-se na hipótese da existência de informações e técnicas superiores às de consenso do mercado que vão possibilitar um ganho superior à média para o gestor que tiver posse destas informações e técnicas e souber usá-las. Já na hipótese de mercados eficientes partimos do pressuposto de que os mercados financeiros processam de maneira rápida e eficiente todas as informações disponíveis sobre os ativos financeiros, ou seja, os preços dos títulos refletem todas as informações das quais os investidores dispõem sobre o valor do título. De acordo com essa hipótese, à medida que novas informações a respeito de um título são disponibilizadas, o preço do ativo ajusta-se rapidamente para que, a qualquer momento, torne-se igual à estimativa consensual do mercado sobre seu valor. Se isto realmente ocorresse no mundo real, não haveria ativos financeiros nem abaixo e nem acima do preço (além disso, a gestão ativa de carteiras seria inútil e ineficiente nesta situação).

Assim, [Moffitt \(2017\)](#) constatou que a relação entre martingais e SNPL's se estabelece da seguinte forma: um SNPL para a série temporal financeira X_t existe se e somente se existir uma “série temporal de informações observáveis” Z_t tal que X_t (em relação a Z_t) não é um martingal. Assim, temos o seguinte:

3.6.3 O TEOREMA FUNDAMENTAL DA GESTÃO ATIVA

Seja $\{X_t : t \geq 0\}$ uma série temporal financeira adaptada à série temporal $\{Z_t : t \geq 0\}$. Temos que $\{X_t : t \geq 0\}$ é um martingal em relação a $\{Z_t : t \geq 0\}$ se e somente se não existir um Sistema de Negociação Potencialmente Lucrativo (SNPL) $f_t(Z_t, Z_{t-1}, \dots)$ para $\{X_t : t \geq 0\}$.

Demonstração: (\implies)

$\{X_t : t \geq 0\}$ é adaptada à $\{Z_t : t \geq 0\} \implies X_t$ é \mathcal{F}_t -mensurável, onde $\mathcal{F}_t = \sigma(Z_s, 0 \leq s \leq t)$, ou seja, $E(X_t | \mathcal{F}_t) = X_t$ para cada $t \in [0, \infty)$.

Além disso, $\{X_t : t \geq 0\}$ é martingal em relação a $\{Z_t : t \geq 0\}$, então

$$E(G_{t,s}(f)) = \sum_{u=t-s+1}^t E[(X_u - X_{u-1})f_{u-1}(Z_{u-1}, Z_{u-2}, \dots) | (Z_{u-1}, Z_{u-2}, \dots)]$$

$$E(G_{t,s}(f)) = \sum_{u=t-s+1}^t E[(X_u - X_{u-1})f_{u-1}(Z_{u-1}, Z_{u-2}, \dots) | \mathcal{F}_{u-1}]$$

pois $\mathcal{F}_{u-1} = \sigma(Z_s, 0 \leq s \leq u-1)$

$$E(G_{t,s}(f)) = \sum_{u=t-s+1}^t [E(X_u | \mathcal{F}_{u-1}) - X_{u-1}] f_{u-1}(Z_{u-1}, Z_{u-2}, \dots)$$

pois X_{u-1} e $f_{u-1}(Z_{u-1}, Z_{u-2}, \dots)$ são \mathcal{F}_{u-1} -mensuráveis.

Sabemos que $E(X_u | \mathcal{F}_{u-1}) = X_{u-1}$ pois $\{X_t : t \geq 0\}$ é martingal em relação à $\{Z_t : t \geq 0\}$.

Assim, temos:

$$E(G_{t,s}(f)) = \sum_{u=t-s+1}^t [X_{u-1} - X_{u-1}] f_{u-1}(Z_{u-1}, Z_{u-2}, \dots)$$

$$\implies E(G_{t,s}(f)) = 0, \text{ para } \forall t, s > 0$$

$$\implies \nexists \text{ SNPL } f_t(Z_t, Z_{t-1}, \dots) \text{ para } \{X_t : t \geq 0\}.$$

Demonstração: (\Leftarrow)

Suponha que $\{X_t : t \geq 0\}$ não é um martingal em relação a $\{Z_t : t \geq 0\}$. Assim, temos o seguinte conjunto:

$$K = \{t \in \mathbb{N}; E(X_{t+1} | Z_t, Z_{t-1}, \dots) = E(X_{t+1} | \mathcal{F}_t) \neq X_t\}$$

Ou seja, K é o conjunto dos instantes de tempo que não obedecem a propriedade martingal. Assim, definiremos um sistema de negociação $f_t(Z_t, Z_{t-1}, \dots)$ para $\{X_t : t \geq 0\}$ da seguinte forma:

$$f_t(Z_t, Z_{t-1}, \dots) = \begin{cases} \text{signal}(E[X_{t+1} | \mathcal{F}_t] - X_t) & \text{se } t \in K \\ 0 & \text{se } t \notin K \end{cases}$$

Já a função sinal é definida da seguinte maneira:

$$\text{senal}(x) = \begin{cases} 1 & \text{se } x > 0 \\ 0 & \text{se } x = 0 \\ -1 & \text{se } x < 0 \end{cases}$$

Assim, temos: $E(G_{t,s}(f)) = \sum_{u=t-s+1}^t [E(X_u|\mathcal{F}_{u-1}) - X_{u-1}]f_{u-1}(Z_{u-1}, Z_{u-2}, \dots)$

Logo, $E(G_{t,s}(f)) = E(G_{t,s}(f))_{(u \in K)} + E(G_{t,s}(f))_{(u \notin K)}$

Sabemos que $E(G_{t,s}(f))_{(u \in K)} > 0$ e $E(G_{t,s}(f))_{(u \notin K)} = 0$

Assim, constatamos que $E(G_{t,s}(f)) > 0, \forall s, t > 0$

$\implies f_t(Z_t, Z_{t-1}, \dots)$ é um SNPL para $\{X_t : t \geq 0\}$.

Em suma, o teorema 3.6.3 nos fornece o seguinte critério teórico para a gestão ativa de carteiras: as séries temporais de preços e retornos que não são martingais em relação às variáveis de condicionamento são capazes de gerar sistemas de negociação potencialmente lucrativos (SNPL's).

3.6.4 PROPOSIÇÃO: Seja $\{X_t : t \geq 0\}$ uma série temporal de preços que possui uma correlação serial ρ previsível, intermitente e não-nula ($|\rho| \neq 0$). Então existe um SNPL para $\{X_t : t \geq 0\}$.

A demonstração desta proposição está no capítulo 2 (na página 39) do livro do [Moffitt \(2017\)](#). Este enunciado é importante pois sintetiza os movimentos observados nos mercados financeiros rotineiramente. Quando os agentes do mercado detectam uma autocorrelação positiva (nesta situação retornos com os mesmos sinais tendem a se repetir no futuro, ou seja, o mercado está numa tendência clara de alta ou baixa) na série de retornos de determinado ativo, eles montam estratégias com o objetivo de obter lucros neste cenário. Já quando os agentes detectam uma autocorrelação negativa também é possível elaborar operações lucrativas pois nesta perspectiva retornos de sinais contrários tendem a se intercalar durante a evolução dos preços (neste caso os operadores montam estratégias de “reversão à média”).

4 MODELAGEM DO MERCADO ATRAVÉS DO HMM

4.1 INTRODUÇÃO

Na área de finanças, os HMM's são geralmente denominados como “modelos de troca de regime” (“regime switching models”) a partir do trabalho fundamental de [Hamilton \(1989\)](#). Desde então, esses modelos têm sido utilizados por um número grande de pesquisadores.

A vantagem em utilizar o HMM é que ele não impõe um conjunto rigoroso de suposições para o seu funcionamento, mas ainda assim obtemos resultados teóricos e computacionais úteis que não são alcançáveis com outros modelos simples. No mundo real das finanças, tais vantagens na modelagem são particularmente valiosas pois nos abrem a possibilidade de aprendermos mais sobre o comportamento dos ativos financeiros.

Além disso, iremos falar neste capítulo sobre a utilidade do HMM na modelagem do mercado de ações. Nesta etapa iremos abordar os principais aspectos de um importante artigo, de autoria de [Rydén, Teräsvirta e Åsbrink \(1998\)](#), que mostra que o HMM reproduz a maioria dos fatos estilizados sobre as séries diárias de retornos estabelecidos por [Granger e Ding \(1994\)](#). Para finalizar, veremos que estes resultados nos permitem ter uma perspectiva exploratória do HMM (tal como o trabalho de [Nguyen \(2017\)](#)) com o objetivo de utilizar o modelo na montagem de estratégias de negociação de ações.

4.2 CARACTERÍSTICAS DO HMM

Uma das principais características do HMM é a sua capacidade de emular o funcionamento de um determinado sistema. Ao analisarmos a matriz de transição A (estudada no capítulo 2) podemos determinar a probabilidade de um sistema permanecer em seu estado atual ou sofrer alguma alteração. No contexto dos mercados financeiros, essas probabilidades são muito úteis para os investidores estimarem a tendência dos mercados num determinado período. Além disso, a modelagem financeira através do HMM tende a ser mais simples que outros métodos utilizados em finanças (que são baseados majoritariamente em equações diferenciais estocásticas).

Assim, utilizaremos o HMM neste trabalho devido à sua capacidade de assimilar a dinâmica oculta dos estados que governam o mercado de ações. Portanto, com o HMM conseguiremos construir indicadores aptos a estimar a tendência dominante (alta, baixa ou estagnação) na evolução do preço de determinada ação num certo período de tempo. Com estes indicadores será possível implementar um sistema de negociação de ações que tenha como objetivo gerar lucro nestas operações.

4.3 FATOS ESTILIZADOS EM FINANÇAS E O HMM

Segundo [Aiube \(2013\)](#), fatos estilizados são “regularidades estatísticas observadas em um grande número de séries financeiras de retornos, a partir de estudos empíricos em diversos mercados”. Em seu artigo sobre fatos estilizados dos retornos diários das ações, [Rydén, Teräsvirta e Åsbrink \(1998\)](#) analisaram a capacidade do HMM de reproduzir quatro propriedades temporais e três propriedades estruturais (relativas às funções de distribuição de probabilidade) dos retornos examinados. Estas propriedades são as seguintes:

(I) Propriedades Temporais (PT):

(PT1): Os retornos r_t não são autocorrelacionados (exceto possivelmente no lag-1).

(PT2): As funções de autocorrelação de $|r_t|$ e r_t^2 decaem lentamente (partindo da primeira autocorrelação), e $\text{corr}(|r_t|, |r_{t-k}|) > \text{corr}(r_t^2, r_{t-k}^2)$. As autocorrelações permanecem positivas em lags muito longos. Além disso, o decaimento é muito mais lento que a taxa exponencial de decaimento dos modelos estacionários ARMA.

(PT3): $\text{corr}(|r_t|, |r_{t-k}|) > \text{corr}(|r_t|^\theta, |r_{t-k}|^\theta)$, onde $\theta \neq 1$. As autocorrelações das potências dos retornos absolutos são maiores quando $\theta = 1$ (esta propriedade é chamada de “efeito Taylor”).

(PT4): As autocorrelações observadas em $\text{Sinal}(r_t)$ são insignificantes.

(II) Propriedades Estruturais (PE):

(PE1): $|r_t|$ e $\text{Sinal}(r_t)$ são independentes.

(PE2): $|r_t|$ possui os mesmos valores para a média e o desvio-padrão.

(PE3): A distribuição marginal de $|r_t|$ é exponencial (após a remoção dos outliers).

O trabalho de [Rydén, Teräsvirta e Åsbrink \(1998\)](#) demonstra que o HMM reproduz a maioria dos fatos estilizados das séries diárias dos retornos estabelecidos por [Granger e Ding \(1994\)](#). A única exceção é a incapacidade do HMM de reproduzir o lento decaimento das funções de autocorrelação de $|r_t|$ e r_t^2 (propriedade PT2). Este lento decaimento é associado a propriedade de “longa memória” da série temporal. Além disso, os valores de $|r_t|$ e r_t^2 são estabelecidos como medidas de volatilidade das séries diárias dos retornos. Um dos principais efeitos causados pela “memória longa” da série temporal é o fenômeno descrito na literatura como “cluster de volatilidade” (isto ocorre quando a volatilidade

dos retornos passados afetam a volatilidade dos retornos futuros por um período consideravelmente longo de tempo; este tópico é aprofundado nos artigos de [Cont \(2001\)](#) e [Nystrup, Madsen e Lindström \(2015\)](#)). Ou seja, a propriedade PT2 nos diz que $|r_t|$ e r_t^2 são autocorrelacionados (logo a série dos retornos possui uma dependência temporal de ordem superior). A incapacidade do HMM em modelar a dependência temporal de ordem superior pode ser atribuída a distribuição de probabilidade implícita do tempo de permanência nos estados ocultos do modelo (distribuição geométrica; cuja propriedade é a “falta de memória”, ou seja, em um determinado momento, o tempo de espera para a mudança de um estado para um estado distinto independe do tempo já gasto no estado anterior). Além disso, a cadeia de Markov associada ao HMM tem na perda de memória a base de caracterização do modelo pois nele o futuro só depende do estado presente.

4.4 HMM E O MERCADO ACIONÁRIO

Como visto anteriormente, os HMMs são capazes de modelar as transições dos estados ocultos a partir da análise dos dados observáveis. Quando fazemos a previsão do preço de uma determinada ação, estamos realizando (ainda que intuitivamente) uma modelagem similar à natureza dos HMMs. O preço de uma determinada ação (dado observável) depende de uma multiplicidade de fatores (exemplos: decisões governamentais, decisões gerenciais da empresa, condições financeiras da empresa, nível de atividade econômica do país, dentre outros fatores) que geram uma dinâmica oculta dos estados que governam o preço deste ativo financeiro. Este trabalho é baseado no artigo produzido pela pesquisadora [Nguyen \(2017\)](#). A referida publicação desenvolveu uma metodologia (que será detalhada no próximo capítulo) que utiliza o HMM para prever os preços das ações e aplica estes resultados num sistema de negociação. Além disso, o HMM será modelado através de sequências de observações múltiplas e independentes. Assim, neste trabalho teremos os seguintes dados observáveis múltiplos: retornos dos preços de fechamento, retornos dos preços de abertura, retornos dos preços mínimos e retornos dos preços máximos das ações selecionadas. Estes dados observáveis serão utilizados para o HMM prever os preços de fechamento futuros das ações. Utilizaremos os retornos como dados observáveis e assumiremos que estas variáveis são independentes entre si (fazendo com que a modelagem do HMM se torne mais simples). Uma diferença importante entre este trabalho e o desenvolvido por [Nguyen \(2017\)](#) é que esta dissertação testará a eficácia do sistema de negociação em três cenários bem distintos (períodos de alta, de baixa e de neutralidade dos preços das ações). Nguyen testou o sistema de negociação em um único período (15/08/2016 a 11/08/2017) para 3 ações distintas (Apple, Facebook e Google) e obteve os seguintes resultados:

(a) Ação selecionada: Apple.

(i) Período de teste: 15/08/2016 a 11/08/2017.

(ii) Valorização da ação no período: 43,84% (método *buy and hold*).

(iii) Lucratividade do sistema de negociação: 31,91%.

(a) Ação selecionada: Facebook.

(i) Período de teste: 15/08/2016 a 11/08/2017.

(ii) Valorização da ação no período: 35,65% (método *buy and hold*).

(iii) Lucratividade do sistema de negociação: 23,53%.

(a) Ação selecionada: Google.

(i) Período de teste: 15/08/2016 a 11/08/2017.

(ii) Valorização da ação no período: 15,40% (método *buy and hold*).

(iii) Lucratividade do sistema de negociação: 24,86%.

Ou seja, podemos concluir que este sistema de negociação provavelmente será lucrativo num período de valorização das ações. Porém, o trabalho de [Nguyen \(2017\)](#) não testa esta metodologia em cenários em que as ações ficam estagnadas ou se desvalorizam. Além disso, a referida pesquisadora em seu artigo compara o desempenho do seu modelo com um *benchmark* heterodoxo (no qual o retorno previsto do dia seguinte tem como referência o retorno do dia atual). Diferentemente de [Nguyen \(2017\)](#), iremos utilizar a rentabilidade do método *buy and hold* (que é uma referência padrão em Finanças) como *benchmark* dos modelos apresentados neste trabalho. Assim, esta dissertação tem o objetivo de testar a eficácia do sistema de negociação baseado no HMM para ações brasileiras em três cenários-padrão do mercado acionário (valorização, desvalorização e estagnação das ações) e verificar a sua viabilidade.

5 METODOLOGIA

5.1 INTRODUÇÃO

A acurácia de qualquer modelo de previsão é um fator importante para determinarmos a sua utilidade na modelagem de fenômenos. No caso específico do HMM, a precisão das previsões irá variar em função do número de estados ocultos utilizados para modelar o HMM. Portanto, para prevermos os preços das ações utilizando o HMM temos que primeiramente determinar o número ideal de estados ocultos do modelo. Neste capítulo descreveremos os dois critérios padrão - o critério de informação de Akaike (AIC) e o critério de informação Bayesiano (BIC) - utilizados para alcançarmos este objetivo. Por fim, iremos desenvolver a metodologia necessária para se construir um sistema de negociação baseado no HMM capaz de prever e negociar o preço das ações.

5.2 SELEÇÃO DOS MODELOS

Escolher o número de estados ocultos para o HMM é uma tarefa crítica. Utilizaremos os dois critérios mais comuns (o AIC e o BIC) para avaliar os desempenhos do HMM com diferentes números de estados escondidos. Estes critérios medem o quão bem um modelo se ajusta a determinados dados, através da avaliação do logaritmo da verossimilhança do HMM junto com o termo de penalização. A verossimilhança L (associada a uma escolha λ dos parâmetros) do HMM é a probabilidade das observações quando os parâmetros do modelo tomam um valor λ . Assim, temos:

$$L = P(O|\lambda)$$

Os dois critérios (AIC e BIC) são adequados para o HMM pois no algoritmo de Baum-Welch o método EM é utilizado para maximizar o logaritmo da verossimilhança do modelo. Assim, AIC e BIC são calculados da seguinte forma:

$$AIC = -2\ln(L) + 2K$$

$$BIC = -2\ln(L) + K \ln(M)$$

Onde L é a verossimilhança do modelo, M é o número de observações e K é o número de parâmetros estimados do modelo. Neste trabalho, assumiremos que a distribuição correspondente a cada estado oculto é uma distribuição gaussiana. Portanto, o número de parâmetros, K , é definido como $K = N^2 + 2N - 1$, onde N é o número de estados ocultos utilizados no HMM. O AIC é uma medida relativa de qualidade entre modelos estatísticos.

Ele estima a perda de informação relativa quando um modelo é usado para representar o processo gerador dos dados. Dado um conjunto de modelos candidatos, o modelo escolhido será aquele que apresentar o menor valor para AIC pois o primeiro termo ($-2 \ln(L)$; que nos mostra o quão bom o modelo se ajusta aos dados) diminui com o aumento do número de estados e o segundo termo ($2K$; chamado de termo de penalidade) aumenta com o crescimento do número de parâmetros. O BIC é um critério de seleção bem semelhante ao AIC, diferenciando-se apenas pelo termo de penalidade ($K \ln(M)$). Comparando com o AIC, o termo de penalidade do BIC é mais rigoroso, ou seja, ele favorece os modelos com menor número de parâmetros.

5.3 PREVISÃO E NEGOCIAÇÃO DAS AÇÕES

Após determinarmos o número ideal de estados ocultos do HMM, iremos utilizar este modelo para prever e negociar o preço das ações. Para realizarmos este procedimento escolheremos os seguintes ativos financeiros: índice Ibovespa e algumas ações (das seguintes empresas: Petrobrás, Vale, Itaú e Ambev). Os períodos de análise serão feitos nos seguintes intervalos: 22/06/2015 a 21/08/2015 (época em que as ações se desvalorizaram), 25/01/2016 a 25/03/2016 (época em que as ações se valorizaram) e 26/02/2018 a 27/04/2018 (época em que os preços das ações ficaram estagnados) Os valores previstos dos ativos serão confrontados com os valores reais para se medir a acurácia e a viabilidade do HMM neste tipo de situação. Além disso, estas previsões serão incorporadas num sistema de negociação de ações com o objetivo de medir a lucratividade destas operações.

5.3.1 PREVISÃO DO PREÇO DAS AÇÕES

Partindo do pressuposto de que os preços das ações repetem os padrões apresentados no passado, nesta etapa iremos basicamente buscar o intervalo de tempo pretérito no qual as ações se comportam de maneira semelhante ao período atual. Com esta informação, poderemos estimar os preços do dia seguinte das ações. Assim, o processo de previsão pode ser dividido resumidamente nas seguintes etapas:

Etapa 1: Determinar o intervalo de tempo no qual o HMM será testado (intervalo $[T_1, T_2]$). Além disso, estabeleceremos o número x de dias da Janela Temporal Retrospectiva (que será utilizada para capturar no passado o intervalo de tempo com a verossimilhança semelhante a do intervalo atual).

Etapa 2: Executar o HMM para o intervalo $[T - (r - 1), T]$ (onde $[T - (r - 1), T] = [T_1 - i, T_2 - i]$ tal que $i \in [1, r]$, onde $i \in \mathbb{N}$; e $r = (T_2 - T_1) + 1$) e calcular a verossimilhança do modelo.

Etapa 3: Encontrar o intervalo do passado (que pertença a Janela Temporal Retrospectiva) que possua uma verossimilhança similar a do intervalo atual ($[T - (r - 1), T]$).

Etapa 4: Utilizar o último retorno da ação no intervalo subsequente ao encontrado na Etapa 3 para prever o retorno da ação no dia seguinte (dia $T + 1$).

Esta forma de se prever o preço das ações baseia-se nos artigos de [Hassan e Nath \(2005\)](#) e de [Nguyen \(2017\)](#). Este trabalho segue a metodologia aplicada em [Nguyen \(2017\)](#) na qual aplicamos o HMM para os retornos dos preços máximos, mínimos, de abertura e de fechamento, que iremos considerar independentes. Assim, a metodologia aplicada nesta dissertação é mais simples que a técnica apresentada em [Hassan e Nath \(2005\)](#) (os autores do artigo aplicaram o HMM diretamente nos preços máximos, mínimos, de abertura e de fechamento, que são dependentes entre si).

Assim, vamos supor que queremos prever o preço de fechamento do dia seguinte da ação A. Ao detalharmos a Etapa 2 citada anteriormente temos o seguinte:

Pegamos a série temporal O_t dos retornos de A para executar o HMM e calcular a sua verossimilhança. Assim, temos:

$$O_t = \left\{ O_t^{(1)}, O_t^{(2)}, O_t^{(3)}, O_t^{(4)}; t \in \mathbb{N}; t \in [T - (r - 1), T] \right\}$$

onde

$O_t^{(1)}$: retorno do preço mínimo de A no tempo t ;

$O_t^{(2)}$: retorno do preço máximo de A no tempo t ;

$O_t^{(3)}$: retorno do preço de abertura de A no tempo t ;

$O_t^{(4)}$: retorno do preço de fechamento de A no tempo t ;

Já na Etapa 3 temos o seguinte:

(i) Movemos o bloco de dados “para trás” em um dia para obter novos dados de observação. Assim, temos:

$$O_t = \left\{ O_t^{(1)}, O_t^{(2)}, O_t^{(3)}, O_t^{(4)}; t \in \mathbb{N}; t \in [T - r, T - 1] \right\}$$

Com estes novos dados, executamos o HMM novamente e obtemos a nova verossimilhança.

(ii) Analogamente, moveremos o bloco de dados “para trás” em um dia sucessivamente até atingirmos o número x de dias da Janela Temporal Retrospectiva. Assim, temos:

$$O_{t,0} = \left\{ O_t^{(1)}, O_t^{(2)}, O_t^{(3)}, O_t^{(4)}; t \in \mathbb{N}; t \in [T - (r - 1), T] \right\}$$

(com a verossimilhança do HMM em $O_{t,0}$ igual a $V(O_{t,0})$);

$$O_{t,1} = \left\{ O_t^{(1)}, O_t^{(2)}, O_t^{(3)}, O_t^{(4)}; t \in \mathbb{N}; t \in [T - (r - 1) - 1, T - 1] \right\}$$

(com a verossimilhança do HMM em $O_{t,1}$ igual a $V(O_{t,1})$);

(...)

$$O_{t,x} = \left\{ O_t^{(1)}, O_t^{(2)}, O_t^{(3)}, O_t^{(4)}; t \in \mathbb{N}; t \in [T - (r - 1) - x, T - x] \right\}$$

(com a verossimilhança do HMM em $O_{t,x}$ igual a $V(O_{t,x})$);

Assim, temos:

$$V(O_{t,j}) = \min_{1 \leq i \leq x} |V(O_{t,0}) - V(O_{t,i})|$$

onde

$$O_{t,j} = \left\{ O_t^{(1)}, O_t^{(2)}, O_t^{(3)}, O_t^{(4)}; t \in \mathbb{N}; t \in [(T - j) - (r - 1), (T - j)] \right\}$$

sendo $T^* = (T - j)$, temos:

$$O_{t,j} = \left\{ O_t^{(1)}, O_t^{(2)}, O_t^{(3)}, O_t^{(4)}; t \in \mathbb{N}; t \in [T^* - (r - 1), T^*] \right\}$$

onde $[T^* - (r - 1), T^*]$ é o intervalo do passado (que pertence a Janela Temporal Retrospectiva) no qual a verossimilhança do HMM da ação A é similar à verossimilhança do intervalo atual $([T - (r - 1), T])$.

Na Etapa 4 faremos o seguinte:

Após encontrarmos o intervalo do passado similar, estimaremos o retorno do preço de fechamento da ação A no tempo $T + 1$ utilizando a seguinte fórmula:

$$O_{t+1}^{(4)} = O_{t^*+1}^{(4)}$$

Assim, o preço de fechamento previsto para a ação A no tempo $T + 1$ é calculado

da seguinte maneira:

$$P_{t+1} = P_t(O_{t+1}^{(4)} + 1)$$

Após realizarmos a primeira previsão para o retorno da ação A (para o dia $T + 1$), atualizaremos a janela de dados (O_t) movendo-a “para frente” um dia. Assim, temos:

$$O_t = \left\{ O_t^{(1)}, O_t^{(2)}, O_t^{(3)}, O_t^{(4)}; t \in \mathbb{N}; t \in [(T + 1) - (r - 1), (T + 1)] \right\}$$

E assim estimaremos o retorno da ação A para o dia $T + 2$. Analogamente, repetiremos este processo de previsão sucessivamente até obtermos todas as previsões para a ação A no intervalo $[T_1, T_2]$. Mediremos a acurácia do HMM neste tipo de problema no próximo capítulo utilizando dois indicadores: o erro percentual absoluto médio e a correlação entre os preços previstos e os preços reais. Além disso, analisaremos posteriormente a viabilidade de se implementar estratégias de negociação utilizando o HMM. A estratégia será feita da seguinte forma: se o HMM prever que o preço de uma determinada ação subirá amanhã, então compraremos esta ação hoje e a venderemos amanhã (vamos considerar que compramos e vendemos a ação pelo preço de fechamento). Se o HMM prever que o preço de uma determinada ação não subirá amanhã, então não faremos nada. Observamos que esta estratégia é similar à adotada por [Nguyen \(2017\)](#) em seu trabalho. Salientamos que iremos considerar que não existem custos de transações nestas operações e que poderemos realizar negociações arbitrariamente fracionadas.

Observamos que a metodologia apresentada neste capítulo será implementada através do programa estatístico R (os detalhes dos códigos desenvolvidos neste trabalho serão apresentados no apêndice A desta dissertação). Além disso, a utilização do pacote `depmixS4`, cuja as funcionalidades são detalhadas no artigo de [Visser, Speekenbrink et al. \(2010\)](#), do software R nos permitirá ajustar e modelar os HMM's de maneira satisfatória.

6 RESULTADOS

Neste capítulo iremos apresentar os resultados obtidos com a metodologia apresentada no capítulo anterior. Escolheremos cinco ativos financeiros (ações das empresas Ambev, Itaú, Petrobrás, Vale e o índice Ibovespa) e mediremos a rentabilidade do sistema de negociação apresentado anteriormente, cuja as variáveis são o número de estados do modelo HMM e o intervalo em dias da Janela Temporal Retrospectiva (estabelecidas a priori com 10, 20, 40 ou 80 dias). Cada célula das tabelas que serão apresentadas nas próximas seções terão os seguintes dados: correlação entre a série de previsões do modelo e a série de preços da ação (primeira linha), rentabilidade do sistema de negociação (segunda linha) e o erro percentual absoluto médio do sistema de negociação em relação à série de preços da ação (terceira linha). Salientamos que as células que forem preenchidas com ** correspondem aos casos nos quais o modelo HMM associado não obteve convergência. Além disso, as células destacadas em negrito em cada tabela são os dados dos modelos que obtiveram os melhores desempenhos em cada cenário. Sendo assim, listaremos a seguir os resultados obtidos.

6.1 CASO 1 - AMBEV

CENÁRIO I : DESVALORIZAÇÃO

Modelo / Janela temporal	10 dias	20 dias	40 dias	80 dias
HMM-2	0,5014	0,4871	0,5151	0,5009
	-7,01%	-5,25%	-4,84%	-3,80%
	1,65%	1,58%	1,68%	1,70%
HMM-3	0,5355	0,6523	0,7525	0,7108
	-2,52%	-2,96%	-2,02%	1,05%
	1,76%	1,41%	1,32%	1,36%
HMM-4	**	**	**	**
	**	**	**	**
	**	**	**	**

Tabela 1 – Desempenho do HMM com a ação ABEV3 num período de desvalorização.

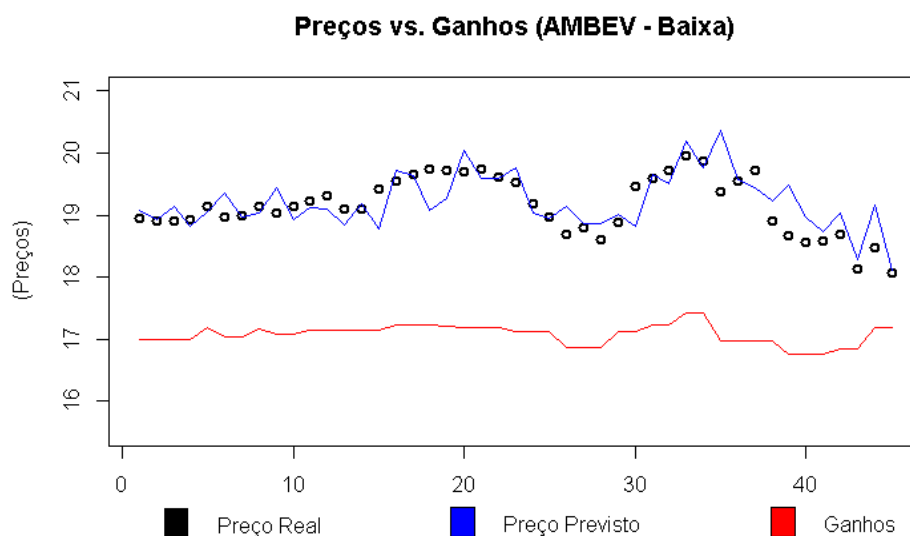


Figura 1 – Desempenho do modelo HMM-3/80d num período de desvalorização (baixa).

(i) Período de avaliação: 22/06/2015 a 21/08/2015.

(ii) Valorização da ação no período (método *buy and hold*): -4,69%.

(iii) Modelo HMM com melhor desempenho no período: HMM-3/80d (HMM de 3 estados com uma Janela Temporal Retrospectiva de 80 dias).

(iv) Rentabilidade do HMM com melhor desempenho (HMM-3/80d): 1,05%.

Observamos que o modelo HMM-3/80d superou a estratégia *buy and hold*. Além disso, a Figura 1 mostra a evolução da riqueza (gráfico vermelho) do agente que utilizar o modelo HMM-3/80d para negociar as ações da Ambev no período.

CENÁRIO II : VALORIZAÇÃO

Modelo / Janela temporal	10 dias	20 dias	40 dias	80 dias
HMM-2	0,5904	0,5948	0,6228	0,6328
	8,84%	17,32%	16,74%	12,35%
	2,16%	2,17%	2,19%	2,05%
HMM-3	**	**	**	**
	**	**	**	**
	**	**	**	**

Tabela 2 – Desempenho do HMM com a ação ABEV3 num período de valorização.

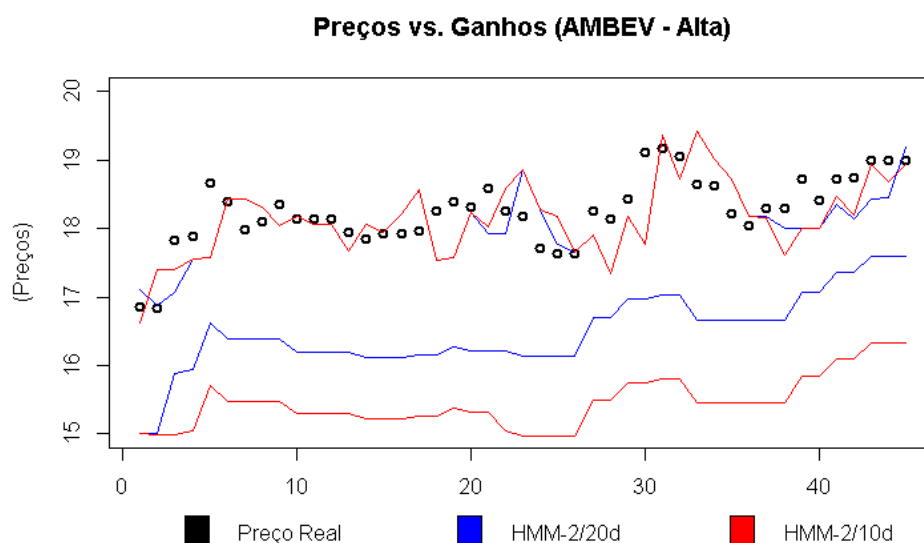


Figura 2 – Comparação entre HMM-2/20d e HMM-2/10d num período de valorização (alta).

(i) Período de avaliação: 25/01/2016 a 25/03/2016.

(ii) Valorização da ação no período (método *buy and hold*): 12,76%.

(iii) Modelo HMM com melhor desempenho no período: HMM-2/20d.

(iv) Rentabilidade do HMM com melhor desempenho (HMM-2/20d): 17,32%.

Observamos que o modelo HMM-2/20d superou a estratégia *buy and hold*. Além disso, a Figura 2 (diferentemente da Figura 1, que descreve a performance do melhor modelo isoladamente) compara os desempenhos dos modelos HMM-2/20d (gráficos azuis, onde a parte superior representa a série de previsão de preços e a inferior os ganhos acumulados com o modelo neste período) e HMM-2/10d (gráficos vermelhos, onde a parte superior representa a série de previsão de preços e a inferior os ganhos acumulados com o modelo neste período). Apesar destes modelos apresentarem correlações similares em relação à série de preços da ação (0,5948 e 0,5904), a previsão acertada na alta do preço do 2º para o 3º dia fez com que o modelo HMM-2/20d acumulasse uma riqueza consideravelmente maior (17,32% de ganho em relação à riqueza inicial) que a alcançada com o modelo HMM-2/10d (8,84% de ganho em relação à riqueza inicial).

6.2 CASO 2 - ITAÚ

CENÁRIO I : DESVALORIZAÇÃO

Modelo / Janela temporal	10 dias	20 dias	40 dias	80 dias
HMM-2	0,8574	0,8718	0,8842	0,8857
	-8,05%	-3,49%	-4,53%	-8,35%
	2,45%	2,26%	2,13%	2,15%
HMM-3	0,8623	0,8636	**	**
	-12,50%	-9,87%	**	**
	2,40%	2,54%	**	**
HMM-4	**	**	**	**
	**	**	**	**
	**	**	**	**

Tabela 3 – Desempenho do HMM com a ação ITUB4 num período de desvalorização.

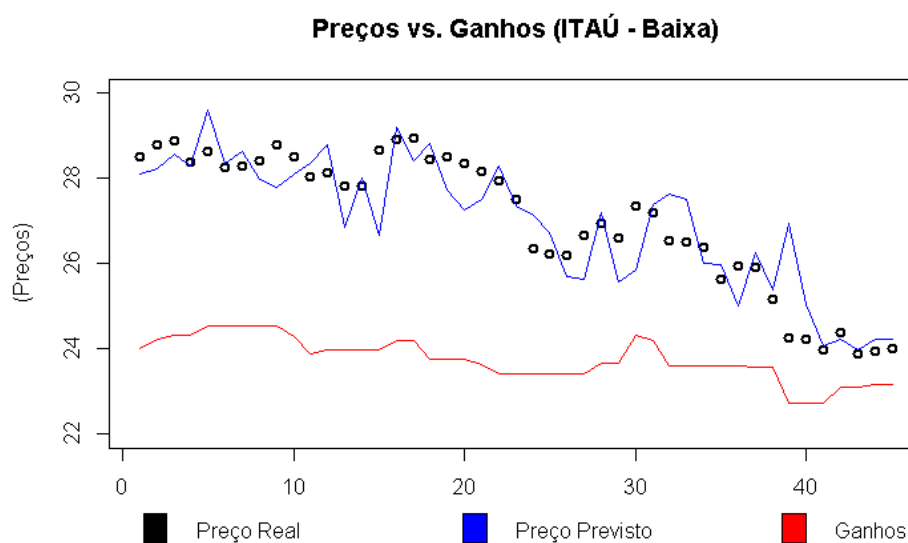


Figura 3 – Desempenho do modelo HMM-2/20d num período de desvalorização (baixa).

(i) Período de avaliação: 22/06/2015 a 21/08/2015.

(ii) Valorização da ação no período (método *buy and hold*): -15,80%.

(iii) Modelo HMM com melhor desempenho no período: HMM-2/20d.

(iv) Rentabilidade do HMM com melhor desempenho (HMM-2/20d): -3,49%.

Observamos que o modelo HMM-2/20d superou a estratégia *buy and hold*. Além disso, a Figura 3 mostra a evolução da riqueza (gráfico vermelho) do agente que utilizar o modelo HMM-2/20d para negociar as ações do Itaú no período.

CENÁRIO II : VALORIZAÇÃO

Modelo / Janela temporal	10 dias	20 dias	40 dias	80 dias
HMM-2	0,8975 30,83% 4,14%	0,8855 31,82% 4,72%	0,8764 25,25% 4,91%	0,8946 26,57% 4,78%
HMM-3	0,8797 33,36% 4,69%	0,8875 16,76% 4,45%	0,9085 12,49% 4,14%	0,9174 15,07% 4,02%
HMM-4	0,9125 16,56% 3,82%	** ** **	** ** **	** ** **
HMM-5	** ** **	** ** **	** ** **	** ** **

Tabela 4 – Desempenho do HMM com a ação ITUB4 num período de valorização.

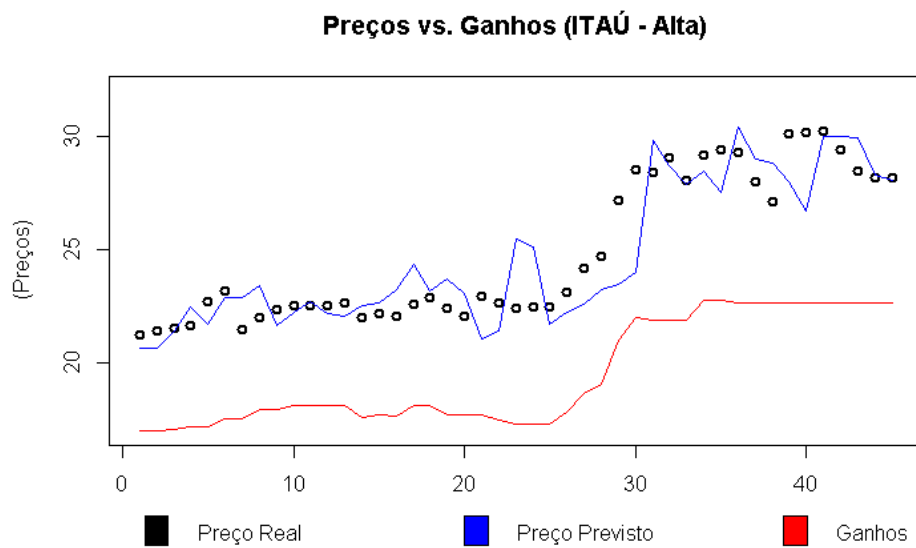


Figura 4 – Desempenho do modelo HMM-3/10d num período de valorização (alta).

(i) Período de avaliação: 25/01/2015 a 25/03/2015.

(ii) Valorização da ação no período (método *buy and hold*): 32,78%.

(iii) Modelo HMM com melhor desempenho no período: HMM-3/10d.

(iv) Rentabilidade do HMM com melhor desempenho (HMM-3/10d): 33,36%.

Observamos que o modelo HMM-3/10d superou a estratégia *buy and hold*. Além disso, a Figura 4 mostra a evolução da riqueza (gráfico vermelho) do agente que utilizar o modelo HMM-3/10d para negociar as ações do Itaú no período.

6.3 CASO 3 - PETROBRÁS

CENÁRIO I : DESVALORIZAÇÃO

Modelo / Janela temporal	10 dias	20 dias	40 dias	80 dias
HMM-2	0,9321	0,9230	0,9082	0,9109
	-17,11%	-26,46%	-32,88%	-33,44%
	3,72%	4,02%	4,30%	4,36%
HMM-3	0,9091	**	**	**
	-18,91%	**	**	**
	4,24%	**	**	**
HMM-4	**	**	**	**
	**	**	**	**
	**	**	**	**

Tabela 5 – Desempenho do HMM com a ação PETR4 num período de desvalorização.

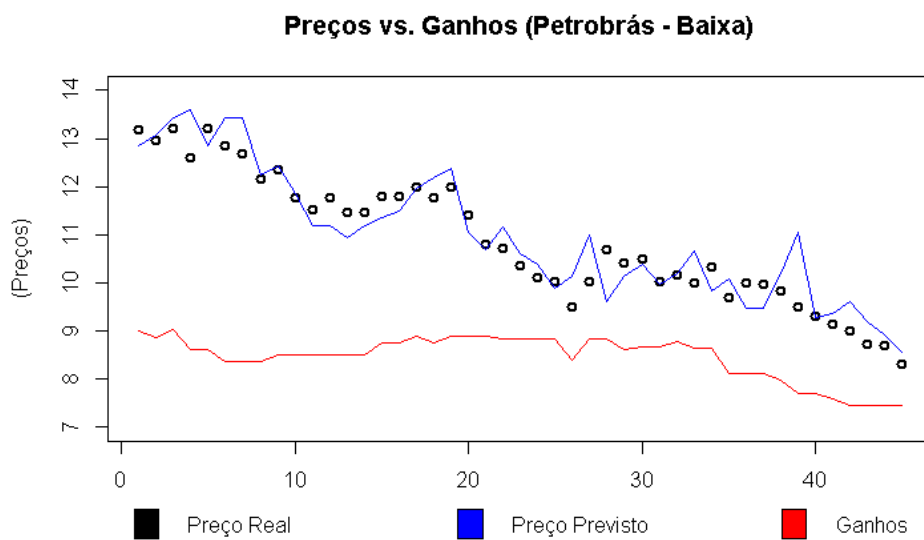


Figura 5 – Desempenho do modelo HMM-2/10d num período de desvalorização (baixa).

(i) Período de avaliação: 22/06/2015 a 21/08/2015.

(ii) Valorização da ação no período (método *buy and hold*): -36,98%.

(iii) Modelo HMM com melhor desempenho no período: HMM-2/10d.

(iv) Rentabilidade do HMM com melhor desempenho (HMM-2/10d): -17,11%.

Observamos que o modelo HMM-2/10d superou a estratégia *buy and hold*. Além disso, a Figura 5 mostra a evolução da riqueza (gráfico vermelho) do agente que utilizar o modelo HMM-2/10d para negociar as ações da Petrobrás no período.

CENÁRIO II : VALORIZAÇÃO

Modelo / Janela temporal	10 dias	20 dias	40 dias	80 dias
HMM-2	0,9194	0,9163	0,9308	0,9308
	55,83%	47,95%	47,95%	47,95%
	7,03%	7,54%	6,91%	6,91%
HMM-3	**	**	**	**
	**	**	**	**
	**	**	**	**

Tabela 6 – Desempenho do HMM com a ação PETR4 num período de valorização.

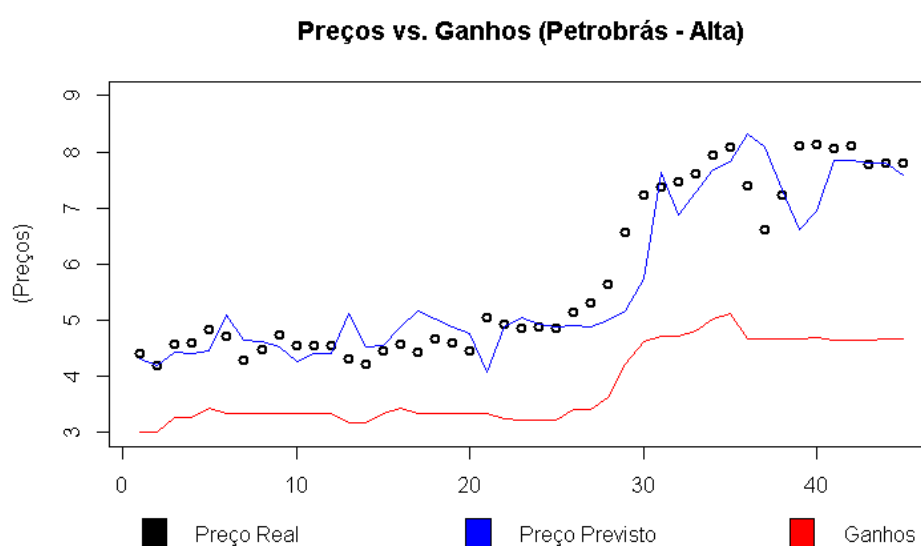


Figura 6 – Desempenho do modelo HMM-2/10d num período de valorização (alta).

- (i) Período de avaliação: 25/01/2015 a 25/03/2015.
- (ii) Valorização da ação no período (método *buy and hold*): 77,09%.
- (iii) Modelo HMM com melhor desempenho no período: HMM-2/10d.
- (iv) Rentabilidade do HMM com melhor desempenho (HMM-2/10d): 55,83%.

Observamos que o modelo HMM-2/10d não superou a estratégia *buy and hold*. Além disso, a Figura 6 mostra a evolução da riqueza (gráfico vermelho) do agente que utilizar o modelo HMM-2/10d para negociar as ações da Petrobrás no período.

6.4 CASO 4 - VALE

CENÁRIO I : DESVALORIZAÇÃO

Modelo / Janela temporal	10 dias	20 dias	40 dias	80 dias
HMM-2	0,6307	0,6019	0,5411	0,5029
	-5,57%	-4,51%	-8,52%	-4,42%
	3,81%	4,32%	4,21%	4,52%
HMM-3	0,6891	0,6879	0,5986	0,5349
	1,61%	3,02%	-6,99%	-7,75%
	3,27%	3,49%	3,94%	3,98%
HMM-4	0,6815	0,6455	0,6463	0,7301
	-8,47%	-6,11%	-3,39%	1,56%
	3,61%	3,78%	3,79%	3,21%
HMM-5	**	**	**	**
	**	**	**	**
	**	**	**	**

Tabela 7 – Desempenho do HMM com a ação VALE3 num período de desvalorização.

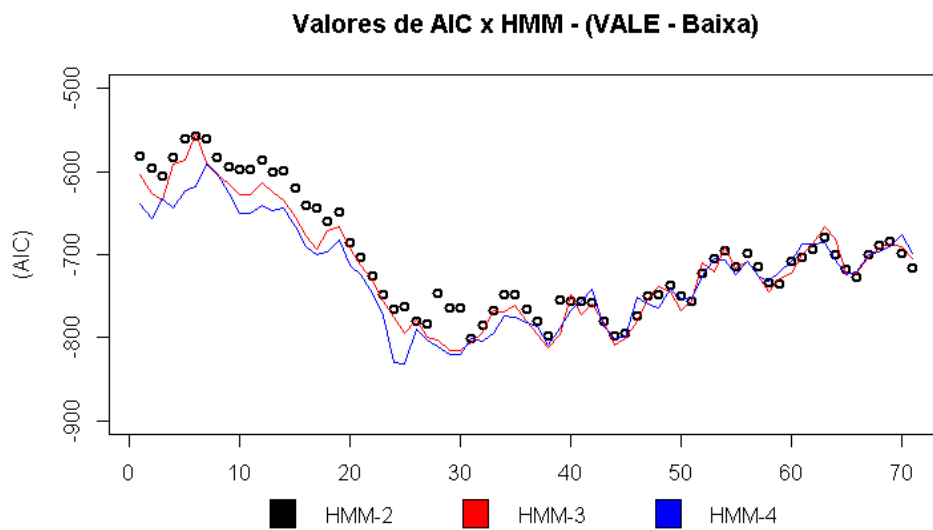


Figura 7 – Valores de AIC para cada HMM num período de desvalorização (baixa).

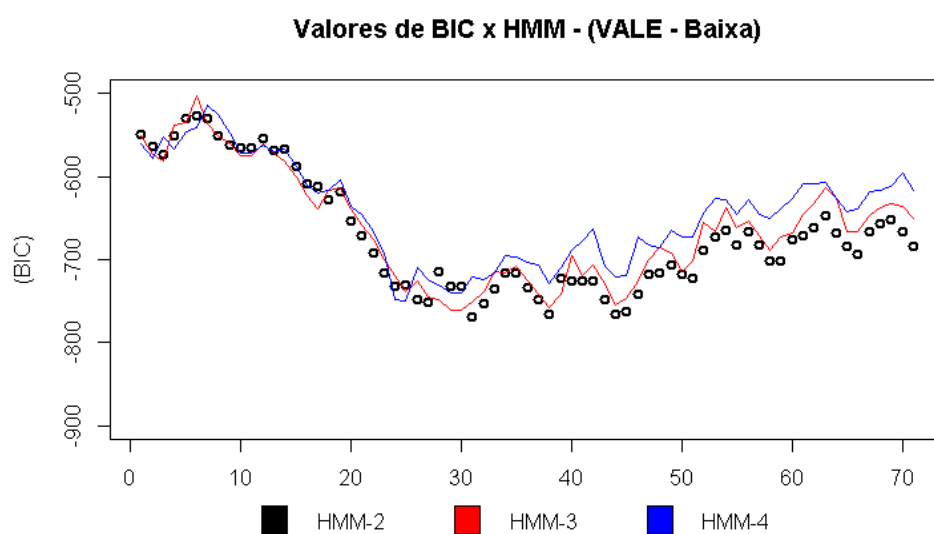


Figura 8 – Valores de BIC para cada HMM num período de desvalorização (baixa).

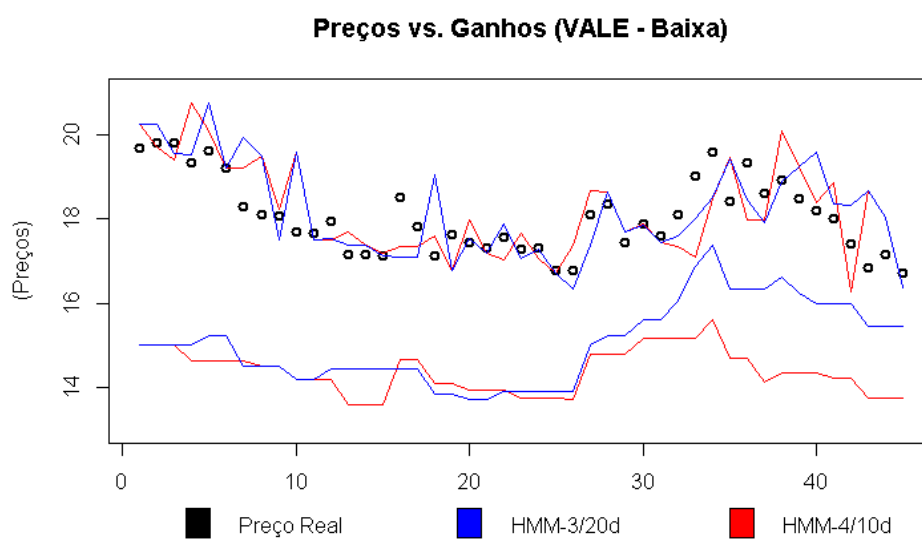


Figura 9 – Comparação entre HMM-3/20d e HMM-4/10d num período de desvalorização (baixa).

(i) Período de avaliação: 22/06/2015 a 21/08/2015.

(ii) Valorização da ação no período (método *buy and hold*): -15,18%.

(iii) Modelo HMM com melhor desempenho no período: HMM-3/20d.

(iv) Rentabilidade do HMM com melhor desempenho (HMM-3/20d): 3,02%.

Observamos que o modelo HMM-3/20d superou a estratégia *buy and hold*. Conforme detalhado na Tabela 7, conseguimos gerar neste cenário 3 classes de modelos distintas e convergentes (HMM-2, HMM-3 e HMM-4). Sendo assim, estamos numa situação propícia para verificar se a performance destes modelos estão consonantes com a determinação do número ideal de estados ocultos do HMM (feita através da análise de AIC e BIC). Além disso, na Figura 7 notamos que os valores de AIC para os modelos HMM-2, HMM-3 e HMM-4 são bem próximos. Salientamos que até a metade do eixo das abcissas os valores de AIC para os modelos HMM-3 e HMM-4 estão num patamar levemente inferior ao valor de AIC para o modelo HMM-2. Já em relação ao BIC na Figura 8, observamos que os valores dos modelos HMM-2 e HMM-3 estão num patamar levemente inferior ao encontrado no HMM-4 na metade final do eixo das abcissas. Conforme já estudamos no capítulo 5, os modelos mais adequados são aqueles que apresentam os menores valores de AIC e BIC. Sendo assim, ao analisarmos em conjunto as Figuras 7 e 8 devemos escolher o modelo HMM-3 para a ação da Vale no período analisado. Esta conclusão é coerente com o resultado empírico (no qual constatamos que o melhor modelo para a ação da Vale no período é o HMM-3/20d). A Figura 9 compara os desempenhos dos modelos HMM-3/20d (gráficos azuis, onde a parte superior representa a série de previsão de preços e a inferior os ganhos acumulados com o modelo neste período) e HMM-4/10d (gráficos vermelhos, onde a parte superior representa a série de previsão de preços e a inferior os ganhos acumulados com o modelo neste período). Apesar destes modelos apresentarem correlações similares em relação à série de preços da ação (0,6879 e 0,6815), as previsões acertadas nos dias 27, 31, 32 e 36 fizeram com que o modelo HMM-3/20d acumulasse uma riqueza consideravelmente maior (3,02% de ganho em relação à riqueza inicial) que a alcançada com o modelo HMM-4/10d (-8,47% de perda em relação à riqueza inicial).

CENÁRIO II : VALORIZAÇÃO

Modelo / Janela temporal	10 dias	20 dias	40 dias	80 dias
HMM-2	0,8639	0,8824	0,8801	0,8801
	16,86%	5,84%	5,84%	5,84%
	7,34%	7,06%	7,12%	7,12%
HMM-3	**	**	**	**
	**	**	**	**
	**	**	**	**

Tabela 8 – Desempenho do HMM com a ação VALE3 num período de valorização.

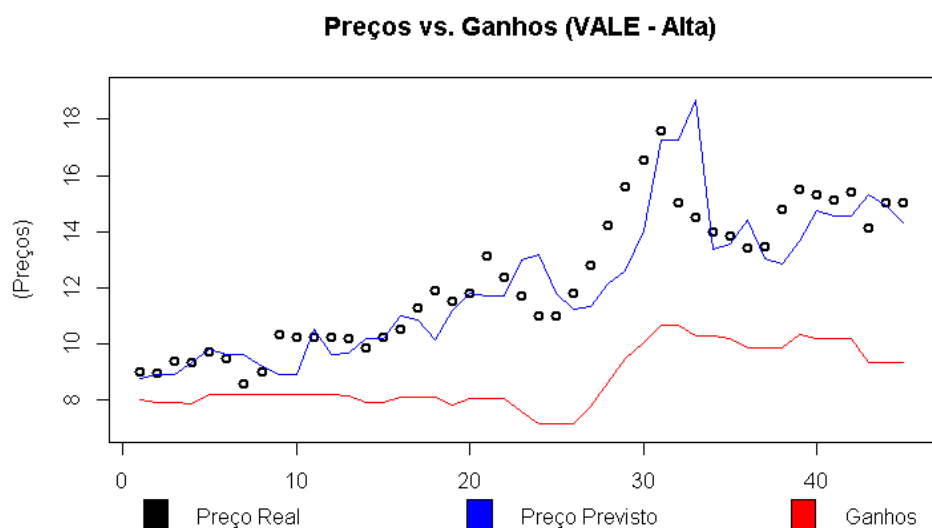


Figura 10 – Desempenho do modelo HMM-2/10d num período de valorização (alta).

- (i) Período de avaliação: 25/01/2015 a 25/03/2015.
- (ii) Valorização da ação no período (método *buy and hold*): 66,67%.
- (iii) Modelo HMM com melhor desempenho no período: HMM-2/10d.
- (iv) Rentabilidade do HMM com melhor desempenho (HMM-2/10d): 16,86%.

Observamos que o modelo HMM-2/10d não superou a estratégia *buy and hold*. Além disso, a Figura 10 mostra a evolução da riqueza (gráfico vermelho) do agente que utilizar o modelo HMM-2/10d para negociar as ações da Vale no período.

6.5 CASO 5 - IBOVESPA

CENÁRIO I : DESVALORIZAÇÃO

Modelo / Janela temporal	10 dias	20 dias	40 dias	80 dias
HMM-2	0,8770 -5,51% 1,76%	0,8712 -11,77% 1,78%	0,8603 -14,75% 1,87%	0,8849 -13,93% 1,74%
HMM-3	0,8790 -3,96% 1,80%	0,8646 0,82% 1,93%	0,8803 -6,51% 1,74%	0,9207 -5,59% 1,37%
HMM-4	0,8899 -5,79% 1,77%	0,8341 -8,96% 2,17%	0,8764 -9,32% 1,71%	0,8894 -3,27% 1,67%
HMM-5	** ** **	** ** **	** ** **	** ** **

Tabela 9 – Desempenho do HMM com o índice Ibovespa num período de desvalorização.

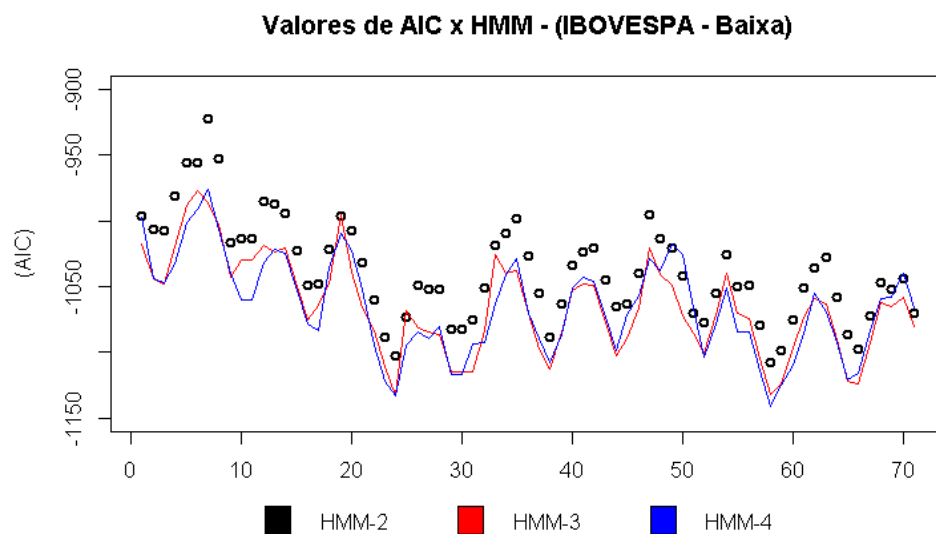


Figura 11 – Valores de AIC para cada HMM num período de desvalorização (baixa).

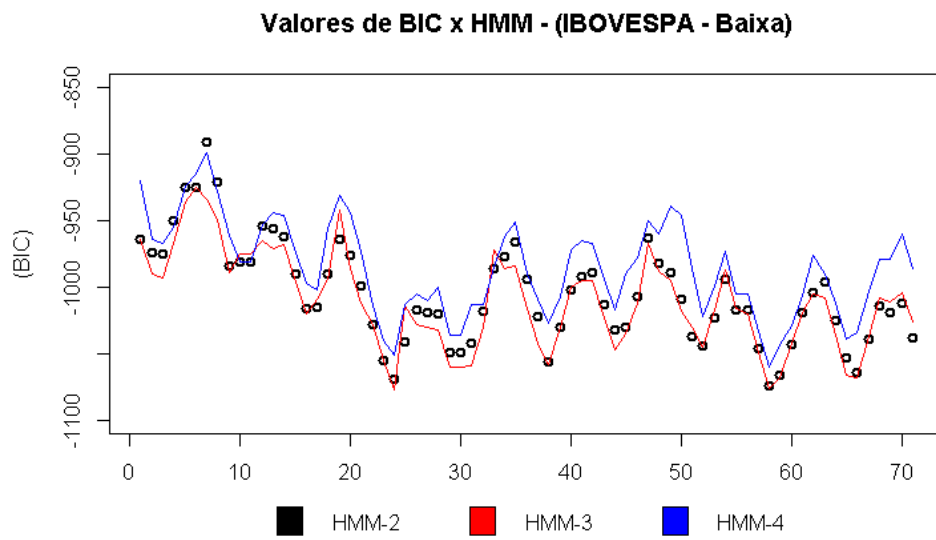


Figura 12 – Valores de BIC para cada HMM num período de desvalorização (baixa).

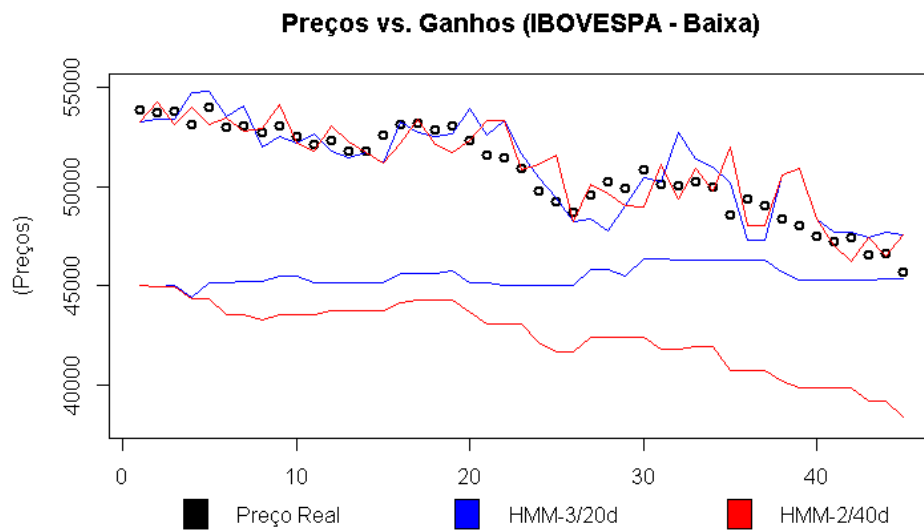


Figura 13 – Comparação entre HMM-3/20d e HMM-2/40d num período de desvalorização (baixa).

(i) Período de avaliação: 22/06/2015 a 21/08/2015.

(ii) Valorização da ação no período (método *buy and hold*): $-15,12\%$.

(iii) Modelo HMM com melhor desempenho no período: HMM-3/20d.

(iv) Rentabilidade do HMM com melhor desempenho (HMM-3/20d): $0,82\%$.

Observamos que o modelo HMM-3/20d superou a estratégia *buy and hold*. Conforme detalhado nas Tabelas 9, 10 e 11, conseguimos gerar nos cenários do Ibovespa no mínimo 3 classes de modelos distintas e convergentes (HMM-2, HMM-3 e HMM-4). Sendo assim, estamos numa situação propícia para verificar se a performance destes modelos estão consonantes com a determinação do número ideal de estados ocultos do HMM (feita através da análise de AIC e BIC). Além disso, na Figura 11 notamos que os valores de AIC para os modelos HMM-2, HMM-3 e HMM-4 são em geral bem próximos. Salientamos que os valores de AIC para os modelos HMM-3 e HMM-4 estão num patamar levemente inferior ao valor de AIC para o modelo HMM-2. Já em relação ao BIC na Figura 12, observamos que os valores dos modelos HMM-2 e HMM-3 estão num patamar levemente inferior ao encontrado no HMM-4. Conforme já dito anteriormente, os modelos mais adequados são aqueles que apresentam os menores valores de AIC e BIC. Sendo assim, ao analisarmos em conjunto as Figuras 11 e 12 devemos escolher o modelo HMM-3 para o índice Ibovespa no período analisado. Esta conclusão é coerente com o resultado empírico (no qual constatamos que o melhor modelo para o índice Ibovespa no período é o HMM-3/20d). A Figura 13 compara os desempenhos dos modelos HMM-3/20d (gráficos azuis, onde a parte superior representa a série de previsão de preços e a inferior os ganhos acumulados com o modelo neste período) e HMM-2/40d (gráficos vermelhos, onde a parte superior representa a série de previsão de preços e a inferior os ganhos acumulados com o modelo neste período). Apesar destes modelos apresentarem correlações similares em relação à série de preços do índice (0,8646 e 0,8603), as previsões acertadas nos dias 20, 23, 24, 29, 30, 34, 42 e 44 fizeram com que o modelo HMM-3/20d acumulasse uma riqueza consideravelmente maior (0,82% de ganho em relação à riqueza inicial) que a alcançada com o modelo HMM-2/40d ($-14,75\%$ de perda em relação à riqueza inicial).

CENÁRIO II : VALORIZAÇÃO

Modelo / Janela temporal	10 dias	20 dias	40 dias	80 dias
HMM-2	0,9302 1,01% 2,80%	0,9197 5,43% 3,07%	0,9354 7,88% 2,74%	0,9260 2,47% 3,13%
HMM-3	0,9258 8,07% 2,86%	0,9281 8,90% 2,91%	0,9349 6,36% 2,90%	0,9304 2,10% 2,98%
HMM-4	0,9227 18,02% 2,98%	0,9289 11,21% 2,86%	0,9265 2,43% 3,07%	0,9477 10,87% 2,54%
HMM-5	0,9258 13,95% 2,90%	0,9222 10,18% 2,85%	** ** **	** ** **
HMM-6	** ** **	** ** **	** ** **	** ** **

Tabela 10 – Desempenho do HMM com o índice Ibovespa num período de valorização.

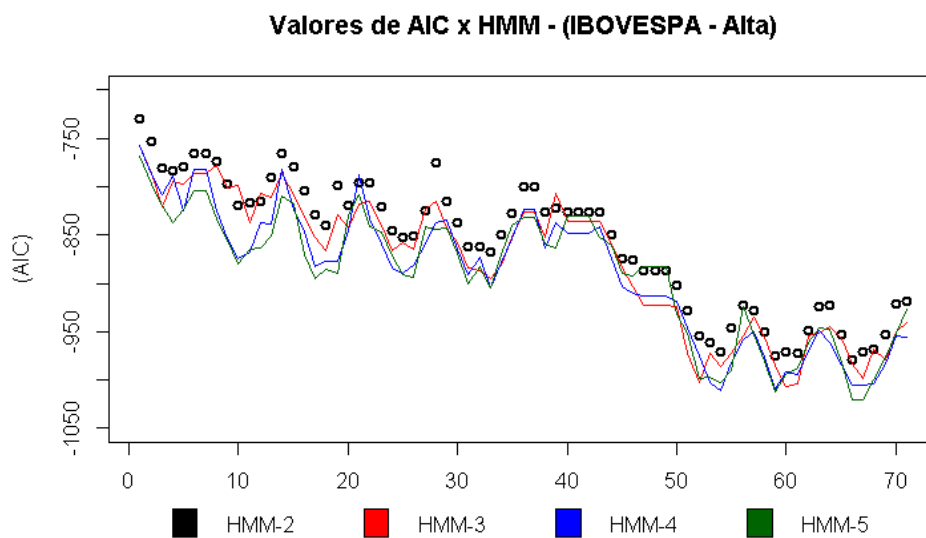


Figura 14 – Valores de AIC para cada HMM num período de valorização (alta).

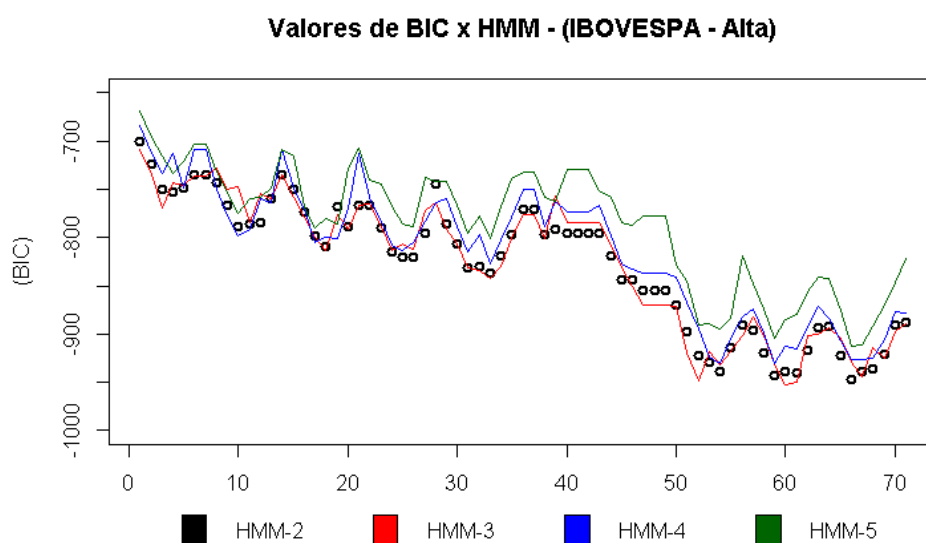


Figura 15 – Valores de BIC para cada HMM num período de valorização (alta).

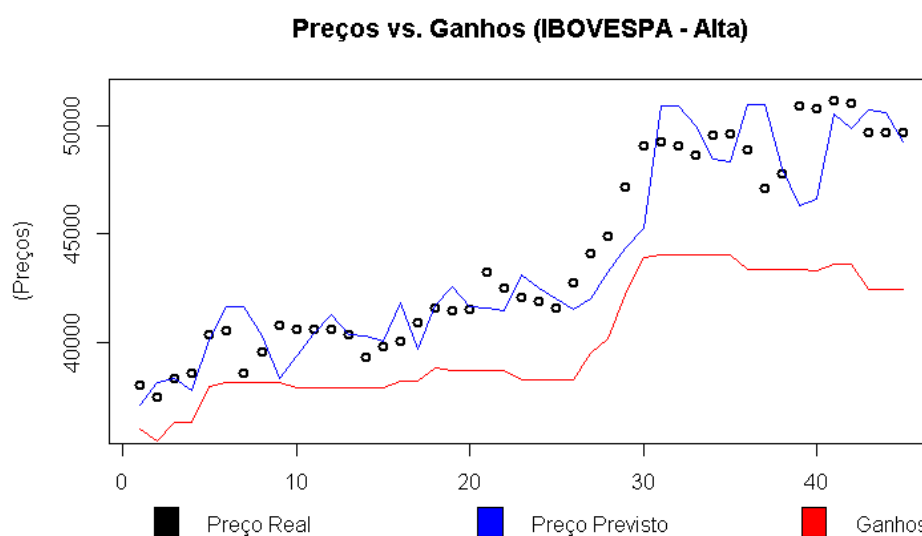


Figura 16 – Desempenho do modelo HMM-4/10d num período de valorização (alta).

(i) Período de avaliação: 25/01/2015 a 25/03/2015.

(ii) Valorização do índice no período (método *buy and hold*): 30,57%.

(iii) Modelo HMM com melhor desempenho no período: HMM-4/10d.

(iv) Rentabilidade do HMM com melhor desempenho (HMM-4/10d): 18,02%.

Observamos que o modelo HMM-4/10d não superou a estratégia *buy and hold*. Além disso, na Figura 14 notamos que os valores de AIC para os modelos HMM-2, HMM-3, HMM-4 e HMM-5 são em geral bem próximos. Salientamos que os valores de AIC para os modelos HMM-3, HMM-4 e HMM-5 estão num patamar levemente inferior ao valor de AIC para o modelo HMM-2. Já em relação ao BIC na Figura 15, observamos que os valores dos modelos HMM-2, HMM-3 e HMM-4 estão num patamar levemente inferior ao encontrado no HMM-5. Conforme já dito anteriormente, os modelos mais adequados são aqueles que apresentam os menores valores de AIC e BIC. Sendo assim, ao analisarmos em conjunto as Figuras 14 e 15 devemos escolher os modelos HMM-3 ou HMM-4 para o índice Ibovespa no período analisado. Esta conclusão é coerente com o resultado empírico (no qual constatamos que o melhor modelo para o índice Ibovespa no período é o HMM-4/10d; também observamos na Tabela 10 que nas janelas temporais de 10 e 20 dias os modelos HMM-3 e HMM-4 sempre superam os modelos HMM-2 e HMM-5). A Figura 16 mostra a evolução da riqueza (gráfico vermelho) do agente que utilizar o modelo HMM-4/10d para negociar o índice Ibovespa no período.

CENÁRIO III : ESTAGNAÇÃO

Modelo / Janela temporal	10 dias	20 dias	40 dias	80 dias
HMM-2	0,6831	0,7138	0,7221	0,6826
	-0,32%	-2,32%	-2,49%	-0,76%
	2,04%	1,89%	1,95%	2,18%
HMM-3	0,6508	0,7129	0,6572	0,6627
	1,57%	3,37%	-3,64%	-4,20%
	2,06%	1,85%	1,98%	1,97%
HMM-4	0,7394	0,6657	0,7036	0,7454
	-0,70%	0,10%	-1,49%	-2,61%
	1,66%	1,94%	1,84%	1,70%
HMM-5	0,6892	0,6813	0,6191	0,6431
	-3,21%	-3,64%	-7,67%	-10,55%
	2,04%	1,95%	2,07%	2,23%
HMM-6	**	**	**	**
	**	**	**	**
	**	**	**	**

Tabela 11 – Desempenho do HMM com o índice Ibovespa num período de estagnação.

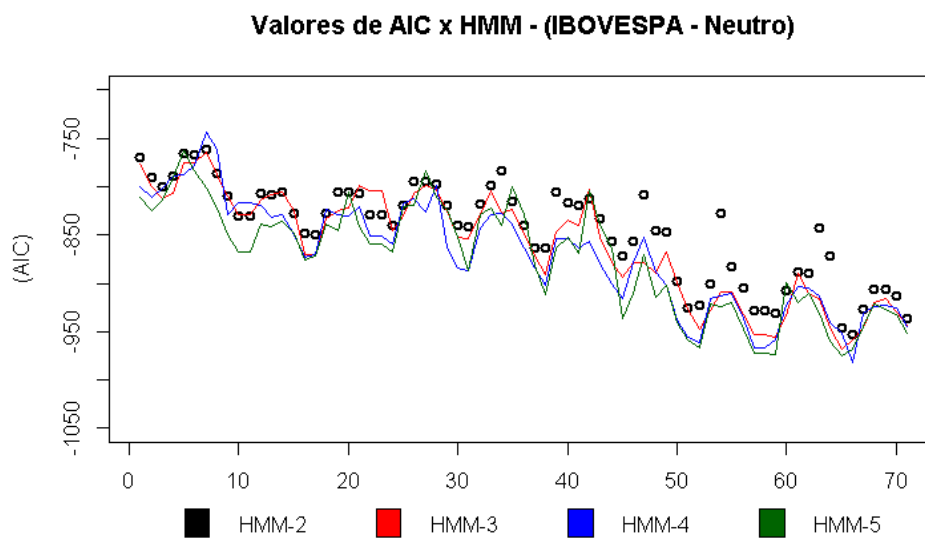


Figura 17 – Valores de AIC para cada HMM num período de estagnação (neutro).

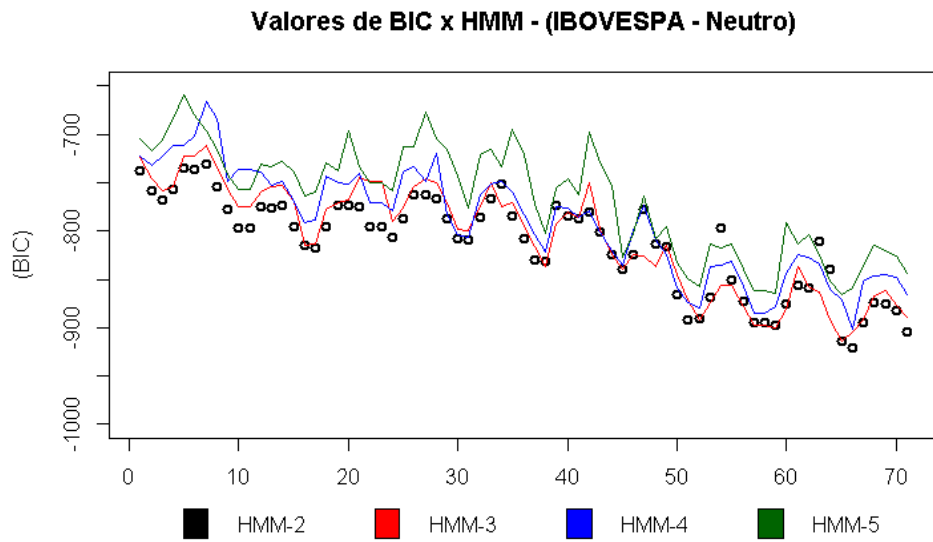


Figura 18 – Valores de BIC para cada HMM num período de estagnação (neutro).

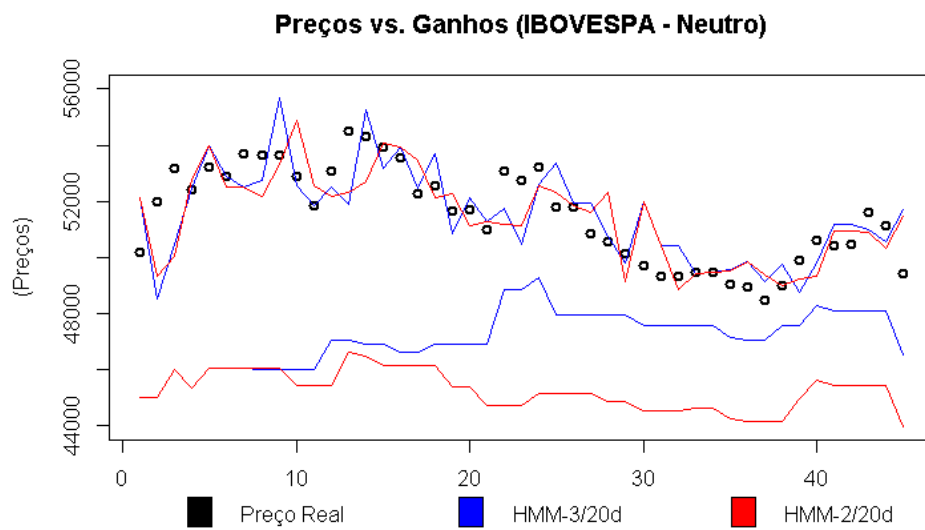


Figura 19 – Comparação entre HMM-3/20d e HMM-2/20d num período de estagnação (neutro).

(i) Período de avaliação: 11/04/2016 a 10/06/2016.

(ii) Valorização do índice no período (método *buy and hold*): $-1,48\%$.

(iii) Modelo HMM com melhor desempenho no período: HMM-3/20d.

(iv) Rentabilidade do HMM com melhor desempenho (HMM-3/20d): $3,37\%$.

Observamos que o modelo HMM-3/20d superou a estratégia *buy and hold*. Além disso, na Figura 17 notamos que os valores de AIC para os modelos HMM-2, HMM-3, HMM-4 e HMM-5 são em geral bem próximos. Salientamos que os valores de AIC para os modelos HMM-3, HMM-4 e HMM-5 estão num patamar levemente inferior ao valor de AIC para o modelo HMM-2. Já em relação ao BIC na Figura 18, observamos que os valores dos modelos HMM-2 e HMM-3 estão num patamar levemente inferior ao encontrado nos modelos HMM-4 e HMM-5. Conforme já dito anteriormente, os modelos mais adequados são aqueles que apresentam os menores valores de AIC e BIC. Sendo assim, ao analisarmos em conjunto as Figuras 17 e 18 devemos escolher o modelo HMM-3 para o índice Ibovespa no período analisado. Esta conclusão é coerente com o resultado empírico (no qual constatamos que o melhor modelo para o índice Ibovespa no período é o HMM-3/20d). A Figura 19 compara os desempenhos dos modelos HMM-3/20d (gráficos azuis, onde a parte superior representa a série de previsão de preços e a inferior os ganhos acumulados com o modelo neste período) e HMM-2/20d (gráficos vermelhos, onde a parte superior representa a série de previsão de preços e a inferior os ganhos acumulados com o modelo neste período). Apesar destes modelos apresentarem correlações similares em relação à série de preços do índice (0,7129 e 0,7138), as previsões acertadas nos dias 17, 18, 20 e 21 fizeram com que o modelo HMM-3/20d acumulasse uma riqueza maior (3,37% de ganho em relação à riqueza inicial) que a alcançada com o modelo HMM-2/20d ($-2,32\%$ de perda em relação à riqueza inicial).

7 CONCLUSÃO

O sistema de negociação apresentado neste trabalho (baseado no modelo HMM e nas Janelas Temporais Retrospectivas) obteve previsões dos preços diários das ações com erros percentuais absolutos médios que oscilaram entre 1,36% e 7,34%. Ou seja, o modelo de previsão apresentado nesta dissertação possui baixa precisão na estimação do valor do preço diário futuro dos ativos financeiros. Porém, em mais de 80% das situações analisadas neste trabalho a correlação entre a série de previsões do preço do modelo ideal e a série de preços da ação foi acima de 70%. Este fato nos sugere que o modelo HMM prevê de maneira razoável a *tendência* nos preços dos ativos financeiros. Tal afirmação foi comprovada empiricamente medindo-se a rentabilidade do sistema de negociação em diversas situações. Os resultados obtidos neste trabalho estão resumidos na tabela abaixo (as células preenchidas com ** representam as situações nas quais os métodos AIC e BIC não foram utilizados pois nestes casos o HMM só convergiu com 2 estados ocultos).

Ativo	Cenário	Melhor Modelo	AIC/BIC	Buy and Hold	Rentabilidade do Modelo
Ambev	baixa	HMM-3/80d	n=2 ou 3	-4,69%	+1,05%
Itau	baixa	HMM-2/20d	**	-15,80%	-3,49%
Petrobras	baixa	HMM-2/10d	**	-36,98%	-17,11%
Vale	baixa	HMM-3/20d	n=3	-15,18%	+3,02%
Ibovespa	baixa	HMM-3/20d	n=3	-15,12%	+0,82%
Ambev	alta	HMM-2/20d	**	+12,76%	+17,32%
Itau	alta	HMM-3/10d	n=2 ou 3	+32,78%	+33,36%
Petrobras	alta	HMM-2/10d	**	+77,09%	+55,83%
Vale	alta	HMM-2/10d	**	+66,67%	+16,86%
Ibovespa	alta	HMM-4/10d	n=3 ou 4	+30,57%	+18,02%
Ibovespa	neutro	HMM-3/20d	n=3	-1,48%	+3,37%

Tabela 12 – Resumo - Desempenho dos modelos

Nos cenários onde havia uma *tendência* de alta nos preços, o sistema de negociação sempre apresentou rentabilidade positiva. Porém, somente em duas das cinco situações estudadas (40%) o sistema de negociação superou a estratégia de *buy and hold* aplicadas aos ativos no período.

Já nos cenários onde havia uma *tendência* de queda nos preços a perspectiva de utilização do sistema de negociação é mais promissora. Nas cinco situações de desvalorização estudadas o modelo superou a estratégia de *buy and hold* (além disso, em três dos cinco cenários analisados o sistema de negociação apresentou uma rentabilidade positiva).

Também observamos que em cada cenário estudado o modelo mais bem su-

cedido conquistou esta condição ao prever melhor “microtendências” e dias singulares (datas nas quais os ativos obtiveram valorizações ou desvalorizações expressivas) no período em comparação com outros modelos. Salientamos que estes eventos tiveram um impacto significativo nos ganhos de cada modelo mas representaram menos de 20% do período de cada cenário. Assim, não somos capazes de excluir a hipótese de que os melhores modelos de cada cenário prevêm os dias singulares e as microtendências por simples acaso. Para afastarmos esta possibilidade, precisaríamos implementar testes estatísticos adequados. Porém, observamos três fatos nas tabelas e gráficos do capítulo de resultados que nos sugerem que os melhores modelos não foram determinados de maneira aleatória.

No primeiro fato verificamos que entre os 11 modelos mais adequados a cada cenário tivemos a seguinte frequência das Janelas Temporais Retrospectivas: 5 vezes para a janela de 10 dias, 5 vezes para a janela de 20 dias, nenhuma vez para a janela de 40 dias e uma vez para a janela de 80 dias. Sabemos que o tamanho desta amostra é inadequado para efetuarmos o teste qui-quadrado (para realizarmos adequadamente este teste precisamos de pelo menos 5 observações em cada categoria e, no caso de poucos grupos, pelo menos 10). Mas, se por hipótese tivéssemos analisado 220 cenários com a distribuição percentual das frequências das janelas parecida com a encontrada na amostra dos 11 modelos (exemplo: $f_{10dias} = 100$, $f_{20dias} = 100$, $f_{40dias} = 10$ e $f_{80dias} = 10$), iríamos através do teste qui-quadrado rejeitar a Hipótese de distribuição equiprovável das Janelas Temporais Retrospectivas do conjunto dos melhores modelos, sinalizando que os melhores sistemas de negociação de cada cenário não são determinados aleatoriamente.

Os demais fatos interessantes observados nos resultados são os seguintes:

(i): entre os 5 cenários de alta estudados neste trabalho, em 4 observamos que os melhores modelos eram compostos pela Janela Temporal Retrospectiva de 10 dias. Já nos 5 cenários de baixa a situação se inverte pois em 4 ocasiões observamos que os melhores modelos não tinham em sua composição a Janela Temporal Retrospectiva de 10 dias. Uma possível explicação para isso é que no período de alta os agentes do mercado estavam com um maior “apetite” ao risco fazendo com que as tendências do mercado fossem melhor captadas pelo modelo num horizonte temporal mais curto. Salientamos que no período de alta analisado neste trabalho (25/01/2016 a 25/03/2016) o mercado estava muito otimista com a perspectiva (que posteriormente se concretizou) da queda do governo de Dilma Rousseff pois o seu sucessor (Michel Temer) sinalizara para o mercado o compromisso em adotar políticas econômicas demandadas pelo setor empresarial (ajuste fiscal, reforma trabalhista e reforma da previdência).

(ii): os HMM’s com os números de estados ideais de cada cenário obtiveram nas Ja-

nelas Temporais Retrospectivas de 10 e 20 dias melhores desempenhos em 85% das vezes quando comparamos com a performance dos modelos que possuíam as mesmas janelas com um número de estados diferentes. Além disso, quando aplicamos os critérios de AIC e BIC para selecionar o número de estados ideal do modelo em determinado cenário, obtivemos resultados que foram empiricamente comprovados através das tabelas de desempenho.

Compreendemos que estes fatos também carecem de testes estatísticos adequados para a sua validação. Porém, sabemos que o indício é a circunstância indicativa da existência ou inexistência de determinado fato. Normalmente, a veracidade de um fato se faz com um conjunto de indícios apoiados uns nos outros, uns complementando e interagindo entre si dentro de uma estrutura lógica. Ou seja, diante dos fatos apresentados nesta dissertação, podemos sugerir que os desempenhos do sistema de negociação apresentados neste trabalho não foram obras do acaso.

Porém, um eventual uso do HMM como parte de uma estratégia de negociação depende de uma análise mais aprofundada das razões de seu desempenho nos mais diversos cenários. No escopo restrito deste trabalho, observamos que métodos simples baseados em HMM superam com folga o *benchmark* básico dado por estratégias *buy and hold* em cenários de desvalorização. Por outro lado, estas estratégias possuem um desempenho irregular em situações de alta dos preços. Diante disso, concluímos que este trabalho possui, em determinado aspecto, resultados numéricos similares aos encontrados por [Nguyen \(2017\)](#) (que testou o modelo HMM em 3 cenários de alta mas em somente um desses casos o modelo superou a estratégia *buy and hold*), porém, entendemos (em oposição à pesquisadora) que o HMM precisa ser mais estudado para conhecermos melhor as forças e fraquezas deste modelo num contexto de negociação de ações no mercado financeiro.

Referências

- AIUBE, F. A. L. **Modelos quantitativos em Finanças com enfoque em commodities**. [S.l.]: Bookman Editora, 2013. Citado 2 vezes nas páginas 14 e 25.
- CONT, R. Empirical properties of asset returns: stylized facts and statistical issues. Taylor & Francis, 2001. Citado na página 26.
- GRANGER, C.; DING, Z. Stylized facts on the temporal and distributional properties of daily data from speculative markets. 1994. Citado 2 vezes nas páginas 24 e 25.
- HAMILTON, J. D. A new approach to the economic analysis of nonstationary time series and the business cycle. **Econometrica: Journal of the Econometric Society**, JSTOR, p. 357–384, 1989. Citado na página 24.
- HASSAN, M. R.; NATH, B. Stock market forecasting using hidden markov model: a new approach. In: IEEE. **Intelligent Systems Design and Applications, 2005. ISDA'05. Proceedings. 5th International Conference on**. [S.l.], 2005. p. 192–196. Citado na página 30.
- IBE, O. **Markov processes for stochastic modeling**. [S.l.]: Newnes, 2013. Citado na página 2.
- KOSKI, T. **Hidden Markov models for bioinformatics**. [S.l.]: Springer Science & Business Media, 2001. v. 2. Citado na página 2.
- MOFFITT, S. D. **The Strategic Analysis of Financial Markets, Volume 1: Framework**. [S.l.]: World Scientific Series in Finance. World Scientific, 2017. Citado 5 vezes nas páginas 1, 14, 20, 21 e 23.
- NGUYEN, N. An analysis and implementation of the hidden markov model to technology stock prediction. **Risks**, Multidisciplinary Digital Publishing Institute, v. 5, n. 4, p. 62, 2017. Citado 7 vezes nas páginas 1, 24, 26, 27, 30, 32 e 58.
- NYSTRUP, P.; MADSEN, H.; LINDSTRÖM, E. Stylised facts of financial time series and hidden markov models in continuous time. **Quantitative Finance**, Taylor & Francis, v. 15, n. 9, p. 1531–1541, 2015. Citado na página 26.
- PETTERS, A. O.; DONG, X. et al. **An Introduction to Mathematical Finance with Applications**. [S.l.]: Springer, 2016. Citado na página 14.
- RYDÉN, T.; TERÄSVIRTA, T.; ÅSBRINK, S. Stylized facts of daily return series and the hidden markov model. **Journal of applied econometrics**, Wiley Online Library, v. 13, n. 3, p. 217–244, 1998. Citado 3 vezes nas páginas 1, 24 e 25.
- STAMP, M. **Introduction to Machine Learning with Applications in Information Security**. [S.l.]: Chapman and Hall/CRC, 2017. Citado na página 2.
- VISSER, I.; SPEEKENBRINK, M. et al. depmixs4: an r package for hidden markov models. **Journal of Statistical Software**, v. 36, n. 7, p. 1–21, 2010. Citado na página 32.

Apêndices

APÊNDICE A – CÓDIGO-FONTE EM R

Neste apêndice iremos apresentar o código-fonte em R desenvolvido durante este trabalho.

```
#####
```

```
#Ajuste de Datas e do tamanho da Janela Temporal Retrospectiva
```

```
library('depmixS4')
library('quantmod')
library('ggplot2')
library('lubridate')
```

```
data_intervalo_1 <- dmy(25012016)
data_intervalo_100 <- dmy(25032016)
```

```
data_intervalo_ref_1 <- dmy(25012016)
data_intervalo_ref_100 <- dmy(25032016)
```

```
for (intv in 1:61) {
  data_intervalo_1 <- data_intervalo_1 - ddays(1)
  data_intervalo_100 <- data_intervalo_100 - ddays(1)
}
```

```
data_intervalo_sub_inicio <- data_intervalo_1
data_intervalo_sub_fim <- data_intervalo_100
```

```
for (intv in 1:10) { #Marcação 10, 20, 40 e 80
  data_intervalo_sub_inicio <- data_intervalo_sub_inicio - ddays(1)
  data_intervalo_sub_fim <- data_intervalo_sub_fim - ddays(1)
}
```

```
data_1 <- data_intervalo_1
data_100 <- data_intervalo_100
```

```
data_S1 <- data_intervalo_1
```

```
data_E1 <- data_intervalo_100

data_intervalo_ref_1
data_intervalo_ref_100

data_intervalo_1
data_intervalo_100

data_intervalo_sub_inicio
data_intervalo_sub_fim

#####

#Execução do algoritmo HMM e cálculo da verossimilhança

data_intervalo_subR_inicio <- data_intervalo_sub_inicio
data_intervalo_subR_fim <- data_intervalo_sub_fim

preco_prev_vet <- seq(1,61,1)

preco_verdade_vet <- seq(1,61,1)

Simul_AIC <- seq(1,71,1)      #Marcação 71, 81, 101 e 141
V_AIC2 <- seq(1,71,1)       #Marcação 71, 81, 101 e 141
V_BIC2 <- seq(1,71,1)       #Marcação 71, 81, 101 e 141
V_LogLik2 <- seq(1,71,1)    #Marcação 71, 81, 101 e 141

for (j in 1:71) {          #Marcação 71, 81, 101 e 141

print(j)

getSymbols( "^BVSP", from=data_intervalo_sub_inicio,
to=data_intervalo_sub_fim)

BVSPRets_op = diff( log( Op( BVSP ) ) )
returns_op = as.numeric(BVSPRets_op)

BVSPRets_cl = diff( log( Cl( BVSP ) ) )
returns_cl = as.numeric(BVSPRets_cl)
```

```
BVSPRets_min = diff( log( Lo( BVSP ) ) )
returns_min = as.numeric(BVSPRets_min)

BVSPRets_max = diff( log( Hi( BVSP ) ) )
returns_max = as.numeric(BVSPRets_max)

returns_op <- na.exclude(returns_op)
returns_cl <- na.exclude(returns_cl)
returns_min <- na.exclude(returns_min)
returns_max <- na.exclude(returns_max)

HMMModel <- data.frame(returns_op, returns_min, returns_max, returns_cl)

head(HMMModel)

colnames(HMMModel) <- c("Abertura","Mínimo","Máximo","Fechamento")

colnames(HMMModel)

HMM <- depmixS4::dpmix(list(Abertura~1,Mínimo~1,Máximo~1,Fechamento~1),
data = HMMModel,nstates=2, family=list(gaussian(),gaussian(),
gaussian(), gaussian()))
set.seed(3)
HMMfit <- depmixS4::fit(HMM, verbose = FALSE)

#HMM3 <- depmixS4::dpmix(list(Abertura~1,Mínimo~1,Máximo~1,Fechamento~1),
data = HMMModel,nstates=3, family=list(gaussian(),gaussian(),
gaussian(), gaussian()))
#set.seed(2)
#HMMfit3 <- depmixS4::fit(HMM3, verbose = FALSE)

HMMstate <- depmixS4::posterior(HMMfit)

#HMMstate3 <- depmixS4::posterior(HMMfit3)

Simul_AIC[j] <- j
V_AIC2[j] <- AIC(HMMfit)
#V_AIC3[j] <- AIC(HMMfit3)
```



```
V_BIC2[j] <- BIC(HMMfit)
#V_BIC3[j] <- BIC(HMMfit3)
V_LogLik2[j] <- logLik(HMMfit)[1]
#V_LogLik3[j] <- logLik(HMMfit3)[1]

data_intervalo_sub_inicio <- data_intervalo_sub_inicio + ddays(1)
data_intervalo_sub_fim <- data_intervalo_sub_fim + ddays(1)

}

data_intervalo_sub_inicio <- data_intervalo_subR_inicio
data_intervalo_sub_fim <- data_intervalo_subR_fim

#####

#Execução do algoritmo HMM e previsão do preço no 1º dia do intervalo

y <- 1

print(y)

getSymbols( "^BVSP", from=data_S1, to=data_E1)

BVSPrets_op = diff( log( Op( BVSP ) ) )
returns_op = as.numeric(BVSPrets_op)

BVSPrets_cl = diff( log( Cl( BVSP ) ) )
returns_cl = as.numeric(BVSPrets_cl)

BVSPrets_min = diff( log( Lo( BVSP ) ) )
returns_min = as.numeric(BVSPrets_min)

BVSPrets_max = diff( log( Hi( BVSP ) ) )
returns_max = as.numeric(BVSPrets_max)

returns_op <- na.exclude(returns_op)
returns_cl <- na.exclude(returns_cl)
returns_min <- na.exclude(returns_min)
returns_max <- na.exclude(returns_max)
```

```
HMMModel <- data.frame(returns_op, returns_min, returns_max, returns_cl)

head(HMMModel)

colnames(HMMModel) <- c("Abertura","Mínimo","Máximo","Fechamento")

colnames(HMMModel)

HMM <- depmixS4::dempix(list(Abertura~1,Mínimo~1,Máximo~1,Fechamento~1),
data = HMMModel,nstates=2, family=list(gaussian(),gaussian(),
gaussian(), gaussian()))
set.seed(3)
HMMfit <- depmixS4::fit(HMM, verbose = FALSE)

#HMM3 <- depmixS4::dempix(list(Abertura~1,Mínimo~1,Máximo~1,Fechamento~1),
data = HMMModel,nstates=3, family=list(gaussian(),gaussian(),
gaussian(), gaussian()))
#set.seed(2)
#HMMfit3 <- depmixS4::fit(HMM3, verbose = FALSE)

HMMstate <- depmixS4::posterior(HMMfit)

#HMMstate3 <- depmixS4::posterior(HMMfit3)

head(HMMstate)

#head(HMMstate3)

post_probs <- depmixS4::posterior(HMMfit)

#post_probs <- depmixS4::posterior(HMMfit3)

D <- abs(V_LogLik2[11]-V_LogLik2[10])
r <- 10

print(V_LogLik2[10])

for (k in 10:1) {
```

```
if ( abs(V_LogLik2[11]-V_LogLik2[k]) < D ) {
  r <- k
  D <- abs(V_LogLik2[11]-V_LogLik2[k])
}
#print(k)
#print(V_LogLik2[k])
#print( abs(V_LogLik2[11]-V_LogLik2[k]) )
}

print(r)
print(D)

print(V_LogLik2[r])

v <- r-1

data_vero_1 <- data_intervalo_sub_inicio + ddays(v)
data_vero_100 <- data_intervalo_sub_fim + ddays(v)

p <- v+1

data_prev_1 <- data_intervalo_sub_inicio + ddays(p)
data_prev_100 <- data_intervalo_sub_fim + ddays(p)

data_vero_1
data_vero_100

data_prev_1
data_prev_100

getSymbols( "^BVSP", from=data_prev_1, to=data_prev_100 + ddays(1))

BVSPrets_prev_cl = periodReturn(BVSP,period='daily')
returns_prev_cl = as.numeric(BVSPrets_prev_cl)

tamanho_cl <- length(returns_prev_cl)

retorno_prev <- returns_prev_cl[tamanho_cl]
```

```
getSymbols( "^BVSP", from=data_S1, to=data_E1)

precos_fech = Cl( BVSP )
precos_fech = as.numeric(precos_fech)
precos_fech <- na.exclude(precos_fech)

tamanho <- length(precos_fech)

ultimo_preco <- precos_fech[tamanho]

preco_prev <- ultimo_preco*(1+retorno_prev)

data_E1_mais1 <- data_E1 + ddays(1)

getSymbols( "^BVSP", from=data_S1, to=data_E1_mais1 + ddays(1))

precos_fech = Cl( BVSP )
precos_fech = as.numeric(precos_fech)
precos_fech <- na.exclude(precos_fech)

tamanho2 <- length(precos_fech)

ultimo_preco <- precos_fech[tamanho2]

preco_verdade <- ultimo_preco

preco_prev

preco_verdade

preco_prev_vet[y] <- preco_prev

preco_verdade_vet[y] <- preco_verdade

data_S1 <- data_1 + ddays(1)
data_E1 <- data_100 + ddays(1)

data_1 <- data_1 + ddays(1)
```

```
data_100 <- data_100 + ddays(1)

data_intervalo_sub_inicio <- data_intervalo_sub_inicio + ddays(1)
data_intervalo_sub_fim <- data_intervalo_sub_fim + ddays(1)

#####

#Execução do algoritmo HMM e previsão dos preços no restante do intervalo

for (i in 1:60) {

  z <- i+1

  print(i)

  getSymbols( "^BVSP", from=data_S1, to=data_E1)

  BVSPrets_op = diff( log( Op( BVSP ) ) )
  returns_op = as.numeric(BVSPrets_op)

  BVSPrets_cl = diff( log( Cl( BVSP ) ) )
  returns_cl = as.numeric(BVSPrets_cl)

  BVSPrets_min = diff( log( Lo( BVSP ) ) )
  returns_min = as.numeric(BVSPrets_min)

  BVSPrets_max = diff( log( Hi( BVSP ) ) )
  returns_max = as.numeric(BVSPrets_max)

  returns_op <- na.exclude(returns_op)
  returns_cl <- na.exclude(returns_cl)
  returns_min <- na.exclude(returns_min)
  returns_max <- na.exclude(returns_max)

  HMMModel <- data.frame(returns_op, returns_min, returns_max, returns_cl)

  head(HMMModel)

  colnames(HMMModel) <- c("Abertura","Mínimo","Máximo","Fechamento")
```

```
colnames(HMMModel)

HMM <- depmixS4::dpmix(list(Abertura~1,Mínimo~1,Máximo~1,Fechamento~1),
  data = HMMModel,nstates=2, family=list(gaussian(),gaussian(),
  gaussian(), gaussian()))
set.seed(3)
HMMfit <- depmixS4::fit(HMM, verbose = FALSE)

#HMM3 <- depmixS4::dpmix(list(Abertura~1,Mínimo~1,Máximo~1,Fechamento~1),
  data = HMMModel,nstates=3, family=list(gaussian(),gaussian(),
  gaussian(), gaussian()))
#set.seed(2)
#HMMfit3 <- depmixS4::fit(HMM3, verbose = FALSE)

HMMstate <- depmixS4::posterior(HMMfit)

#HMMstate3 <- depmixS4::posterior(HMMfit3)

head(HMMstate)

#head(HMMstate3)

post_probs <- depmixS4::posterior(HMMfit)

#post_probs <- depmixS4::posterior(HMMfit3)

q1 <- i+11          #Marcação 11, 21, 41 e 81
q2 <- q1-1
q3 <- i+1

D <- abs(V_LogLik2[q1]-V_LogLik2[q2])
r <- q2

print(V_LogLik2[q2])

for (k in q2:q3) {

if ( abs(V_LogLik2[q1]-V_LogLik2[k]) < D ) {
```

```
r <- k
D <- abs(V_LogLik2[q1]-V_LogLik2[k])
}
#print(k)
#print(V_LogLik2[k])
#print( abs(V_LogLik2[q1]-V_LogLik2[k]) )
}

print(r)
print(D)

print(V_LogLik2[r])

v <- r-1

data_vero_1 <- data_intervalo_sub_inicio + ddays(v)
data_vero_100 <- data_intervalo_sub_fim + ddays(v)

p <- v+1

data_prev_1 <- data_intervalo_sub_inicio + ddays(p)
data_prev_100 <- data_intervalo_sub_fim + ddays(p)

data_vero_1
data_vero_100

data_prev_1
data_prev_100

getSymbols( "^BVSP", from=data_prev_1, to=data_prev_100 + ddays(1))

BVSPrets_prev_cl = periodReturn(BVSP,period='daily')
returns_prev_cl = as.numeric(BVSPrets_prev_cl)

tamanho_cl <- length(returns_prev_cl)

retorno_prev <- returns_prev_cl[tamanho_cl]

getSymbols( "^BVSP", from=data_S1, to=data_E1)
```

```
precos_fech = Cl( BVSP )
precos_fech = as.numeric(precos_fech)
precos_fech <- na.exclude(precos_fech)

tamanho <- length(precos_fech)

ultimo_preco <- precos_fech[tamanho]

preco_prev <- ultimo_preco*(1+retorno_prev)

data_E1_mais1 <- data_E1+ddays(1)

getSymbols( "^BVSP", from=data_S1, to=data_E1_mais1 + ddays(1))

precos_fech = Cl( BVSP )
precos_fech = as.numeric(precos_fech)
precos_fech <- na.exclude(precos_fech)

tamanho2 <- length(precos_fech)

ultimo_preco <- precos_fech[tamanho2]

preco_verdade <- ultimo_preco

preco_prev

preco_verdade

preco_prev_vet[z] <- preco_prev

preco_verdade_vet[z] <- preco_verdade

data_S1 <- data_1 + ddays(1)
data_E1 <- data_100 + ddays(1)

data_1 <- data_1 + ddays(1)
data_100 <- data_100 + ddays(1)
```



```
data_intervalo_sub_inicio <- data_intervalo_sub_inicio + ddays(1)
data_intervalo_sub_fim <- data_intervalo_sub_fim + ddays(1)
}
```

```
#####
```

```
#Correlação entre os preços reais e os preços previstos
```

```
for (cont in 1:8) {
  aux <- (7*cont)-1
  preco_verdade_vet[aux] <- NA
  preco_prev_vet[aux] <- NA
}
```

```
for (cont2 in 1:8) {
  aux <- 7*cont2
  preco_verdade_vet[aux] <- NA
  preco_prev_vet[aux] <- NA
}
```

```
preco_verdade_vet <- na.excluíde(preco_verdade_vet)
preco_prev_vet <- na.excluíde(preco_prev_vet)
```

```
plot(preco_prev_vet, lwd=2, xlab='', ylab='(Preços)',
      main='Preço Previsto (Preto) vs. Preço Verdadeiro (Vermelho)',
      ylim=c(44000,55000))
lines(preco_verdade_vet, col='red', ylim=c(44000,55000))
```

```
cor(preco_verdade_vet,preco_prev_vet)
```

```
cor.test(preco_verdade_vet,preco_prev_vet)
```

```
preco_prev_vet2 <- preco_prev_vet
```

```
#####
```

```
#Evolução da Riqueza ao longo do intervalo
```

```
tam <- length(preco_prev_vet2)
```

```
riq2_vet <- seq(1,tam,1)

tam2 <- tam-1

money <- 150

riq2 <- money

riq2_vet[1] <- riq2

for (tr in 1:tam2) {
  if ( preco_prev_vet2[tr+1] > preco_prev_vet2[tr] ) {
    num <- (preco_verdade_vet[tr+1]-preco_verdade_vet[tr])
    ret_stock <- (num/preco_verdade_vet[tr])
    riq2 <- riq2*(1+ret_stock)
  }
  riq2_vet[tr+1] <- riq2
}

riq2

gperc2 <- (riq2-money)/money

num <- (preco_verdade_vet[tam]-preco_verdade_vet[1])
retfinal_stock <- (num/preco_verdade_vet[1])

gperc2

retfinal_stock

#plot(preco_verdade_vet, lwd=2, xlab='', ylab='(Preços)',
main='Preço Previsto (Azul) vs. Preço Verdadeiro (Preto)',
ylim=c(140,200))
#lines(preco_prev_vet, col='blue', ylim=c(140,200))
#lines(riq2_vet, col='red', ylim=c(140,200))

#plot(riq2_vet, lwd=2, xlab='', ylab='(Riqueza)',
main='Riqueza Acumulada', ylim=c(750,1050))
```

```
#lines(riq2_vet, col='red', ylim=c(750,1050))

#####

#Cálculo do erro percentual absoluto médio

M4 <- 0

for (k in 1:45) {

M1 <- abs(preco_verdade_vet[k]-preco_prev_vet[k])
M2 <- preco_verdade_vet[k]
M3 <- M1/M2
M4 <- M4 + M3
}

M5 <- M4/45

M5
```