

INSTITUTO NACIONAL DE MATEMÁTICA  
PURA E APLICADA

PROJETO FINAL DE CURSO

---

**Modelos de Arbitragem  
Estatística para o Mercado  
Acionário Brasileiro**

---

*Autor:*

Rafael Pires Ferreira  
Gonçalves

*Orientador:*

Sergei Vieira

11 de agosto de 2011

# Agradecimentos

Agradeço primeiramente a Deus por guiar meus caminhos e me dar saúde para poder seguir minha trajetória.

A toda minha família que sempre se sacrificou e torceu por mim.

Também agradeço ao Jorge Zubelli, que me ajudou muito durante todo o curso, e ao meu orientador Sergei Vieira com suas sugestões ao meu trabalho.

Aos meus amigos de turma do IMPA, Leonardo, Daniel, Henrique, Arthur, Fábio e Maristela agradeço a amizade construída nestes últimos anos.

Aos amigos de outros cursos do IMPA, Guillermo, Wanderson, Tertuliano, Marcelo, Juan Pablo, Élsio, Vinícius pela ajuda e apoio.

# Resumo

Os mercados financeiros, mesmo nos países mais desenvolvidos, até hoje não são completamente eficientes, permitindo dessa forma arbitragem entre os ativos. Um dos métodos mais populares de explorar esse evento é através de operações do tipo arbitragem estatística, no qual o processo de investimento está todo baseado em modelos matemáticos. A estratégia baseia-se na compra/venda de ativos toda vez que o diferencial de preços entre eles distorcer muito.

Os ativos usados no presente trabalho foram ações de empresas listadas na Bolsa de Valores de São Paulo, sendo que na montagem dos pares de ativos optou-se por utilizar sempre pares de ativos de um mesmo setor. O método usado para definir a escolha dos pares foi o de análise por componentes principais, que busca encontrar quais são os componentes principais responsáveis por explicar a maior parte da variação de todo o grupo de dados.

Dado o comportamento aleatório das séries financeiras, o modelo linear de reversão à média de Ornstein-Uhlenbeck foi o escolhido para tentar descrever o comportamento das variáveis financeiras. Adicionalmente foi incorporado no modelo o componente de aversão a risco dos agentes, tendo-se obtido dessa forma o portfólio ótimo a partir da maximização da função utilidade do agente. Ainda na parte final do trabalho foi testado um modelo não-linear com tangente hiperbólica que leva em consideração o fato do investidor ficar mais cético para valores muito distantes da média do diferencial de preços.

Além disso, foram testados também a implementação do modelo para dados de alta frequência, assim como a comparação dos resultados dentre os diversos setores de ações, e uma análise da relação risco retorno obtida para cada caso.

**Palavras-chave:** Arbitragem estatística, análise por componentes principais, Ornstein-Uhlenbeck

# Abstract

Financial markets, even in more developed countries, are still not fully efficient, thus allowing arbitrage between assets. One of the most popular methods to explore this is through operations such as statistical arbitrage, in which the entire investment process is based on mathematical models. The strategy is based on purchase/sale of assets on each time the price differential between them distorts much.

The assets used in this study were stocks listed on the São Paulo Stock Exchange and the assembly of the assets pairs used, correspond to companies of the same industry. The method used to define the choice of pairs was the principal component analysis, which seeks to find the main components responsible for explaining most of the variation of the data group.

Given the random behavior of financial series, the linear mean reverting Ornstein-Uhlenbeck model was chosen to try to describe the behavior of financial variables. Furthermore, it was added to the model an agent risk aversion parameter, thus obtaining the optimal portfolio that maximizes the agent utility function. At the end of this study, it was tested a nonlinear model with hyperbolic tangent that takes into account the fact that investors become more skeptical as mispricing becomes too large.

Also, it was tested the implementation of the model for high frequency data, as well as a comparison of results among the various stock sectors, and an analysis of the risk return obtained for each case.

**Keywords:** Statistical arbitrage, principal component analysis, Ornstein-Uhlenbeck

# Conteúdo

<b>1</b>	<b>Revisão dos Conceitos Matemáticos Utilizados</b>	<b>11</b>
1.1	Filtração . . . . .	11
1.2	Processo Estocástico . . . . .	11
1.3	Movimento Browniano . . . . .	11
1.4	Filtro Browniano . . . . .	12
1.5	Martingal . . . . .	12
1.6	Integral Estocástica . . . . .	13
1.7	Fórmula de Itô . . . . .	13
1.8	Regra do Produto . . . . .	14
1.9	Equação de Hamilton-Jacobi-Bellman (HJB) . . . . .	14
<b>2</b>	<b>Análise por Componentes Principais</b>	<b>18</b>
2.1	Descrição da Análise por Componentes Principais . . . . .	18
2.2	Descrição dos Dados Utilizados . . . . .	21
2.3	Interpretação das Componentes Principais . . . . .	22
2.4	Análise da Série de Dados . . . . .	24
<b>3</b>	<b>Modelo Linear de Reversão à Média</b>	<b>30</b>
3.1	Descrição do Processo . . . . .	30
3.2	Calibração do Resíduo de Cointegração . . . . .	31
3.3	Geração dos Sinais . . . . .	35
3.4	Ajustando o sinal pelo Volume Diário . . . . .	40
3.5	Restringindo o Número de Pares . . . . .	41
3.6	Trabalhando com Dados de Alta Frequência . . . . .	42
<b>4</b>	<b>Escolha do Portfólio dado um Modelo com Aversão ao Risco</b>	<b>47</b>
4.1	Modelando as Preferências dos Investidores . . . . .	47
4.2	Escolha do Portfólio Ótimo . . . . .	48
4.3	Características da Demanda de <i>Hedging</i> Intertemporal . . . . .	55
4.4	Principais Premissas do Modelo Não-Linear de Reversão à Média com Tangente Hiperbólica . . . . .	61
4.5	Contrastando com a dinâmica do OU . . . . .	62
4.6	Comparação dos Resultados: $N_{OU}$ versus $N_{TANH}$ . . . . .	63
<b>5</b>	<b>Conclusão</b>	<b>69</b>

<b>A</b>	<b>Resolução de equações diferenciais</b>	<b>74</b>
<b>B</b>	<b>Códigos do MatLab</b>	<b>82</b>
B.1	Componetes Principais . . . . .	82
B.2	Construção dos Sinais . . . . .	83
B.3	Calibração do Modelo de Reversão à Média . . . . .	85
B.4	Escolha do Portfólio Ótimo Segundo o Modelo com Tangente Hiperbólica . . . . .	85

# Lista de Figuras

2.1	Primeiro autovetor ordenado pelo tamanho do coeficiente. . . . .	23
2.2	Segundo autovetor ordenado pelo tamanho do coeficiente. . . . .	23
2.3	Terceiro autovetor ordenado pelo tamanho do coeficiente. . . . .	24
2.4	Autovalores da matriz de correlação dos retornos das ações. . . . .	25
2.5	Percentual da variância explicada pelos maior autovetor. . . . .	26
3.1	O <i>spread</i> , os sinais e o <i>payoff</i> , respectivamente, para BBDC4 x ITUB4. . . . .	37
3.2	O <i>spread</i> , os sinais e o <i>payoff</i> , respectivamente, para ITUB4 x GGBR4. . . . .	38
3.3	Resultados considerando diversos valores para o custo de corretagem. . . . .	39
3.4	<i>Payoffs</i> dos diversos portfólios comparados com o Ibovespa. . . . .	40
3.5	<i>Payoffs</i> dos portfólios com e sem ajuste pelo volume diário. . . . .	41
3.6	<i>Payoffs</i> dos diversos portfólios, com ajuste pelo volume diário, comparados com o Ibovespa. . . . .	42
3.7	<i>Payoffs</i> dos portfólios Top 10, Top 15 e Top 20. . . . .	44
3.8	<i>Payoffs</i> para um intervalo de cinco minutos. . . . .	44
3.9	<i>Payoffs</i> para diferentes intervalos de tempo para BBDC4 x ITUB4. . . . .	45
3.10	<i>Payoffs</i> para diferentes intervalos de tempo para CSNA3 x GGBR4. . . . .	46
4.1	Demanda total, $N$ , e evolução da diferença de sinais entre o fator míope e a demanda de <i>hedging</i> intertemporal para CYRE3 x RSID3, com $\gamma = 0,5$ . . . . .	56
4.2	Demanda total, $N$ , e evolução da diferença de sinais entre o fator míope e a demanda de <i>hedging</i> intertemporal para CYRE3 x RSID3, com $\gamma = 1,5$ . . . . .	56
4.3	<i>Payoffs</i> dos setores para o caso $\gamma = 1,5$ e $\gamma = 0,5$ . . . . .	57
4.4	<i>Payoffs</i> dos diversos portfólios comparados com o Ibovespa, com ajuste de 500 dias. . . . .	58
4.5	<i>Payoffs</i> dos diversos portfólios comparados com o Ibovespa, com ajuste de 700 dias. . . . .	59

4.6	<i>Payoffs</i> dos diversos portfólios comparados com o Ibovespa, com ajuste de 1000 dias. . . . .	59
4.7	Demanda total, $N$ , para BBDC4 x ITUB4 ( $k$ elevado) e CYRE3 x RSID3 ( $k$ moderado), com ajuste de 700 dias e $\gamma = 1,5$ . . . . .	60
4.8	<i>Payoffs</i> considerando custos de corretagem e alavancagem, com ajuste de 500 dias e $\gamma = 0,5$ . . . . .	60
4.9	Demanda total, $N$ , para os diversos modelos. . . . .	62
4.10	Demanda total, $N$ , para o modelo de OU com $\gamma = 0,5$ e da tangente hiperbólica para BBDC4 x ITUB4 ( $k$ elevado) e CYRE3 x RSID3 ( $k$ moderado). . . . .	66
4.11	<i>Payoffs</i> dos diversos portfólios comparados com o Ibovespa, com ajuste de 500 dias. . . . .	67
4.12	<i>Payoffs</i> dos diversos portfólios comparados com o Ibovespa, com ajuste de 700 dias. . . . .	67
4.13	<i>Payoffs</i> dos diversos portfólios comparados com o Ibovespa, com ajuste de 1000 dias. . . . .	68



# Lista de Tabelas

2.1	Relação de todas as empresas estudadas ao longo do trabalho separadas por setores. . . . .	27
2.2	Ativos ordenados por ordem decrescente de coeficiente na primeira componente principal. . . . .	28
2.3	Ativos ordenados por ordem decrescente de coeficiente na segunda componente principal. . . . .	28
2.4	Ativos ordenados por ordem decrescente de coeficiente na terceira componente principal. . . . .	29
2.5	Ativos ordenados por ordem decrescente de coeficiente na primeira componente principal, com os respectivos segmentos de atuação. . . . .	29
3.1	Comparação dos resultados com e sem ajuste pelo volume diário. . . . .	43
4.1	Comparação dos resultados para diferentes valores de $\gamma$ , com ajuste de 700 dias. . . . .	57
4.2	Comparação dos resultados do OU com $\gamma = 0,5$ e da tangente hiperbólica, com ajuste de 700 dias. . . . .	68

# Introdução

O termo arbitragem estatística compreende estratégias nas quais buscam-se obter lucro a partir de anomalias no diferencial de preços entre pares de ativos [25]. Nesse tipo de estratégia, o processo de investimento está todo baseado em modelos matemáticos, excluindo dessa maneira uma análise qualitativa com relação aos fundamentos econômicos por parte dos ativos envolvidos [3].

A estratégia consiste em assumir uma posição comprada em um ativo e vendida no outro, toda vez que o desvio de preços alcançar valores muito distantes da média. Quanto maior a magnitude do desvio de preços, maior a probabilidade deste corrigir e convergir para a média e conseqüentemente maior a possibilidade de ganhos [27]. Embora, não seja possível prever quando o desvio vai convergir, ou, até mesmo se este não vai piorar antes de melhorar [16].

A enorme liquidez do mercado acionário, o que reduz custos de transação e facilita a implementação das posições, é o local preferido para as operações de arbitragem estatística. Nesse caso, a escolha dos pares de ativos pode aparecer sob o formato de pares de ações, ações versus índice ou carteiras de ações. As estratégias envolvendo pares de ações, mais conhecidas como *pairs trading*, podem envolver ações do mesmo setor (intra-setorial) ou ações de setores distintos (inter-setorial).

Na escolha dos pares de ações a serem analisados os métodos mais usuais são os de correlação e cointegração, como apresentado em [9] e [24]. O uso da cointegração na construção de estratégias com ações também pode ser encontrado em [1]. Outro método bastante utilizado no estudo dos ativos financeiros é o de análise por componentes principais [21]. Seu objetivo é a obtenção de um número pequeno de combinações lineares a fim de transformar as variáveis originais que estavam correlacionadas entre si, em um novo conjunto de variáveis não correlacionadas. Eliminando assim parte da informação redundante e facilitando o surgimento da informação de boa qualidade [15].

O mercado acionário é composto por séries temporais que usualmente apresentam um comportamento de passeio aleatório. Trabalhos como o de Gatevet *al* [14] usaram uma abordagem baseada na menor distância entre os pares de ativos para explorar anomalias nesse processo. No trabalho de Baronyan [5] foi analisado o comportamento do diferencial de preços usando para isso diferentes regimes de volatilidade. O modelo linear de reversão à

média de Ornstein-Uhlenbeck, atualmente, é o mais utilizado para descrever a dinâmica das variáveis financeiras, com excelentes resultados encontrados no mercado americano, como exposto em [3].

Uma característica ausente nos trabalhos citados anteriormente é a análise do grau de aversão ao risco dos agentes. Esse é um tópico importante dado que todos os investidores são em algum nível avessos ao risco, o que os distingue é o seu grau de aversão. Nesse sentido, quanto menor a aversão do investidor ao risco, maior a propensão de buscar operações que tenham maiores riscos desde que sejam recompensados devidamente na forma de ganhos.

No presente trabalho, a partir da definição dos pares de ativos pela análise dos componentes principais, implementou-se a metodologia proposta por Avellaneda *et al* [3] para análise do comportamento do *spread*, usando para isso dados do mercado acionário brasileiro. Posteriormente foram testadas algumas mudanças na implementação do modelo proposto, assim como foram utilizados dados de alta frequência.

Para incorporar o grau de aversão ao risco no comportamento dos preços foi obtido o portfólio ótimo a partir da maximização da função utilidade do agente no modelo de Ornstein-Uhlenbeck, conforme proposto por Jurek *et al* [16]. Na parte final do trabalho foi testado um modelo não-linear com tangente hiperbólica, apresentado no trabalho de Alsayed e McGroarty [2], que incorpora o fato de que o investidor fica mais cético para valores muito distantes da média do diferencial de preços, optando dessa maneira por assumir posições menores quando comparadas com o modelo de Ornstein-Uhlenbeck.

Dessa maneira, o trabalho está estruturado da seguinte forma. No Capítulo 1 fazemos uma breve revisão dos principais conceitos matemáticos utilizados ao longo do trabalho. No Capítulo 2 demonstramos como foi feita a escolha dos pares de ativos através da análise dos componentes principais. No Capítulo 3 apresentamos o modelo linear de Ornstein-Uhlenbeck de reversão à média, e apresentamos um método para aproveitar oportunidades de arbitragem baseados em sinais gerados pelo diferencial de preços entre os pares de ações.

No Capítulo 4 analisamos o comportamento do investidor frente ao risco dado um modelo linear de aversão ao risco, e mostramos uma estratégia de arbitragem baseado no estudo do portfólio ótimo. Consideramos ainda o caso do investidor em um modelo não-linear com tangente hiperbólica, e comparamos as duas estratégias em termos de resultados. Finalmente no Capítulo 5 apresentamos a conclusão do trabalho.

# Capítulo 1

## Revisão dos Conceitos Matemáticos Utilizados

Este capítulo apresenta um resumo de alguns dos conceitos matemáticos que foram utilizados para provar resultados dos capítulos posteriores. Mais detalhes e as demonstrações dos resultados encontrados abaixo podem ser encontrados em [4] e [13].

### 1.1 Filtração

Dados um conjunto não-vazio  $\Omega$ , um número positivo  $T$  e uma família  $\{\mathcal{F}(t)\}_{t \in [0, T]}$  de  $\sigma$ -álgebras de  $\Omega$ . Chamamos essa família de filtração, se para todo par de números  $t, s$ , com  $0 \leq s \leq t \leq T$ , se cumpre que  $\mathcal{F}(s) \subset \mathcal{F}(t)$ .

### 1.2 Processo Estocástico

Dado um espaço de probabilidade  $(\Omega, \mathcal{F}, \mathbb{P})$ , chamamos de processo estocástico a qualquer família  $\{X_t\}_{t \in I}$  de variáveis aleatórias. Neste trabalho, o conjunto de índices  $I$  normalmente vai ser  $[0, +\infty)$  ou um intervalo fechado  $[0, T]$ , onde  $T > 0$ . Por outro lado, dada uma filtração  $\{\mathcal{F}(t)\}_{t \in [0, T]}$ , onde todas as  $\sigma$ -álgebras  $\mathcal{F}(t)$  estão contidas em  $\mathcal{F}$ , dizemos que o processo  $\{X_t\}_{t \in [0, T]}$  é adaptado à filtração se para todo  $t \in [0, T]$ , a variável  $X_t$  é  $\mathcal{F}(t)$ -mensurável.

### 1.3 Movimento Browniano

Seja o espaço de probabilidade  $(\Omega, \mathcal{F}, \mathbb{P})$ . Chamamos de Movimento Browniano um processo  $\{W(t)\}_{t > 0}$  com trajetória contínua, tal que:

- $W(0) = 0$ , P-qc
- Para todo  $0 \leq s < t$  se satisfaz  $W(t) - W(s) \sim N(0, t - s)$  (“*incrementos estacionários*”).

- Para todo  $0 \leq v \leq u \leq s \leq t$ , as variáveis  $W(u) - W(v)$  e  $W(t) - W(s)$  são independentes (“*incrementos independentes*”).

## 1.4 Filtro Browniano

Seja  $(\Omega, \mathcal{F}, \mathcal{P})$  um espaço de probabilidade no qual está definido um movimento Browniano  $W(t)$ ,  $t \geq 0$ . Uma filtração do movimento Browniano é dada por um filtro  $\{\mathcal{F}(t)\}$  de sub  $\sigma$ -álgebras de  $\mathcal{F}$  tal que

- O movimento browniano  $\{W(t)\}$  é um processo adaptado ao filtro.
- Para  $0 \leq t < u$ , o incremento  $W(u) - W(t)$  é independente de  $\mathcal{F}(t)$ .

## 1.5 Martingal

Dados um espaço de probabilidade  $(\Omega, \mathcal{F}, P)$ , um número positivo fixo  $T$  e uma filtração  $\{\mathcal{F}(t)\}_{t \in [0, T]}$ , com  $\mathcal{F}(t) \subset \mathcal{F}$ . Considere um processo estocástico adaptado  $X(t)$ , dizemos que  $X(t)$  é um martingal, se:

- $\mathbb{E}[|X(t)|] < \infty, \forall t \in [0, T]$ , e
- $\mathbb{E}[X(t)|\mathcal{F}(s)] = X(s)$ , P-qc, para todo  $s, t \in [0, T]$ , com  $s \leq t$ .

É importante notar que um movimento browniano uni-dimensional  $W(t)$  é um martingal em relação ao Filtro Browniano.

Demonstração: Dado  $t > 0$ , temos que  $W(t) = W(t) - W(0) \sim N(0, t)$ , logo

$$\mathbb{E}[|W(t)|] = \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{+\infty} |x| e^{-\frac{1}{2} \frac{x^2}{t}} dx < +\infty$$

Também temos que se  $s \leq t$ , então

$$\begin{aligned} E[W(t)|\mathcal{F}(s)] &= E[W(t) - W(s) + W(s)|\mathcal{F}(s)] \\ &= E[W(t) - W(s)|\mathcal{F}(s)] + E[W(s)|\mathcal{F}(s)] \\ &= E[W(t) - W(s)] + W(s) \\ &= W(s) \text{ P - qc.} \end{aligned}$$

## 1.6 Integral Estocástica

Considerando um processo estocástico simples  $\{X(t)\}_{t \in [0, T]}$ , quer dizer, um processo para o qual existe uma partição  $\{0 = t_0 < t_1 < \dots < t_{n-1} < t_n = T\}$  do intervalo  $[0, T]$  tal que para todo  $i = 1, 2, \dots, n$  tem-se  $X(t) = X(t_{i-1})$  para todo  $t \in [t_{i-1}, t_i)$ . A integral estocástica  $I(X, t)$  para  $t \in [0, T]$  é definida por:

$$I(X, t) := \int_0^t X(s) dW(s) := \sum_{i=1}^n X(t_i) [W(t_i \wedge t) - W(t_{i-1} \wedge t)]$$

onde  $a \wedge b := \min\{a, b\}$ .

A partir desta definição de integral estocástica (também chamada de Integral de Itô) para processos simples é possível definir a Integral Estocástica para qualquer outro processo de quadrado integrável adaptado à filtração browniana.

Para a integral de Itô temos as seguintes propriedades:

- $\{I(X, t)\}_{t \in [0, T]}$  é um martingal com respeito a filtração  $\{\mathcal{F}(t)\}_{t \in [0, T]}$ . Em particular,  $\mathbb{E}[I(X, t) | \mathcal{F}(0)] = I(X, 0) = 0, \forall t \in [0, T]$ .
- (Isometria de Itô) Para todo  $t \in [0, T]$  vale:

$$\mathbb{E} \left[ \left( \int_0^t X(s) dW(s) \right)^2 \right] = \mathbb{E} \left[ \int_0^t X^2(s) ds \right].$$

- A variação quadrática acumulada até o tempo  $t$  é dada por,

$$[I, I] = \int_0^t \Delta^2(s) ds$$

- Os caminhos de  $I(X, t)$  são contínuos.
- Para cada  $t$ , a variável  $I(X, t)$  é  $\mathcal{F}(t)$ -mensurável, isto é, o processo  $\{I(X, t)\}_{t \in [0, T]}$  é adaptado.

## 1.7 Fórmula de Itô

Dado um movimento Browniano uni-dimensional  $W(t)$  e um processo estocástico  $X(t)$  descrito por:

$$X(t) = X(0) + \int_0^t K(s) ds + \int_0^t H(s) dW(s)$$

onde os processos  $K(t)$  e  $H(t)$  são adaptados com respeito a uma filtração browniana  $\mathcal{F}(t)$  definidas sob um espaço de probabilidade  $(\Omega, \mathcal{F}, P)$  e uma função  $f : R^2 \rightarrow R$  tal que  $f(t, x)$  é uma vez continuamente diferenciável em  $t$  e duas vezes continuamente diferenciável em  $x$  (diz-se que  $f \in C^{1,2}$ ). Então, para todo  $t \geq 0$  temos que:

$$\begin{aligned} f(t, X(t)) &= f(0, X(0)) + \int_0^t f_t(s, X(s))ds + \int_0^t f_x(s, X(s))dX(s) \\ &+ \frac{1}{2} \int_0^t f_{xx}(s, X(s))dX(s)dX(s) \end{aligned}$$

o que pode ser expresso em forma diferencial como:

$$df(t, X(t)) = f_t(t, X(t))dt + f_x(t, X(t))dX(t) + \frac{1}{2}f_{xx}(t, X(t))dX(t)dX(t)$$

onde  $dX(t)dX(t)$  é a variação quadrática de  $X(t)$ .

$$f(T, W(T)) = f(0, W(0)) + \int_0^T f_t(t, W(t))dt + \int_0^T f_x(t, W(t))dW(t) + \frac{1}{2} \int_0^T f_{xx}(t, W(t))dt$$

## 1.8 Regra do Produto

Dados dois processos estocásticos  $X(t)$  e  $Y(t)$  temos que:

$$X(t)Y(t) = X(0)Y(0) + \int_0^t X(s)dY(s) + \int_0^t Y(s)dX(s) + \int_0^t dX(s)dY(s)$$

o que expresso em forma diferencial é

$$d(X(t)Y(t)) = X(t)dY(t) + Y(t)dX(t) + dX(t)dY(t)$$

## 1.9 Equação de Hamilton-Jacobi-Bellman (HJB)

Muitas vezes temos um processo  $x(t)$  (por exemplo o portfólio) que vai evoluir no tempo, cujo comportamento pode ser alterado por meio de decisões tomadas pelos agentes. Este tipo de problemas são conhecidos como problemas de controle [?] e podem ser modelados através do seguinte sistema de controle:

$$\begin{cases} x'(t) = b(t, x(t), u(t)), & \text{para } t \in [0, T], \\ x(0) = x_0 \end{cases} \quad (1.1)$$

onde com a função  $u(\cdot) : [0, T] \rightarrow U$ , chamada de controle, podemos influenciar o processo  $x(t)$  e  $x_0 \in R$  é a condição inicial,  $T > 0$ , e  $b : [0, T] \times R \times U \rightarrow R$ .

Por outro lado é lógico pensar que, segundo as decisões tomadas, algumas trajetórias de  $x(t)$  serão melhores do que outras. Isto é expresso por meio de uma função de custo associado, que é definida por:

$$J(u(\cdot)) = \int_0^T f(t, x(t), u(t))dt + h(x(T)) \quad (1.2)$$

Salvo algumas condições temos que nosso sistema de controle admite uma única solução  $x(\cdot) \in C([0, T]; R)$ .

Agora, nosso problema consiste em minimizar a função  $J(\cdot)$  sobre todos os possíveis controles  $u(\cdot)$ . Para isso, vamos olhar o sistema de controle sobre intervalos de tempo. Seja  $(s, y) \in [0, T] \times R$ . Consideramos o seguinte sistema de controle sobre  $[s, T]$ :

$$\begin{cases} x'(t) = b(t, x(t), u(t)), & \text{para } t \in [s, T], \\ x(s) = y \end{cases}$$

e consideramos a seguinte função de custo dependente das condições iniciais de  $(s, y)$

$$J(s, y, u(\cdot)) = \int_s^T f(t, x(t), u(t))dt + h(x(T)) \quad (1.3)$$

Logo definimos a seguinte função, chamada de função valor:

$$V(s, y) = \inf\{J(s, y, u(\cdot))\} \quad (1.4)$$

onde o ínfimo é tomado sobre todos os controles  $u(\cdot)$  no intervalo  $[s, T]$  e condições iniciais  $(s, y) \in [s, T] \times R$ . Note-se, que para qualquer  $y \in R$  se satisfaz que  $V(T, y) = h(y)$ .

Denotemos por  $V[s, T]$  o conjunto de todos os controles  $u(\cdot)$  no intervalo de tempo  $[s, T]$ . Agora suponhamos que temos  $0 \leq s < s' \leq T$  e  $x \in R$ . Então fixando um controle  $\hat{u} \in V[s, T]$ , denotemos por  $\hat{x}(t)$  o processo gerado por ele. Logo se usamos um outro controle  $u \in V[s', T]$  e o processo  $x(t)$  tal que  $x(s') = \hat{x}(s')$ , temos que

$$V(s, x) \leq \int_s^{s'} f(t, \hat{x}(t), \hat{u}(t))dt + \int_{s'}^T f(t, x(t), u(t))dt + h(x(T))$$

Logo depois, tomando o ínfimo sobre todos os controles  $u \in V[s', T]$  temos que

$$V(s, x) \leq \int_s^{s'} f(t, \hat{x}(t), \hat{u}(t))dt + V(s', \hat{x}(s'))$$

e desde que o controle  $\hat{u} \in V[s, T]$  seja escolhido arbitrariamente, temos que

$$V(s, x) \leq \inf_{\hat{u} \in V[s, T]} \left\{ \int_s^{s'} f(t, \hat{x}(t), \hat{u}(t))dt + V(s', \hat{x}(s')) \right\}$$



De fato é possível provar a igualdade acima sob certas hipóteses [29]. Esta relação é chamada de *Princípio de Otimalidade de Bellman*, que diz que para todo  $0 \leq s \leq s' \leq T$  e todo  $y \in R$  satisfaz-se a seguinte igualdade.

$$V(s, y) = \inf_{u \in V[s, T]} \left\{ \int_s^{s'} f(t, x(t), u(t)) dt + V(s', x(s'); u(\cdot)) \right\} \quad (1.5)$$

Agora, usando este princípio, vamos supor que  $V \in C^1([0, T] \times R)$ . Dado  $t$ , pegando  $T > t' > t$ , temos pelo Princípio de Otimalidade que

$$V(t, x) = \inf_u \left\{ \int_t^{t'} f(s, x(s), u(s)) ds + V(t', x(t')) \right\} \quad (1.6)$$

Se  $V$  é diferenciável temos que

$$V(t', x(t')) = V(t, x(t)) + \frac{\partial V(t, x(t))}{\partial t} (t' - t) + \frac{\partial V(t, x(t))}{\partial x} x'(t) (t' - t) + o(t' - t) \quad (1.7)$$

Logo, usando esta relação na equação acima e dividindo por  $t' - t > 0$  temos que

$$0 = V_t(t, x) + \inf_u \left\{ \frac{1}{t' - t} \int_t^{t'} f(s, x(s), u(s)) ds + V_x(t, x) x'(t) + \frac{o(t' - t)}{(t' - t)} \right\} \quad (1.8)$$

E pegando o limite quando  $t' \downarrow t$ , supondo que o limite pode ser permutado com o ínfimo, temos

$$0 = V_t(t, x) + \inf_u \{ f(s, x(t), u(t)) + V_x(t, x) x'(t) \} \quad (1.9)$$

e como  $x'(t) = b(t, x(t), u(t))$ , e desde que  $V(T, x) = h(x)$ , vemos que  $V$  é solução para o seguinte problema de equações diferenciais de primeira ordem:

$$\begin{cases} v_t + \inf_{u \in U} \{ b(t, x, u) v_x + f(t, x, u) \} = 0 \\ v|_{t=T} = h(x), (t, x) \in [0, T] \times R \end{cases} \quad (1.10)$$

Esta equação é chamada de *Equação de Hamilton-Jacobi-Bellman*. Estas idéias podem ser estendidas para o caso de um controle estocástico.

Seja  $W_t$ , um movimento browniano. Consideremos o seguinte sistema de controle estocástico:

$$\begin{cases} dx(t) = b(t, x(t), u(t)) + \sigma(t, x(t), u(t)) dW_t, t \in [0, T] \\ x(0) = x_0 \end{cases} \quad (1.11)$$

com a função de custo

$$J(u(\cdot)) = \mathbb{E} \left[ \int_0^T f(t, x(t), u(t)) dt + h(x(T)) \right] \quad (1.12)$$

Defina  $U[0, T] = \{u(\cdot) \text{ é mensurável em } [0, T], \text{ e } \{F_t\}_{t \geq 0} \text{ - adaptado}\}$ , onde  $\{F_t\}_{t \geq 0}$  é a filtração browniana.

O controle estocástico ótimo do problema é dado pela minimização da função anterior sobre  $U[0, T]$ . Para isso, consideramos novamente processos sobre subintervalos de tempo. Seja  $(s, y) \in [0, T] \times R$ , consideramos o controle do sistema sobre  $[s, T]$ :

$$\begin{cases} dx(t) = b(t, x(t), u(t)) + \sigma(t, x(t), u(t)) dW_t, t \in [s, T] \\ x(s) = y \end{cases} \quad (1.13)$$

Novamente, a função custo, que depende do controle  $u(\cdot) \in U[s, T]$ , é definida por:

$$J(s, y; u(\cdot)) = \mathbb{E} \left[ \int_s^T f(t, x(t), u(t)) dt + h(x(T)) \right] \quad (1.14)$$

e definindo a função valor

$$V(t, x) = \inf_{u(\cdot) \in U[s, T]} \{J(t, x; u(\cdot))\}$$

temos a versão estocástica da equação de Hamilton-Jacobi-Bellman.

Supondo que a função valor  $V$  pertence a  $C^{1,2}([0, T])$  temos que ela satisfaz a seguinte equação

$$\begin{cases} v_t + \inf_{u \in U} \left\{ \frac{1}{2} \sigma^2(t, x, u) v_{xx} + b(t, x, u) v_x + f(t, x, u) \right\} = 0 \\ v|_{t=T} = h(x), (t, x) \in [0, T] \times R \end{cases} \quad (1.15)$$

Este resultado será utilizado no capítulo 4, na determinação da alocação ótima do portfólio para os modelos desenvolvidos.

# Capítulo 2

## Análise por Componentes Principais

### 2.1 Descrição da Análise por Componentes Principais

Consideremos uma massa de dados descrita por diversas variáveis. O objetivo principal da Análise de Componentes Principais (PCA) é a obtenção de um pequeno número de combinações lineares (componentes principais) de um conjunto de variáveis, que retenham o máximo possível da informação contida nas variáveis originais [15].

Essa técnica já é amplamente utilizada nos campos da física, identificação facial, estudos de tipologia dos solos e outros. No presente trabalho a usaremos para auxiliar na escolha dos pares de ativos que serão usados no *long-short*. Isto é motivado pelo fato que em trabalhos anteriores de Plerou *et al* [21] mostraram bons resultados para o estudo dos dados no mercado acionário americano.

Supondo que temos os preços de  $N$  ativos  $\{S_i(t)\}$  com retornos  $R_i(t)$ . No tempo  $t_0$  definimos a *primeira componente principal* como sendo o vetor unitário

$$v^{(1)} = (v_1^{(1)}, v_2^{(1)}, \dots, v_N^{(1)})$$

que maximiza a variância da combinação linear

$$Z^{(1)} = \sum_{i=1}^N v_i^{(1)} R_i$$

Uma vez que algumas ações são mais voláteis do que outras, é conveniente trabalhar com retornos padronizados.

$$Y_i = \frac{R_i - \bar{R}_i}{\bar{\sigma}_i} \quad (2.1)$$

onde  $\bar{R}_i$  é a média do retorno  $R_i$  e  $\bar{\sigma}_i^2$  é a sua variância.

Assim nossa primeira componente principal vai ser o vetor unitário  $v^{(1)} = (v_1^{(1)}, v_2^{(1)}, \dots, v_N^{(1)})$  tal que maximiza a variância da combinação linear

$$Z^{(1)} = \sum_{i=1}^N v_i^{(1)} Y_i$$

Chamaremos de segunda componente principal ao vetor unitário

$$v^{(2)} = (v_1^{(2)}, v_2^{(2)}, \dots, v_N^{(2)})$$

tal que a variância da combinação

$$Z^{(2)} = \sum_{i=1}^N v_i^{(2)} Y_i$$

seja máxima e que as variáveis  $Z^{(1)}$  e  $Z^{(2)}$  sejam não-correlacionadas. Em geral, definimos a  $k$ -ésima componente principal como sendo o vetor unitário  $v^{(k)}$  que faz com que a combinação

$$Z^{(k)} = \sum_{i=1}^N v_i^{(k)} Y_i$$

tenha variância máxima e seja não correlacionada com as combinações

$$Z^{(1)}, Z^{(2)}, \dots, Z^{(k-1)}.$$

Assim, vamos obter agora uma caracterização dos componentes principais que permita obtê-los de forma fácil. Notemos que desde que as médias das variáveis  $Y_i$  são todas iguais a zero, pegando um vetor  $v = (v_1, v_2, \dots, v_N)$  qualquer e fazendo  $Z = \sum_{i=1}^N v_i Y_i$ , temos que a média de  $Z$  é zero, logo

$$\begin{aligned} \text{Var}(Z) &= \mathbb{E}[Z^2] \\ &= \mathbb{E}[(\sum_{i=1}^N v_i Y_i)^2] \\ &= \mathbb{E}[(\sum_{i=1}^N v_i Y_i)(\sum_{j=1}^N v_j Y_j)] \\ &= \sum_{i=1}^N \sum_{j=1}^N v_i v_j \mathbb{E}[Y_i Y_j] \\ &= \langle \rho v, v \rangle \geq 0 \end{aligned} \tag{2.2}$$

onde  $\rho$  é a matriz de correlação das variáveis  $Y_1, Y_2, \dots, Y_N$ .

Da própria definição da matriz de correlação, temos que ela é simétrica e da desigualdade anterior segue-se que é semi-definida positiva. Então da álgebra linear (ver [18, 20, 23]) se sabe que esta matriz tem todos os seus autovalores não negativos e ainda mais, pelo teorema espectral, existe uma base ortonormal de  $\mathbb{R}^N$  formada por autovetores da matriz.

Esta observação é importante, dada a reformulação que pode-se efetuar do problema de determinação das componentes principais. Com efeito, da

relação (2.2) é fácil ver que a primeira componente  $v^{(1)}$  é solução do problema de otimização quadrática

$$\text{Max}\langle \rho v, v \rangle \text{ sujeito a } \|v\|^2 = 1$$

logo, pelas condições de Karush-Kuhn-Tucker (KKT), deve existir  $\lambda$  tal que

$$0 = 2\rho v^{(1)} + \lambda v^{(1)}$$

e fazendo  $\lambda_1 = -\lambda/2$  temos  $\rho v^{(1)} = \lambda_1 v^{(1)}$ , assim temos que  $v^{(1)}$  é um autovetor da matriz de correlação. Além do mais desde que a matriz  $\rho$  é semi-definida positiva temos que  $\lambda_1 \geq 0$  e se  $\lambda \geq 0$  for outro autovalor da matriz de correlação, pegando um autovetor unitário  $v$  associado temos que

$$\text{Var}(Z^{(1)}) = \lambda_1 = \langle \lambda_1 v^{(1)}, v^{(1)} \rangle = \langle \rho v^{(1)}, v^{(1)} \rangle \geq \langle \rho v, v \rangle = \langle \lambda v, v \rangle = \lambda$$

de onde vemos que de fato a primeira componente principal é um autovetor unitário correspondente ao maior autovalor da matriz de correlação.

De forma similar é fácil ver que a  $k$ -ésima componente principal  $v^{(k)}$  é solução do problema

$$\text{Max}\langle \rho v, v \rangle \text{ sujeito a } \|v\|^2 \leq 1 \text{ e } \langle v, v^{(i)} \rangle = 0, i = 1, 2, \dots, k-1$$

de onde se pode ver que  $v^{(k)}$  é um autovetor associado ao  $k$ -ésimo maior autovalor  $\lambda_k$  da matriz de correlação  $\rho$  e que  $\text{Var}(Z^{(k)})$ .

Também da mesma definição do problema anterior vemos que o conjunto de componentes principais  $v^{(1)}, v^{(2)}, \dots, v^{(N)}$  constituem um sistema ortonormal de vetores, logo, pegando a matriz  $A$  cujas colunas são os componentes principais temos que  $\rho A = A \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_N)$  e desde que  $A$  é ortogonal se segue que  $\rho = A \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_N) A^T$ , logo temos que

$$N = \text{tr}(\rho) = \lambda_1 + \lambda_2 + \dots + \lambda_N$$

e em particular

$$N \geq \lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \dots \geq \lambda_N \geq 0.$$

Para nossa aplicação usamos dados históricos do preço de fechamento das ações em uma seção transversal de  $N$  ações voltando  $M$  dias. Para simplicidade de exposição, a seção transversal é assumida ser idêntica ao universo de investimento, embora este não precisa ser o caso na prática. Vamos representar os dados de retorno existentes, em qualquer data  $t_0$ , voltando  $M+1$  dias como uma matriz

$$R_{ik} = \frac{S_i(t_0 - (k-1)) - S_i(t_0 - k)}{S_i(t_0 - k)}, \quad k = 1, \dots, M, \quad i = 1, \dots, N, \quad (2.3)$$

e os retornos padronizados podem ser aproximados por

$$Y_{ik} = \frac{R_{ik} - \bar{R}_i}{\bar{\sigma}_i} \quad (2.4)$$

onde

$$\bar{R}_i = \frac{1}{M} \sum_{k=1}^M R_{ik} \quad (2.5)$$

e

$$\bar{\sigma}_i^2 = \frac{1}{M-1} \sum_{k=1}^M (R_{ik} - \bar{R}_i)^2 \quad (2.6)$$

e a matriz de correlação empírica dos dados é definida por

$$\rho_{ij} = \frac{1}{M-1} \sum_{k=1}^M Y_{ik} Y_{jk}, \quad (2.7)$$

Nessa parte do trabalho, será sempre usado uma janela de tempo para estimação da matriz de correlação de 252 dias anteriores ao período de *trading*.

## 2.2 Descrição dos Dados Utilizados

Os dados utilizados em todo o trabalho foram retirados do Bloomberg e correspondem as séries de preços das sessenta e uma ações, elas estão expostas por seu código na tabela 2.1, com maiores volumes de negócios listadas na Bolsa de Valores de São Paulo no período de janeiro de 2005 a dezembro de 2010.

O critério de escolha das ações foi ter negociado diariamente na média um valor superior a dez milhões de reais. Essa restrição com relação a liquidez das ações foi definida uma vez que ações muito pouco líquidas não só trazem um risco maior em situações de *stress* como também podem afetar o retorno da estratégia dado a dificuldade de compra/venda do ativo (elevado descompasso entre o preço de compra e o preço de venda), e a dificuldade de aluguel da ação o que encarece a taxa cobrada.

Nem todas as ações possuem dados desde 2005, visto que durante os anos de 2007 e 2008 houve uma grande quantidade de ações que fizeram a abertura de capital. Nesses casos, o período de observação para esses dados foi menor. Embora, para todos os ativos analisados, houveram negociações para todos os dias, desde que ocorreu a oferta inicial de ações.

Foram utilizados para construção dos modelos os preços de fechamento e o volume financeiro negociado no dia para cada ativo.

## 2.3 Interpretação das Componentes Principais

Da própria definição dos componentes principais, é de se esperar que ativos com comportamentos similares tenham coeficientes também similares no componente principal, e que aqueles que são mais representativos tenham coeficientes maiores. Portanto, usaremos estes coeficientes para determinar os pares que serão arbitrados, sendo o único problema a definição de qual componente principal que será usado neste processo, embora haja uma suspeita inicial de que o principal componente seja o mais indicado

No caso do mercado acionário americano (Plerou *et al* [21]) viram que embora a maior parte dos componentes não contem informação relevante (de fato elas tem o mesmo comportamento dos autovetores de uma matriz aleatória), em geral os componentes principais relativos aos maiores autovalores (2% do total) estes sim contem informação relevante que além do mais são estáveis ao longo do tempo.

A análise mostra também que a primeira componente representa a influência do mercado que é comum a todas as ações, além disso os outros componentes que são bem comportados (não tem comportamento de autovetores de matrizes aleatórias) mostram a existência de correlações entre as ações que tem características similares, exemplo, valor de mercado, ações do mesmo setor, empresas da mesma região geográfica, e outras.

Essa característica acima, vale para os vetores de alto escalão, mas a medida que descemos no espectro para a os autovetores de ruído, a propriedade que os coeficientes próximos correspondem a empresas no mesmo setor é menos verdade e a coerência tende a não se manter.

Como apontado por Laloux *et al* [19], para o caso do primeiro componente principal, todos os coeficientes de  $v_i^{(1)}$ ,  $i = 1, 2, \dots, N$  são positivos. Assim, o autoportfólio tem um peso positivo

$$Q_i^{(1)} = \frac{v_i^{(1)}}{\bar{\sigma}_i}.$$

Notamos que esses pesos são inversamente proporcionais à volatilidade. Esta ponderação é consistente com a capitalização de mercado, uma vez que empresas de maior capitalização tendem a ter volatilidades menores.

Como pode ser observado na figura 2.1 os coeficientes da primeira componente principal são todos positivos, e desde que os componentes principais são ortogonais, os outros componentes tem de ter entradas negativas, como é observado nas figuras 2.2 e 2.3.

Os coeficientes dos autovetores foram classificados em ordem decrescente:

$$v_{n_1}^{(j)} \geq v_{n_2}^{(j)} \geq \dots \geq v_{n_N}^{(j)}, \quad (2.8)$$

o termo  $n_i$  representa a ordem das empresas de acordo com o tamanho de seus coeficientes.

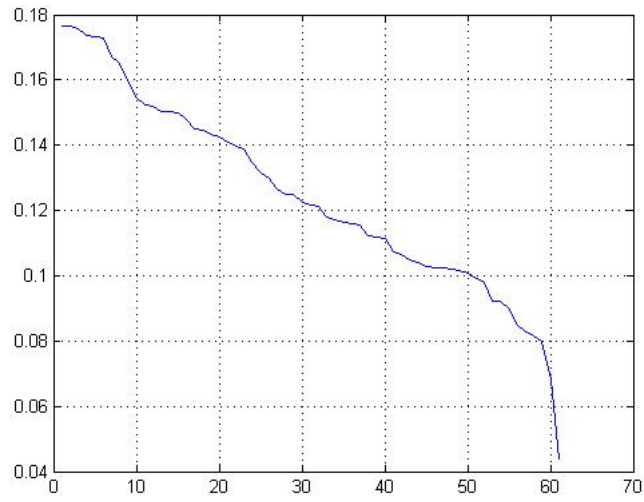


Figura 2.1: Primeiro autovetor ordenado pelo tamanho do coeficiente.

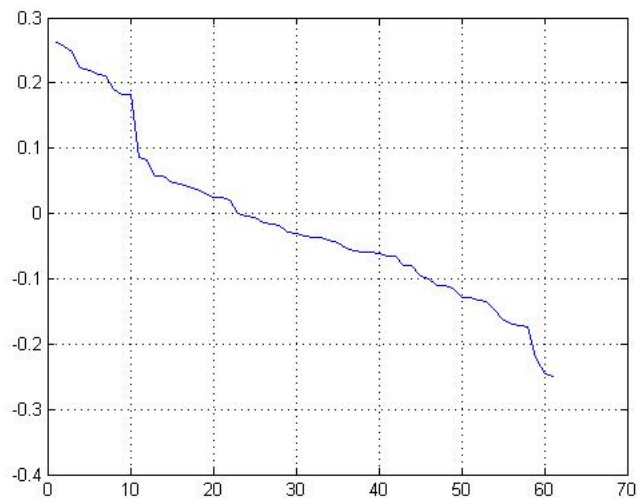


Figura 2.2: Segundo autovetor ordenado pelo tamanho do coeficiente.



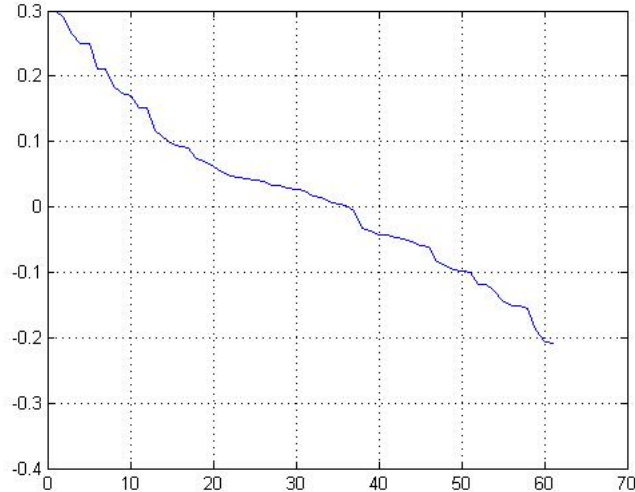


Figura 2.3: Terceiro autovetor ordenado pelo tamanho do coeficiente.

## 2.4 Análise da Série de Dados

Na tabela 2.2 mostra-se a distribuição dos ativos por ordem de importância (valor do coeficiente correspondente) no primeiro componente principal. Vemos que há uma predominância de empresas ligadas a setores de siderurgia, mineração, financeiro, óleo e gás que de fato são os setores que possuem maior peso no Ibovespa.

A fim de ressaltar esse fato, Laloux *et al* [19], faz uma regressão do autoportfólio principal (o portfólio gerado pelo primeiro componente principal fazendo uma ponderação dos ativos considerados) com cada ativo que o compõe e a partir daí o autor pega os resíduos e faz uma nova análise por componentes principais dessa forma ele pôde constatar que os dez principais ativos de cada novo componente principal resultante eram ou do mesmo setor ou tinham valores de mercado similares. Podemos reparar nas tabelas 2.3 e 2.4 a predominância de empresas dos mesmo setor nos segundo e terceiro componente principal.

Conforme comentado anteriormente, o principal componente representa a influência do mercado que é comum a todos os ativos, e de fato podemos reparar na figura 2.4 que o principal componente tem um peso bem maior que os outros.

No trabalho de [21], os autores concluíram que o principal autovalor da matriz de correlação representa a influência do risco de mercado que incide sobre todas as empresas. Pode-se constatar na figura 2.5, que o valor da variância explicada pelo principal autovalor varia ao longo do tempo. De tal forma, que em períodos de baixa volatilidade esse percentual da variância explicada diminui, enquanto que este aumenta em períodos de alta volatili-

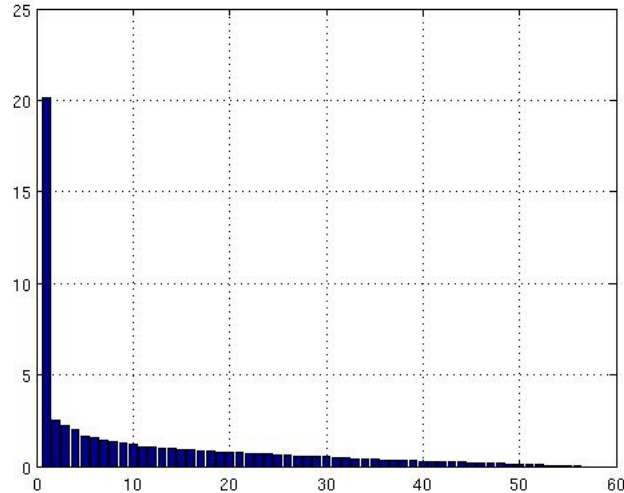


Figura 2.4: Autovalores da matriz de correlação dos retornos das ações.

dade, como durante a crise econômica que ocorreu no segundo semestre de 2008.

Esses resultados são similares ao observado por [3] para o mercado americano. Estes autores ainda concluíram que em períodos onde o número de autovetores necessários para explicar um determinado nível de variância é comparativamente menor, são também períodos nos quais o modelo de reversão à média costuma apresentar resultados melhores.

Finalmente a fim de escolher os pares a serem arbitrados as ações foram classificadas por setores e em cada um deles foi aplicado a análise de componentes principais. É de se esperar, ao se trabalhar com pares do mesmo setor que estão sujeitos aos mesmos fatores econômicos, que a correlação seja maior do que empresas de setores diferentes, e portanto a descrição qualitativa por análise dos componentes principais deva funcionar melhor.

Fazendo uma análise dos setores de materiais básicos e de consumo pode-se perceber que empresas da mesma área de atuação ficam próximas no primeiro componente principal. Dessa maneira, na hora de construir os pares, utilizou-se apenas ações situadas nas proximidades no autovetor. Isso faz com que mesmo dentro de um mesmo setor, exemplo, materiais básicos, evita-se construir pares envolvendo uma mineradora, VALE5, com uma empresa do ramo de papel e celulose, SUZB5.

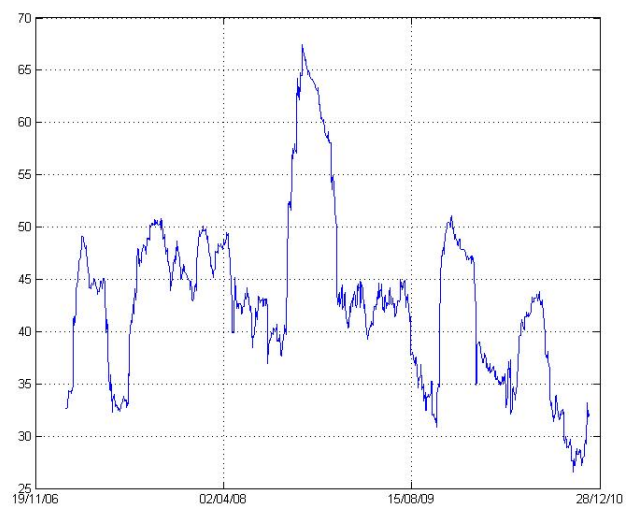


Figura 2.5: Percentual da variância explicada pelos maior autovetor.

Empresas Estudadas Divididas por Setores		
Aeroespacial e Transporte	Construção Civil	Serviços Financeiros
CCR (CCRO3) Embraer (EMBR3) Gol (GOLL4) LLX (LLXL3) Localiza (RENT3) TAM (TAMM4)	BR Malls (BRML3) Cyrela (CYRE3) Duratex (DTEX3) Gafisa (GFSA3) MRV (MRVE3) PDG (PDGR3) Rossi (RSID3)	Banco do Brasil (BBAS3) Bradesco (BBDC4) Itaúsa (ITSA4) Itaú (ITUB4) Redecard (RDCD3) BVMF (BVMF3)
Materiais Básicos	Petróleo e Gás	Telecomunicações
Bradespar (BRAP4) CSN (CSNA3) Gerdau (GGBR4) Gerdau (GOAU4) MMX (MMXM3) Usiminas (USIM3) Usiminas (USIM5) Vale (VALE3) Vale (VALE5) Fibria (FIBR3) Suzano (SUZB5)	Braskem (BRKM5) OGX (OGXP3) Petrobrás (PETR3) Petrobrás (PETR4) Ultrapar (UGPA4) Cosan (CSAN3)	Brasil Telecom (BRTO4) NET (NETC4) TIM (TCSL4) Telemar (TMAR5) Telemar (TNLP4) Vivo (VIVO4)
Consumo	Utilidades Públicas	
Ambev (AMBV4) BR Foods (BRFS3) Souza Cruz (CRUZ3) JBS (JBSS3) Marfrig (MRFG3) B2W (BTOW3) Hypermarcas (HYPE3) Lojas Americanas (LAME4) Lojas Renner (LREN3) Natura (NATU3) Pão de Açúcar (PCAR5)	Cesp (CESP6) Cemig (CMIG4) CPFL (CPFE3) Copel (CPLE6) Eletropaulo (ELPL6) Light (LIGT3) Sabesp (SBSP3) Tractebel (TBLE3)	

Tabela 2.1: Relação de todas as empresas estudadas ao longo do trabalho separadas por setores.

Empresa	Setor
GGBR4	Materiais Básicos
VALE5	Materiais Básicos
VALE3	Materiais Básicos
BRAP4	Materiais Básicos
BBDC4	Serviços Financeiros
ITUB4	Serviços Financeiros
ITSA4	Serviços Financeiros
CSNA3	Materiais Básicos
GOAU4	Materiais Básicos
PETR4	Petróleo e Gás
PETR3	Petróleo e Gás
USIM5	Materiais Básicos
BVMF3	Serviços Financeiros
GFSA3	Construção Civil
CYRE3	Construção Civil

Tabela 2.2: Ativos ordenados por ordem decrescente de coeficiente na primeira componente principal.

Empresa	Setor
CMIG4	Utilidades Públicas
CPLE6	Utilidades Públicas
CPFE3	Utilidades Públicas
TBLE3	Utilidades Públicas
NATU3	Consumo
BRTO4	Telecomunicações
PCAR5	Consumo
CESP6	Utilidades Públicas
TMAR5	Elétrico
CRUZ3	Consumo
TNLP4	Telecomunicações
LIGT3	Utilidades Públicas
CCRO3	Aeroespacial e Transporte
ELPL6	Utilidades Públicas
VIVO4	Telecomunicações
AMBV4	Consumo

Tabela 2.3: Ativos ordenados por ordem decrescente de coeficiente na segunda componente principal.

Empresa	Setor
RSID3	Construção Civil
PDGR3	Construção Civil
MRVE3	Construção Civil
GFS3	Construção Civil
GFS3	Construção Civil
LREN3	Consumo
LAME4	Consumo
BTOW3	Consumo
LLXL3	Aeroespacial e Transporte
GOLL4	Aeroespacial e Transporte
RENT3	Aeroespacial e Transporte
DTEX3	Consumo e Varejo
CSAN3	Petróleo e Gás
CCRO3	Aeroespacial e Transporte
MMXM3	Materiais Básicos

Tabela 2.4: Ativos ordenados por ordem decrescente de coeficiente na terceira componente principal.

Principais Componentes por Setor			
Materiais Básicos	Segmento	Consumo	Segmento
VALE3	Mineração	LAME4	Varejo
VALE5	Mineração	BTOW3	Varejo
GGBR4	Siderurgia	LREN3	Vestuário
BRAP4	Mineração	JBSS3	Frigorífico
CSNA3	Siderurgia	PCAR5	Supermercado
GOAU4	Siderurgia	AMBV4	Bebidas
USIM5	Siderurgia	BRFS3	Alimentos
USIM3	Siderurgia	NATU3	Cosméticos
SUZB5	Papel e Celulose	MRFG3	Frigorífico
FIBR3	Papel e Celulose	CRUZ3	Cigarros
MMXM3	Mineração	HYPE3	Farmacêutico

Tabela 2.5: Ativos ordenados por ordem decrescente de coeficiente na primeira componente principal, com os respectivos segmentos de atuação.

# Capítulo 3

## Modelo Linear de Reversão à Média

### 3.1 Descrição do Processo

Dado que nos capítulo anterior estudamos como escolher os pares de ativos a serem observados, neste capítulo apresentaremos uma estratégia para determinar os momentos que podem ser adequados para abrir ou fechar uma posição nestes ativos. Com este intuito seguiremos a estratégia proposta por Avellaneda *et al* [3], que será detalhada a seguir.

Consideremos dois ativos cujos preços são  $P_t$  e  $Q_t$  que tem características muito similares, como os pares escolhidos seguindo a estratégia descrita no capítulo anterior. Dessa maneira, se espera que os retornos destes ativos tenham uma dependência forte entre eles que pode ser expressa pela equação

$$\log(P_t/P_{t_0}) = \alpha(t - t_0) + \beta \log(Q_t/Q_{t_0}) + X_t \quad (3.1)$$

onde o processo  $X_t$ , que chamaremos de *resíduo*, pode ser estacionário ou de reversão à média. Dado que normalmente o valor de  $\alpha$  é pequeno, conclui-se que a grande parte das informações sobre as discordâncias dos retornos dos ativos vem de  $X_t$ , e esse fato pode ser usado para arbitragem entre os ativos.

Vamos supor agora o processo  $X_t$  é um processo de Ornstein-Uhlenbeck para determinar os sinais que nos permitam conhecer os momentos adequados de fazer os investimentos. Esta suposição é feita levando em conta o comportamento oscilatório do processo  $X_t$  em torno da média. Na parte seguinte do trabalho não será demonstrado formalmente se esta suposição é de fato correta, apenas o eventual ganho ou perda nas operações poderá nos dar uma intuição de se o modelo é adequado ou não.

Assim nosso processo deve de satisfazer a seguinte equação.

$$dX(t) = k(m - X(t))dt + \sigma dW(t), \quad k > 0. \quad (3.2)$$

Este processo é estacionário e auto-regressivo com defasagem 1 (modelo AR-1). Em particular o incremento  $dX(t)$  tem média igual a zero e média condicional igual a

$$\mathbb{E}\{dX(t)|X(s), s \leq t\} = k(m - X(t))dt \quad (3.3)$$

Esta média condicional, ou previsão dos retornos diários esperados do processo residual, é positivo ou negativo de acordo com o sinal de  $m - X(t)$ . Consideramos que os parâmetros  $\sigma, k, m$  e  $\sigma$  da equação diferencial estocástica variam lentamente em relação aos incrementos do movimento browniano  $dW(t)$  na janela de tempo de interesse.

Nas simulações, assim como adotado por Avellaneda *et al*[3] estimamos o processo residual das nossas ações em uma janela de tempo de 60 dias, assumindo implicitamente que os parâmetros são constantes ao longo da janela. Nós aceitamos esta hipótese para as ações nas quais a velocidade de reversão à média (estimativa do  $k$ ) é suficientemente elevada e rejeitamos para as ações com um baixa velocidade de reversão à média. De acordo com a equação (3.1), o retorno diário esperado é de:

$$\alpha dt + k(m - X(t))dt \quad (3.4)$$

O segundo termo corresponde à previsão do modelo para o retorno baseado na posição do processo estacionário  $X(t)$ , ele prevê um retorno negativo se  $X(t)$  é suficientemente elevado e um retorno positivo, se  $X(t)$  é suficientemente baixo. O parâmetro  $k$  é chamado a velocidade de reversão à média. Se  $k \gg 1$ , a ação reverte rapidamente a sua média e o efeito do *drift* é desprezível.

## 3.2 Calibração do Resíduo de Cointegração

Supondo que o resíduo de cointegração é um processo de reversão à média daremos um método para determinar seus parâmetros, para isso, em primeiro lugar resolveremos a equação (3.2), para posteriormente, através de uma discretização, determinar seus parâmetros. A maior parte dos resultados obtidos para este processo servem, de fato para um processo de OU geral.

Usaremos o fator integrante  $e^{ks}$ , cujo diferencial é dado por:

$$de^{ks} = ke^{ks}ds + 0dW(s) \quad (3.5)$$

Aplicando a Regra do Produto temos:

$$\begin{aligned} d(e^{ks}X(s)) &= de^{ks}X(s) + e^{ks}dX(s) \\ &= X(s)ke^{ks}ds + e^{ks}[k(m - X(s))ds + \sigma dW(s)] \\ &= X(s)ke^{ks}ds + e^{ks}k(m - X(s))ds + e^{ks}\sigma dW(s) \\ &= X(s)ke^{ks}ds + e^{ks}kmds - e^{ks}kX(s)ds \\ &\quad + e^{ks}\sigma dW(s) \\ &= e^{ks}kmds + e^{ks}\sigma dW(s) \end{aligned}$$



Integrando ambas as partes de 0 a  $t$  temos

$$\begin{aligned} \underbrace{\int_0^t d(e^{ks}X(s))}_{e^{kt}X(t) - X(0)} &= \underbrace{\int_0^t e^{ks}kmds}_{m(e^{kt} - 1)} + \int_0^t e^{ks}\sigma dW(s) & (3.6) \\ e^{ks}X(s) \Big|_0^t &= km \frac{e^{ks}}{k} \Big|_0^t + \int_0^t e^{ks}\sigma dW(s) \\ e^{kt}X(t) - X(0) &= m(e^{kt} - 1) + \sigma \int_0^t e^{ks}dW(s) \\ e^{kt}X(t) &= X(0) + m(e^{kt} - 1) + \sigma \int_0^t e^{ks}dW(s) \end{aligned}$$

dividindo ambos os lados por  $e^{kt}$  obtemos a solução da equação (3.2), dada por

$$X(t) = e^{-kt}X(0) + m(1 - e^{-kt}) + \sigma e^{-kt} \int_0^t e^{ks}dW(s). \quad (3.7)$$

Logo temos,

$$\begin{aligned} E[X(t)] &= E\left[e^{-kt}X(0) + m(1 - e^{-kt}) + \sigma e^{-kt} \int_0^t e^{ks}dW(s)\right] & (3.8) \\ &= E[e^{-kt}X(0) + m(1 - e^{-kt})] + E\left[\sigma e^{-kt} \int_0^t e^{ks}dW(s)\right]. \end{aligned}$$

Como  $E\left[\sigma e^{-kt} \int_0^t e^{ks}dW(s)\right] = 0$ , pois a integral estocástica de uma função não-aleatória tem média zero e pelo fato de ser  $e^{-kt}X(0) + m(1 - e^{-kt})$  constante, temos que:

$$E[X(t)] = e^{-kt}X(0) + m(1 - e^{-kt}). \quad (3.9)$$

Ainda,

$$E^2[X(t)] = e^{-2kt}X^2(0) + m^2(1 - e^{-kt})^2 + 2e^{-kt}X(0)m(1 - e^{-kt}). \quad (3.10)$$

Temos também

$$\begin{aligned} E[X^2(t)] &= E\left[e^{-2kt}X^2(0) + m^2(1 - e^{-kt})^2 + \sigma^2 e^{-2kt} \left[\int_0^t e^{ks}dW(s)\right]^2 \right. \\ &\quad + 2[e^{-kt}X(0)m(1 - e^{-kt}) + e^{-kt}X(0)\sigma e^{-kt} \int_0^t e^{ks}dW(s) \\ &\quad \left. + m(1 - e^{-kt})\sigma e^{-kt} \int_0^t e^{ks}dW(s)\right]. \end{aligned}$$

Pela Isometria de Itô, temos que

$$\left[\int_0^t e^{ks}dW(s)\right]^2 = \int_0^t [e^{ks}]^2 ds = \int_0^t e^{2ks} ds = \frac{1}{2k}(e^{2kt} - 1) \quad (3.11)$$

e novamente usando o fato que toda matriz estocástica de funções não estocásticas tem média zero e como os demais termos são constantes temos:

$$E[X^2(t)] = e^{-2kt} X^2(0) + m^2(1 - e^{-kt})^2 + \sigma^2 e^{-2kt} \frac{1}{2k} (e^{2kt} - 1) + 2e^{-kt} X(0)m(1 - e^{-kt}) \quad (3.12)$$

Com isso temos que,

$$\begin{aligned} \text{Var}[X(t)] &= E^2[X(t)] - E[X^2(t)] \\ &= \sigma^2 e^{-2kt} \frac{1}{2k} (e^{2kt} - 1) \end{aligned}$$

$$\text{Var}[X(t)] = \frac{\sigma^2}{2k} (1 - e^{-2kt}) \quad (3.13)$$

Com isso, segue que  $X(t)$  tem distribuição normal com média e variância definidos pelas equações (3.9) e (3.13), isto é,  $X(t) \sim N(e^{-2kt} X(0) + m(1 - e^{-2kt}), \frac{\sigma^2}{2k} (1 - e^{-2kt}))$ .

Posteriormente, caracterizamos o processo  $X(t)$  definido pela equação (3.7) como uma versão contínua do processo Auto-Regressivo de primeira ordem, AR(1) em tempo discreto, da forma

$$X(t) = a + bX(t - \Delta t) + \epsilon_t \quad (3.14)$$

com  $a = m(1 - e^{-k\Delta t})$ ,  $b = e^{-k\Delta t}$  e  $\epsilon_t$  é um ruído branco gaussiano.

Dado  $X(t)$  definido pela equação (3.7), temos que em  $t - \Delta t$ :

$$X(t - \Delta t) = e^{-(t-\Delta t)} X(0) + m(1 - e^{-k(t-\Delta t)}) + \sigma e^{-k(t-\Delta t)} \int_0^{t-\Delta t} e^{ks} dW(s), \quad (3.15)$$

logo,

$$e^{-k\Delta t} X(t - \Delta t) = e^{-kt} X(0) + m(e^{-k\Delta t} - e^{-kt}) + \sigma e^{-kt} \int_0^{t-\Delta t} e^{ks} dW(s). \quad (3.16)$$

Fazendo  $X(t) - e^{-k\Delta t} X(t - \Delta t)$  temos:

$$\begin{aligned} X(t) - e^{-k\Delta t} X(t - \Delta t) &= m(1 - e^{-k\Delta t}) + \sigma e^{-kt} \int_{t-\Delta t}^t e^{ks} dW(s) \\ &= m(1 - e^{-k\Delta t}) + \epsilon_t \end{aligned}$$

onde,  $\epsilon_t \sim N(0, \sigma_\epsilon^2)$ .

Como  $\epsilon_t$  tem média zero, segue que  $\sigma_\epsilon^2 = E[\epsilon_t^2]$ .

Ainda, como  $\epsilon_t = \sigma e^{-kt} \int_{t-\Delta t}^t e^{ks} dW(s)$ , por definição, temos que:

$$\begin{aligned} \sigma_\epsilon^2 &= E \left[ \left( \sigma e^{-kt} \int_{t-\Delta t}^t e^{ks} dW(s) \right)^2 \right] \\ &= E \left[ \sigma^2 e^{-2kt} \left( \int_{t-\Delta t}^t e^{ks} dW(s) \right)^2 \right] \end{aligned} \quad (3.17)$$

Pela Isometria de Itô, temos:

$$\begin{aligned}
\sigma_\epsilon^2 &= E \left[ \sigma^2 e^{-2kt} \int_{t-\Delta t}^t (e^{ks})^2 ds \right] \\
&= E \left[ \sigma^2 e^{-2kt} \int_{t-\Delta t}^t e^{2ks} ds \right] \\
&= \sigma^2 e^{-2kt} \left( \frac{e^{2ks}}{2k} \Big|_{t-\Delta t}^t \right) \\
&= \sigma^2 \frac{e^{-2kt}}{2k} (e^{2kt} - e^{2k(t-\Delta t)}) \\
&= \frac{\sigma^2}{2k} (1 - e^{-2k\Delta t})
\end{aligned} \tag{3.18}$$

logo,  $\epsilon_t \sim N(0, \frac{\sigma}{2k}(1 - e^{-2k\Delta t}))$ .

Com isso, temos:

$$X(t) = m(1 - e^{-\eta\Delta t}) + e^{-k\Delta t}X(t - \Delta t) + \epsilon_t, \tag{3.19}$$

portanto, temos que  $X(t)$  é uma representação contínua de um AR(1) da forma

$$X(t) = a + bX(t - \Delta t) + \epsilon_t \tag{3.20}$$

com  $a = m(1 - e^{-k\Delta t})$ ,  $b = e^{-k\Delta t}$  e  $\epsilon_t$  um ruído branco gaussiano.

Finalmente, temos que

$$k = -\frac{\log(b)}{\Delta t}, \quad b \neq 1 \text{ e } b > 0 \tag{3.21}$$

também, como  $a = m(1 - e^{-k\Delta t})$ , temos que

$$m = \frac{a}{1 - b} \tag{3.22}$$

ainda, pela equação (3.18) vemos que

$$\begin{aligned}
\sigma^2 &= \frac{2k\sigma_\epsilon^2}{1 - e^{-2k\Delta t}} \\
&= \frac{-2 \log(b)}{(1 - b^2)\Delta t} \sigma_\epsilon^2 \\
&= \frac{2 \log(b)}{(b^2 - 1)\Delta t} \sigma_\epsilon^2
\end{aligned}$$

e assim

$$\sigma = \sigma_\epsilon \sqrt{\frac{2 \log(b)}{(b^2 - 1)\Delta t}} \tag{3.23}$$

Assim vemos que os parâmetros do processo de OU podem ser estimados a partir do processo regressivo AR(1).

### 3.3 Geração dos Sinais

Considerando que só temos dados discretos dos preços dos nossos ativos, é conveniente usar uma versão discreta

$$R_n^P = \beta_0 + \beta R_n^Q + \epsilon_n, \quad n = 1, 2, \dots, 60 \quad (3.24)$$

da equação

$$\frac{dP_t}{P_t} = \alpha dt + \beta \frac{dQ_t}{Q_t} + dX_t$$

que por sua vez é a versão diferencial da equação (3.1).

Para fazer nossas estimações consideramos que os retornos são ordenados cronologicamente, e  $R_{60}^P$  é o último retorno observado, baseado na variação dos preços de fechamento. Agora, definimos processos auxiliares

$$X_k = \sum_{j=1}^k \epsilon_j, \quad k = 1, 2, \dots, 60, \quad (3.25)$$

que pode ser visto como uma versão discreta de  $X(t)$ , o processo de OU que estamos estimando. Observe que a regressão “força” os resíduos a terem média zero, por isso temos  $X_{60} = 0$ . A estimativa dos parâmetros OU é feito resolvendo a regressão abaixo.

$$X_{n+1} = a + bX_n + \zeta_{n+1}, \quad n = 1, \dots, 59 \quad (3.26)$$

Logo, usando os resultado obtidos acima temos que

$$m = \frac{a}{1-b}$$

$$\Delta t = -\frac{\log(b)}{k}$$

$$\sigma = \sigma_\epsilon \sqrt{\frac{2 \log(b)}{(b^2 - 1)\Delta t}} = \sigma_\epsilon \sqrt{\frac{2k}{1 - b^2}}$$

e desde que  $\text{Var}(\zeta) = \sigma_\epsilon^2$  se segue

$$\sigma = \sqrt{\frac{2k \text{Var}(\zeta)}{1 - b^2}}$$

Por outro lado, das equações (3.9 ) e (3.13) temos que para valores grandes de  $t$

$$\mathbb{E}[X(t)] = m \quad \text{e} \quad \text{Var}[X(t)] = \frac{\sigma^2}{2k} = \frac{\text{Var}(\zeta)}{1 - b^2}$$

denotamos o desvio padrão de  $X(t)$  por  $\sigma_{eq}$ , logo

$$\sigma_{eq} = \sqrt{\frac{\text{Var}(\zeta)}{1-b^2}}$$

Optou-se por trabalhar apenas com ações que tivessem velocidades de reversão à média superiores a  $k > 252/30$ . Isto porque ações com baixa reversão à média tendem a ter um comportamento similar ao ruído branco, o que faz com que o modelo não funcione adequadamente. Definimos o sinal como

$$s = \frac{X(t) - m}{\sigma_{eq}} \quad (3.27)$$

por outro lado, tomando em conta que  $X(t) = X_{60} = 0$ , temos

$$s = -\frac{m}{\sigma_{eq}} = -\frac{a\sqrt{1-b^2}}{(1-b)\sqrt{\text{Var}(\zeta)}}. \quad (3.28)$$

Esta variável mede a distância para o equilíbrio do resíduo cointegrado em unidades do desvio-padrão. O sinal de troca baseado no processo de reversão à média é

$$\begin{aligned} \text{comprar para abrir se } s_i &< -\bar{s}_{ca} \\ \text{vender para abrir se } s_i &> +\bar{s}_{va} \\ \text{fechar posição vendida se } s_i &< +\bar{s}_{cf} \\ \text{fechar posição comprada se } s_i &> -\bar{s}_{vf} \end{aligned} \quad (3.29)$$

onde os valores de corte são determinados empiricamente. Entrar em uma operação significa, por exemplo, a compra das ações consideradas caras e a venda das ações que estão baratas. Da mesma forma, fechando uma posição significa venda da barata e compra da cara.

Uma vez que todas as quantidades são expressas em variáveis adimensionais, esperamos que os valores de cortes para  $\bar{s}_{ca}$ ,  $\bar{s}_{va}$ ,  $\bar{s}_{cf}$ ,  $\bar{s}_{vf}$  possam ser válidas para diferentes ações. Os valores de corte foram selecionados de forma empírica, de acordo com o proposto por Avellaneda *et al* [3]. Com base nesta análise, uma boa escolha de pontos de corte é:

$$\begin{aligned} \bar{s}_{ca} &= \bar{s}_{va} = 1.25 \\ \bar{s}_{cf} &= 0.75 \\ \bar{s}_{vf} &= 0.50 \end{aligned} \quad (3.30)$$

Assim, abrimos posição quando o sinal excede 1,25 em valor absoluto. Fechamos uma posição comprada quando o sinal atinge -0,50. O racional para abrir posição quando o sinal está longe do equilíbrio ocorre pois só pretendemos montar posições quando acreditamos ter detectado um valor anormal para o resíduo cointegrado.

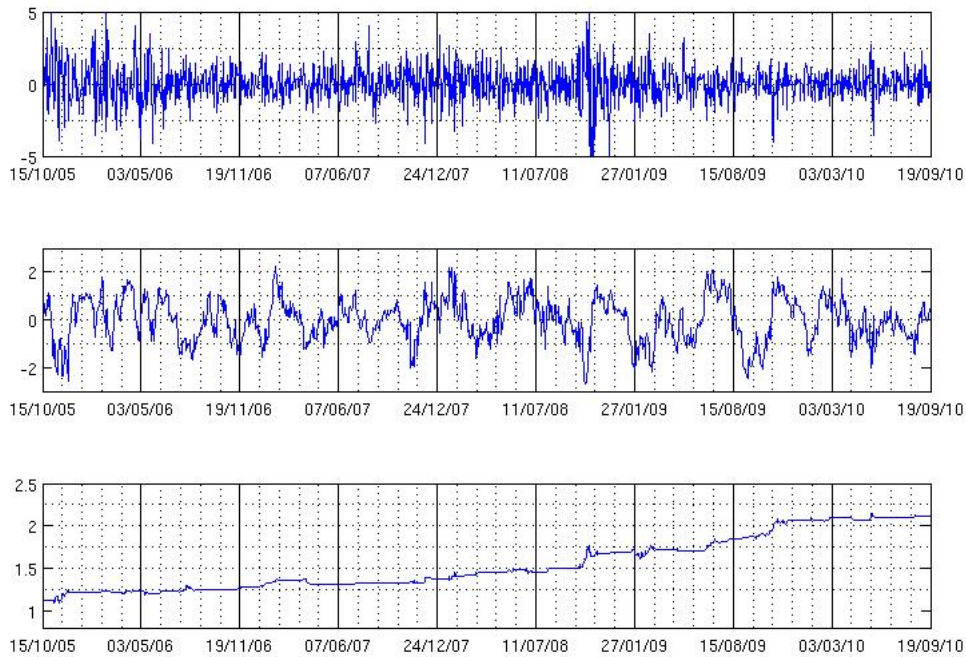


Figura 3.1: O *spread*, os sinais e o *payoff*, respectivamente, para BBDC4 x ITUB4.

Nos resultados foram usados uma janela de duzentos e cinquenta e dois dias para a estimação dos pares e sessenta dias para estimação dos sinais. Além disso, foi considerado taxa de juros,  $r$ , igual a zero, e que sempre que um sinal é formado toda riqueza é investida no *spread*.

Nas figuras 3.2 e 3.1, mostramos a evolução do diferencial dos retornos, dos sinais e do ganho acumulado (*payoff*) para dois pares de ações, um do setor de serviços financeiros e o outro envolvendo uma empresa de serviços financeiros e outra do setor de materiais básicos. Embora, a estratégia inter-setorial (ações de setores diferentes) tenha apresentado um resultado superior que a intra-setorial (ações do mesmo setor), a volatilidade também foi maior, 2,09% contra 6,55%.

Dado que o objetivo do trabalho é apresentar uma estratégia consistente em termos de retornos, mas também com uma baixa volatilidade, o fato anterior reforça a postura já discutida anteriormente de se trabalhar apenas com pares intra-setoriais ao longo do trabalho.

No presente trabalho o único custo considerado foi o de corretagem. O custo padrão de corretagem da Bovespa é de 0,5% sobre qualquer ordem de compra/venda, entretanto uma prática comum no mercado é a devolução pelas instituições financeira de parte dessa corretagem. Para fundos de ações

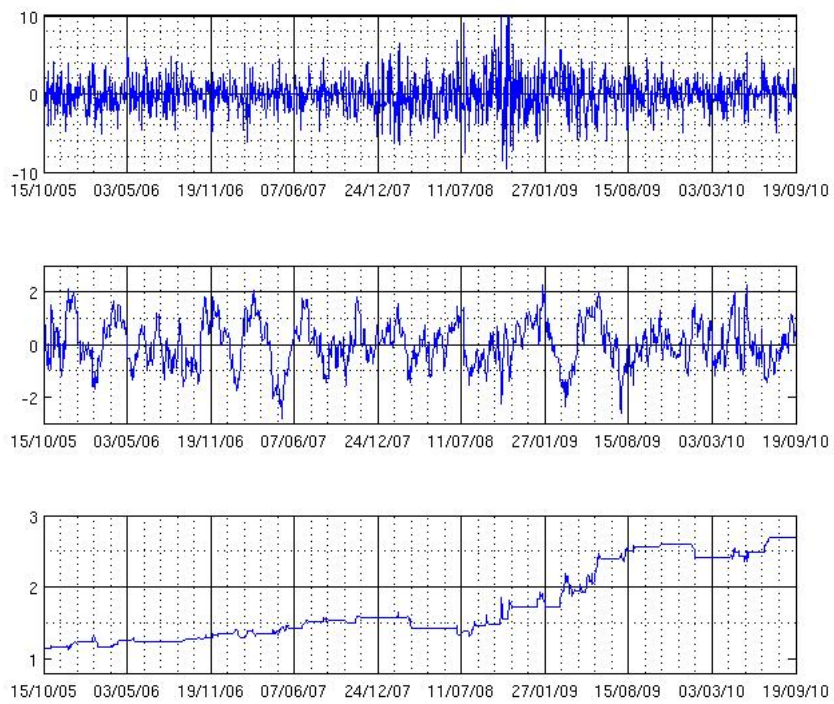


Figura 3.2: O *spread*, os *sinais* e o *payoff*, respectivamente, para ITUB4 x GGBR4.

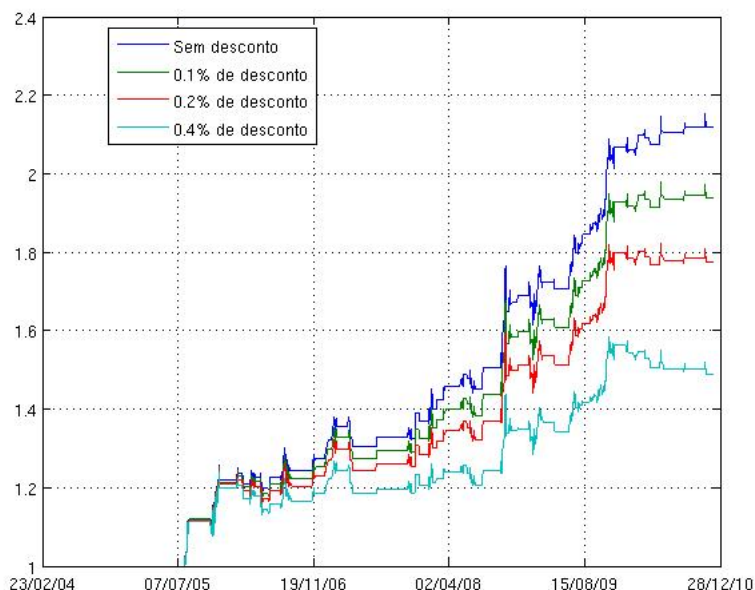


Figura 3.3: Resultados considerando diversos valores para o custo de corretagem.

com cerca de dez milhões de reais, essa devolução alcança 80%, podendo chegar a 90-95% para fundos maiores. No trabalho foi considerado uma devolução de 60%, o que tornaria o desconto que incide em toda transação de compra/venda em 0,2%. Como no *long-short* o processo de compra/venda é simultâneo, seria 0,2% de desconto por transação. O gráfico 3.3 faz um teste de sensibilidade com esses descontos.

Outro custo que existe, mas não foi considerado no trabalho é o de aluguel de ações, que dado que trabalhou-se apenas com as ações mais líquidas, deve ficar na região de 3-5% ao ano.

Assim, como mostra o gráfico 3.4 foi testado a estratégia de arbitragem dos pares com os ativos do mesmo setor, conforme classificação da tabela 2.1. No período observado, o setor com melhor desempenho, já considerando os custos de corretagem, foi o de construção civil, seguido de perto por materiais básicos e consumo. Dos oito setores, três tiveram desempenho superiores ao do Ibovespa no período, sendo que em nenhum caso em nenhum caso houve perda. Além disso, os portfólios não apresentaram perdas significativas durante o período da crise financeira de 2008, e apresentaram volatilidades três vezes menores que as encontradas em uma ação.



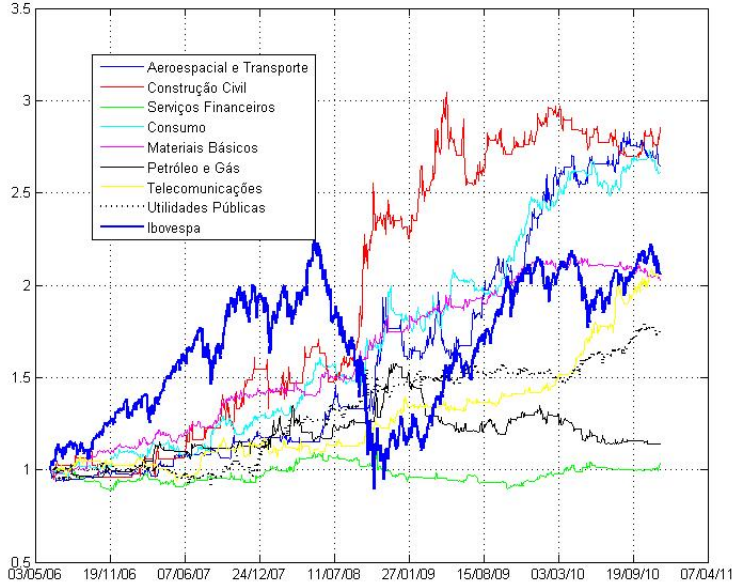


Figura 3.4: *Payoffs* dos diversos portfólios comparados com o Ibovespa.

### 3.4 Ajustando o sinal pelo Volume Diário

Na construção dos sinais a única variável observável usada foi o retorno das ações. Entretanto, existem outras variáveis, tais como, volume de negócios, relação preço/lucro, divulgação de notícias envolvendo as empresas, que podem estar interferindo no sinal, por isso uma melhor análise destas poder ser útil na análise final.

Sendo que a estratégia usada para analisar esse efeito foi relacionar o volume negociado no dia, com a média dos últimos trinta dias, criando-se dessa maneira um novo vetor de retornos.

$$\bar{R}_t = \left( \frac{S_t - S_{t-\Delta t}}{S_{t-\Delta t}} \right) \left( \frac{\frac{1}{k} \sum_{j=0}^{k-1} V_{t-j\Delta t}}{V_t} \right) = R_t \left( \frac{\frac{1}{k} \sum_{j=0}^{k-1} V_{t-j\Delta t}}{V_t} \right) \quad (3.31)$$

onde  $\bar{R}_t$  é o novo vetor de retornos ajustado para o volume diário negociado,  $S_t$  é o preço do ativo e  $R_t$  é o retorno sem o ajuste no tempo  $t$ .

A principal razão ao fazer essa adaptação nos retornos diários observados é evitar que variações anormais no volume negociado, não disparem sinais incorretos de compra ou venda de uma ação. Dessa maneira para volumes negociados no dia muito distantes da média nos últimos trinta dias o retorno observado nesse dia será menor que para o caso do retorno simples no qual não é feito esse tipo de ajuste. De forma inversa vai ocorrer para volumes muito baixos.

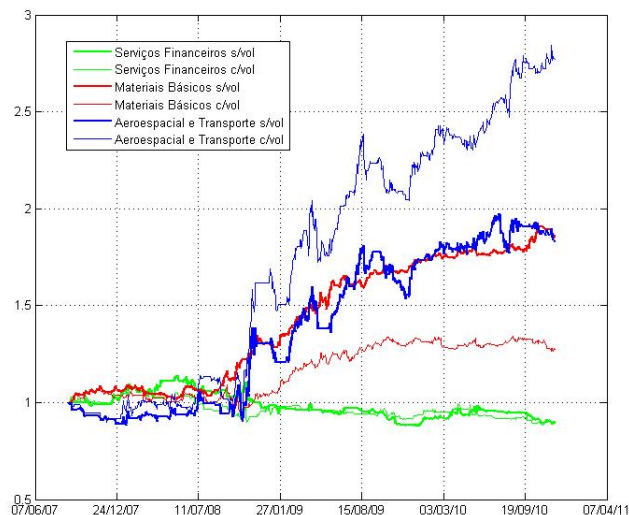


Figura 3.5: *Payoffs* dos portfólios com e sem ajuste pelo volume diário.

Em uma primeira análise foi feita a comparação para três setores distintos dos sinais com e sem ajuste pelo volume diário, e conforme pode ser visto não foi constatada uma unanimidade em termos de resultados. Dado que dos três casos, o modelo com ajuste pelo volume se comportou melhor para um caso e pior para os outros dois, conforme mostrado na figura 3.5.

Quando fazemos a comparação de todos os setores verificamos que o retorno dos sinais ajustados pelo volume foram maiores que os do modelo sem ajuste. Embora, esses retornos tenham sido acompanhados por uma volatilidade maior, ainda sim o retorno ajustado foi melhor, conforme mostrado na tabela 3.1. O retorno ajustado foi calculado como sendo a diferença entre o retorno observado e o retorno do Ibovespa do período dividido pela volatilidade.

### 3.5 Restringindo o Número de Pares

Até o momento a estratégia utilizada era abrir um par, dado um determinado sinal, e desfazer assim que o par tivesse atingido o objetivo. Para os portfólios setoriais usados isso significava, testar determinado número de pares indicados pela análise do componente principal, e investir caso houvesse oportunidade, e deixar o investimento parado, caso não houvesse oportunidade. Ao final do processo, o ganho do portfólio era a média de todos os pares no período.

Dessa forma, tentando melhorar essa estratégia, foi feita uma mudança. Agora, são analisados todos os pares de todos os setores, são quantificados seus respectivos sinais, colocados em ordem decrescente, e os pares de ativos

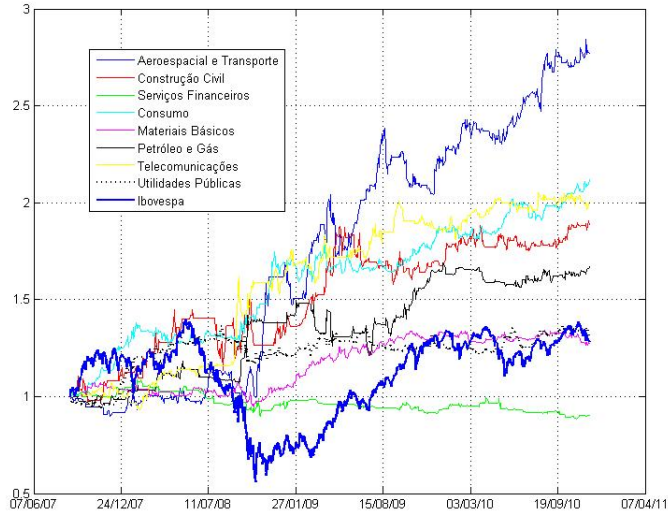


Figura 3.6: *Payoffs* dos diversos portfólios, com ajuste pelo volume diário, comparados com o Ibovespa.

com maiores sinais são escolhidos para montar posição. Essa verificação é feita todo dia. Nesse modelo um par pode ser desfeito antes mesmo de ter seu objetivo atingido, desde que haja outro par em uma situação melhor. Chamamos dessa maneira de portfólio *Top 10* para o grupo de pares de ativos com os dez maiores desvios em relação à média. Os portfólios *Top 15* e *Top 20* seguem a mesma lógica do anterior, mudando apenas o número de pares. O resultado como pode ser visto na figura 3.7, foi bastante consistente.

### 3.6 Trabalhando com Dados de Alta Frequência

Considerando que os resultados encontrados para o modelo linear de reversão à média apresentado nas seções anteriores usando dados diários foram satisfatórios, decidiu-se fazer um teste para dados de alta frequência. Nesse caso são usados observações com intervalos que variam de minutos até alguns poucos segundos. Trabalhos anteriores como os de Dunis *et al* [12] mostraram bons resultados com dados alta frequência para o mercado europeu.

No nosso caso foram usados dados com intervalos de cinco, dez, trinta e sessenta minutos. Esse tipo de informação o Bloomberg não fornece para intervalos muito longos de tempo, no máximo seis meses (para esse trabalho o período foi de 12/Jul/2010 até 21/Jan/2011), mas que totalizaram até cerca de onze mil observações para os dados com intervalo de cinco minutos. Como ocorria nos dados diários a janela de tempo de calibração dos parâmetros também foi de sessenta, só que em vez de dias usou-se cinco minutos.

Conforme foi comentado anteriormente para os dados diários considera-

Comparativo dos Resultados						
Setor	Sinais s/ Ajuste Volume			Sinais com Ajuste Volume		
	<i>Payoff</i>	Vol.	Ret. Ajust.	<i>Payoff</i>	Vol.	Ret. Ajust.
Aeroespacial e Transporte	24,25%	4,95%	1,02	34,80%	8,50%	3,24
Construção Civil	26,33%	6,93%	1,03	20,50%	6,01%	2,24
Serviços Financeiros	0,81%	2,39%	-7,70	-2,93%	2,49%	-4,53
Consumo	23,84%	3,98%	1,17	24,58%	4,27%	4,02
Materiais Básicos	17,08%	2,89%	-0,74	7,39%	2,68%	-0,23
Petróleo e Gás	2,91%	4,05%	-4,02	16,27%	6,07%	1,45
Telecomunicações	17,89%	3,65%	-0,36	22,64%	5,72%	2,70
Utlidades Públicas	13,24%	3,04%	-1,96	8,04%	4,44%	0,01

Tabela 3.1: Comparação dos resultados com e sem ajuste pelo volume diário.

se um desconto de 80% sobre a corretagem padrão de 0,5% da Bovespa. Entretanto, para dados de alta frequência esse desconto aumenta dado que o número de negócios aumenta de maneira bastante consistente. Nesses caso, o desconto varia entre 99% a 99,9%. Para o presente trabalho o valor considerado foi de 99%.

Nessa parte do trabalho não foram considerados os portfólios setoriais, dada a eventual dificuldade que se teria em montar de forma simultânea todos os pares dos setores. Por isso, optou-se por fazer um teste com pares que possuem as maiores liquidez, ou seja, que apresentam os maiores volumes diários de negócios para aplicar o modelo, como mostrado na figura 3.8. Como pode-se reparar, mesmo considerando os custos de corretagem, em todos os casos foram positivos, oscilando entre 20% e 50%.

Posteriormente foi feita uma análise para os pares anteriores com a evolução dos ganhos para os diferentes intervalos de tempo, como mostrado nos gráficos 3.9 e 3.10. Em ambos os casos há uma gradual piora dos ganhos conforme se aumenta o intervalos de tempo dos dados. Também é importante comentar que para todos os casos a volatilidade encontrada foi inferior a 1%, bastante inferior aos dados diários, na qual o mínimo encontrado costuma ser próximo a 2% podendo chegar até 8%.

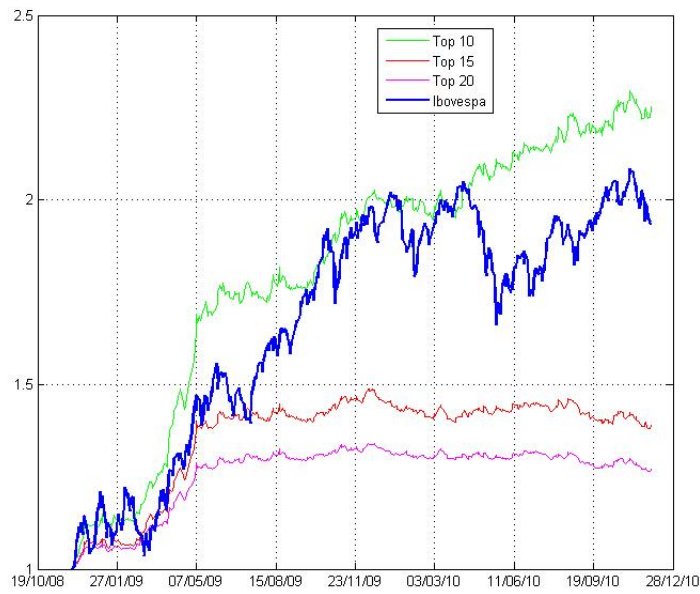


Figura 3.7: *Payoffs* dos portfólios Top 10, Top 15 e Top 20.

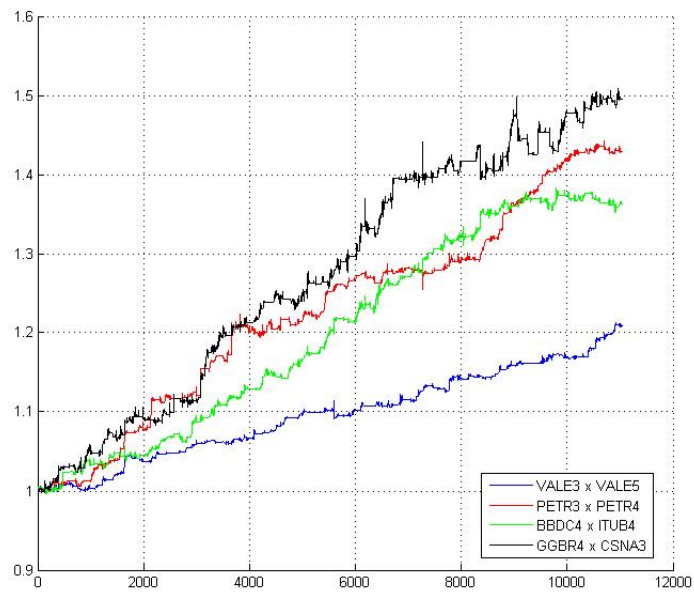


Figura 3.8: *Payoffs* para um intervalo de cinco minutos.

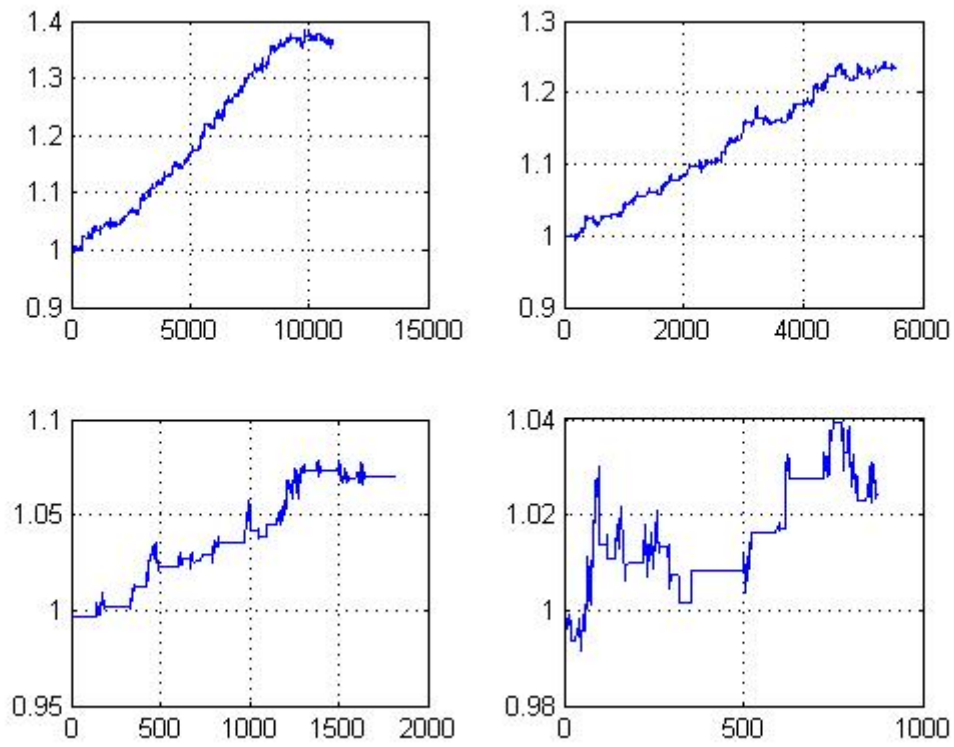


Figura 3.9: *Payoffs* para diferentes intervalos de tempo para BBDC4 x ITUB4.

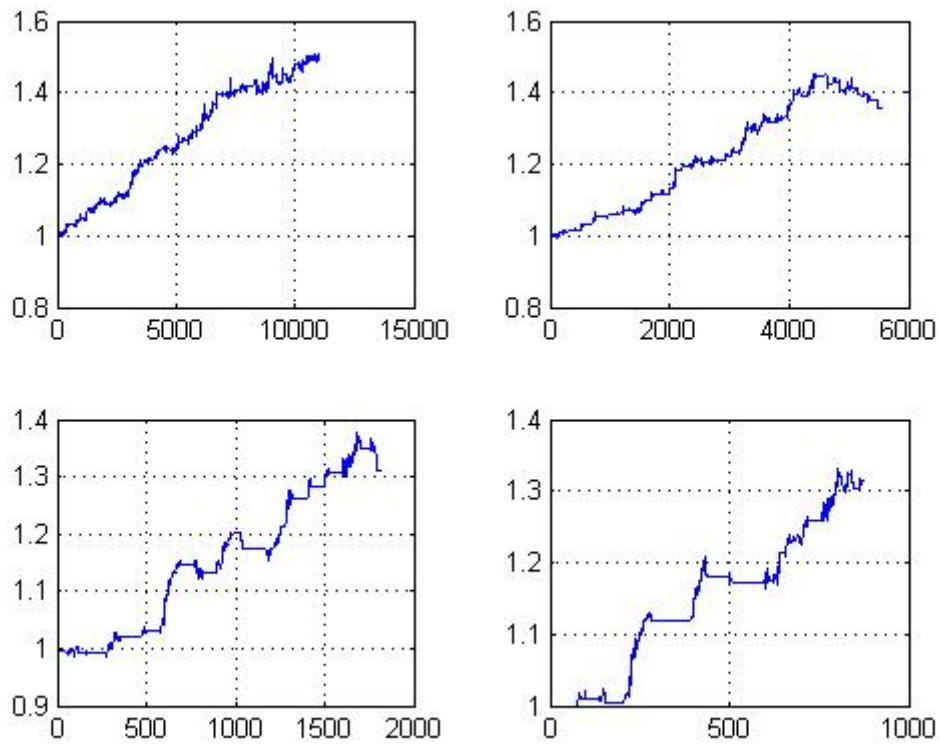


Figura 3.10: *Payoffs* para diferentes intervalos de tempo para CSNA3 x GGBR4.

# Capítulo 4

## Escolha do Portfólio dado um Modelo com Aversão ao Risco

### 4.1 Modelando as Preferências dos Investidores

O trabalho até o momento preocupou-se exclusivamente em analisar a escolha dos pares e o melhor momento de comprar/vender. Ficou faltando, portanto, para essa última parte do trabalho falarmos sobre a escolha do portfólio ótimo. De modo geral, uma possibilidade a ser analisada para essa questão seria considerar que o arbitrador, sempre investe toda sua riqueza no *spread*.

Entretanto, seria natural considerar que dado que o arbitrador sabe que o processo de diferencial de preços segue um modelo de reversão à média, esperar que este quando perceber que o *spread* se afasta muito da média assumir posições maiores que a sua riqueza para aproveitar essa oportunidade de ganho, valendo-se da possibilidade de alavancagem que o mercado possibilita.

Por outro lado, o processo de reversão à média não mostra quando o diferencial de preços vai reverter à média nem se há a possibilidade de piorar antes de melhorar. Houveram momentos específicos na história provocados por eventos únicos de ausência de liquidez, no qual fundos alavancados tiveram perdas que em alguns casos foram irreversíveis como discutido no trabalho de Khandani e Lo [17].

Jurek e Yang [16] desenvolveram um modelo de valor relativo em um horizonte finito do tempo, na qual a preocupação principal do modelo refere-se a estrutura de preferências dos investidores e o processo de caracterização do evolução do valor da oportunidade de arbitragem.

Para modelar as preferências do investidor vamos usar a família de funções de utilidade, dependentes de um parâmetro  $\gamma$ , bastante comum em economia.

$$u_{\gamma}(w) = \begin{cases} \frac{w^{1-\gamma} - 1}{1 - \gamma} & \text{se } \gamma \neq 1 \\ \log(w) & \text{se } \gamma = 1 \end{cases}$$



Note que o caso  $\gamma = 1$  na definição acima é um caso limite da definição dos outros casos, pois, para  $w > 0$  temos, usando a Regra de l'Hôpital, que

$$\lim_{\gamma \rightarrow 1} u_\gamma(w) = \lim_{\gamma \rightarrow 1} \frac{w^{1-\gamma} - 1}{1 - \gamma} = \lim_{\gamma \rightarrow 1} \frac{-w^{1-\gamma} \log(w)}{-1} = \log(w) = u_1(w)$$

Por outro lado, para todo  $\gamma$ , temos que

$$u'_\gamma(w) = w^{-\gamma}$$

o que indica que estas funções são crescentes ( $u'_\gamma(\cdot) > 0$ ), o que descreve a fato que o investidor nunca está satisfeito. Também, desde que

$$u''_\gamma(w) = -\gamma w^{-\gamma-1}$$

temos que, para  $\gamma > 0$  estas funções são côncavas, o que expressa o fato que o investidor é averso ao risco. Pode-se ver que quanto maior for o parâmetro  $\gamma$  a função de utilidade é mais côncava, o que pode ser interpretado como uma maior aversão ao risco. O caso  $\gamma = 0$  se reduz a  $u_0(w) = w - 1$ , que representa um investidor neutro ao risco. Não consideramos o caso em que  $\gamma < 0$ , pois neste caso a função  $u_\gamma(\cdot)$  é convexa, o que descreve investidores propensos ao risco.

Dessa forma, considere a estrutura alternativa de preferência para o arbitrador em tempo contínuo. Assume-se que o agente tem aversão ao risco constante e que maximiza a utilidade descontada da riqueza terminal. A função valor do arbitrador no tempo  $t$  - denotado por  $V_t$  - assume a forma:

$$V_t = \sup E_t \left[ e^{-\beta(T-t)} u_\gamma(X_T) \right] = \sup E_t \left[ e^{-\beta(T-t)} \frac{X_T^{1-\gamma}}{1-\gamma} \right] \quad (4.1)$$

onde o supremo é tomado sobre todas as possíveis estratégias de investimento. Note-se também que na equação acima temos tirado termo “-1” do numerador da função utilidade, isto foi feito por comodidade, dado que isso não muda em nada a descrição das preferências do investidor.

Em seu trabalho Xiong [28] considerou apenas arbitradores com utilidade na forma de logaritmo, dessa forma não abordando uma espectro maior de níveis da aversão ao risco. Já no trabalho de Jurek e Yang [16] eles abordaram diferentes níveis de aversão ao risco criando um novo componente, chamado por eles de *demand de hedging intertemporal*. Este novo termo desempenha, nas atividades de arbitragem, um papel importante determinando quando os arbitrados especulam o desvio de preços.

## 4.2 Escolha do Portfólio Ótimo

A oportunidade de arbitragem nesse modelo, dada pelo diferencial de preços entre pares de ações relacionadas, expõe o arbitrador a duas dimensões importantes de riscos: (i) incerteza sobre o momento em que o desvio nos preços

será eliminado (risco de horizonte); (ii) a possibilidade do desvio nos preços divergir arbitrariamente ainda mais longe de sua média antes da convergência (risco de divergência). Para capturar essas duas formas de risco, modela-se o desvio nos preços usando o processo de reversão à média de OU, visto anteriormente.

Nesse modelo, o arbitrador pode investir em um ativo livre de risco e no *spread* com reversão á média. Denotando os preços dos ativos por  $B_t$  e  $S_t$ , respectivamente. Suas dinâmicas são dadas pelas equações abaixo:

$$dB_t = rB_t dt \quad (4.2)$$

$$dS_t = k(\bar{S} - S_t)dt + \sigma dW_t \quad (4.3)$$

Dado que o valor do *spread* representa o preço de uma carteira *long-short*, comprando (vendendo) uma unidade do *spread* é equivalente a comprar (vender) uma unidade do ativo sobrevalorizado e vender (comprar) uma unidade do ativo subvalorizado.

Se denotarmos o número de unidades do *spread* e do ativo livre de risco na posse do agente por  $N_t$  e  $M_t$ , respectivamente, a descrição do processo de riqueza pode ser escrita como:

$$dX_t = N_t dS_t + M_t dB_t \quad (4.4)$$

Para facilitar a leitura das equações abaixo excluimos o tempo no momento de escrever os processos, logo, a evolução da riqueza satisfaz:

$$\begin{aligned} dX &= N dS + \frac{W - NS}{B} dB \\ &= (r(X - NS) + k(\bar{S} - S)N)dt + \sigma N dW \end{aligned} \quad (4.5)$$

Vemos que o processo de riqueza depende do preço do ativo  $S_t$ , que não podemos controlar, e do processo  $N_t$  que é a posição que o investidor toma, o que sim podemos controlar, logo a estratégia de investimento fica modelada por este processo. Assim, na definição da função valor do arbitrador, o supremo tem de ser tomado sobre todos possíveis caminhos de  $N_t$ . Por comodidade, vamos expressar agora a função valor dependendo de  $\tau = T - t$ .

$$V(X(\tau), S(\tau), \tau) = \begin{cases} \max_{N(S(\tau), \tau)} E_t [e^{-\beta\tau} \log X_T | N(\cdot)] & \gamma = 1 \\ \max_{N(S(\tau), \tau)} E_t [e^{-\beta\tau} \frac{X_T^{1-\gamma}}{1-\gamma} | N(\cdot)] & \gamma \neq 1 \end{cases} \quad (4.6)$$

Pelo Lema de Itô, temos:

$$\begin{aligned}
dV &= V_S dS + V_W dS dX - \frac{\partial V}{\partial \tau} dt + \frac{1}{2} V_{SS} dS^2 + \frac{1}{2} V_{XX} dX^2 + V_{SX} dX dS \quad (4.7) \\
&= V_S (k(\bar{S} - S) dt + \sigma dW) + V_X ((r(X - NS) + k(\bar{S} - S)N) dt + \sigma N dW) + \\
&\quad - \frac{\partial V}{\partial \tau} dt + \frac{1}{2} V_{SS} \sigma^2 dt + \frac{1}{2} V_{XX} N^2 \sigma^2 dt + V_{SX} N \sigma^2 dt
\end{aligned}$$

Para resolver este problema de controle estocástico usamos o bem conhecido critério de otimalidade dado pela equação Hamilton-Jacobi-Bellman (HJB),  $E_t[dV(\cdot)] = \beta V$ , portanto se torna:

$$\begin{aligned}
\beta V &= \max_N V_S k(\bar{S} - S) + V_X (r(X - NS) + k(\bar{S} - S)N) + \quad (4.8) \\
&\quad - \frac{\partial V}{\partial \tau} + \frac{1}{2} V_{SS} \sigma^2 + \frac{1}{2} V_{XX} N^2 \sigma^2 + V_{SX} N \sigma^2
\end{aligned}$$

Logo, supondo que o máximo acima é atingido no ponto  $N$ , que é a quantidade ótima investida no *spread*, temos, pela condição necessária de primeira ordem, que

$$\begin{aligned}
0 &= \frac{\partial}{\partial N} (V_S k(\bar{S} - S) + V_X (r(X - NS) + k(\bar{S} - S)N) \\
&\quad - \frac{\partial V}{\partial \tau} + \frac{1}{2} V_{SS} \sigma^2 + \frac{1}{2} V_{XX} N^2 \sigma^2 + V_{SX} N \sigma^2) \\
&= (k(\bar{S} - S) - rS) V_X + N \sigma^2 V_{XX} + \sigma^2 V_{SX}
\end{aligned}$$

e isolando  $N$  na equação acima, temos que a quantidade ótima a investir é

$$N = - \left( \frac{V_X}{V_{XX}} \right) \left( \frac{k(\bar{S} - S) - rS}{\sigma^2} \right) - \frac{V_{SX}}{V_{XX}} \quad (4.9)$$

Dentro do primeiro termo, o primeiro dado é o fator de aversão ao risco absoluto da função de valor, e o segundo dado é o índice de Sharpe instantâneo do *spread* dividido por seu desvio padrão. O segundo termo representa a demanda de *hedging* intertemporal.

Para um investidor com aversão ao risco constante e com utilidade cobrindo apenas a riqueza terminal, vamos supor que a função valor do investidor tem o seguinte formato.

$$V(X, S, \tau) = \begin{cases} e^{-\beta\tau} \log X + e^{-\beta\tau} (A(\tau)S^2 + B(\tau)S + C(\tau)) & \gamma = 1 \\ e^{-\beta\tau} \frac{X^{1-\gamma}}{1-\gamma} + \exp(A(\tau)S^2 + B(\tau)S + C(\tau)) & \gamma \neq 1 \end{cases} \quad (4.10)$$

Adicionalmente, a fim de cumprir as condições de contorno no valor da função, impõe-se que os coeficientes das funções,  $A(\tau)$ ,  $B(\tau)$ ,  $C(\tau)$ , satisfaçam a seguinte condição de contorno:  $A(0) = B(0) = C(0) = 0$ .

Agora analisaremos o problema de derivar uma expressão para as funções  $A(\tau)$ ,  $B(\tau)$  e  $C(\tau)$ .

**Caso 1: Investidor com utilidade logarítmica** ( $\gamma = 1$ ) Substituindo a regra proposta do portfólio na equação Hamilton-Jacobi-Bellman (HJB), junto com as derivadas relevantes da função obtemos a seguinte condição:

$$\begin{aligned} \beta V = & S^2 \left( \frac{(k+r)^2}{2\sigma^2} - 2kA(\tau) - A'(\tau) \right) + S \left( -kB(\tau) - \frac{k(k+r)\bar{S}}{\sigma^2} + \right. \\ & \left. + 2k\bar{S}A(\tau) - B'(\tau) \right) + r + \sigma^2 A(\tau) + \frac{1}{2}k\bar{S} \left( 2B(\tau) + \frac{k\bar{S}}{\sigma^2} \right) - C'(\tau) + \beta V \end{aligned} \quad (4.11)$$

como a condição de não-arbitragem deve-se manter para todos os valores de  $S$ , cada um dos coeficientes no polinômio acima com respeito a  $S$  devem ser iguais a zero em todos os pontos no tempo. Isto nos leva a um sistema de três EDPs para os três coeficientes das funções  $A(\tau)$ ,  $B(\tau)$ ,  $C(\tau)$ :

$$A'(\tau) = -2kA(\tau) + \frac{(k+r)^2}{2\sigma^2} \quad (4.12)$$

$$B'(\tau) = -kB(\tau) - \frac{k(k+r)\bar{S}}{\sigma^2} + 2k\bar{S}A(\tau) \quad (4.13)$$

$$C'(\tau) = r + \sigma^2 A(\tau) + \frac{1}{2}k\bar{S} \left( 2B(\tau) + \frac{k\bar{S}}{\sigma^2} \right) \quad (4.14)$$

Então após de resolver este sistema (ver apêndice A) obtemos:

$$A(\tau) = \frac{(k+r)^2(1-e^{-2k\tau})}{4k\sigma^2} \quad (4.15)$$

$$B(\tau) = -\frac{(k+r)(k-r+e^{-k\tau}(k+r))(1-e^{-k\tau})}{2k\sigma^2} \bar{S} \quad (4.16)$$

$$\begin{aligned} C(\tau) = & -\left( \frac{k+r}{2k} \right)^2 \sinh(k\tau)e^{-k\tau} \\ & + \frac{(k+r)(k-3r+e^{-k\tau}(k+r))(1-e^{-k\tau})}{4k\sigma^2} \bar{S}^2 \\ & + \left( \frac{k^2+6kr+r^2}{4k} + \frac{1}{2} \left( \frac{r\bar{S}}{\sigma} \right)^2 \right) \tau \end{aligned} \quad (4.17)$$

## **Caso 2: Investidor Geral** ( $\gamma \neq 1$ )

Substituindo a função proposta junto com a correspondente regra do portfólio na condição de otimalidade e recolhendo valores potenciais de  $S$ ,

temos:

$$\begin{aligned} \beta = & S^2 \cdot \left[ \frac{2\sigma^2}{\gamma} \cdot A(\tau)^2 - 2 \left( \frac{\kappa}{\gamma} + r \left( \frac{1-\gamma}{\gamma} \right) \right) \cdot A(\tau) + \frac{1}{2} \left( \frac{1-\gamma}{\gamma} \right) \left( \frac{\kappa+r}{\sigma} \right)^2 - A'(\tau) \right] \\ & + S \cdot \left[ \left( \frac{2}{\gamma} \cdot A(\tau) - \left( \frac{\kappa+r}{\sigma^2} \right) \left( \frac{1-\gamma}{\gamma} \right) \right) \kappa \bar{S} + \left( \frac{2\sigma^2}{\gamma} \cdot A(\tau) - \left( \frac{\kappa}{\gamma} + r \cdot \left( \frac{1-\gamma}{\gamma} \right) \right) \right) \cdot B(\tau) - B'(\tau) \right] \\ & + (1-\gamma)r + \sigma^2 \cdot A(\tau) + \frac{\sigma^2}{2\gamma} \cdot B(\tau)^2 + \left( \frac{1}{\gamma} \cdot B(\tau) + \left( \frac{1}{2\sigma^2} \right) \left( \frac{1-\gamma}{\gamma} \right) \kappa \bar{S} \right) \kappa \bar{S} + \beta - C'(\tau). \end{aligned} \quad (4.18)$$

Desde que a equação de otimalidade seja satisfeita para todos os valores  $S$ , os coeficientes das diferentes potências de  $S$  devem ser todos iguais a zero. Isto nos leva a um sistema de três equações diferenciais ordinárias para as funções dos coeficientes  $\{A(\tau), B(\tau), C(\tau)\}$ , as quais podem ser resolvidas sequencialmente para obter a expressão final da função. O sistema de EDOs incluem a equação de Ricatti e é dada por:

$$A'(\tau) = \frac{2\sigma^2}{\gamma} A(\tau)^2 - 2 \left( \frac{\kappa}{\gamma} + r \left( \frac{1-\gamma}{\gamma} \right) \right) A(\tau) + \frac{1}{2} \left( \frac{1-\gamma}{\gamma} \right) \left( \frac{\kappa+r}{\sigma} \right)^2$$

$$B'(\tau) = - \left( \frac{\kappa}{\gamma} + r \left( \frac{1-\gamma}{\gamma} \right) - \frac{2\sigma^2}{\gamma} A(\tau) \right) B(\tau) + \left( \frac{2}{\gamma} A(\tau) - \left( \frac{\kappa+r}{\sigma^2} \right) \left( \frac{1-\gamma}{\gamma} \right) \right) \kappa \bar{S}$$

$$C'(\tau) = \sigma^2 A(\tau) + \frac{\sigma^2}{2\gamma} B(\tau)^2 + \frac{\kappa \bar{S}}{\gamma} B(\tau) + (1-\gamma)r + \left( \frac{1}{2\sigma^2} \right) \left( \frac{1-\gamma}{\gamma} \right) (\kappa \bar{S})^2$$

A primeira equação diferencial no sistema pertence à classe de equações de Ricatti, as quais tem à forma geral:

$$\frac{dA(\tau)}{d\tau} = c_1 A(\tau)^2 + 2c_2 A(\tau) + c_3$$

onde  $c_1 = \frac{2\sigma^2}{\gamma}$ ,  $c_2 = - \left( \frac{\kappa}{\gamma} + r \left( \frac{1-\gamma}{\gamma} \right) \right)$  e  $c_3 = \frac{1}{2} \left( \frac{1-\gamma}{\gamma} \right) \left( \frac{\kappa+r}{\sigma} \right)^2$ .

Como pode-se ver no apêndice A, as soluções dependem de duas condições, o sinal do discriminante

$$\Delta = 4 \left[ \frac{\kappa^2 - (1-\gamma)r^2}{\gamma} \right]$$

e se o coeficiente do risco de aversão relativo é maior ou igual que a unidade, esta última condição está relacionada com domínios que funções hiperbólicas inversas aparecendo na constante de integração, elas tendem a reprimir as

quantidades investidas, dependendo da propensão ao risco (aversão de riscos) nas preferências dos agentes, como se pode ver na fórmula da solução final.

O significado matemático de se  $\gamma$  é maior que a unidade é imediato, desde que esta condição seja suficiente para assegurar que o discriminante é positivo. Entretanto, esta condição não é necessária. Acontece que o sinal do discriminante é unicamente determinado pela relação entre o coeficiente de aversão ao risco e a razão dos dois parâmetros do modelo,  $\kappa$  - a razão de aversão média - e  $r$  - a taxa livre de risco. Em resumo temos chegado à seguinte solução geral para a equação de Ricatti:

$$A(\tau) = \begin{cases} -\frac{c_2}{c_1} + \frac{\sqrt{-\Delta}}{2c_1} \tan\left(\frac{\sqrt{-\Delta}}{2}\tau + \tan^{-1}\left(\frac{2c_2}{\sqrt{-\Delta}}\right)\right) & \gamma \in (0, \gamma_0) \\ -\frac{c_2}{c_1} \left(1 + \frac{1}{c_2\tau-1}\right) & \gamma = \gamma_0 \\ -\frac{c_2}{c_1} + \frac{\sqrt{\Delta}}{2c_1} \coth\left(-\frac{\sqrt{\Delta}}{2}\tau + \coth^{-1}\left(\frac{2c_2}{\sqrt{\Delta}}\right)\right) & \gamma \in (\gamma_0, 1) \\ -\frac{c_2}{c_1} + \frac{\sqrt{\Delta}}{2c_1} \tanh\left(-\frac{\sqrt{\Delta}}{2}\tau + \tanh^{-1}\left(\frac{2c_2}{\sqrt{\Delta}}\right)\right) & \gamma \in (1, \infty) \end{cases}$$

Note que se  $\gamma_0 < 0$ , ou seja, quando a razão da reversão média excede a taxa livre de risco, somente as soluções com  $\tanh$  e  $\coth$  são aplicáveis. Com estas soluções em mãos podemos proceder a resolver a segunda EDO do sistema para  $B(\tau)$ .

Utilizando a notação já introduzida, a EDO para  $B(\tau)$  pode ser expressada da seguinte forma:

$$\frac{dB(\tau)}{d\tau} = c_2B(\tau) + c_1A(\tau)B(\tau) + c_4A(\tau) + c_5$$

onde as duas constantes adicionais são definidas como:

$$c_4 = \frac{2\kappa\bar{S}}{\gamma}$$

$$c_5 = -\left(\frac{\kappa+r}{\sigma^2}\right)\left(\frac{1-\gamma}{\gamma}\right)\kappa\bar{S}$$

Assim como com  $A(\tau)$ , a solução da EDO para  $B(\tau)$  tomará diferentes formas dependendo da localização do modelo no espaço dos parâmetros. Se pode mostrar que a solução da EDO de acima é dada por:

$$B(\tau) = \begin{cases} \frac{c_4\phi_1(\tau) + \frac{4\phi_2(\tau)}{\sqrt{-\Delta}}\left(c_5 - \frac{c_4}{c_1}\right)\cos\left(\frac{\sqrt{-\Delta}}{2}\tau - \tan^{-1}\left(\frac{c_2}{\sqrt{-\Delta}}\right)\right)}{c_1\sqrt{-\Delta}} & \gamma \in (0, \gamma_0) \\ \left(\frac{c_1c_5(c_2\tau-2)-c_2^2c_4}{2c_1(c_2\tau-1)}\right)\tau & \gamma = \gamma_0 \\ \frac{4(c_2c_5 - c_3c_4 + (c_3c_4 - c_2c_5)\cosh(\frac{\sqrt{\Delta}}{2}\tau)) + 2c_5\sqrt{\Delta}\sinh(\frac{\sqrt{\Delta}}{2}\tau)}{\Delta\cosh(\frac{\sqrt{\Delta}}{2}\tau) - 2c_2\sqrt{\Delta}\sinh(\frac{\sqrt{\Delta}}{2}\tau)} & \gamma \in (\gamma_0, \infty) \end{cases}$$

onde  $\phi_1(\tau)$  e  $\phi_2(\tau)$  são dadas por:

$$\phi_1(\tau) = \sqrt{-\Delta}\left(\cos\left(\frac{\sqrt{-\Delta}}{2}\tau\right) - 1\right) + 2c_2\sin\left(\frac{\sqrt{-\Delta}}{2}\tau\right)$$

$$\phi_2(\tau) = \tanh^{-1} \left( \tan \left( \frac{1}{4} \left( \sqrt{-\Delta} \tau - 2 \tanh^{-1} \left( \frac{2c_2}{\sqrt{-\Delta}} \right) \right) \right) \right) + \tanh^{-1} \left( \tan \left( \frac{1}{2} \tanh^{-1} \left( \frac{2c_2}{\sqrt{-\Delta}} \right) \right) \right)$$

Finalmente conhecendo já as soluções de  $A$  e  $B$  procedemos a resolver a EDO para  $C(\tau)$ .

Esta última EDO pode parecer simples de resolver, dado que só precisamos fazer uma integração, a dificuldade aparece pelo fato de ter de integrar produtos cruzados e funções trigonométricas hiperbólicas. Embora soluções analíticas sejam possíveis para as integrais de todas as funções aparecendo linearmente ( $A(\tau)$  e  $B(\tau)$ ), o termo quadrático,  $B(\tau)^2$ , não tem uma solução com expressão fechada. Isto nos impede de derivar uma expressão analítica para  $C(\tau)$  no caso geral, quando  $\bar{S} \neq 0$ . Entretanto, isso não é problema, pois não precisamos da função  $C(\tau)$  para determinar o portfólio ótimo  $N$ .

No caso especial quando  $\bar{S} = 0$  pode se ver, sem dificuldade que  $c_4 = c_5 = 0$  o que leva a  $B(\tau) = 0$  para todo  $\tau$ . Este resultado faz que a última EDO seja simples de resolver, neste caso temos

$$C(\tau) = \begin{cases} \frac{1}{2} \left( (\kappa + 3r(1 - \gamma))\tau + \gamma \log \left( \frac{\sqrt{\frac{-\Delta}{c_1 c_3}}}{2 \cos \left( -\frac{\sqrt{-\Delta}}{2} \tau - \tan^{-1} \left( \frac{2c_2}{\sqrt{-\Delta}} \right) \right)} \right) \right) & \gamma \in (0, \gamma_0) \\ \frac{1}{2} \left( (\kappa + 3r(1 - \gamma))\tau + \gamma \log \left( \frac{1}{1 - c_2 \tau} \right) \right) & \gamma = \gamma_0 \\ \frac{1}{2} \left( (\kappa + 3r(1 - \gamma))\tau + \gamma \log \left( \frac{\sqrt{\frac{\Delta}{c_1 c_3}}}{2 \sinh \left( \frac{\sqrt{\Delta}}{2} \tau - \coth^{-1} \left( \frac{2c_2}{\sqrt{\Delta}} \right) \right)} \right) \right) & \gamma \in (\gamma_0, 1) \\ \frac{1}{2} \left( (\kappa + 3r(1 - \gamma))\tau + \gamma \log \left( \frac{\sqrt{\frac{\Delta}{-c_1 c_3}}}{2 \cosh \left( \frac{\sqrt{\Delta}}{2} \tau - \tanh^{-1} \left( \frac{2c_2}{\sqrt{\Delta}} \right) \right)} \right) \right) & \gamma \in (1, \infty) \end{cases}$$

Agora que temos uma expressão para a função valor do investidor, substituindo ela na equação (4.9) temos que a alocação ótima do portfólio de um agente com aversão ao risco constante com utilidade definida sobre a riqueza terminal, é dada por:

$$N = \begin{cases} \left( \frac{k(\bar{S} - S) - rS}{\sigma^2} \right) X & \gamma = 1 \\ \left( \frac{k(\bar{S} - S) - rS}{\gamma \sigma^2} + \frac{2A(\tau)S + B(\tau)}{\gamma} \right) X & \gamma \neq 1 \end{cases} \quad (4.19)$$

Nestes resultados temos permitido que o *spread*,  $S_t$ , tenha uma média arbitrária de longo prazo  $\bar{S}$ , mas em aplicações onde os valores dos ativos subjacentes são substitutos perfeitos a média de longo prazo do *spread* médio será igual a zero ( $\bar{S} = 0$ ), o que apresenta algumas vantagens no momento de calcular o portfólio ótimo. Neste caso temos visto que  $c_4 = c_5 = 0$ , o que faz com que  $B(\tau) = 0$ , para todos os valores de  $\tau$ .

### 4.3 Características da Demanda de *Hedging* Intertemporal

Na função estratégia ótima, (4.19), pode-se observar dois componentes diferenciados. O primeiro componente,  $(\frac{k(\bar{S}-S)-rS}{\gamma\sigma^2}X)$ , representa, segundo Jurek e Yang [16], a demanda míope e responde somente a variações instantâneas no desvio dos preços, enquanto que o segundo componente  $(\frac{2A(\tau)S+B(\tau)}{\gamma}X)$  representa a demanda de *hedging* intertemporal cujo sinal e magnitude dependem da aversão a risco do arbitrador e dos parâmetros do processo que descreve a oportunidade de arbitragem.

Aliado a isso, da fórmula de  $B(\tau)$  vemos que ele está fortemente influenciado pelas constantes  $c_4$  e  $c_5$ , que por sua vez são proporcionais à média do *spread*, de onde pode-se interpretar o termo  $B(\tau)$  como um ajuste para a média de longo prazo do *spread*.

Simplificando para o caso de  $\bar{S} = 0$ , vemos que para investidores que são mais (menos) avessos ao risco que a utilidade logarítmica,  $A(\tau)$  é não-positivo (não-negativo) e decrescente (crescente) em  $\tau$ . Intuitivamente, como a arbitragem reverte a média, uma divergência no diferencial de preços coincide com uma melhora na oportunidade de investimento instantânea. Isso indica que a demanda de *hedging* diminui a alocação míope dos arbitadores tolerantes ao risco ( $\gamma < 1$ ), e amplifica a alocação míope dos arbitadores avessos ao risco ( $\gamma > 1$ ).

Entretanto, essa diferença de sinais entre o fator míope e a demanda de *hedging* intertemporal para o caso do investidor menos avesso ao risco e a igualdade de sinais para o investidor mais avesso ao risco só vale para o caso de  $\bar{S} = 0$ . No presente trabalho, essa hipótese apenas foi usada para facilitar o entendimento da parte teórica. Pois, analisando os dados refutou-se essa hipótese de que de  $\bar{S} = 0$ . Dessa forma o investidor, independente do grau de aversão ao risco, vai apresentar efeito míope e demanda de *hedging* intertemporal com sinais que variam ao longo do período, podendo ser iguais ou diferentes, conforme apresentado nas figuras 4.1 e 4.2.

Na construção do diferencial de preços  $S_t$ , tomamos o *spread* nos preços de fechamento diário para cada par. Por exemplo, denotando esses preços por  $P_{i,t}$  para  $i = \{1, 2\}$  implica:

$$S_t = P_{1,t} - P_{2,t} \quad (4.20)$$

O procedimento de calibração envolve a calibração do processo de OU para os parâmetros de reversão, volatilidade, e nível natural  $\{\hat{k}, \hat{\sigma}, \hat{S}\}$  respectivamente, durante o período de treinamento. Foi usado o procedimento dos Mínimos Quadrados para estimar os parâmetros do OU. Uma vez obtidas as estimativas para  $\{\hat{k}, \hat{\sigma}, \hat{S}\}$ , substituímos em  $N_{OU}$ .

No gráfico 4.3 fazemos uma comparação dos resultados dos setores de petróleo e gás e serviços financeiros para um investidor menos avesso ao risco



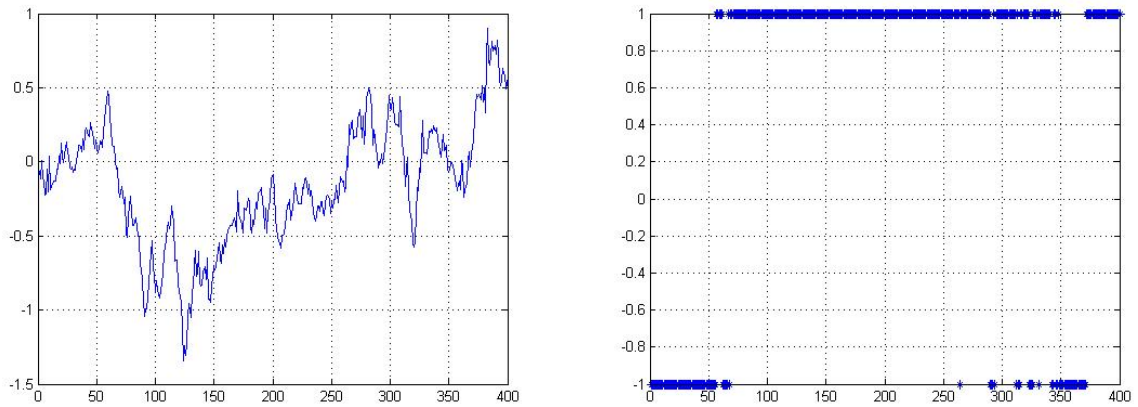


Figura 4.1: Demanda total,  $N$ , e evolução da diferença de sinais entre o fator míope e a demanda de *hedging* intertemporal para CYRE3 x RSID3, com  $\gamma = 0,5$ .

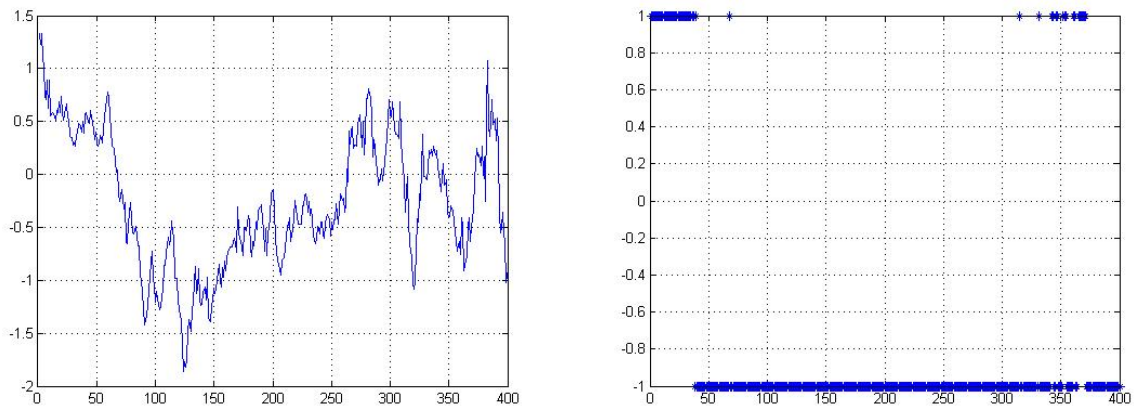


Figura 4.2: Demanda total,  $N$ , e evolução da diferença de sinais entre o fator míope e a demanda de *hedging* intertemporal para CYRE3 x RSID3, com  $\gamma = 1,5$ .

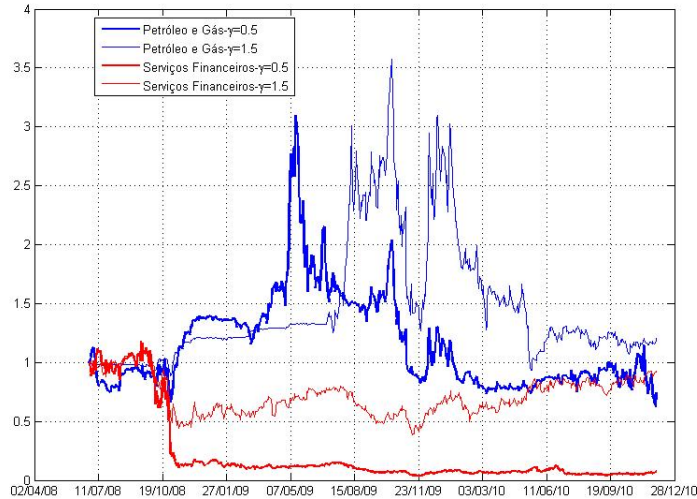


Figura 4.3: *Payoffs* dos setores para o caso  $\gamma = 1,5$  e  $\gamma = 0,5$ .

( $\gamma = 0.5$ ), e para um investidor avesso ao risco ( $\gamma = 1,5$ ). Em ambos, os casos foram encontradas perdas para o período observado. Mais do que isso, para os casos onde  $\gamma = 0.5$ , o prejuízo não só foi maior, como a princípio a volatilidade também aparenta ter sido. Para tirar essa dúvida foi feito na tabela 4.1 um comparativo com o resultado de todos os setores para valores diferentes de  $\gamma$ . De um modo geral os resultados não foram bons, dado que a maioria dos setores apresentou perdas, aliado a volatilidades muito elevadas. Podemos ainda constatar o resultado claramente inferior para o modelo com  $\gamma = 1,5$ . Foi usado o mesmo custo de corretagem do capítulo anterior para dados diários.

Comparativo dos Resultados						
Setor	Investidor com ( $\gamma = 0.5$ )			Investidor com ( $\gamma = 1,5$ )		
	<i>Payoff</i>	Vol.	Ret. Ajust.	<i>Payoff</i>	Vol.	Ret. Ajust.
Aeroespacial e Transporte	14,96%	15,29%	0,91	-17,79%	41,29%	-0,46
Construção Civil	-3,21%	11,92%	-1,69	-16,65%	21,04%	-0,84
Serviços Financeiros	-1,52%	15,30%	-0,17	-41,09%	24,50%	-1,72
Consumo	14,27%	12,30%	1,08	-23,32%	19,84%	-1,23
Materiais Básicos	0,16%	11,16%	-0,08	3,25%	15,07%	0,15
Petróleo e Gás	3,86%	17,94%	0,16	-5,83%	16,83%	-0,41
Telecomunicações	-4,50%	9,60%	-0,58	-37,32%	20,74%	-1,85
Utlidades Públicas	21,39%	18,75%	1,08	21,43%	31,15%	0,65

Tabela 4.1: Comparação dos resultados para diferentes valores de  $\gamma$ , com ajuste de 700 dias.

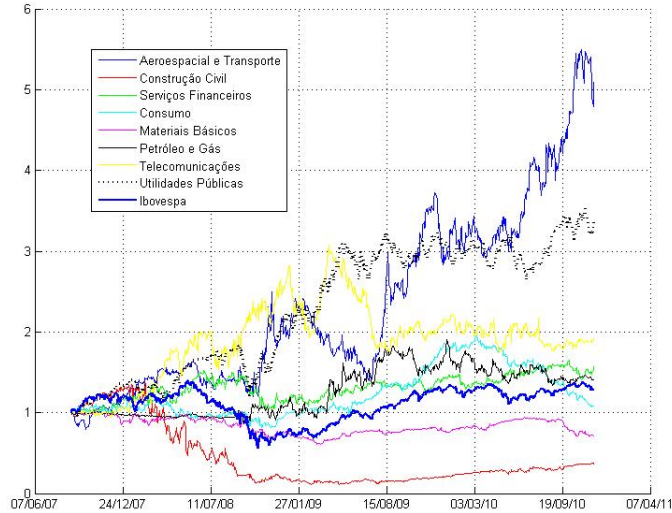


Figura 4.4: *Payoffs* dos diversos portfólios comparados com o Ibovespa, com ajuste de 500 dias.

Ainda pode-se observar que esse modelo apresentado tem uma enorme sensibilidade com relação ao período de calibração dos parâmetros. Pois, conforme pode ser observado nas figuras 4.4, 4.5 e 4.6, variando os períodos de calibração (foram testados 500, 700 e 1000 dias) o desempenho dos portfólios muda completamente. Havendo casos onde um determinado setor apresentava perdas sob determinada condição, e passou a apresentar ganhos quando os parâmetros foram estimados novamente.

Outro ponto observado nesse modelo foi que para pares de ativos que possuem um velocidade de reversão à média muito elevada,  $k$ , e consequentemente uma baixa volatilidade,  $\sigma$ , a posição que maximiza a riqueza assume valores muito elevados, podendo passar de oitenta vezes a riqueza, e na prática graus de alavancagem dessa magnitude são impossíveis de serem implementados, devido a exigência de margens de garantias exigidas pelas entidades reguladoras do mercado. Na figura 4.7 ilustramos essa situação mostrando o caso de um par do setor de serviços financeiros que apresenta elevado  $k$  e um par do setor de construção civil que apresenta um  $k$  moderado.

Uma possível saída para limitar esse excesso de alavancagem seria limitar o tamanho da posição. O gráfico 4.8 mostra como ficaria o resultado do setor de petróleo e gás caso uma limitação de até duas vezes a riqueza atual fosse aplicada. Essa solução, embora seja uma saída simples para o problema, do ponto de vista matemático está incorreta, dado que no processo de otimização da riqueza essa restrição não foi considerada. Caso, ela tivesse sido considerada, bem provavelmente teríamos encontrado resultados bem diferentes em relação aos atuais.

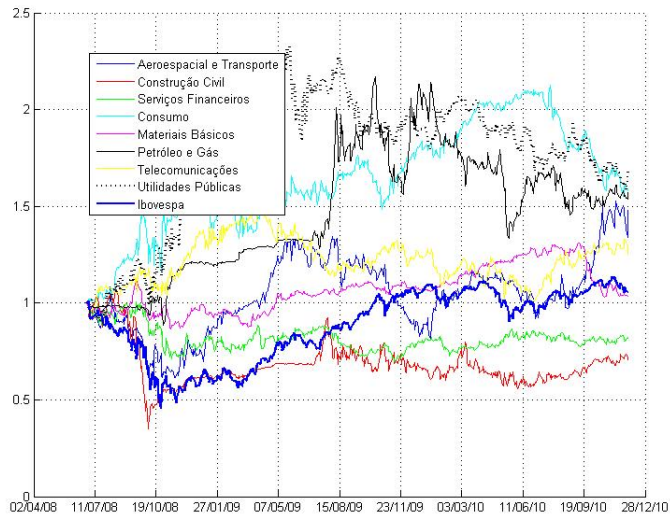


Figura 4.5: *Payoffs* dos diversos portfólios comparados com o Ibovespa, com ajuste de 700 dias.

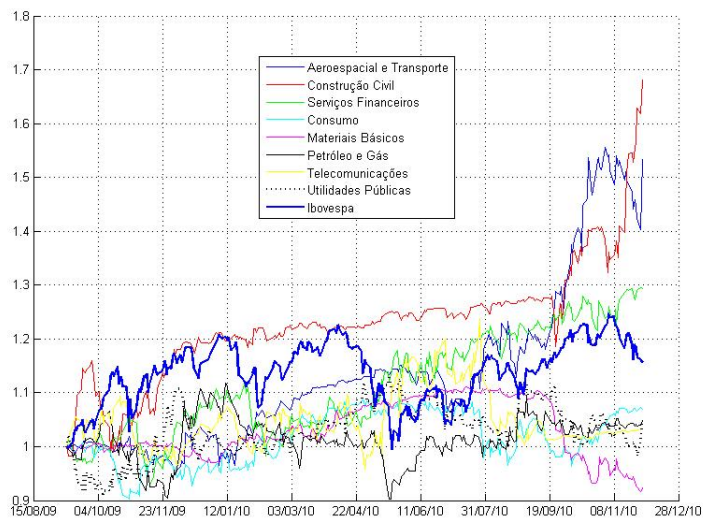


Figura 4.6: *Payoffs* dos diversos portfólios comparados com o Ibovespa, com ajuste de 1000 dias.



Figura 4.7: Demanda total,  $N$ , para BBDC4 x ITUB4 ( $k$  elevado) e CYRE3 x RSID3 ( $k$  moderado), com ajuste de 700 dias e  $\gamma = 1,5$ .

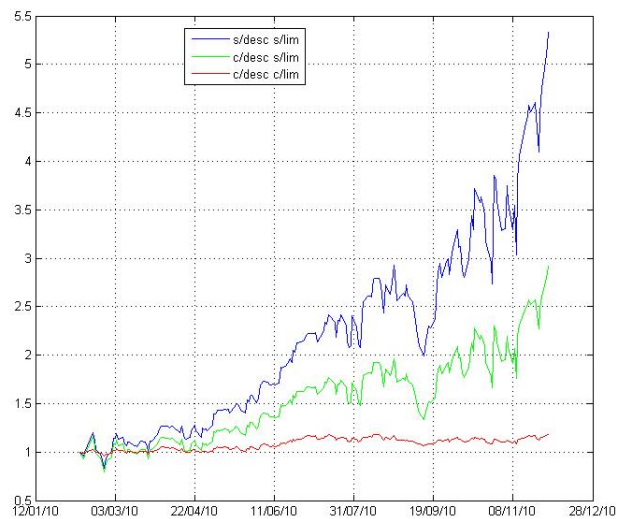


Figura 4.8: *Payoffs* considerando custos de corretagem e alavancagem, com ajuste de 500 dias e  $\gamma = 0,5$ .

Além disso, dada uma determinada magnitude do *spread*,  $S_t$ , o grau de agressividade com o qual arbitadores procuram negociar contra o diferencial de preços depende do tempo remanescente para a data de avaliação,  $\tau$ . Em geral, conforme  $\tau$  aumenta a magnitude da demanda de *hedging* aumenta, dado que o arbitrador tem mais tempo para explorar a variação nessa oportunidade de investimento.

Nas figuras abaixo, para uma riqueza inicial de 100 e um  $\gamma = 1,5$ , ou seja, para o caso do investidor mais avesso ao risco que o implícito na utilidade logarítmica foi analisado a evolução da alocação ótima,  $N$ , assim como o ganho acumulado para um determinado par de ativos, considerando que o arbitrador está sempre abrindo/fechando posição sempre que o *spread* se afasta da média. Novamente não foram considerados custos de transação, e  $r = 0$ . Repare que de acordo com o modelo a alocação ótima no período pode chegar a cerca de cinco vezes a riqueza atual do agente. Sendo que esse grau de alavancagem na prática não é possível.

## 4.4 Principais Premissas do Modelo Não-Linear de Reversão à Média com Tangente Hiperbólica

No trabalho de Alsayed e McGroarty [2] foi proposto um processo estocástico não-linear de reversão à média, cuja força de reversão à média enfraquece dado o nível do diferencial de preços.

Dessa maneira, quando o diferencial dos preços está perto de seu nível natural, a reversão é forte em relação a perturbação aleatória, gradualmente empurrando o diferencial de preços para seu nível natural. Se, no entanto, o diferencial de preços é relativamente grande, a força da reversão é mais fraca em relação às perturbações aleatórias, potencialmente permitindo permanência do diferencial de preços distante de seu nível natural. A figura 4.10 ilustra bem a alocação ótima para os diferentes modelos discutidos até agora, dado uma situação hipotética para determinado nível dos *spreads*.

Similarmente ao feito no capítulo anterior para o modelo de OU, vamos considerar um universo consistindo de dois ativos. O primeiro é o diferencial de preços  $S_t$  que descreve a evolução da oportunidade de arbitragem, o segundo é o  $B_t$  ativo livre de risco. A dinâmica destes, respectivamente, são dadas pelas seguintes equações:

$$dS_t = -\frac{k}{c} \tanh(c(S_t - \bar{S}))dt + \sigma dW_t, \quad (4.21)$$

$$dB_t = rB_t dt \quad (4.22)$$

onde  $\bar{S}$  é o nível natural do diferencial de preços,  $r$  é a taxa de risco-livre,  $W_t$  é um movimento Browniano com respeito à medida da probabilidade do

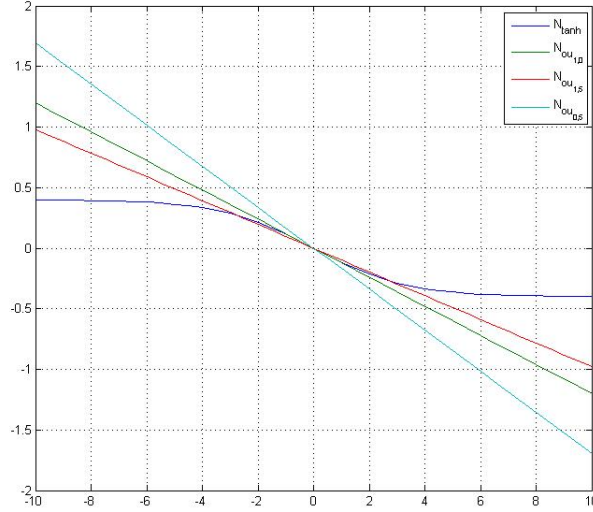


Figura 4.9: Demanda total,  $N$ , para os diversos modelos.

mundo real,  $k > 0$  é o parâmetro de reversão média, e  $\sigma > 0$  é o parâmetro de volatilidade. Do mesmo modo consideramos o tempo de arbitragem finito,  $0 \leq t \leq T < \infty$ . O parâmetro  $c > 0$  mede a não-linearidade no componente de reversão à média do processo descrito pela equação (4.21).

## 4.5 Contrastando com a dinâmica do OU

O parâmetro  $c$  especifica a não-linearidade de reversão à média no nosso processo estocástico (4.21). No caso especial em que  $c = 0$ , a dinâmica do diferencial de preços (4.21) se reduz a:

$$\lim_{c \rightarrow 0} \left[ -\frac{k}{c} \tanh(c(S_t - \bar{S})) dt \right] = -k(S_t - \bar{S}) dt. \quad (4.23)$$

O lado direito da equação (4.23) é precisamente a componente de reversão à média de um processo OU com parâmetros  $k$  e  $\sigma$  e  $A(\tau) = B(\tau) = 0$ . Isso reduz (4.21) a:

$$dS_t \rightarrow -k(S_t - \bar{S}) dt + \sigma dW_t, \quad (4.24)$$

o que implica que o modelo fica reduzido precisamente ao especificado em Jurek e Yang [16].

Nota-se que forçando  $S_t$  muito acima do seu nível natural obtemos:

$$\lim_{S_t \rightarrow \infty} \left[ -\frac{k}{c} \tanh(c(S_t - \bar{S})) dt \right] \rightarrow -\frac{k}{c} dt. \quad (4.25)$$

O lado direito da equação (4.25) implica que a intensidade da reversão à média que puxa  $S_t$  para seu nível natural tem enfraquecido no limite. Isto reduz a dinâmica do diferencial de preços (4.21) para:

$$dS_t \rightarrow -\frac{k}{c}dt + \sigma dW_t. \quad (4.26)$$

Assim vemos que para valores de  $S_t$  muito afastados da média a dinâmica do processo depende apenas dos valores de  $c$  e  $k$ .

## 4.6 Comparação dos Resultados: $N_{OU}$ versus $N_{TANH}$

De acordo com Alsayed e McGroarty [2] foi assumido nesse modelo a utilidade logarítmica, por dois motivos principais. Primeiro, em tempo discreto, os autores citam o trabalho de Breiman [8] que mostrou que investidores que assumem utilidade logarítmica vão, com probabilidade um, ter uma performance superior a qualquer outra estratégia, sem ocorrer falência. Por outro lado Xiong [28] sugere que o teorema de Brieman [8] vale em tempo contínuo. A segunda razão para assumir utilidade logarítmica refere-se a tratabilidade do problema de otimização estocástica resultante em Boguslavsky e Boguslavskaya [7]. Dessa forma foi obtido a seguinte estratégia para um investimento ótimo.

Um arbitrador que maximiza sua utilidade logarítmica diante das oportunidades de investimento (4.21) e (4.22) deve investir

$$N = \left( \frac{-\frac{k}{c} \tanh[c(S - \bar{S})] - r(S - \bar{S})}{\sigma^2} \right) X. \quad (4.27)$$

Para vermos isto, primeiro substituímos (4.21) e (4.22) na equação de restrição orçamentária para obter:

$$\begin{aligned} dX &= NdS + \frac{X - N(S - \bar{S})}{B} dB \\ &= \left( -\frac{k}{c} N \tanh(c(S - \bar{S})) + r(X - N(S - \bar{S})) \right) dt + \sigma N dW. \end{aligned} \quad (4.28)$$

Agora expandindo a função da utilidade esperada e usando o Lema de Itô, temos:

$$dV = \frac{\partial V}{\partial S} dS + \frac{\partial V}{\partial X} dX + \frac{\partial^2 V}{\partial S^2} (dS)^2 + \frac{\partial^2 V}{\partial X^2} (dX)^2 + \frac{\partial^2 V}{\partial S \partial X} (dS \cdot dX) - \frac{\partial V}{\partial \tau} dt. \quad (4.29)$$



Substituindo (4.21) e (4.29), temos:

$$\begin{aligned}
dV &= \frac{\partial V}{\partial X} \left( \left( -\frac{k}{c} N \tanh(c(S - \bar{S})) + r(X - N(S - \bar{S})) \right) dt + \sigma N dW \right) \\
&+ \frac{\partial V}{\partial S} \left( -\frac{k}{c} N \tanh(c(S - \bar{S})) dt + \sigma dW \right) + \frac{1}{2} \sigma^2 \frac{\partial^2 V}{\partial S^2} dt + \frac{1}{2} \sigma^2 N^2 \frac{\partial^2 V}{\partial X^2} dt \\
&+ \frac{1}{2} \sigma^2 N \frac{\partial^2 V}{\partial S \partial X} dt - \frac{\partial V}{\partial \tau} dt
\end{aligned} \tag{4.30}$$

Logo, utilizamos o principio de otimalidade de Bellman  $E_t(dV) = 0$  para derivar a equação de HJB para a otimização deste problema:

$$\begin{aligned}
0 &= \frac{\partial V}{\partial X} \left( -\frac{k}{c} N \tanh(c(S - \bar{S})) + r(X - N(S - \bar{S})) \right) \\
&+ \frac{\partial V}{\partial S} \left( -\frac{k}{c} N \tanh(c(S - \bar{S})) \right) + \frac{1}{2} \sigma^2 \frac{\partial^2 V}{\partial S^2} + \frac{1}{2} \sigma^2 N^2 \frac{\partial^2 V}{\partial X^2} + \frac{1}{2} \sigma^2 N \frac{\partial^2 V}{\partial S \partial X} - \frac{\partial V}{\partial \tau}.
\end{aligned} \tag{4.31}$$

junto com a condição terminal:

$$V(S, X, 0) = 0 \tag{4.32}$$

A equação (4.31) é a equação de HJB para o nosso problema de otimização. A fim de obter a política ótima, maximizamos o lado direito de (4.31) em relação a  $N$ . Dessa forma, temos:

$$0 = \frac{\partial V}{\partial X} \left( -\frac{k}{c} \tanh(c(S - \bar{S})) - r(S - \bar{S}) \right) + \sigma^2 N \frac{\partial^2 V}{\partial X^2} + \frac{1}{2} \sigma^2 \frac{\partial^2 V}{\partial S \partial X}, \tag{4.33}$$

e finalmente resolvendo para  $N$  obtemos:

$$N = - \left( \frac{V_X}{V_{XX}} \right) \left( \frac{-\frac{k}{c} \tanh(c(S - \bar{S})) - r(S - \bar{S})}{\sigma^2} \right) - \left( \frac{V_{SX}}{V_{XX}} \right). \tag{4.34}$$

Postulamos a seguir uma tentativa de solução, em termos de uma forma funcional para  $V(S, X, t)$ , que resolve a equação HJB (4.31), e é dada pela seguinte solução:

$$V(S, X, \tau) = \log(X) + f(S, \tau). \tag{4.35}$$

Aqui  $f$  é uma função arbitrária de  $S$  e  $\tau$ .

A condição suficiente para otimalidade do portfólio é a concavidade da função com relação a variável  $X$  [7]. Substituindo as derivadas parciais relevantes na política de portfólio ótima em (4.34), temos:

$$N = \left( \frac{-\frac{k}{c} \tanh(c(S - \bar{S})) - r(S - \bar{S})}{\sigma^2} \right) X, \tag{4.36}$$

e substituindo (4.36) junto com as derivadas parciais relevantes na equação de HJB (4.31) obtemos uma simplificação da equação de HJB, que vem dada por:

$$\begin{aligned}
0 = & \left( -\frac{k}{c} \tanh(c(S - \bar{S})) \right) f_S - f_\tau + \frac{1}{2} \sigma^2 f_{SS} \\
& - \frac{\left( \frac{k}{c} \tanh(c(S - \bar{S})) - r(S - \bar{S}) \right)^2}{2\sigma^2} \\
& + r \left( 1 - \frac{\left( \frac{k}{c} \tanh(c(S - \bar{S})) - r(S - \bar{S}) \right) (S - \bar{S})}{\sigma^2} \right) \\
& + \frac{k \tanh(c(S - \bar{S})) \left( -\frac{k}{c} \tanh(c(S - \bar{S})) - r(S - \bar{S}) \right)}{\sigma^2 c}.
\end{aligned} \tag{4.37}$$

Assim, a equação de HJB é reduzida em dimensionalidade, e agora só depende de  $S$  e  $\tau$ . Dado que o portfólio ótimo não depende de  $f$  temos que apenas fazer o pressuposto técnico de que existe uma função  $f(S, \tau)$  duas vezes diferenciável em  $S$ , uma vez em  $\tau$ , e que satisfaz a equação (4.37).

Por outro lado é interessante observar que quando o  $c$  decresce para zero em (4.36) obtemos a correspondente estratégia OU ótima, a qual rotulamos  $N_{OU}$ :

$$\begin{aligned}
N_{OU} &= \lim_{c \rightarrow 0} \left( \frac{-\frac{k}{c} \tanh[c(S - \bar{S})] - r(S - \bar{S})}{\sigma^2 c} \right) X \\
&= \left( \frac{-(k + r)(S - \bar{S})}{\sigma^2} \right) X.
\end{aligned} \tag{4.38}$$

Para um conjunto de parâmetros e  $\bar{S} = 0$ ,  $N_{OU}$  gera as especificações em Jurek e Yang [16]. Fazendo  $r = 0$  produz as especificações em Boguslavsky e Boguslavskaya [7].

Agora que têm caracterizado a estratégia de portfólio ótimo (4.36) e exploramos suas propriedades, estamos prontos para implementar nossos resultados com os dados históricos. Fazemos isso testando a eficácia empírica da estratégia  $N_{TANH}$  (4.36) contra a estratégia  $N_{OU}$  (4.38). Adotamos como coeficiente de não-linearidade  $c = \frac{1}{2\delta}$  no  $N_{TANH}$ .

No gráfico 4.10 podemos comparar as posições distintas assumidas para um par com  $k$  elevado e outro par com  $k$  moderado para o modelo de OU e da tangente hiperbólica. Para o caso do  $k$  moderado o comportamento é similar, enquanto que para o caso do  $k$  elevado, eles diferem bastante.

A tabela 4.2 faz um comparativo dos portfólios setoriais para o caso de modelo OU com  $\gamma = 0,5$ , que apresentou melhor resultado na seção anterior, com o modelo da tangente hiperbólica. Uma coisa que ficou clara foi a redução na volatilidade gerada pelo modelo da tangente hiperbólica quando comparada com o modelo anterior. Embora, essa redução também tenha gerado redução nos retornos para alguns casos, quando é feita uma

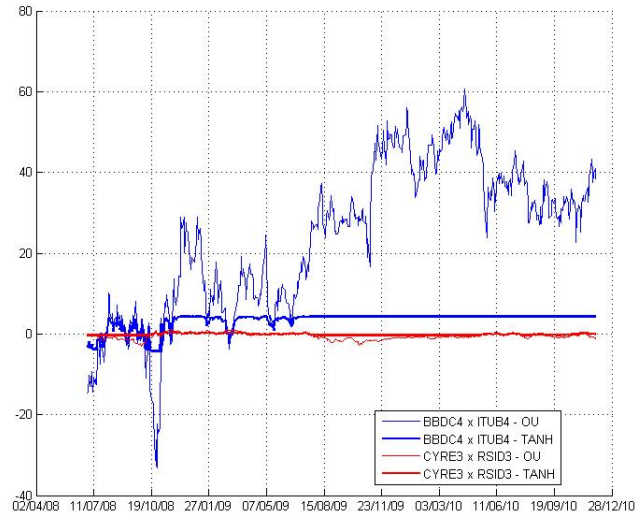


Figura 4.10: Demanda total,  $N$ , para o modelo de OU com  $\gamma = 0,5$  e da tangente hiperbólica para BBDC4 x ITUB4 ( $k$  elevado) e CYRE3 x RSID3 ( $k$  moderado).

avaliação pelo retorno ajustado pode-se constatar uma melhora na qualidade dos resultados gerados pela tangente hiperbólica.

Assim, como ocorreu para o caso do modelo OU, esse modelo da tangente hiperbólica também é bem sensível ao período de calibração, como pode ser constatado nas figuras 4.11, 4.12 e 4.13. Além disso, embora o modelo da tangente hiperbólica assuma posições menores quando comparadas com o modelo OU, ainda sim para pares de ativos com elevada velocidade de reversão à média, ele pode assumir posições de até oito vezes a riqueza, ainda bastante distantes da realidade.

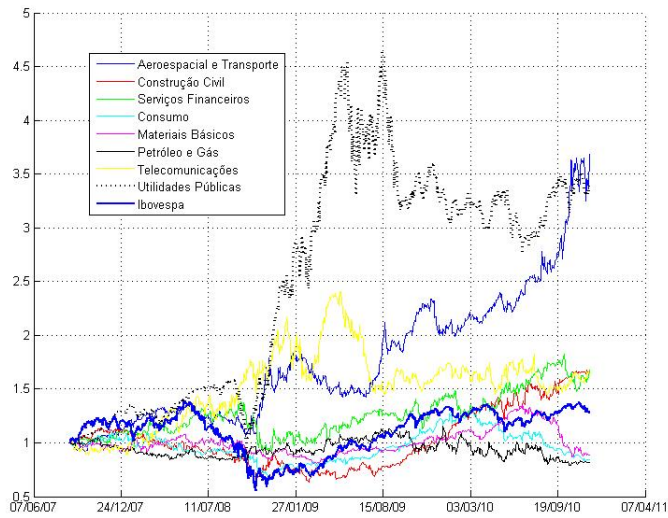


Figura 4.11: *Payoffs* dos diversos portfólios comparados com o Ibovespa, com ajuste de 500 dias.

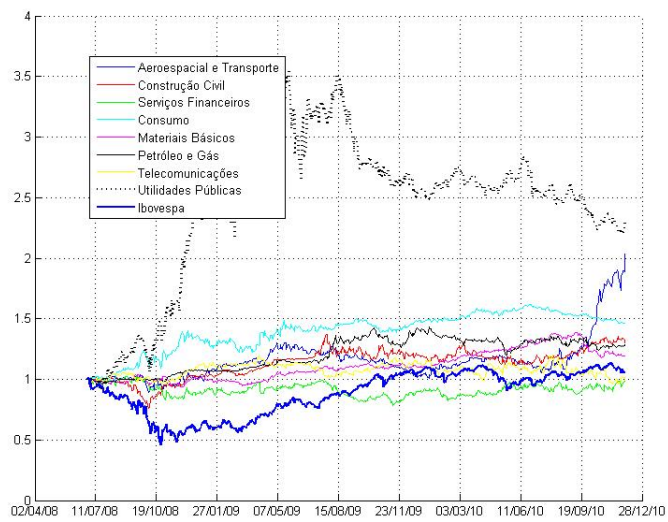


Figura 4.12: *Payoffs* dos diversos portfólios comparados com o Ibovespa, com ajuste de 700 dias.

Comparativo dos Resultados						
Setor	OU com ( $\gamma = 0.5$ )			Tangente Hiperbólica		
	<i>Payoff</i>	Vol.	Ret. Ajust.	<i>Payoff</i>	Vol.	Ret. Ajust.
Aeroespacial e Transporte	14,96%	15,29%	0,91	12,14%	6,76%	1,64
Construção Civil	-3,21%	11,92%	-1,69	4,42%	4,36%	0,77
Serviços Financeiros	-1,52%	15,30%	-0,17	-1,30%	4,67%	-0,50
Consumo	14,27%	12,30%	1,08	6,87%	4,81%	1,21
Materiais Básicos	0,16%	11,16%	-0,08	1,44%	3,52%	0,11
Petróleo e Gás	3,86%	17,94%	0,16	4,19%	3,17%	0,99
Telecomunicações	-4,50%	9,60%	-0,58	-1,11%	3,63%	-0,59
Utlidades Públicas	21,39%	18,75%	1,08	10,76%	9,17%	1,06

Tabela 4.2: Comparação dos resultados do OU com  $\gamma = 0,5$  e da tangente hiperbólica, com ajuste de 700 dias.

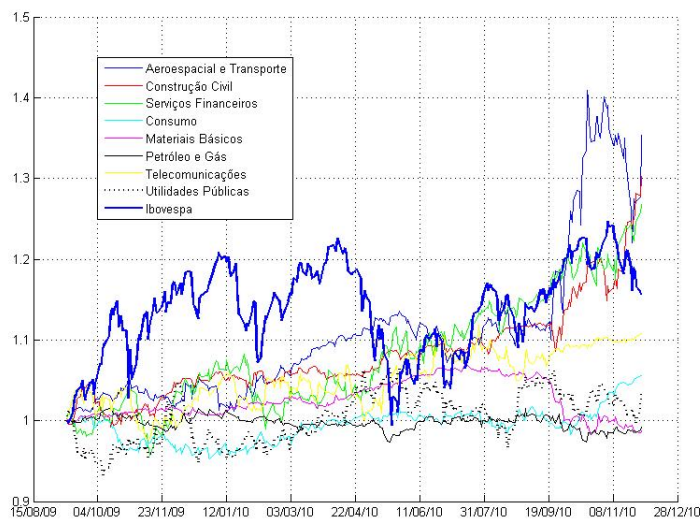


Figura 4.13: *Payoffs* dos diversos portfólios comparados com o Ibovespa, com ajuste de 1000 dias.

# Capítulo 5

## Conclusão

Neste estudo investigou-se modelos de arbitragem estatística entre pares de ativos relacionados, buscando-se aproveitar anomalias no *spread* dos ativos. Dessa maneira, o primeiro passo foi a escolha dos pares via análise dos componentes principais, que mostrou que embora a maior parte dos autovetores tenham um comportamento que se assemelham de uma matriz aleatória, existem alguns componentes principais, uma parcela bem pequena, que de fato apresentam informações relevantes. Assim, o critério de seleção para escolha dos pares foi feita a partir da proximidade das ações no componente principal, dentro de um mesmo setor de atividade.

O modelo de Ornstein-Uhlenbeck de reversão à média foi usado para gerar sinais que sinalizam o quanto o *spread* estava distante da média em unidades de desvio padrão. A partir desses sinais foi estabelecido um critério para compra/venda de determinado ativo que proporcionou excelentes resultados para o caso do mercado brasileiro, gerando retornos anualizados que chegaram a alcançar em alguns casos mais de 30,0%, para um período de cinco anos, com uma baixa volatilidade. Nesse caso foi considerado que o investidor tinha uma riqueza inicial, e que toda vez que houvesse uma oportunidade aplicava-se toda sua riqueza no diferencial de preços sem usar para isso a possibilidade de alavancagem.

Posteriormente foi feito um sinal ajustado pela média do volume negociado. O objetivo desse método era evitar que variações anormais no volume negociado em determinado dia, pudessem estar interferindo na geração dos sinais. Esse método apresentou resultados bastantes consistentes, com *payoffs* superiores aos do modelo anterior, embora com uma volatilidade ligeiramente superior.

Foi implementado também um método no qual todos os pares eram quantificados, seus respectivos sinais colocados em ordem decrescente, e os pares de ativos com maiores sinais eram escolhidos para montar posição, com a verificação sendo feita diariamente. Nesse modelo um par pode ser desfeito antes mesmo de ter seu objetivo atingido, desde que haja outro par em uma situação melhor. O retorno para esse modelo foi positivo, chegando a alcançar

um ganho de até duas vezes o patrimônio em um período de dois anos. Logicamente, a concentração excessiva em um pequeno número de pares traz como consequência um aumento na volatilidade e uma maior dificuldade de implementação da estratégia dependendo do tamanho do capital investido.

A parte restante do trabalho visava responder a questão da escolha do portfólio ótimo e para isso foi apresentado um modelo linear de reversão à média construído a partir do modelo de OU onde as preferências dos investidores estão refletidas no processo. Dessa maneira investidores com nível de aversão ao risco distintas vão ter diferentes comportamentos dado o mesmo diferencial de preços, onde foi introduzido um novo parâmetro chamado de demanda de *hedging* intertemporal que proporciona esse comportamento diferenciado.

Neste caso o portfólio foi otimizado a partir do processo da riqueza do agente e provou-se que quanto mais avesso ao risco um arbitrador mais ele vai investir sua riqueza conforme o *spread* se distancia da média, dado que ele sabe se tratar de um processo de reversão à média. Nesse modelo, está implícito também que o agente usa de alavancagem para poder maximizar sua riqueza, dado que de acordo com o nível do *spread* o agente assume posições que chegavam até oitenta vezes maiores que sua riqueza disponível, embora essas posições na prática não sejam possíveis de serem implementadas. Ainda, foi observado que quanto maior a velocidade de reversão à média e consequentemente menor a volatilidade do *spread*, maior a posição ótima.

Esse modelo apresentou resultados piores em termos de retornos quando comparado com o modelo de geração de sinais, além de apresentar uma volatilidade excessivamente alta em função do grau de alavancagem considerado. Por essa maneira foi considerada como uma possível saída fazer uma restrição de alavancagem de até duas vezes a riqueza. Entretanto, por essa restrição não ter sido considerada previamente na solução do processo de otimização, do ponto de vista matemático está errada.

No último método foi usado um modelo de equilíbrio parcial que buscava principalmente capturar o risco de quebra de circuito inerente nas operações de arbitragem. De acordo com este modelo, quanto maior o diferencial de preços, mais cético fica o investidor com relação a oportunidade de arbitragem. Esse modelo apresentou os mesmos problemas do modelo linear de OU citados anteriormente de excesso de alavancagem e volatilidade, embora em níveis bem inferiores. Esse modelo também apresentou resultados piores que o modelo dos sinais, entretanto melhores que o modelo anterior. Para esses últimos dois modelos também ficou provada a sensibilidade do processo ao período de calibração dos parâmetros, dado que variação do número de dias da calibração provoca mudanças significativas no desempenho dos portfólios dentro dos setores.

# Bibliografia

- [1] Alexander, C.; Giblin, I.; Weddington, W. (2001) COINTEGRATION AND ASSET ALLOCATION: A NEW ACTIVE HEDGE FUND STRATEGY
- [2] Alsayed, H.; McGroarty, F. (2010) OPTIMAL PORTFOLIO SELECTION IN NONLINEAR ARBITRAGE SPREADS
- [3] Avellaneda, M.; Lee, J. (2009) STATISTICAL ARBITRAGE IN THE U.S. EQUITIES MARKET
- [4] Baumeister, J.; Leitão, A. (2002) INTRODUÇÃO À TEORIA DE CONTROLE E PROGRAMAÇÃO DINÂMICA
- [5] Baronyan, S. (2010) INVESTIGATION OF STOCHASTIC PAIRS TRADING STRATEGIES UNDER DIFFERENT VOLATILITY REGIMES
- [6] Bowen, D.; Hutchinson, M.; O’Sullivan, N. (2010) HIGH FREQUENCY EQUITY PAIRS TRADING: TRANSACTION COSTS, SPEED OF EXECUTION AND PATTERNS IN RETURNS
- [7] Boguslavsky, M.; Boguslavskaya, E. (2004) ARBITRAGE UNDER POWER
- [8] Breiman, L. (1961) OPTIMAL GAMBLING SYSTEM FOR FAVORABLE GAMES
- [9] Caldeira, J. (2010) ARBITRAGEM ESTATÍSTICA E ESTRATÉGIA LONG-SHORT PAIRS TRADING, ABORDAGEM DA COINTEGRAÇÃO APLICADA A DADOS BRASILEIROS
- [10] Caldeira, J.; Portugal, M. (2009) ESTRATÉGIAS LONG-SHORT, NEUTRA AO MERCADO, E INDEX-TRACKING BASEADAS EM PORTFÓLIOS COINTEGRADOS
- [11] Dickey, D.; Fuller, W. (1981) LIKELIHOOD RATIO STATISTICS FOR AUTOREGRESSIVE TIME SERIES WITH A UNIT ROOT – ECONOMETRICA, 49: 1057-1072



- [12] Dunis, C.; Giorgioni, G.; Laws, J.; Rudy, J. (2010) STATISTICAL ARBITRAGE AND HIGH-FREQUENCY DATA WITH AN APPLICATION TO EUROSTOXX 50 EQUITIES
- [13] Follmer, H.; Schied, A. (2004) STOCHASTIC FINANCE - AN INTRODUCTION IN DISCRETE TIME
- [14] Gatev, E.; Goetzmann, G.; Rouwenhorst, K. (2006) PAIRS TRADING: PERFORMANCE OF A RELATIVE VALUE ARBITRAGE RULE. THE REVIEW OF FINANCIAL STUDIES, 19:797–827
- [15] Jolliffe, I. (2002) PRINCIPAL COMPONENT ANALYSIS
- [16] Jurek, J.; Yang, H. (2009) DYNAMIC PORTFOLIO SELECTION IN ARBITRAGE
- [17] Khandani, A.; Lo, A. (2007) WHAT HAPPENED TO THE QUANTS IN AUGUST 2007?
- [18] Lages, E. (1995) ÁLGEBRA LINEAR
- [19] Laloux, L.; Cizeau, P.; Potters, M. (2000) RANDOM MATRIX THEORY AND FINANCIAL CORRELATIONS – INTERNATIONAL JOURNAL OF THEORETICAL AND APPLIED FINANCE, 3: 391-397
- [20] Meucci, A. (2005) RISK AND ASSET ALLOCATION
- [21] Plerou, V.; Gopikrishnan, P.; Rosenow, B.; Amaral, L. N.; Guhr, T.; Stanley, H. E. (2002) RANDOM MATRIX APPROACH TO CROSS CORRELATIONS IN FINANCIAL DATA - PHYS. REV., E 65, 066126: 1-18
- [22] Stanley, H.; Plerou, V.; Gabaix, X. (2008) A STATISTICAL PHYSICS VIEW OF FINANCIAL FLUCTUATIONS: EVIDENCE FOR SCALING AND UNIVERSALITY - PHYSICA, A 387: 3967-3981
- [23] Strang, G. (1988) LINEAR ALGEBRA
- [24] Takimoto, E. (2007) PAIRS TRADING IN THE BRASILIAN STOCK MARKET
- [25] Thomaidis, N.; Kondakis, N. (2006) AN INTELLIGENT STATISTICAL ARBITRAGE TRADING SYSTEM
- [26] Velayutham, A.; Lukman, D.; Chiu, J.; Modarresi, K. (2010) HIGH-FREQUENCY TRADING
- [27] Vidyamurthy, G. (2004) PAIRS TRADING, QUANTITATIVE METHODS AND ANALYSIS

- [28] Xiong, W. (2001) CONVERGENCE TRADING WITH WEALTH EFFECTS: AN AMPLIFICATION MECHANISM IN FINANCIAL MARKETS – JOURNAL OF FINANCIAL ECONOMICS 62 (2001) 247-292
- [29] Yong, J.; Zhou, X. (1999) STOCHASTIC CONTROLS, HAMILTONIAN SYSTEMS AND HJB EQUATIONS

# Apêndice A

## Resolução de equações diferenciais

**Caso 1: Investidor com utilidade logarítmica** ( $\gamma = 1$ ) Para resolver a primeira equação usamos o termo integrante  $e^{2k\tau}$ , de onde obtemos

$$A'(\tau)e^{2k\tau} + (e^{2k\tau})'A(\tau) = \frac{(k+r)^2}{2\sigma^2}e^{2k\tau}$$

logo, integrando

$$A(\tau)e^{2k\tau} - A(0) = \int_0^\tau [A(x)e^{2kx}]' dx = \int_0^\tau \frac{(k+r)^2}{2\sigma^2}e^{2kx} dx = \frac{(k+r)^2}{4k\sigma^2}(e^{2k\tau} - 1)$$

e usando a condição  $A(0) = 0$  obtemos

$$A(\tau) = \frac{(k+r)^2(1 - e^{-2k\tau})}{4k\sigma^2}$$

Usando esta solução vamos achar  $B(\tau)$ . Da equação (4.13) temos que

$$\begin{aligned} B'(\tau) &= -kB(\tau) - \frac{k(k+r)\bar{S}}{\sigma^2} + 2k\bar{S}A(\tau) \\ &= -kB(\tau) - \frac{k(k+r)\bar{S}}{\sigma^2} + 2k\bar{S}\frac{(k+r)^2(1 - e^{-2k\tau})}{4k\sigma^2} \\ &= -kB(\tau) + \frac{(k+r)\bar{S}}{\sigma^2}\left(-k + \frac{1}{2}(k+r)(1 - e^{-2k\tau})\right) \end{aligned}$$

logo, usando o fator integrante  $e^{k\tau}$  temos que

$$(e^{k\tau}B(\tau))' = e^{k\tau}B'(\tau) + ke^{k\tau}B(\tau) = \frac{(k+r)\bar{S}}{\sigma^2}\left(-ke^{k\tau} + \frac{1}{2}(k+r)(e^{k\tau} - e^{-k\tau})\right)$$

e integrando, usando o fato  $B(0) = 0$ , temos que

$$\begin{aligned} e^{k\tau}B(\tau) &= \frac{(k+r)\bar{S}}{\sigma^2} \left[ -\int_0^\tau ke^{kx} dx + (k+r) \int_0^\tau \sinh(kx) dx \right] \\ &= \frac{(k+r)\bar{S}}{\sigma^2} \left[ -e^{k\tau} + 1 + \frac{(k+r)}{k}(\cosh(k\tau) - \cosh(0)) \right] \\ &= \frac{(k+r)\bar{S}}{\sigma^2} \left[ -e^{k\tau} + 1 + \frac{(k+r)}{k} \left( \frac{e^{k\tau} + e^{-k\tau}}{2} - 1 \right) \right] \end{aligned}$$

de onde se segue que

$$\begin{aligned}
B(\tau) &= \frac{(k+r)\bar{S}}{\sigma^2} \left[ -1 + e^{-k\tau} + \frac{(k+r)}{k} \left( \frac{1+e^{-2k\tau}}{2} - e^{-k\tau} \right) \right] \\
&= \frac{(k+r)\bar{S}}{\sigma^2} \left[ -1 + e^{-k\tau} + \frac{(k+r)(1-e^{-k\tau})^2}{2k} \right] \\
&= \frac{(k+r)(1-e^{-k\tau})}{\sigma^2} \left[ -1 + \frac{(k+r)(1-e^{-k\tau})}{2k} \right] \bar{S}
\end{aligned}$$

de onde obtemos:

$$B(\tau) = -\frac{(k+r)(k-r+e^{-k\tau}(k+r))(1-e^{-k\tau})}{2k\sigma^2} \bar{S}$$

Finalmente, dado que temos  $A(\tau)$  e  $B(\tau)$ , estamos em condições de determinar  $C(\tau)$ . Da equação (4.14) temos que

$$\begin{aligned}
C'(\tau) &= r + \sigma^2 A(\tau) + \frac{1}{2} k \bar{S} \left( 2B(\tau) + \frac{k\bar{S}}{\sigma^2} \right) \\
&= r + \frac{(k+r)^2(1-e^{-2k\tau})}{4k} + \frac{1}{2} k \bar{S} \left( \frac{k\bar{S}}{\sigma^2} - \frac{(k+r)(k-r+e^{-k\tau}(k+r))(1-e^{-k\tau})}{k\sigma^2} \bar{S} \right) \\
&= r + \frac{(k+r)^2(1-e^{-2k\tau})}{4k} + \frac{1}{2} \left( \frac{k\bar{S}}{\sigma} \right)^2 - \frac{1}{2} \frac{(k+r)(k-r+e^{-k\tau}(k+r))(1-e^{-k\tau})}{\sigma^2} \bar{S}^2
\end{aligned}$$

Logo, integrando temos que

$$\begin{aligned}
C(\tau) &= \frac{(k+r)^2}{4k} \int_0^\tau (1-e^{-2kx}) dx - \frac{(k+r)\bar{S}^2}{2\sigma^2} \int_0^\tau (k-r+e^{-kx}(k+r))(1-e^{-kx}) dx \\
&\quad + \left[ r + \frac{1}{2} \left( \frac{k\bar{S}}{\sigma} \right)^2 \right] \tau \\
&= \frac{(k+r)^2}{4k} \left[ \tau + \frac{e^{-2k\tau}-1}{2k} \right] - \frac{(k+r)\bar{S}^2}{2\sigma^2} \left[ (k-r) \int_0^\tau (1-e^{-kx}) dx + (k+r) \int_0^\tau (e^{-kx} - e^{-2kx}) dx \right] \\
&\quad + \left[ r + \frac{1}{2} \left( \frac{k\bar{S}}{\sigma} \right)^2 \right] \tau \\
&= \frac{(k+r)^2}{4k} \left[ \tau - \frac{e^{-k\tau} \sinh(k\tau)}{k} \right] - \frac{(k+r)\bar{S}^2}{2\sigma^2} \left[ (k-r) \left( \tau + \frac{e^{-k\tau}-1}{k} \right) + (k+r) \left( \frac{1-e^{-k\tau}}{k} + \frac{e^{-2k\tau}-1}{2k} \right) \right] \\
&\quad + \left[ r + \frac{1}{2} \left( \frac{k\bar{S}}{\sigma} \right)^2 \right] \tau \\
&= -\left( \frac{k+r}{2k} \right)^2 e^{-k\tau} \sinh(k\tau) - \frac{(k+r)\bar{S}^2}{2\sigma^2} \left[ \frac{(k-r)(e^{-k\tau}-1)}{k} + \frac{(k+r)(e^{-k\tau}-1)^2}{2k} \right] \\
&\quad + \left[ r + \frac{1}{2} \left( \frac{k\bar{S}}{\sigma} \right)^2 + \frac{(k+r)^2}{4k} - \frac{(k^2-r^2)\bar{S}^2}{2\sigma^2} \right] \tau \\
&= -\left( \frac{k+r}{2k} \right)^2 e^{-k\tau} \sinh(k\tau) + \frac{(k+r)(1-e^{-k\tau})\bar{S}^2}{2\sigma^2} \left[ \frac{(k-r)}{k} + \frac{(k+r)(e^{-k\tau}-1)}{2k} \right] \\
&\quad + \left[ r + \frac{(k+r)^2}{4k} + \frac{r^2\bar{S}^2}{2\sigma^2} \right] \tau
\end{aligned}$$

de onde se segue que

$$\begin{aligned}
C(\tau) = & -\left(\frac{k+r}{2k}\right)^2 \sinh(k\tau)e^{-k\tau} \\
& + \frac{(k+r)(k-3r+e^{-k\tau}(k+r))(1-e^{-k\tau})}{4k\sigma^2} \bar{S}^2 \\
& + \left(\frac{k^2+6kr+r^2}{4k} + \frac{1}{2}\left(\frac{r\bar{S}}{\sigma}\right)^2\right)\tau
\end{aligned}$$

### Caso 2: Investidor Geral ( $\gamma \neq 1$ )

Vamos resolver a equação de Ricatti,

$$\frac{dA(\tau)}{d\tau} = c_1A(\tau)^2 + 2c_2A(\tau) + c_3$$

onde  $c_1 = \frac{2\sigma^2}{\gamma}$ ,  $c_2 = -\left(\frac{k}{\gamma} + r\left(\frac{1-\gamma}{\gamma}\right)\right)$  e  $c_3 = \frac{1}{2}\left(\frac{1-\gamma}{\gamma}\right)\left(\frac{k+r}{\sigma}\right)^2$ ,

Desde que  $c_1 \neq 0$ , podemos fazer uma mudança de variável

$$v(\tau) = c_1A(\tau) \tag{A.1}$$

de onde temos que

$$v'(\tau) = c_1A'(\tau) \tag{A.2}$$

$$\begin{aligned}
v'(\tau) &= c_1A'(\tau) \\
&= c_1[c_1A(\tau)^2 + 2c_2A(\tau) + c_3] \\
&= (c_1A(\tau))^2 + 2c_2(c_1A(\tau)) + c_1c_3 \\
&= (v(\tau))^2 + 2c_2(v(\tau)) + c_1c_3
\end{aligned}$$

e fazendo mais uma mudança de variável

$$\frac{\mu'(\tau)}{\mu(\tau)} = -v(\tau) \tag{A.3}$$

temos que

$$\begin{aligned}
-\left(\frac{\mu'(\tau)}{\mu(\tau)}\right)' &= (v(\tau))' = (v(\tau))^2 + 2c_2(v(\tau)) + c_1c_3 \\
-\left[\frac{\mu''(\tau)\mu(\tau) - (\mu'(\tau))^2}{(\mu(\tau))^2}\right] &= \left[\frac{-\mu'(\tau)}{\mu(\tau)}\right]^2 + 2c_2\left[\frac{-\mu'(\tau)}{\mu(\tau)}\right] + c_1c_3 \\
-\frac{\mu''(\tau)}{\mu(\tau)} &= -2c_2\frac{\mu'(\tau)}{\mu(\tau)} + c_1c_3
\end{aligned}$$

e assim

$$\mu''(\tau) - 2c_2\mu'(\tau) + c_1c_3\mu(\tau) = 0 \tag{A.4}$$

cuja equação característica é

$$\lambda^2 - 2c_2\lambda + c_1c_3 = 0$$

e cujo discriminante é

$$\begin{aligned} \Delta &= (-2c_2)^2 - 4c_1c_3 \\ &= 4[c_2^2 - c_1c_3] \\ &= 4\left[\left(\frac{k}{r} + r\left(\frac{1-\gamma}{\gamma}\right)\right)^2 - \frac{2\sigma^2}{\gamma} \frac{1}{2}\left(\frac{1-\gamma}{\gamma}\right)\left(\frac{k+r}{\sigma}\right)^2\right] \\ &= \frac{4}{\gamma^2}\left[k^2\gamma - (1-\gamma)r^2 + r^2(1-\gamma)^2\right] \\ &= 4\left[\frac{k^2 - (1-\gamma)r^2}{\gamma}\right] \end{aligned}$$

logo, as soluções são da forma

$$\lambda = \frac{2c_2 \pm \sqrt{\Delta}}{2} \quad (\text{A.5})$$

Agora, nota-se que

$$\Delta < 0 \Leftrightarrow k^2 < (1-\gamma)r^2 \Leftrightarrow \gamma < 1 - \left(\frac{\kappa}{r}\right)^2$$

$$\Delta = 0 \Leftrightarrow \gamma = 1 - \left(\frac{\kappa}{r}\right)^2$$

$$\Delta > 0 \Leftrightarrow \gamma > 1 - \left(\frac{\kappa}{r}\right)^2$$

Logo, fazendo  $\gamma_0 = 1 - \left(\frac{\kappa}{r}\right)^2$  temos que se  $\gamma < \gamma_0$ , então  $\Delta < 0$  e as soluções da equação característica são

$$\lambda_1 = \frac{2c_2 + \sqrt{-\Delta}i}{2}$$

e

$$\lambda_2 = \frac{2c_2 - \sqrt{-\Delta}i}{2}$$

A solução geral da equação (A.4) é

$$\mu(\tau) = k_1e^{\lambda_1\tau} + k_2e^{\lambda_2\tau}$$

onde  $k_1$  e  $k_2$  são constantes à determinar.

Mas, como

$$\mu'(\tau) = \lambda_1k_1e^{\lambda_1\tau} + \lambda_2k_2e^{\lambda_2\tau}$$

temos que

$$v(\tau) = -\frac{\mu'(\tau)}{\mu(\tau)} = -\frac{\lambda_1k_1e^{\lambda_1\tau} + \lambda_2k_2e^{\lambda_2\tau}}{k_1e^{\lambda_1\tau} + k_2e^{\lambda_2\tau}}$$

e usando a condição  $A(0) = 0$ , temos que  $\lambda_1 k_1 + \lambda_2 k_2 = 0$  e assim  $k_2 = -\frac{\lambda_1 k_1}{\lambda_2}$ . Logo vemos que

$$\begin{aligned}
A(\tau) &= \frac{v(\tau)}{c_1} = -\frac{1}{c_1} \frac{\lambda_1 \lambda_2 [e^{\lambda_1 \tau} - e^{\lambda_2 \tau}]}{\lambda_2 e^{\lambda_1 \tau} - \lambda_1 e^{\lambda_2 \tau}} \\
&= -\frac{1}{c_1} \frac{c_1 c_3 [e^{\lambda_1 \tau} - e^{\lambda_2 \tau}]}{\lambda_2 e^{\lambda_1 \tau} - \lambda_1 e^{\lambda_2 \tau}} = -\frac{c_3 e^{c_2 \tau} [e^{\frac{i\sqrt{-\Delta}\tau}{2}} - e^{\frac{-i\sqrt{-\Delta}\tau}{2}}]}{e^{c_2 \tau} [\lambda_2 e^{\frac{i\sqrt{-\Delta}\tau}{2}} - \lambda_1 e^{\frac{-i\sqrt{-\Delta}\tau}{2}}]} \\
&= -\frac{c_3 \operatorname{sen}\left(\frac{-\sqrt{-\Delta}\tau}{2}\right)}{c_2 \operatorname{sen}\left(\frac{-\sqrt{-\Delta}\tau}{2}\right) - \frac{\sqrt{-\Delta}}{2} \cos\left(\frac{\sqrt{-\Delta}\tau}{2}\right)} \\
&= -\frac{4c_1 c_3 \operatorname{sen}\left(\frac{-\sqrt{-\Delta}\tau}{2}\right)}{4c_1 \left[ \frac{\sqrt{-\Delta}}{2} \cos\left(\frac{\sqrt{-\Delta}\tau}{2}\right) - c_2 \operatorname{sen}\left(\frac{\sqrt{-\Delta}\tau}{2}\right) \right]} \\
&= \frac{1}{4c_1} \left[ \frac{(4c_2^2 - \Delta) \operatorname{sen}\left(\frac{\sqrt{-\Delta}\tau}{2}\right)}{\frac{\sqrt{-\Delta}}{2} \cos\left(\frac{\sqrt{-\Delta}\tau}{2}\right) - c_2 \operatorname{sen}\left(\frac{\sqrt{-\Delta}\tau}{2}\right)} \right] \\
&= \frac{1}{2c_1} \left[ \frac{\left(\frac{4c_2^2}{\sqrt{-\Delta}} + \sqrt{-\Delta}\right) \operatorname{sen}\left(\frac{\sqrt{-\Delta}\tau}{2}\right)}{\cos\left(\frac{\sqrt{-\Delta}\tau}{2}\right) - \frac{2c_2}{\sqrt{-\Delta}} \operatorname{sen}\left(\frac{\sqrt{-\Delta}\tau}{2}\right)} \right] \\
&= \frac{1}{2c_1} \left[ -2c_2 + \frac{\sqrt{-\Delta} \operatorname{sen}\left(\frac{\sqrt{-\Delta}\tau}{2}\right) + 2c_2 \cos\left(\frac{\sqrt{-\Delta}\tau}{2}\right)}{\cos\left(\frac{\sqrt{-\Delta}\tau}{2}\right) - \frac{2c_2}{\sqrt{-\Delta}} \operatorname{sen}\left(\frac{\sqrt{-\Delta}\tau}{2}\right)} \right] \\
&= \frac{1}{2c_1} \left[ -2c_2 + \frac{\sqrt{-\Delta} \left( \operatorname{sen}\left(\frac{\sqrt{-\Delta}\tau}{2}\right) + \frac{2c_2}{\sqrt{-\Delta}} \cos\left(\frac{\sqrt{-\Delta}\tau}{2}\right) \right)}{\cos\left(\frac{\sqrt{-\Delta}\tau}{2}\right) - \frac{2c_2}{\sqrt{-\Delta}} \operatorname{sen}\left(\frac{\sqrt{-\Delta}\tau}{2}\right)} \right]
\end{aligned}$$

E fazendo  $\theta = \tan^{-1}\left(\frac{2c_2}{\sqrt{-\Delta}}\right)$  temos

$$\begin{aligned}
A(\tau) &= \frac{1}{2c_1} \left[ -2c_2 + \frac{\sqrt{-\Delta} \left( \sin\left(\frac{\sqrt{-\Delta}\tau}{2}\right) + \tan\theta \cos\left(\frac{\sqrt{-\Delta}\tau}{2}\right) \right)}{\cos\left(\frac{\sqrt{-\Delta}\tau}{2}\right) - \tan\theta \sin\left(\frac{\sqrt{-\Delta}\tau}{2}\right)} \right] \\
&= \frac{1}{2c_1} \left[ -2c_2 + \frac{\sqrt{-\Delta} \sin\left(\frac{\sqrt{-\Delta}\tau}{2} + \theta\right)}{\cos\left(\frac{\sqrt{-\Delta}\tau}{2} + \theta\right)} \right] \\
&= -\frac{c_2}{c_1} + \frac{\sqrt{-\Delta}}{2c_1} \tan \left[ \frac{\sqrt{-\Delta}\tau}{2} + \tan^{-1} \left( \frac{2c_2}{\sqrt{-\Delta}} \right) \right]
\end{aligned}$$

Agora, veremos o caso em que  $\gamma = \gamma_0$ . Neste caso  $\Delta = 0$  e temos que a equação característica tem  $c_2$  como a única solução de multiplicidade 2. A solução  $\mu(\tau)$  tem a forma geral

$$\mu(\tau) = k_1 e^{c_2\tau} + k_2 \tau e^{c_2\tau}$$

Logo,

$$\mu'(\tau) = k_1 c_2 e^{c_2\tau} + k_2 e^{c_2\tau} + k_2 c_2 \tau e^{c_2\tau}$$

e então

$$v(\tau) = -\frac{\mu'(\tau)}{\mu(\tau)} = -\frac{(k_1 c_2 + k_2 + k_2 c_2 \tau) e^{c_2\tau}}{(k_1 + k_2 + k_2 \tau) e^{c_2\tau}} = -\frac{k_1 c_2 + k_2 + k_2 c_2 \tau}{k_1 + k_2 + k_2 \tau}$$

E como

$$A(\tau) = \frac{v(\tau)}{c_1} = -\frac{k_1 c_2 + k_2 + k_2 c_2 \tau}{c_1 (k_1 + k_2 + k_2 \tau)}$$

E usando o fato que  $A(0) = 0$ , temos que

$$k_1 c_2 + k_2 = 0$$

Logo,

$$\begin{aligned}
A(\tau) &= -\frac{k_1 c_2^2 \tau}{c_1 (k_1 - k_2 c_2 \tau)} = \frac{c_2^2 \tau}{c_1 (1 - c_2 \tau)} = \frac{c_2}{c_1} \left[ \frac{c_2 \tau}{1 - c_2 \tau} \right] \\
&= \frac{c_2}{c_1} \left[ -1 + \frac{1}{1 - c_2 \tau} \right] = -\frac{c_2}{c_1} \left[ 1 + \frac{1}{c_2 \tau - 1} \right]
\end{aligned}$$

Finalmente temos o caso em que  $\gamma > \gamma_0$ . Neste caso temos que  $\Delta > 0$  e portanto as soluções da equação característica são:



$$\lambda_1 = c_2 + \frac{\sqrt{\Delta}}{2} \quad \text{e} \quad \lambda_2 = c_2 - \frac{\sqrt{\Delta}}{2}$$

A solução geral de  $\mu(\tau)$  é

$$\mu(\tau) = k_1 e^{\lambda_1 \tau} + k_2 e^{\lambda_2 \tau}$$

logo

$$A(\tau) = \frac{v(\tau)}{c_1} = -\frac{1}{c_1} \frac{\mu'(\tau)}{\mu(\tau)} = -\frac{1}{c_1} \frac{k_1 \lambda_1 e^{\lambda_1 \tau} + k_2 \lambda_2 e^{\lambda_2 \tau}}{k_1 e^{\lambda_1 \tau} + k_2 e^{\lambda_2 \tau}}$$

E usando a condição  $A(0) = 0$ , temos que  $k_1 \lambda_1 + k_2 \lambda_2 = 0$  e assim

$$\begin{aligned} A(\tau) &= \frac{1}{c_1} \frac{\lambda_1 \lambda_2 [e^{\lambda_1 \tau} - e^{\lambda_2 \tau}]}{\lambda_2 e^{\lambda_1 \tau} - \lambda_1 e^{\lambda_2 \tau}} = -\frac{c_3 e^{c_2 \tau} [e^{\frac{\sqrt{\Delta} \tau}{2}} - e^{-\frac{\sqrt{\Delta} \tau}{2}}]}{e^{c_2 \tau} [\lambda_2 e^{\frac{\sqrt{\Delta} \tau}{2}} - \lambda_1 e^{-\frac{\sqrt{\Delta} \tau}{2}}]} \\ &= -\frac{2c_3 \sinh\left(\frac{\sqrt{\Delta} \tau}{2}\right)}{c_2 \left(e^{\frac{\sqrt{\Delta} \tau}{2}} - e^{-\frac{\sqrt{\Delta} \tau}{2}}\right) - \frac{\sqrt{\Delta}}{2} \left(e^{\frac{\sqrt{\Delta} \tau}{2}} + e^{-\frac{\sqrt{\Delta} \tau}{2}}\right)} \\ &= \frac{4c_1 c_3}{4c_1 \left(\frac{\sqrt{\Delta}}{2} \cosh\left(\frac{\sqrt{\Delta} \tau}{2}\right) - c_2 \sinh\left(\frac{\sqrt{\Delta} \tau}{2}\right)\right)} \sinh\left(\frac{\sqrt{\Delta} \tau}{2}\right) \\ &= \frac{(4c_2^2 - \Delta)}{4c_1 \left(\frac{\sqrt{\Delta}}{2} \cosh\left(\frac{\sqrt{\Delta} \tau}{2}\right) - c_2 \sinh\left(\frac{\sqrt{\Delta} \tau}{2}\right)\right)} \sinh\left(\frac{\sqrt{\Delta} \tau}{2}\right) \\ &= \frac{1}{2c_1} \left[ \frac{\left(\frac{4c_2^2}{\sqrt{\Delta}} - \sqrt{\Delta}\right) \sinh\left(\frac{\sqrt{\Delta} \tau}{2}\right)}{\left(\cosh\left(\frac{\sqrt{\Delta} \tau}{2}\right) - \frac{2c_2}{\sqrt{\Delta}} \sinh\left(\frac{\sqrt{\Delta} \tau}{2}\right)\right)} \right] \\ &= \frac{1}{2c_1} \left[ -2c_2 + \frac{\sqrt{\Delta} \left(\frac{2c_2}{\sqrt{\Delta}} \cosh\left(\frac{\sqrt{\Delta} \tau}{2}\right) - \sinh\left(\frac{\sqrt{\Delta} \tau}{2}\right)\right)}{\left(\cosh\left(\frac{\sqrt{\Delta} \tau}{2}\right) - \frac{2c_2}{\sqrt{\Delta}} \sinh\left(\frac{\sqrt{\Delta} \tau}{2}\right)\right)} \right] \end{aligned}$$

Agora, temos dois casos. O primeiro caso é quando  $\left(\frac{2c_2}{\sqrt{\Delta}}\right) < 1$ , isto é

$$\frac{4c_2^2}{\Delta} < 1$$

que é equivalente a

$$4c_2^2 < 4c_2^2 - 4c_1 c_3$$

e isto acontece se e somente se

$$\frac{2\sigma^2}{\gamma} \frac{1}{2} \left( \frac{1-\gamma}{\gamma} \right) \left( \frac{k+r}{\sigma} \right) < 0$$

que é equivalente a que  $\gamma > 1$ . Neste caso, fazendo

$$\theta = \tanh^{-1} \left( \frac{2c_2}{\sqrt{\Delta}} \right)$$

$$\begin{aligned} A(\tau) &= \frac{1}{2c_1} \left[ -2c_2 + \frac{\sqrt{\Delta} \left( \tanh(\theta) \cosh \left( \frac{\sqrt{\Delta}\tau}{2} \right) - \sinh \left( \frac{\sqrt{\Delta}\tau}{2} \right) \right)}{\cosh \left( \frac{\sqrt{\Delta}\tau}{2} \right) - \tanh(\theta) \sinh \left( \frac{\sqrt{\Delta}\tau}{2} \right)} \right] \\ &= -\frac{c_2}{c_1} + \frac{\sqrt{\Delta}}{2c_1} \frac{\sinh \left( \theta - \frac{\sqrt{\Delta}\tau}{2} \right)}{\cosh \left( \theta - \frac{\sqrt{\Delta}\tau}{2} \right)} \\ &= -\frac{c_2}{c_1} + \frac{\sqrt{\Delta}}{2c_1} \tanh \left( -\frac{\sqrt{\Delta}\tau}{2} + \tanh^{-1} \left( \frac{2c_2}{\sqrt{\Delta}} \right) \right) \end{aligned}$$

Por outro lado no caso em que  $\gamma_0 < \gamma < 1$ , temos que  $\left| \frac{2c_2}{\sqrt{\Delta}} \right| > 1$  e então fazendo  $\theta = \coth^{-1} \left( \frac{2c_2}{\sqrt{\Delta}} \right)$ , temos

$$\begin{aligned} A(\tau) &= \frac{1}{2c_1} \left[ -2c_2 + \frac{\sqrt{\Delta} \left( \coth(\theta) \cosh \left( \frac{\sqrt{\Delta}\tau}{2} \right) - \sinh \left( \frac{\sqrt{\Delta}\tau}{2} \right) \right)}{\cosh \left( \frac{\sqrt{\Delta}\tau}{2} \right) - \coth(\theta) \sinh \left( \frac{\sqrt{\Delta}\tau}{2} \right)} \right] \\ &= -\frac{c_2}{c_1} + \frac{\sqrt{\Delta}}{2c_1} \frac{\cosh \left( \theta - \frac{\sqrt{\Delta}\tau}{2} \right)}{\sinh \left( \theta - \frac{\sqrt{\Delta}\tau}{2} \right)} \\ &= -\frac{c_2}{c_1} + \frac{\sqrt{\Delta}}{2c_1} \coth \left( -\frac{\sqrt{\Delta}\tau}{2} + \coth^{-1} \left( \frac{2c_2}{\sqrt{\Delta}} \right) \right) \end{aligned}$$

# Apêndice B

## Códigos do MatLab

### B.1 Componentes Principais

```
function [ind,pares,coef]=pcaport(precos,autovetor)
Autovetor: escolhe o autovetor desejado
Precos: preços dos ativos em ordem crescente de tempo
[tempo,ativos]=size(precos);
```

Calculando os retornos

```
R=zeros(tempo-1,ativos);

for i=1:ativos,
    R(:,i)=(precos(2:end,i)-precos(1:end-1,i))./precos(1:end-1,i);
end;
```

Calculando a média (mR) e variância (s2R) dos retornos

```
mR=zeros(1,ativos);
s2R=zeros(1,ativos);

for i=1:ativos,
    mR(i)=sum(R(:,i))/(tempo-1);
    s2R(i)=(R(:,i)-mR(i))'*(R(:,i)-mR(i));
    s2R(i)=s2R(i)/(tempo-2);
end;
```

Normalizando os retornos

```
Y=zeros(tempo-1,ativos);
for i=1:ativos,
```

```

    Y(:,i)=(R(:,i)-mR(i))/sqrt(s2R(i));
end;

```

Calculando a matriz de covariância

```

Mcov=zeros(ativos);
for i=1:ativos,
    for j=1:ativos,
        Mcov(i,j)=(Y(:,i)'*Y(:,j))/(tempo-2);
    end;
end;

```

Fazendo a análise de componentes principais da matriz de covariância

```

[coef,~,~]=pcacov(Mcov);

if sum(coef(:,1))<0, coef=-coef; end;

[~,ind]=sort(coef(:,autovetor),'descend');

```

```

n=floor(ativos/2);
pares=zeros(n,2);
for i=1:n,
    pares(i,:)=[ind(2*i-1),ind(2*i)];
end;

```

## B.2 Construção dos Sinais

```
function s=sinal1(P1,P2,janela)
```

P1: vetor coluna de preços da primeira ação em ordem crescente de tempo

P2: idem.

Tamanho: número de dados nos vetores de retorno

Janela: número de dias usados para o ajuste

```

tamanho=length(P1)-1;
R1=P1(2:end)./P1(1:end-1)-1;
R2=P2(2:end)./P2(1:end-1)-1;
s=zeros(tamanho-janela+1,1);

```

```
betas=zeros(tamanho-janela+1,1);
```

```

for i=janela:length(R2),
    RS=R1(i-(janela-1):i);

```

```

RI=R2(i-(janela-1):i);

n=janela;
sx=sum(RI);
sy=sum(RS);
sxy=sum(RI.*RS);
sxx=sum(RI.*RI);
beta=(n*sxy-sx*sy)/(n*sxx-sx^2);
alpha=(sy-beta*sx)/n;

raux=zeros(length(RI),2);
raux(:,2)=RI;
raux(:,1)=ones(length(RI),1);
b=regress(RS,raux);
alpha=b(1);
beta=b(2);

epsilon=RS-alpha-beta*RI;

S=zeros(1,janela);
S(1)=epsilon(1);

for j=2:janela,
    S(j)=S(j-1)+epsilon(j);
end;
n=length(S)-1;
sx=sum(S(1:end-1));
sy=sum(S(2:end));
sxy=sum(S(1:end-1).*S(2:end));
sxx=sum(S(1:end-1).*S(1:end-1));
beta=(n*sxy-sx*sy)/(n*sxx-sx^2);
alpha=(sy-beta*sx)/n;

Sy=S(2:end);
Sx=S(1:end-1);
raux=zeros(length(Sy),2);
raux(:,2)=Sx';
raux(:,1)=ones(length(Sy),1);
b=regress(Sy',raux);
betas(i-janela+1)=S(end);
alpha=b(1);
beta=b(2);

```

```

zeta=S(2:end)-alpha-beta*S(1:end-1);
mzeta=sum(zeta)/length(zeta);
varzeta=(zeta-mzeta)*(zeta-mzeta)';
varzeta=varzeta/(length(zeta));
m=alpha/(1-beta);
sigmaeq=sqrt(varzeta/(1-beta^2));
% mudando o sinal
s(i-(janela-1))=-alpha.*sqrt(1+beta)./(sqrt((1-beta)*varzeta));

end;

media=sum(betas)/length(betas);

```

### B.3 Calibração do Modelo de Reversão à Média

```

function [mu,sigma,lambda] = ou_calibrate_ls(S,delta)
n = length(S)-1;

Sx = sum( S(1:end-1) );
Sy = sum( S(2:end) );
Sxx = sum( S(1:end-1).^2 );
Sxy = sum( S(1:end-1).*S(2:end) );
Syy = sum( S(2:end).^2 );

a = ( n*Sxy - Sx*Sy ) / ( n*Sxx - Sx^2 );
b = ( Sy - a*Sx ) / n;
sd = sqrt( (n*Syy - Sy^2 - a*(n*Sxy - Sx*Sy) )/n/(n-2) );

lambda = -log(a)/delta;
mu = b/(1-a);
sigma = sd * sqrt( -2*log(a)/delta/(1-a^2) );
end

```

### B.4 Escolha do Portfólio Ótimo Segundo o Modelo com Tangente Hiperbólica

```

S1=par(:,1);
S2=par(:,2);
Wtanh=zeros(dias+1,1);
Wou=zeros(dias+1,1);

```

```

S1=S1(end-(janela+dias):end);
S2=S2(end-(janela+dias):end);
S=S1-S2;
mmax=max([S1(1),S2(1)]);
S=S/mmax;
S_mostra=S(1:janela);

R1=S1(2:end)./S1(1:end-1)-1;
R2=S2(2:end)./S2(1:end-1)-1;
c=1/(2*sigma);

aux=zeros(1,dias);
aux2=zeros(1,dias);

for i=1:dias,
N_t=(-lambda*tanh(c*(S(janela+i)-mu) )/c-r*(S(janela+i)-mu))*Wtanh(i)/(sigma^2);

NOU_t=(-lambda*(S(janela+i)-mu) -r*(S(janela+i)-mu))*Wou(i)/(sigma^2);

end;

```