

**Leandro Guimarães Marques Alvim**

**Estratégias de Negociação de Ativos  
Financeiros Utilizando Agendamento por  
Intervalos Ponderados**

**Tese de Doutorado**

Tese apresentada ao Programa de Pós-graduação em Informática do Departamento de Informática da PUC-Rio como requisito parcial para obtenção do título de Doutor Informática.

Orientador: Prof. Ruy Luiz Milidiú

Rio de Janeiro  
Março de 2013



**Leandro Guimarães Marques Alvim**

**Estratégias de Negociação de Ativos  
Financeiros utilizando Agendamento por  
Intervalos Ponderados**

Tese apresentada como requisito parcial para a obtenção do grau de Doutor pelo Programa de Pós-graduação em Informática do Departamento de Informática do Centro Técnico e Científico da PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada.

**Prof. Ruy Luiz Milidiú**

Orientador

Departamento de Informática – PUC-Rio

**Prof. Marcus Vinícius Soledade Poggi de Aragão**

Departamento de Informática – PUC-Rio

**Prof. Marco Antonio Casanova**

Departamento de Informática – PUC-Rio

**Prof. Geraldo Bonorino Xexéo**

UFRJ

**Prof. Julio Cesar Duarte**

IME

**Prof. Ronaldo Ribeiro Goldschmidt**

UFRRJ

**Prof. José Eugenio Leal**

Coordenador Setorial do Centro Técnico Científico  
PUC-Rio

Rio de Janeiro, 14 de Março de 2013

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

### **Leandro Guimarães Marques Alvim**

Graduou-se pela Universidade Federal Fluminense (UFF) em Ciência da Computação. Concluiu seu mestrado em Informática no Núcleo de Computação Eletrônica da Universidade Federal do Rio de Janeiro (UFRJ) na área de Inteligência Computacional. Ganhou bolsa de doutorado da Faperj Aluno Nota 10 no ano de 2011. É professor concursado pela Universidade Federal Rural do Rio de Janeiro (UFRRJ), no Departamento de Tecnologias e Linguagens, Campus Nova Iguaçu.

#### Ficha Catalográfica

Alvim, Leandro Guimarães Marques

Estratégias de Negociação de Ativos Financeiros Utilizando Agendamento por Intervalos Ponderados / Leandro Guimarães Marques Alvim; orientador: Ruy Luiz Milidiú. — Rio de Janeiro : PUC–Rio, Departamento de Informática, 2013.

v., 80 f: il. (color) ; 30 cm

1. Tese (doutorado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Informática, 2013.

Inclui bibliografia.

1. Informática – Teses. I. Milidiú, Ruy Luiz. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Informática. III. Título.

CDD: 004

## Agradecimentos

Ao meu orientador professor Ruy Luiz Milidiú pelo apoio e incentivo para a realização deste trabalho.

Ao CNPq, CAPES, Faperj e à PUC-Rio, pelos auxílios concedidos, sem os quais este trabalho não poderia ter sido realizado.

À minha família por todo apoio emocional.

Aos meus colegas do laboratório LEARN da PUC-Rio pela amizade e críticas construtivas ao trabalho.

## Resumo

Alvim, Leandro Guimarães Marques; Milidiú, Ruy Luiz.  
**Estratégias de Negociação de Ativos Financeiros Utilizando Agendamento por Intervalos Ponderados.**  
Rio de Janeiro, 2013. 80p. Tese de Doutorado — Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

Há diferentes tipos de investidores que compõem o mercado financeiro e produzem oportunidades de mercado em diferentes escalas de tempo. Isto evidencia uma estrutura heterogênea de mercado. Nesta tese conjecturamos que podem haver oportunidades mais preditivas do que outras, o que motiva a investigação e a construção de estratégias multirresolução. Para estratégias multirresolução há abordagens que utilizam a decomposição de séries temporais para a operação em resoluções distintas ou propostas para a construção de conjuntos de dados de acordo com decisões de negociação multirresolução. As demais estratégias, em sua maioria, são de resolução única. Nesta tese, abordamos dois problemas, maximização de retorno acumulado e maximização de retorno acumulado com o risco controlado, e propomos uma abordagem computacionalmente eficiente para a construção de estratégias multirresolução, a partir da resolução do problema de Agendamento de Intervalos Ponderados. Nossa metodologia consiste em dividir o dia de mercado em intervalos, especializar traders por intervalo e associar um prêmio a cada trader. Para o problema de maximização de retorno acumulado, o prêmio de cada trader corresponde ao retorno acumulado entre dias para o intervalo de operação associado. Para o problema de maximização de retorno acumulado com controle do risco, o prêmio de cada trader corresponde ao retorno acumulado dividido pelo risco para o intervalo de operação associado. Diferentemente do problema anterior, empregamos um conjunto de traders por intervalo e utilizamos o método de Média-Variância, de Markowitz, para encontrar pesos ótimos para conjunto de traders de forma a controlar o risco. Conjecturamos aqui que o controle do risco por intervalo acarreta no controle do risco global da estratégia para o dia. Para a sinalização das ordens de compra e venda, nossos traders utilizam detectores de oportunidades. Estes detectores utilizam algoritmos de Aprendizado de Máquina que processam informações de indicadores de análise técnica e dados de preço e volume. Realizamos experimentos para dez ativos de maior liquidez da BMF&Bovespa para um período de um ano. Nossa estratégia de Composição de um Time de Traders (CTT) apresenta 0,24% de lucro médio diário e 77,24% de lucro anual,

superando em 300% e 380%, respectivamente, uma estratégia de resolução única. Para os custos adotados, a estratégia CTT é viável a partir de US\$ 50.000,00. Para o problema de maximização do retorno acumulado com risco controlado, a estratégia de Composição de Carteiras por Intervalos (CCI) apresenta em média 0,179% de lucro diário e 55,85% de lucro anual, superando o método de Média-Variância de Markowitz. Para os custos adotados, a estratégia CCI é viável a partir de US\$ 2.000.000,00. As principais contribuições desta tese são: abordagem por Agendamentos de Intervalos Ponderados para a construção de estratégias e o emprego do modelo de Média-Variância para compor uma carteira de traders ao invés da tradicional abordagem por ativos.

### **Palavras-chave**

Intraday Traders; Agendamento de Intervalos Ponderados;  
Finança Computacional; Aprendizado de Máquina; Markowitz

## Abstract

Alvim, Leandro Guimarães Marques; Milidiú, Ruy Luiz (Advisor)  
. **Weighted Interval Scheduling Resolution for Building Financial Market Trading Strategies**. Rio de Janeiro, 2013. 80p. DSc. Thesis — Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

There are different types of investors who make up the financial market and produce market opportunities at different time scales. This indicates a heterogeneous market structure. In this thesis, we conjecture that may have more predictive opportunities than others, what motivates research and construction of we denominate multiresolution optimal strategies. For multiresolution strategies there are time series decomposition approaches for operating at different resolutions or proposals for dataset construction according to multiresolution trading optimal decisions. The other approaches, are single resolution. Thus, we address two problems, maximizing cumulative returns and maximizing cumulative returns with risk control. Here, we propose solving the Weighted Interval Scheduling problem to build multiresolution strategies. Our methodology consists of dividing the market day into time intervals, specialize traders by interval and associate a prize to each trader. For the cumulative return maximization problem, the prize corresponds to cumulative returns between days for the associated trader operation interval. For the cumulative return maximization problem with risk control each trader prize corresponds to cumulative return divided by risk with associated operation interval. In order to control the risk, we employ a set of traders by interval and apply the Markowitz Mean-Variance method to find optimal weight for set of traders. Here, we conjecture that controlling each interval risk leads to the overall risk control of the day. For signaling buy and sell orders, our traders use opportunity detectors. These detectors correspond to Machine Learning algorithms that process technical analysis indicators, price and volume data. We conducted experiments for ten of the most liquid BMF&Bovespa stocks to a one year span. Our Trading Team Composition strategy results indicates an average of 0.24% daily profit and a 77.24% anual profit, exceeding by 300% and 380%, respectively, a single resolution strategy. Regarding operational costs, CTT strategy is viable from US\$ 50,000. For the cumulative return maximization problem under risk control, our Portfolio Composition by Intervals strategy results indicates an average of 0.179% daily profit and a 55.85% anual profit, exceeding a Markowitz Mean-Variance method.

Regarding operational costs, CCI strategy is viable from US\$ 2,000,000. Our main contributions are: the Weighted Interval Scheduling approach for building multiresolution strategies and a portfolio composition of traders instead of stocks performances.

## **Keywords**

Intraday Traders; Weighted Interval Scheduling; Computational Finance; Machine Learning; Markowitz;



# Sumário

1	Introdução	<b>13</b>
1.1	Problema	13
1.2	Metodologia	14
1.3	Resumo dos Resultados	14
1.4	Contribuições	15
1.5	Organização da Tese	16
2	Revisão da Literatura	<b>18</b>
2.1	Construção de Preditores do Mercado Financeiro	18
2.2	Construção de Estratégias Multirresolução	19
2.3	Seleção de Carteiras	19
2.4	Considerações Finais	20
3	Conceitos Básicos	<b>21</b>
3.1	Carteira de Ativos	21
3.2	Modelo de Média-Variância	22
3.3	Agendamento de Intervalos Ponderados	25
3.4	Máquina de Vetores de Suporte	27
3.5	Mínimos Quadrados Parciais	32
3.6	Considerações Finais	36
4	Detecção de Oportunidades	<b>38</b>
4.1	A Tarefa	38
4.2	Engenharia de Atributos	40
4.3	Métricas	45
4.4	Considerações Finais	46
5	Construção Eficiente de Estratégias de Negociação	<b>47</b>
5.1	Trader Atômico	47
5.2	Composição de um Time de Traders	48
5.3	Composição de Carteiras por Intervalos	50
5.4	Considerações Finais	53
6	Sistema de Trading	<b>54</b>
6.1	Premissas de Mercado	54
6.2	Custos Operacionais	55
6.3	Métricas de Desempenho	55
6.4	Arquitetura do Sistema de Trading	55
6.5	Considerações Finais	58
7	Avaliação Empírica	<b>60</b>
7.1	Conjunto de Dados	61
7.2	Detectors de Oportunidade	62
7.3	Baselines	63
7.4	Composição de um Time de Traders	64

7.5	Composição de Carteiras por Intervalos	69
7.6	Consolidação dos Resultados	73
8	Conclusões	<b>74</b>
8.1	Visão Geral	74
8.2	Contribuições	75
8.3	Trabalhos Futuros	76

## Lista de figuras

3.1	Exemplo de Fronteira Eficiente.	23
3.2	Exemplo do Problema de Agendamento de Intervalos Ponderados.	26
3.3	SVM - Margem.	29
3.4	PLS - diagrama de blocos.	33
4.1	Exemplo de oportunidades possíveis.	39
5.1	Exemplo de distribuição de operações para dois dias de mercado.	50
5.2	Exemplo de solução para o problema de Agendamento de Intervalos Ponderados.	51
5.3	Exemplo de carteiras de traders por intervalo.	53
6.1	Etapa de detecção de oportunidades.	56
6.2	Etapa de construção das estratégias de negociação.	58
7.1	CTT - Comparativo.	66
7.2	CTT - Desempenho das estratégias BLS1 e CTT ao longo do dia.	67
7.3	Simulação das estratégias CTT e BLS1 com custos operacionais.	68
7.4	CTT - Custos operacionais.	68

## Lista de tabelas

1.1	Lucros com CTT para maximização de retorno acumulado.	14
1.2	Lucros com CCI para maximização do retorno com risco controlado.	15
3.1	Exemplo de carteira de ativos.	22
7.1	Informações do conjunto de dados.	61
7.2	Retorno médio por ativo.	62
7.3	BLS1 - Média dos lucros diários.	64
7.4	CTT - Resultados de lucro.	65
7.5	CTT - Risco.	65
7.6	CTT - Lucro por unidade risco.	65
7.7	CTT - Número médio de operações ao dia.	65
7.8	CCI - Soluções viáveis com relação ao total de carteiras.	70
7.9	BLS2 - Desempenho.	70
7.10	CCI - Risco	71
7.11	CCI - Lucro.	71
7.12	CCI - Operações.	72
7.13	CCI - Lucro.	72
7.14	CCI - Corretagem.	73

# 1

## Introdução

O mercado financeiro é composto por investidores que operam e analisam o mercado de maneira distinta. Por exemplo, existem investidores que operam com frequência horária, outros com frequência diária e demais. Investidores de alta frequência aproveitam-se das pequenas variações de mercado investindo grandes quantias de dinheiro, enquanto investidores de baixa frequência aproveitam-se das grandes variações de mercado e por consequência sem a necessidade de grandes quantias de dinheiro. Há diferentes tipos de investidores que compõem o mercado financeiro e produzem oportunidades de mercado em diferentes escalas de tempo. Isto evidencia uma estrutura heterogênea de mercado. Nesta tese, conjecturamos que podem haver oportunidades mais previsíveis do que outras, o que motiva a investigação e a construção de estratégias multirresolução. Definimos estratégia multirresolução como uma estratégia que opera em diferentes escalas de tempo.

Para estratégias multirresolução, a literatura é reduzida. O enfoque mais comum (Zhang et al., 2001) utiliza a abordagem por Wavelets, em que é feita uma decomposição de uma série temporal dos preços ou retornos do ativo, para posterior associação de traders especializados por resolução e ativo. Outro enfoque (Boyarshinov, 2005), sugere a construção de um conjunto de dados, a partir de operações de traders. Para isto, seleciona-se um subconjunto de operações que geram o maior retorno acumulado para o dia ou retorno acumulado com risco controlado por índices de razão retorno sobre risco, como por exemplo índice de Sharpe. O conjunto de dados possui apenas exemplos que participam das operações selecionadas. As demais estratégias, em sua maioria, são de resolução única, tal que operam em um único ativo por vez (Gomide e Milidiú, 2010) ou em vários ativos simultaneamente (da Silva e Laber, 2011).

### 1.1

#### Problema

Nesta tese, abordamos dois problemas relacionados ao mercado financeiro: maximização de retorno acumulado e maximização de retorno acumu-

lado com o risco controlado. Reformulamos estes dois problemas e propomos um método computacionalmente eficiente para a construção de estratégias multirresolução, a partir da resolução do problema de Agendamento de Intervalos Ponderados.

## 1.2 Metodologia

Nossa metodologia consiste em dividir o dia de mercado em intervalos, especializar traders por intervalo e associar um prêmio para cada trader. Para o problema de maximização de retorno acumulado, o prêmio de cada trader corresponde ao retorno acumulado entre dias para o intervalo de operação associado. Para o problema de maximização de retorno acumulado com controle do risco, o prêmio de cada trader corresponde ao retorno acumulado dividido pelo risco para o intervalo de operação associado. Diferentemente do problema anterior, empregamos um conjunto de traders por intervalo e utilizamos o método de Média-Variância, de Markowitz, para encontrar pesos ótimos para o conjunto de traders, de forma a controlar o risco. Conjecturamos aqui que o controle do risco por intervalo acarreta no controle do risco global da estratégia para o dia. Para a sinalização das ordens de compra e venda, nossos traders utilizam detectores de oportunidades. Estes detectores correspondem a algoritmos de Aprendizado de Máquina que processam informações de indicadores de análise técnica e dados de preço e volume.

## 1.3 Resumo dos Resultados

A seguir, apresentamos um resumo dos nossos resultados para os dois problemas abordados nesta tese, a saber:

1. maximização de retorno acumulado e
2. maximização de retorno acumulado com risco controlado.

Estratégia	Lucro Diário	Lucro Anual
BLS1	0,06%	16,07%
CTT	0,24%	77,26%

Tabela 1.1: Lucros com CTT para maximização de retorno acumulado.

Primeiramente, apresentamos na Tabela 1.1, os resultados da estratégia de Composição de um Time de Traders (CTT) para o problema de maximização de retorno acumulado. O baseline BLS1, representa a abordagem tradicional para elaboração de estratégias de maximização de retorno, em que

é selecionada em um conjunto de validação a estratégia mais lucrativa, que opera em um único ativo e em uma única resolução por vez.

Nossos resultados, descritos na Tabela 1.1, indicam que a estratégia CTT apresenta ganhos de 300% e de 380%, com relação ao baseline BLS1, para as métricas de retorno diário médio e lucro anual, respectivamente. Para os custos adotados, a estratégia CTT é viável a partir de US\$ 50.000,00.

Estratégia	Lucro Diário (%)	Lucro Anual (%)
BLS2	-0,139	-34,34
CCI	0,179	55,85

Tabela 1.2: Lucros com CCI para maximização do retorno com risco controlado.

Na Tabela 1.2, apresentamos um resumo dos nossos resultados da estratégia de Composição de Carteiras por Intervalos (CCI) para o problema de maximização do retorno com risco controlado. Os resultados correspondem à média dos lucros de todos os experimentos para cada estratégia. O baseline BLS2, corresponde ao método de Média-Variância de Markowitz para seleção de carteiras.

Podemos observar, na Tabela 1.2, que a estratégia BLS2 apresenta prejuízo. Isto porque, BLS2 depende dos desempenhos dos ativos, que em sua maioria apresentam retorno negativo no período examinado. Desta forma, a estratégia CCI supera o baseline BLS2 para as métricas de retorno médio diário e lucro anual, respectivamente. Para os custos adotados, a estratégia CCI é viável a partir de US\$ 2.000.000,00. Isto porque, para reduzir o risco, a estratégia CCI diversifica seu capital para cada intervalo ao longo do dia. Esta diversificação gera um número maior de operações simultâneas, o que contribui para um aumento do custo de corretagem. Cabe observar que, diferentemente da estratégia CCI, a estratégia CTT opera em um ativo por vez para cada intervalo do dia e a diversificação de capital ocorre apenas quando há um novo intervalo.

## 1.4 Contribuições

Nesta seção, apresentamos as nossas principais contribuições desta tese, a saber:

1. construção de estratégias computacionalmente eficientes, a partir da resolução do problema de Agendamento de Intervalos Ponderados.
2. estimativa do retorno de carteiras, a partir dos desempenhos de cada trader que opera ao longo do dia.

### 3. o Sistema de *Trading* proposto.

A primeira e principal contribuição está relacionada à construção eficiente de estratégias, multiresolução e multiativo, para a tarefa de negociação ao longo do dia de mercado. Abordamos dois tipos de problema: maximização do retorno e a maximização do retorno com risco controlado. Propomos duas novas formulações para o esses dois problemas e ambas equivalem ao problema de Agendamento de Intervalos Ponderados, que pode ser resolvido de forma computacionalmente eficiente. Ressaltamos que a abordagem por Agendamento de Intervalos Ponderados para o problema de maximização do retorno acumulado está publicada (Alvim e Milidiú, 2013) na prestigiosa *Decision Support Systems*<sup>1</sup>.

Nossa segunda contribuição consiste em aplicar métodos de seleção de carteira utilizando-se das informações dos desempenhos de cada trader. Enquanto a abordagem padrão, para métodos de seleção de carteira, é estimar o retorno e o risco de uma carteira a partir dos desempenhos de ativos, nesta tese estimamos o retorno e o risco em função dos lucros médios e dos riscos obtidos a partir das decisões de compra e venda de cada trader. Desta forma, a estratégia fica menos dependente do desempenho de um ativo e mais dependente dos desempenhos dos traders. Como é mais fácil controlar o desempenho de um trader do que de um ativo, podemos ajustar e selecionar melhor nossos traders.

A terceira e última contribuição está na proposta de uma arquitetura de um sistema de *Trading* que é dividida em seis etapas: treinamento e validação dos detectores de oportunidade; agrupamento das decisões de compra e venda dos Traders por intervalos do dia; aplicação de métodos de seleção de carteira, se o risco for informado como parâmetro; avaliação dos desempenhos por intervalos do dia; resolução do problema de Agendamento de Intervalos Ponderados, por programação dinâmica, em um conjunto de validação; e aplicação da solução obtida em um conjunto de teste. Nosso sistema integra métodos de otimização e de Aprendizado de Máquina em uma única arquitetura.

## 1.5

### Organização da Tese

Esta tese está organizada da seguinte forma. No capítulo 2, investigamos os principais trabalhos relacionados a estratégias para negociação ao longo do dia de mercado. Descrevemos primeiramente os principais trabalhos que empregam Aprendizado de Máquina para a tarefa de negociação ao longo do dia de mercado. Adicionalmente, investigamos os principais trabalhos

<sup>1</sup>periódico internacional Qualis A1



relacionados à construção de estratégias multiativo e multirresolução para a tarefa de negociação ao longo do dia de mercado.

No capítulo 3, descrevemos os conceitos básicos relacionados aos problemas e soluções abordados nesta tese. Primeiramente, definimos o conceito de carteira de ativos e o cálculo do retorno esperado de uma carteira. Após isto, introduzimos o conceito de risco como correlação entre ativos e descrevemos a formulação de Markowitz para seleção de carteiras. Por fim, descrevemos o problema de Agendamento de Intervalos Ponderados, que está diretamente relacionado à construção de estratégias multiativo e multirresolução, descritos no capítulo seguinte.

No capítulo 4, descrevemos nossa abordagem de Aprendizado de Máquina para a detecção de oportunidades do mercado financeiro. Primeiramente, descrevemos a tarefa de detecção de oportunidades para o dia de mercado. Por fim, descrevemos nossa engenharia de atributos para geração dos detectores de oportunidades e a metodologia empregada para ajuste de seus parâmetros dos detectores de oportunidades.

No capítulo 5, propomos duas estratégias, multirresolução e multiativo, para a tarefa de negociação ao longo do dia de mercado. Para a primeira estratégia, Composição de um Time de Traders, investigamos o problema maximização do retorno e propomos uma função objetivo para o problema de Agendamento de Intervalos Ponderados. Para segunda estratégia, Composição de Carteiras por Intervalos, investigamos a maximização do retorno com controle do risco. Propomos a incorporação do risco por intervalo, resolvendo métodos de seleção de carteira para cada intervalo, e reformulamos a função objetivo para o problema de Agendamento de Intervalos Ponderados.

No capítulo 6, apresentamos a arquitetura do nosso sistema de trading, destacando as premissas de mercado, os custos operacionais, as métricas de desempenho para as estratégias de negociação.

No capítulo 7, descrevemos, primeiramente, nosso conjunto de dados e período utilizado para os testes. Adicionalmente, descrevemos nossa metodologia e apresentamos dois baselines, um para cada estratégia proposta. Ao final do capítulo, apresentamos os resultados para as duas estratégias propostas e os baselines. Para cada estratégia proposta apresentamos resultados de lucro, risco e o impacto dos custos operacionais.

No capítulo 8, apresentamos nossas conclusões e trabalhos futuros.

## 2

### Revisão da Literatura

Dividimos os trabalhos revisados em três categorias: a aplicação de Aprendizado de Máquina para o mercado financeiro, a construção de estratégias multirresolução e a seleção de carteiras. Estas três categorias abrangem os conceitos centrais combinados nesta tese.

#### 2.1

##### Construção de Preditores do Mercado Financeiro

Nesta seção, examinamos a contribuição de algoritmos de Aprendizado de Máquina Supervisionado para a predição do comportamento do mercado financeiro.

Em (Atsalakis e Valavanis, 2009a), os autores comparam os resultados 150 artigos relacionados à predição do comportamento do mercado financeiro para os algoritmos estatísticos clássicos: Auto Regressivo, Média Móvel Auto Regressiva, Média Móvel Auto Regressiva Integrada, Heterocedasticidade Condicional Auto Regressiva e Volatilidade Estocástica. O estudos indicam resultados promissores para a tarefa.

Em (Atsalakis e Valavanis, 2009b), os mesmos autores realizam um estudo comparativo para mais de 100 artigos relacionados também à predição do comportamento do mercado financeiro, a partir de algoritmos de Aprendizado de Máquina Supervisionado. Os autores indicam que algoritmos de Aprendizado de Máquina Supervisionado produzem melhores resultados em relação aos algoritmos estatísticos clássicos para a tarefa de predição do comportamento de mercado.

Em (Abraham et al., 2003), é investigada a predição da alta e queda de preços para o dia seguinte, para um período de sete anos do índice Nasdaq-100 e para um período de quatro anos do índice NIFTY. Os autores comparam o poder preditivo dos seguintes algoritmos de Aprendizado de Máquina Supervisionado: Redes Neurais Artificiais, a partir do algoritmo de treinamento de Levenberg-Marquardt, Máquina de Vetores de Suporte (SVM), Takagi-Sugeno Neural-Nebuloso e *Difference Boosting Neural Network*. Os autores reportam resultados promissores para todos os algoritmos e constatam

melhores resultados para o SVM.

Em (Huang et al., 2005), é investigada a predição semanal de alta e queda do índice NIKKEI-225 para o período de um ano. Os autores comparam o poder preditivo dos seguintes algoritmos de Aprendizado de Máquina Supervisionado: Linear Discriminant Analysis, Quadratic Discriminant Analysis e Elman Backpropagation Neural Networks. Os resultados indicam 75% de taxa de acerto de alta e queda de preços para o SVM cujo resultado foi o melhor dentre os demais.

## 2.2

### Construção de Estratégias Multirresolução

Nesta seção, descrevemos os principais trabalhos que utilizam estratégias de negociação multirresolução. Uma estratégia é dita multirresolução quando utiliza diferentes frequências de operação.

Em (Zhang et al., 2001), os autores utilizam wavelets para decompor uma série temporal em diferentes resoluções. Os autores treinam uma Rede Neural Multilayer Perceptron para cada resolução da série temporal, de forma a prever o comportamento individual de cada resolução. Para gerar as previsões finais, é associado um peso a cada previsão individual, de acordo com uma métrica de previsão, e um modelo perceptron aprende estes pesos. Ao final, o modelo de previsão é utilizado por um sistema de trading que opera em múltiplas resoluções, de acordo com os pesos aprendidos.

Em (Boyarshinov, 2005), é feito um estudo teórico de construção de estratégias ótimas de agentes negociadores de mercado para aprender atributos relevantes à construção de conjuntos de dados para Aprendizado de Máquina. Este estudo aborda os problemas de maximização de retorno e maximização de retorno sujeito a indicadores de Sterling e Sharpe, que indicam o lucro por unidade de risco. Os autores propõem a aplicação do algoritmo Envoltória Convexa para encontrar um conjunto de informações relevantes para estes problemas e gerar o conjunto de dados. Adicionalmente, os autores propõem um algoritmo combinatório determinístico, baseado nas informações encontradas, para classificação das decisões de compra e venda. Apesar da qualidade teórica do estudo, não há resultados empíricos e os autores não analisam o impacto dos custos operacionais para as operações dos traders.

## 2.3

### Seleção de Carteiras

Nesta seção, revisamos os principais trabalhos sobre métodos de seleção de carteiras, que estão relacionados a esta tese. Adicionalmente, revisamos as

métricas de risco para seleção de carteiras adotadas nesta tese.

Em (Hellström, 2001), os autores investigam uma formulação alternativa ao problema de seleção de carteiras a partir do ranqueamento de ações de acordo com o retorno e com o índice de Shape. Os autores concluem que é mais fácil fazer a predição da ordenação de ativos do que a predição de retorno individual. Para isto, geram um conjunto de dados compostos por ordenações de ativos ao longo do tempo e treinam um Perceptron para aprender a ranquear os ativos. Os resultados indicam que a estratégia proposta supera o índice de mercado utilizado como comparativo.

Em (da Silva e Laber, 2011), os autores investigam métodos de seleção de carteira de Média-Variância de Markowitz (Markowitz, 1952) e variações deste método para dados diários do BMF&Bovespa. Os autores analisam o impacto dos custos transacionais, estimativas de risco e retorno e a frequência da revisão de carteiras para os diferentes métodos de seleção de carteira. Os resultados empíricos indicam que métodos de seleção de carteira de Markowitz e variações deste são eficazes para o controle do risco.

Em (Cuoco et al., 2008), os autores propõem Valor em Risco como métrica para avaliação e controle do risco de carteiras. Os autores mostram empiricamente que a exposição do investidor ao risco, sujeito a restrições de Valor em Risco, é sempre inferior a exposição do investidor sem restrições.

## 2.4

### Considerações Finais

Neste capítulo revisamos os trabalhos mais relevantes de acordo com o que será abordado nesta tese. Na primeira seção indicamos estudos que elucidam ampla utilização e relevância de algoritmos estatísticos clássicos e de Aprendizado de Máquina para a tarefa de predição do mercado financeiro e para estratégias de trading. Ainda nesta seção, para a mesma tarefa, conclui-se que algoritmos de Aprendizado de Máquina apresentam qualidade superior comparados aos algoritmos estatísticos clássicos. Na seção seguinte, descrevemos alguns trabalhos relacionados a estratégias multirresolução, tema pouco abordado na literatura, porém promissor. Por fim, na última seção, descrevemos trabalhos relacionados ao processo de composição de carteiras que tem por conceito base a diversificação de ativos de forma a minimizar o risco do capital investido. A partir desta seção, podemos concluir que tais métodos são eficazes quanto a redução do risco e ainda são fontes de estudo.

## 3 Conceitos Básicos

Neste capítulo, descrevemos os conceitos fundamentais para o entendimento desta tese. Na primeira seção, definimos uma carteira de ativos e estabelecemos uma formulação matemática para o problema de otimização de carteiras. Estes conceitos serão utilizados pela estratégia de maximização do lucro com controle do risco, proposta no capítulo 5. Na seção seguinte, descrevemos o problema de Agendamento de Intervalos Ponderados, sua formulação e resolução. Veremos, no capítulo 5, que as formulações das estratégias propostas nesta tese podem ser transformadas na formulação do problema de Agendamento por Intervalos Ponderados e resolvidas de forma computacionalmente eficiente. Por fim, nas duas últimas seções, descrevemos dois algoritmos de Aprendizado de Máquina que serão utilizados para prever oportunidades de mercado e auxiliarão nas tomadas de decisão das estratégias propostas nesta tese.

### 3.1 Carteira de Ativos

Uma carteira é uma combinação de ativos, tal que para cada ativo é definida uma porcentagem do investimento total. Uma carteira possui duas propriedades fundamentais, que são utilizadas em seu processo de otimização da carteira: retorno e risco. O retorno de uma carteira é calculado como a soma dos rendimentos individuais de cada ativo, ponderados pela respectiva porcentagem de investimento. Já para o cálculo do risco, existem diversas formulações. As mais comuns são a correlação entre ativos, e o Valor em Risco Condicional.

Formalmente, podemos descrever o retorno  $R$  de uma carteira da seguinte forma. Seja  $n$  o número total de ativos de uma carteira,  $w$  o vetor que contém os percentuais de investimento em cada ativo e  $r$  o vetor que contém os retornos de cada investimento, então o retorno total de uma carteira é dado por

$$R = \sum_{i=1}^n w_i r_i = w^t r. \quad (3-1)$$

Ativo	Código	Peso	Rendimento
Petrobras	PETR4	40%	10%
Vale do Rio Doce	VALE5	20%	-12%
Gerdau	GOAU5	25%	5%
VIVO	VIVO4	15%	13%

Tabela 3.1: Exemplo de carteira de ativos.

Na Tabela 3.1, ilustramos um exemplo de carteira com quatro ativos. A coluna peso representa a proporção de investimento para o ativo e a coluna rendimento representa o rendimento até o momento. Para o exemplo da Tabela 3.1, o retorno total da carteira é obtido por

$$R = 0,4 \cdot 0,1 - 0,2 \cdot 0,12 + 0,25 \cdot 0,05 + 0,15 \cdot 0,13 = 0,048.$$

Uma das grandes vantagens de construir uma carteira está na diversificação do investimento em ativos. Através da diversificação de ativos podemos obter riscos inferiores a qualquer um dos ativos isolados (Markowitz, 1952). Outra grande vantagem da utilização de carteiras é encontrar uma carteira ótima, ajustada ao perfil do investidor. Por exemplo, um investidor avesso ao risco pode preferir investir em um grupo de ativos cujo risco é menor e esperar um retorno menor. Já um investidor mais agressivo, pode esperar obter retornos maiores investindo em grupos de ativos mais arriscados.

### 3.2

#### Modelo de Média-Variância

O modelo de Média-Variância, proposto por Markowitz (Markowitz, 1952), possibilita encontrar uma carteira ótima de ativos levando em consideração a relação retorno e risco. É denominado modelo de Média-Variância, pois estimamos o retorno, a partir de uma média dos retornos históricos dos ativos, com o risco é sendo determinado pela variância e a covariância entre ativos. Como alternativa à covariância, podemos utilizar o coeficiente de correlação, que é uma padronização da covariância. A correlação geralmente é mais adequada, pois seus valores estão contidos no intervalo  $[-1, 1]$ .

Em termos de risco, uma correlação positiva significa investir em ativos que crescem juntos positivamente e, se um deles apresentar queda, é provável que o outro também apresente. Para a correlação zero o comportamento tende a ser aleatório. Já uma correlação negativa não é risco para a carteira, pois os ativos se comportam de maneiras opostas. Por exemplo, para um par de ativos que apresentem correlação negativa, se um apresentar uma queda o

outro tenderá a uma alta.

Outra grande contribuição do trabalho de Markowitz está na Fronteira Eficiente. A Fronteira Eficiente representa um conjunto de combinações de ativos que atingem o maior nível de retorno possível para um certo risco ou que, para um certo retorno, atingem o menor risco possível. O interessante é que, com a Fronteira Eficiente, é possível achar uma combinação de ativos cujo risco é menor que o ativo de menor risco. Uma carteira é dita eficiente se esta faz parte da Fronteira Eficiente.

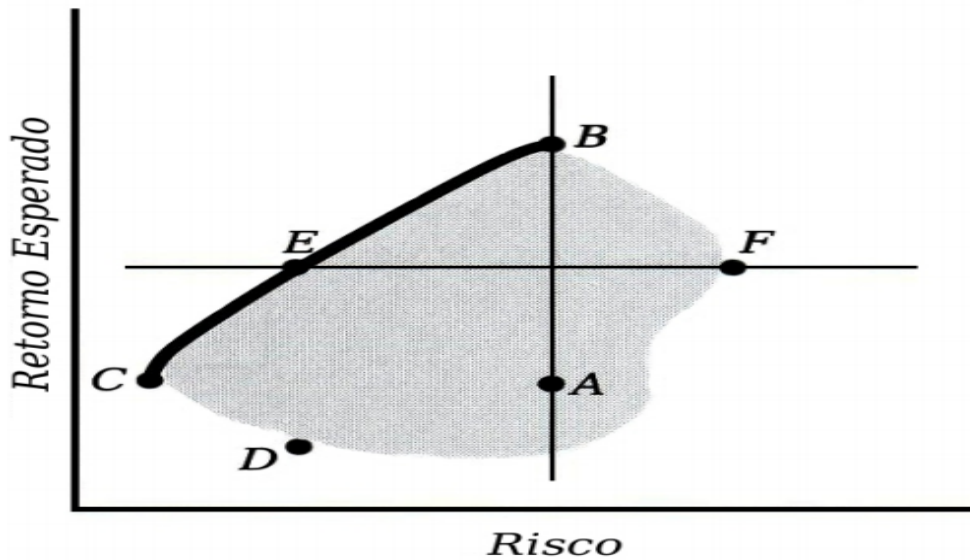


Figura 3.1: Exemplo de Fronteira Eficiente.

Na figura 3.1, ilustramos um exemplo (da Silva e Laber, 2011) de Fronteira Eficiente. A parte cinza representa todas as possíveis carteiras. Os pontos C, B e E representam, cada um, uma carteira eficiente. Já os pontos D, A e F representam carteiras não-eficientes. Markowitz reconhece que investidores possuem diferentes tolerâncias ao risco e necessidades de retorno distintas. Para o exemplo da figura 3.1, se o investidor deseja um risco menor e está com a carteira B, então este pode trocar sua carteira para C. Da mesma forma, se o investidor deseja um retorno maior, este pode trocar a carteira C para a carteira B. De acordo com a teoria da Fronteira Eficiente, pode-se determinar uma carteira ótima de ativos de acordo com parâmetros informados pelo investidor. Estes parâmetros podem envolver risco ou retorno esperado.

Existem diversas formulações possíveis para Markowitz (da Silva e Laber, 2011). A seguir descrevemos formalmente a formulação abordada nesta tese: maximização do retorno sujeito a um risco máximo como restrição.

Seja  $w$  o vetor de pesos da carteira,  $r$  o vetor do retorno esperado da carteira, composto por estimativas de retornos de cada ativo;  $S$  a matriz de correlação do histórico de retorno estimado dos ativos e  $\alpha$  e  $\beta$  vetores que representam os limites mínimos e máximos de investimento para cada ativo. Temos a seguinte formulação:

$$\text{Maximizar } w^t r \quad (3-2)$$

Sujeito a

$$w^t S w \leq \theta \quad (3-3)$$

$$\sum_{i=1}^n w_i = 1 \quad (3-4)$$

$$w_i \geq 0; \forall i \quad (3-5)$$

$$\alpha \leq w_i \leq \beta; \forall i \quad (3-6)$$

A função objetivo (3-2) representa o lucro total da carteira para um vetor de pesos  $w$ . A restrição (3-3) impõe que uma solução seja limitada a um risco máximo  $\theta$ . A restrição (3-4) garante que a soma dos pesos de  $w$  seja sempre igual a um. A restrição (3-5) evita que qualquer peso de  $w$  seja negativo. A restrição (3-6) garante um limite superior  $\beta$  e inferior  $\alpha$  aos pesos de  $w$ , forçando a diversificação. Os valores  $\beta$  e  $\alpha$  são determinados empiricamente em que descrevemos adiante, no capítulo 7.

### 3.2.1

#### Métricas de Risco para Seleção de Carteiras

Nesta subseção, descrevemos duas métricas de risco, Valor em Risco (Var) (Allen e Powell, 2007) e Valor em Risco Condicional (CVar) (Rockafellar e Uryasev, 2000), utilizadas para avaliar o risco de métodos de seleção de carteiras.

Valor em Risco é uma das métricas de risco mais importantes e mais utilizadas (Allen e Powell, 2007) para métodos de seleção de carteiras. Devido à sua simplicidade, Var é adotado em diversas entidades financeiras que efetuam transações em larga escala e também por seguradoras, bancos de varejo e até mesmo por empresas não-financeiras (Allen e Powell, 2007). Var representa o prejuízo mínimo esperado de uma estratégia ou ativo para um dado período de tempo e para um dado intervalo de confiança. Por exemplo, um 95%-Var de R\$1.000,00 significa que a cada 100 operações teremos 5 perdas superiores a R\$1.000,00. Para obter o Var dos desempenhos de uma estratégia,



assume-se que a distribuição é normal, e obtém-se o percentil dos desempenhos para um dado período de tempo (Allen e Powell, 2007). No exemplo descrito anteriormente, conhecer que uma estratégia apresenta 5 perdas superiores a R\$1.000,00 não é totalmente seguro, pois as perdas podem exceder muito acima deste valor. Algo mais útil e seguro seria estimar a média das perdas.

Como descrito anteriormente, um dos problemas da métrica Var é que o conhecimento das chances de prejuízo, a partir de um determinado valor, não é suficiente para se ter uma real ideia da grandeza das perdas ao longo do tempo (Allen e Powell, 2007). Assim, adicionalmente, emprega-se a métrica Valor em Risco Condicional (Cvar) (Rockafellar e Uryasev, 2000) que é comumente denominada de Média das Perdas Excedidas. Seja  $\beta$ -Var o valor com probabilidade  $\beta$  que não excederá  $\beta$ -Var, CVar é a média de  $(1 - \beta)100\%$  das perdas. Por exemplo, para um Var com  $\beta = 95\%$  de confiança, CVar corresponde a média de  $5\%$  das perdas excedidas.

### 3.3

#### Agendamento de Intervalos Ponderados

Nesta seção, descrevemos o problema de Agendamento de Intervalos Ponderados (Kleinberg e Tardos, 2005, Chawla, 2009) que utilizamos para a tarefa de negociação ao longo do dia de mercado. Para melhor entendimento do problema, revisamos um problema mais simples denominado de Agendamento de Intervalos.

O Agendamento de Intervalos é um problema de otimização cujo objetivo é alocar o maior número de tarefas para um dado período de tempo sem que haja sobreposição destas. Podemos descrever o Agendamento de Intervalos da seguinte forma (Chawla, 2009): seja  $T$  um conjunto de tarefas tal que cada tarefa  $i \in T$  possui um tempo inicial  $s_i$ , um tempo final  $f_i$  e tempo final sempre maior que tempo inicial  $f_i > s_i$ , queremos encontrar um subconjunto  $S$  que possua o maior número de tarefas para um período. Seja  $n$  o total de intervalos de um agendamento, a solução ótima para o problema pode ser facilmente encontrada em tempo  $O(n \log n)$  por um algoritmo guloso que seleciona as tarefas pela ordem do tempo final  $f_i$  (Chawla, 2009).

Se adicionarmos prêmios a cada tarefa do problema de Agendamento de Intervalos, o problema passa a se chamar Agendamento de Intervalos Ponderados.

Na figura 3.2, ilustramos um exemplo com quatro tarefas distribuídas ao longo do tempo. Cada linha representa uma tarefa. Os valores entre parênteses representam o identificador da tarefa e o valor acima de cada linha representa o prêmio associado a respectiva tarefa. Para este exemplo a solução ótima

para o problema de Agendamento de Intervalos Ponderados é a combinação das tarefas um e quatro, cuja soma dos prêmios é a maior possível e equivale a 35. Se formos analisar, da esquerda para a direita, a solução ótima para cada subproblema, os ótimos para os instantes 1,2 e 3 equivalem a 10, 15 e 35.

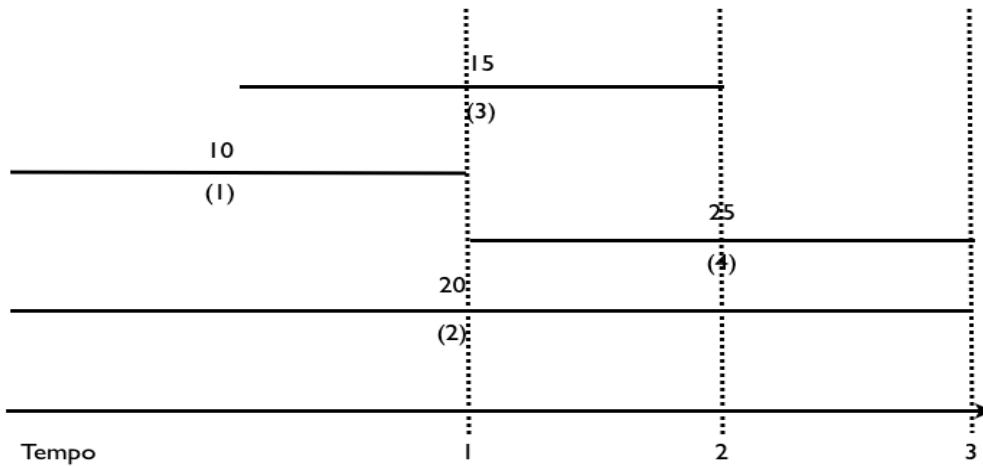


Figura 3.2: Exemplo do Problema de Agendamento de Intervalos Ponderados.

Podemos descrever o problema de Agendamentos de Intervalos Ponderados da seguinte forma (Chawla, 2009): seja  $T$  um conjunto de tarefas tal que cada tarefa  $i \in T$  possui um prêmio associado  $w_i$ , um tempo inicial  $s_i$ , um tempo final  $f_i$  sempre maior que tempo inicial  $f_i > s_i$ . Queremos encontrar um subconjunto de tarefas  $S$  não-sobrepostas cuja soma dos prêmios seja máxima. Formalmente, podemos descrever o problema como

$$S = \underset{S \subseteq T; \forall i, j \in S, se f_i \leq s_j ou f_j \leq s_i}{\text{arg max}} \sum_{i \in S} w_i \quad (3-7)$$

Diferentemente do Agendamento de Intervalos, o Agendamento de Intervalos Ponderados não pode ser resolvido por um algoritmo guloso. Para resolver o problema de Agendamento de Intervalos Ponderados emprega-se a estratégia de dividir e conquistar, descrita a seguir (Chawla, 2009).

Seja  $premio(i)$  a maior soma de todos os prêmios de tarefas não-sobrepostas a partir de  $i$ , tal que  $1 \leq i \leq n$ . Temos que:

$$premio(i) = \text{Máximo}\{w_i + premio(j), premio(i + 1)\}$$

tal que  $j$  representa o índice da primeira tarefa que inicia após o término da tarefa  $i$ . Para encontrar a soma do agendamento ótimo temos o seguinte algoritmo de Programação Dinâmica:

1. Ordene as tarefas em ordem crescente pelo tempo inicial  $s_i$ .

2. Para  $1 \leq i \leq n$ , encontre o menor  $j > i$  tal que a tarefa  $j$  não sobreponha a tarefa  $i$ .
3. Para  $i = n$  até 1, calcule  $premio(i)$  a partir da equação (3-7) e armazene em memória.
4. O prêmio ótimo do agendamento corresponde ao  $premio(1)$ .

Para restaurar a solução final, devemos recuperar os valores dos prêmios e as opções que levaram às melhores soluções.

O tempo de execução  $T(n)$  do algoritmo corresponde a:

$$T(n) = O(n \log n) + O(n \log n) + O(n) = O(n \log n) \quad (3-8)$$

tal que os termos  $O(n \log n)$  correspondem à ordenação e a  $n$  buscas das tarefas que iniciam após a tarefa  $i$ ,  $O(n)$  corresponde à resolução dos  $n$  subproblemas em tempo constante.

### 3.4

#### Máquina de Vetores de Suporte

Máquina de Vetores de Suporte (SVM) corresponde ao estado-da-arte para algoritmos classificadores. O SVM foi desenvolvido no início da década de 90 por Vapnik (Boser et al., 1992) e colaboradores do laboratório AT&T. SVM é amplamente utilizado em diversas áreas como na área de bioinformática, para classificar genes em ativos ou inativos com relação a um tecido canceroso (Guyon et al., 2002); na área de processamento de linguagem natural, para categorizar e-mails como *spam* ou *não-spam* (Drucker et al., 1999) e demais áreas. Na área de finanças, por exemplo, o SVM apresenta excelente desempenho para predição do comportamento de séries temporais de ativos do mercado financeiro (Huang et al., 2005, Abraham et al., 2003).

A principal contribuição de Vapnik e colaboradores foi a introdução do conceito de classificação de máxima margem, que descreveremos adiante. Abaixo, descrevemos as principais vantagens do emprego do classificador SVM.

*Convexidade da função objetivo de sua formulação.*

A aplicação do SVM para um dado problema implica na otimização de uma função quadrática que possui apenas um ótimo local, sendo que este é global. Operar em uma função objetivo com este tipo de característica é uma grande vantagem. Isto porque o excesso de ótimos locais, muito comum em problemas, dificulta o processo de convergência do algoritmo de aprendizado para o ótimo global.

*Desempenho independente do número de atributos.*

Essa é uma das grandes vantagens para a escolha do SVM para certas tarefas. Seu desempenho é quadrático e está em função apenas do número de exemplos. Por exemplo, para representar documentos, DNA e demais exemplos que exigem milhares de atributos, o desempenho do SVM continua quadrático em função dos exemplos, pois independe do número de atributos.

*Elevada capacidade de generalização.*

Além da minimização do risco empírico, comum em algoritmos de treinamento de Aprendizado de Máquina que visam reduzir os erros de classificação ao longo do treinamento, o SVM minimiza o risco estrutural. Este risco estrutural é minimizado a partir do conceito de Máxima Margem. O hiperplano de separação de classes fica em uma região mais distante da grande concentração de exemplos de cada classe, o que reduz o risco de uma futura classificação errônea.

*Pode resolver problemas não-lineares de classificação.*

Como a formulação dual, descrita pela equação (3-24), depende de produtos entre exemplos e classes, o problema não-linear pode ser transformado em linear, a partir do emprego de funções de Kernel. Desta forma, a função objetivo mantém a convexidade da função objetivo de sua formulação.

*Número de parâmetros reduzido.*

O SVM linear apresenta apenas um parâmetro ( $C$ ), descrito a seguir, para ser calibrado. No caso do emprego de funções de Kernel, adiciona-se apenas os parâmetros da função desejada.

### Classificação de Máxima Margem

Seja um problema de classificação binária e um classificador linear cujo hiperplano de separação  $h$  é ilustrado na figura 3.3. Cada ponto desta figura representa um exemplo, com dois atributos; a cor de um ponto corresponde à classe que este pertence; e um ponto circulado representa um Vetor de Suporte tal que, estes correspondem aos exemplos mais relevantes com relação à separação de exemplos entre duas classes. Geometricamente, são os pontos mais próximos da classe oposta.

Denominamos Margem  $m(h)$ , a distância entre os Vetores de Suporte e o hiperplano  $h$  de separação. Numa representação geométrica denotamos  $\mathbf{x}_+$  e  $\mathbf{x}_-$  os exemplos, de classes opostas, mais próximos do hiperplano de separação;  $\mathbf{w}$  o vetor do hiperplano  $h$  com norma  $\|\mathbf{w}\| = \sqrt{\mathbf{w}^T \mathbf{w}}$ , que corresponde aos pesos dos dois atributos de cada exemplo; e  $\hat{\mathbf{w}}$  o vetor unitário na direção de  $\mathbf{w}$ . Os hiperplanos de contorno  $h_+$  e  $h_-$ , paralelos a  $h$ , são representados como

$$h_+ = \mathbf{w}^T \mathbf{x}_+ + b = 1 \quad (3-9)$$

para os exemplos positivos e

$$h_- = \mathbf{w}^T \mathbf{x}_- + b = -1 \tag{3-10}$$

para exemplos negativos. Nosso objetivo, é encontrar uma fórmula que descreva a distância da margem apenas em função do vetor  $\mathbf{w}$ , pois queremos encontrar valores de  $\mathbf{w}$  que maximizem a margem. Para isto, se projetarmos o vetor  $(\mathbf{x}_+ - \mathbf{x}_-)$  na direção do vetor  $\mathbf{w}$ , normal aos hiperplanos separadores podemos obter a distância entre o ponto  $x_+$  e o hiperplano  $h_-$ , a mesma entre os dois hiperplanos  $h_+$  e  $h_-$ , a partir da equação (3-11).

$$\frac{\mathbf{w}}{\|\mathbf{w}\|} \cdot (\mathbf{x}_+ - \mathbf{x}_-) \tag{3-11}$$

Subtraindo a equação (3-9) da equação (3-10) temos

$$\mathbf{w} \cdot (\mathbf{x}_+ - \mathbf{x}_-) = 2 \tag{3-12}$$

e substituindo o resultado da equação (3-12) na equação (3-11) temos a distância  $d$  entre os dois hiperplanos  $h_+$  e  $h_-$  definida como

$$d = \frac{2}{\|\mathbf{w}\|} \tag{3-13}$$

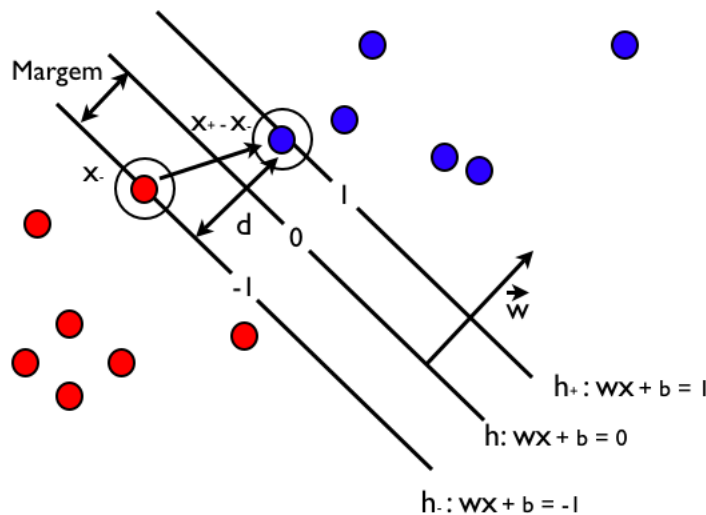


Figura 3.3: SVM - Margem.

## O Problema Primal

Aqui, descrevemos duas formulações, Margem Rígida e Margem Flexível, para problema de classificação de máxima margem. Como descrito anteriormente, para a classificação de Máxima Margem, o SVM busca maximizar  $\frac{2}{\|\mathbf{w}\|}$ , ou seja, minimizar  $\frac{1}{2}\|\mathbf{w}\|^2$ . No entanto, otimizar uma função módulo é equivalente a otimizar a função quadrática descrita como

$$\begin{aligned} & \text{Minimizar}_{\mathbf{w},b} \frac{1}{2}\|\mathbf{w}\|^2 \\ & \text{Sujeito a} \\ & y_i(\mathbf{w}^T \mathbf{x}_i + b) \geq 1, \forall i = 1, \dots, n. \end{aligned} \tag{3-14}$$

O problema de otimização obtido é quadrático, cuja solução apresenta teoria matemática ampla e já estabelecida (Dostl, 2009). Como a função objetivo é convexa e os pontos que satisfazem as restrições compõem um conjunto convexo, o problema apresenta um único mínimo global. Esta formulação, apresentada na equação (3-14), é conhecida como Margem Rígida. Isto porque possui restrições rigorosas, não permitindo nenhum erro na separação das duas classes. No entanto, para certos casos, em que existem *outliers* ou em que o problema é de maior grau de dificuldade, permitir certos erros na classificação possibilita uma melhora na generalização do classificador. Em vista disto, queremos maximizar a margem possibilitando a que alguns exemplos permaneçam além da margem. Desta forma temos a nova formulação, adaptada para

$$\begin{aligned} & \text{Minimizar}_{\mathbf{w},b} \frac{1}{2}\|\mathbf{w}\|^2 \\ & \text{Sujeito a} \\ & y_i(\mathbf{w}^T \mathbf{x}_i + b) \geq 1 - \epsilon_i, \forall i = 1, \dots, n. \end{aligned} \tag{3-15}$$

em que  $0 \leq \epsilon_i \leq 1$  é a margem de erro. Quando  $\epsilon_i < 0$  tem-se uma Variável de Folga, que corresponde a um exemplo dentro da margem de erro. Quando  $\epsilon_i > 1$ , então temos um exemplo que ultrapassou a margem. Com isto, o objetivo agora consiste em classificar os exemplos de modo a maximizar a margem e minimizar a quantidade de Variáveis de Folga. Para isto, devemos introduzir uma penalização  $C \sum_{i=1} \epsilon_i$  na restrição, levando ao problema de otimização

$$\text{Minimizar}_{\mathbf{w},b} \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1} \epsilon_i$$

$$\text{Sujeito a} \tag{3-16}$$

$$y_i(\mathbf{w}^T \mathbf{x}_i + b) \geq 1 - \epsilon_i, \epsilon_i \geq 0$$

Este tipo de formulação é chamado de Margem Flexível e foi introduzido por Cortes em (Cortes e Vladimir, 1995). Para resolver este problema, devemos considerar as duas variáveis livres  $\mathbf{w}$  e  $b$  sujeitas às restrições. Veremos a seguir, a versão dual para este problema. A versão dual é resolvida de forma mais simples que a forma primal.

### O Problema Dual

A formulação para o problema anterior, descrita na equação (3-16), pode ser relaxada a partir do emprego de Multiplicadores de Lagrange, como descrita na equação (3-17)

$$\begin{aligned} L(\mathbf{w}, b, \epsilon, \alpha, \beta) = & \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^n \epsilon_i - \\ & - \sum_{i=1}^n \alpha_i (y_i(\mathbf{w}^T \mathbf{x}_i + b) - 1 + \epsilon_i) - \sum_{j=1}^n \beta_j \epsilon_j, \end{aligned} \tag{3-17}$$

tal que,  $\alpha$  e  $\beta$  são os Multiplicadores de Lagrange.

Neste ponto, podemos dizer que resolver o problema descrito na equação (3-16) equivale a resolver  $p^*$ .

$$p^* = \text{Maximizar}_{\alpha, \beta} (\text{Minimizar}_{\mathbf{w}, b, x} L(\mathbf{w}, b, \epsilon, \alpha, \beta)) \tag{3-18}$$

$$d^* = \text{Minimizar}_{\mathbf{w}, b, x} (\text{Maximizar}_{\alpha, \beta} L(\mathbf{w}, b, \epsilon, \alpha, \beta)) \tag{3-19}$$

Pela teoria da dualidade forte, temos que  $p^* = d^*$ . Logo, podemos resolver o problema original a partir da resolução do dual  $d^*$ . Assim, devemos resolver primeiramente o problema de minimização interno. Desta forma, a partir de derivadas parciais, temos as seguintes equações

$$\frac{\partial L}{\partial b} = 0 \Rightarrow \sum_{i=1}^n \alpha_i y_i = 0 \tag{3-20}$$

$$\frac{\partial L}{\partial \mathbf{w}} = 0 \Rightarrow \mathbf{w} = \sum_{i=1}^n \alpha_i y_i x_i \tag{3-21}$$

$$\frac{\partial L}{\partial \epsilon} = 0 \Rightarrow \alpha_i + \beta_i = C \tag{3-22}$$

Aqui, a solução é ótima quando satisfaz às condições de Karush-Kuhn-Tucker (KKT)

$$\begin{aligned}\alpha_i y_i (\mathbf{w}^T \mathbf{x}_i + b) - 1 + \epsilon_i &= 0 \\ \beta_i \epsilon_i &= 0.\end{aligned}\tag{3-23}$$

A partir do emprego das equações (3-20), (3-21), (3-22) e (3-23) na equação (3-17) teremos a nova formulação

$$\begin{aligned}\text{Maximizar } & \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n y_i y_j \alpha_i \alpha_j \mathbf{x}_i \mathbf{x}_j \\ \text{Sujeito a } & \\ & \sum_{i=1}^n y_i \alpha_i = 0, \\ & 0 \leq \alpha_i \leq C.\end{aligned}\tag{3-24}$$

em que cada  $\alpha_i$  está associado a um exemplo. As condições KKT, descritas na equação (3-23), podem ser reescritas em três casos como

$$\begin{aligned}\text{se } \alpha_i = 0 & \text{ então } y_i (\mathbf{w}^T x_i + b) \geq 1, \text{ erro de classificação;} \\ \text{se } \alpha_i = C & \text{ então } y_i (\mathbf{w}^T x_i + b) \leq 1, \text{ vetor de suporte na margem;} \\ \text{se } 0 \leq \alpha_i \leq C & \text{ então } y_i (\mathbf{w}^T x_i + b) = 1, \text{ vetor de suporte dentro da margem.}\end{aligned}\tag{3-25}$$

Agora, os vetores de suporte incluem não somente os exemplos na margem, mas os exemplos além da margem. Como podemos notar, esta formulação, descrita na equação (3-24), torna a resolução do problema mais simples, pois depende apenas dos lagrangeanos  $\alpha_i$  e  $\alpha_j$ . A equação é quadrática em termos da variável  $\alpha$  e pode ser resolvida por métodos como Newton Rhapsom ou Gradiente Descendente. No entanto, uma abordagem comum para SVM, denominada Otimização Mínima Sequencial, pode ser encontrada em (Platt, 1999). Nesta, duas variáveis  $\alpha$  são otimizadas por iteração e as demais fixadas até que todas as restrições sejam satisfeitas.

### 3.5

#### Mínimos Quadrados Parciais

Introduzido por Herman Wold, o algoritmo de regressão linear Mínimos Quadrados Parciais (PLS) (Wold et al., 1983) foi desenvolvido inicialmente na área de econometria para entender a complexa relação entre diversas variáveis econômicas, de entrada e saída, na ausência de um modelo teórico. PLS obteve maior sucesso na área quimiometria (Haaland e Thomas, 1988a, Haaland e Thomas, 1988b). Tal sucesso alcançado induziu o emprego deste algoritmo em outras áreas como propaganda, monitoração de processos, processamento de imagens (Tenenhaus, 1998) e finanças (Alvim et al., 2010).

Algumas das vantagens do emprego deste algoritmo são: eficácia mesmo



em situações em que o número total de observações é inferior ao número total de atributos; pode ser empregado para reduzir o número total de atributos do problema, a partir de um resumo destes por variáveis latentes; e, por fim, é um algoritmo de fácil implementação e cujo treinamento é relativamente rápido com relação aos demais algoritmos de treinamento para Aprendizado de Máquina.

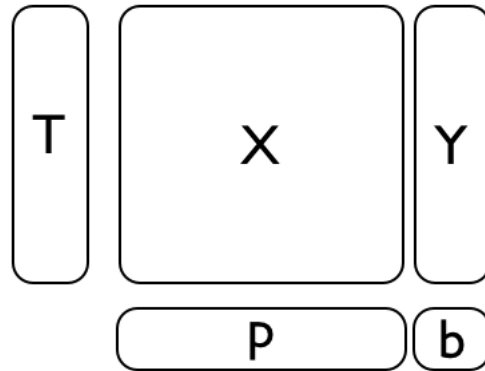


Figura 3.4: PLS - diagrama de blocos.

Formalmente, podemos descrever o algoritmo PLS da seguinte forma. Seja  $X$  uma matriz com  $n$  observações e  $m$  atributos tal que, cada linha desta matriz corresponde a uma observação e cada coluna corresponde a um atributo;  $Y$  um vetor coluna com  $n$  elementos em que, cada um corresponde a uma variável dependente associada a observação de mesma linha;  $T$  um vetor coluna com  $n$  elementos;  $P$  um vetor linha com  $m$  elementos; e  $b$  um escalar. O conceito fundamental do PLS consiste em construir um modelo que resuma, simultaneamente, dados de entrada  $X$  e dados de saída  $Y$ , em uma matriz de correlação, descrita na equação (3-26), a partir do que denominados de variável latente  $T$ .

$$X^T Y. \quad (3-26)$$

Na figura 3.4, ilustramos o conceito chave do algoritmo PLS. Nesta, a variável latente  $T$  armazena a informação tanto de  $X$  quanto  $Y$  e a variável  $b$  e o vetor linha  $P$  armazenam as informações de  $X$  e de  $Y$ , respectivamente. Desta forma, um resumo de  $X$  e  $Y$  é feito a partir dos produtos  $T \cdot b$  e  $T \cdot P$ . Para melhor esclarecimento do algoritmo, dividiremos este em duas etapas: Determinação da Estrutura Latente e Predição.

### 3.5.1

#### Determinação da Estrutura Latente

Nesta fase, o PLS aprende com os dados de treinamento. Podemos dividir esta fase em quatro subfases: Inicialização, Determinação do Vetor de Carga, Determinação da Variável Latente, Determinação dos Coeficientes de Regressão e Estimativa do Submodelo. Na fase de Inicialização, são efetuadas duas cópias dos dados:  $X_{k=0}$ , cópia de  $X$  e o mesmo para  $Y$ , em que  $k$  o índice da  $k$ -ésima iteração do algoritmo. Na fase de Determinação do Vetor de Carga, é necessário primeiramente, encontrar os autovetores ou Vetores de Carga  $W$ . Para isto, existem dois casos. Seja  $Y_{(p,q)}$  em que  $q = 1$ ,  $W_k$  pode ser obtido a partir de

$$X_{k-1}^T \cdot Y_{k-1}.$$

Neste caso estamos tratando do algoritmo PLS1 (Renteria e Milidiú, 2003). Já para o segundo caso, em que  $q > 1$ , que corresponde ao algoritmo PLS2 (Renteria e Milidiú, 2003), deve ser utilizado um método de extração de autovetores para a matriz

$$(X^T Y) \cdot (Y^T X).$$

Note que a matriz da equação (3-26) é multiplicada pela sua transposta para que seja quadrada. Isto assegura que a matriz possui autovalores reais e positivos. Para a extração utiliza-se o método de Mínimos Quadrados Parciais Não-Linear Iterativo (NIPALS) (Golub e Van Loan, 1996), a partir do Método das Potências. No entanto, nesta tese abordamos o algoritmo PLS1.

Na fase seguinte, fase da Determinação da Variável Latente  $T_k$ , projeta-se os dados de entrada  $X$  no vetor  $W$  obtendo

$$T_k = X_{k-1} \cdot W_k.$$

Posteriormente, determinamos os Coeficientes de Regressão. Como a matriz  $X_k$  é estimada a partir de

$$X_k = T_k \cdot P_k,$$

para obter  $P_k$  devemos isolá-lo. Para isto, basta fazer

$$(T_k^T T_k)^{-1} T_k^T X_k = (T_k^T T_k)^{-1} (T_k^T T_k) \cdot P_k$$

Como o resultado de  $(T_k^T T_k)$  é um número, temos que

$$P_k = \frac{T_k^T}{T_k^T T_k} X_k$$

Para obter  $b_k$ , em que

$$Y_k = T_k \cdot b_k$$

o mesmo procedimento, como feito para o cálculo de  $P_k$ , pode ser feito, levando a

$$b_k = \frac{T_k^T}{T_k^T T_k} Y_k.$$

Por fim temos a fase de Estimativa do Submodelo, que procede da seguinte forma

$$X_{k+1} = X_k - T_k \cdot P_k,$$

e

$$Y_k = Y_k - T_k \cdot b_k.$$

Os submodelos estimados de  $X$  e  $Y$  são descartados para que, na próxima iteração do algoritmo, outros submodelos sejam estimados com o resíduo de  $X$  e de  $Y$ . Após a fase de estimativas, o processo se repete até um número  $z$  de iterações desejadas. Cabe ressaltar que, efetuar  $z$  iterações é o mesmo que utilizar  $z$  variáveis latentes.. No algoritmo 1, descrevemos o pseudo-código do PLS1 completo para a determinação da estrutura latente.

---

**Algorithm 1** PLS1 - Determinação da estrutura latente.

---

```

 $X_1 \leftarrow X; Y_1 \leftarrow Y;$ 
for  $k=1$  to  $z$  do
   $W_k \leftarrow X_k^T Y$ 
   $W_k \leftarrow \frac{W_k}{(W_k W_k^T)^{1/2}}$  //normalização
   $T_k \leftarrow X_k W_k$  //variável latente.
  //cálculo dos coeficientes de regressão.
   $b_k \leftarrow \frac{T_k^T}{T_k^T T_k} \cdot Y_k$ 
   $P_k \leftarrow \frac{T_k^T}{T_k^T T_k} \cdot X_k$ 
  //descarte residual.
   $X_{k+1} \leftarrow X_k - T_k P_k$ 
   $Y_{k+1} \leftarrow Y_k - T_k b_k$ 
end for

```

---

### 3.5.2 Predição

Após a determinação da estrutura latente, deve-se, para a fase de Predição, capturar os vetores de carga e os coeficientes de regressão aprendidos. Com estes dados, poderemos estimar cada  $Y'$  a partir de um novo conjunto de dados  $X'$  e dos submodelos aprendidos. Assim, a fase de Predição pode ser descrita em três etapas: determinação da Vetor de Pontuação  $T'_k$ , com o Vetor de Carga  $W_k$  aprendido; Predição de um novo  $Y'_{k+1}$  na forma

$$Y'_{k+1} = Y'_k + T'_k \cdot P_k$$

e por fim o descarte do resíduo, como descrito no algoritmo de treino 1, mas somente para  $X_{k+1}$ . Este processo se repete para o mesmo número  $z$  de fatores aprendidos na fase de determinação de estrutura latente. No algoritmo 2, descrevemos o pseudo-código do PLS1 completo para a predição.

---

#### Algorithm 2 PLS1 - Predição.

---

```

 $X'_1 \leftarrow X'; Y'_1 \leftarrow 0;$ 
for  $i=1$  to  $k$  do
   $T'_k \leftarrow X'_k W_k$ 
  //predição residual.
   $Y'_{k+1} \leftarrow Y'_k + T'_k b_k$ 
  //descarte residual.
   $X'_{k+1} \leftarrow X'_k - T'_k P_k$ 
end for

```

---

### 3.6 Considerações Finais

Neste capítulos vimos os conceitos básicos para o entendimento desta tese. Na primeira seção deste capítulo, definimos uma carteira, risco e lucro desta e, por fim, a formulação de Média-Variância proposta por Markowitz. Concluimos que a formulação apresentada encontra o compromisso ótimo entre risco e lucro. Na seção seguinte, apresentamos o problema de Agendamento por Intervalos Ponderados e sua formulação matemática. Concluimos que este problema pode ser resolvido por um algoritmo computacionalmente eficiente ( $O(n \log n)$ ). Por fim, nas duas últimas seções, apresentamos um algoritmo regressor denominado Mínimos Quadrados Parciais e um algoritmo classificador denominado Máquina de Vetores de Suporte. Vimos algumas das vantagens do emprego de ambos os algoritmos e sua aplicabilidade em problemas de diversas áreas como no mercado financeiro, em processamento de imagens e demais.

No capítulo 4, a seguir, veremos como utilizar estes algoritmos para detectar oportunidades lucrativas de mercado.

## 4 Detecção de Oportunidades

Neste capítulo, apresentamos primeiramente a tarefa de detecção de oportunidades de negociação. Na segunda seção, descrevemos os atributos utilizados para a construção dos detectores de oportunidades, baseados em indicadores de análise técnica. Ainda nesta seção, descrevemos a técnica Janela Deslizante, muito utilizada em problemas com dados sequenciais. Na última seção, descrevemos a métrica adotada para o ajuste de parâmetros dos detectores de oportunidades.

### 4.1 A Tarefa

Dizemos que ocorre uma *oportunidade*, para negociação de um ativo  $s$  em um instante de tempo  $t$ , quando há uma previsão de aumento significativo no preço do ativo  $s$  para um instante de tempo  $t + \delta$ . Ocorrendo a oportunidade, uma estratégia de negociação simples é emitir ordens de compra para o ativo  $s$  no tempo  $t$  e emitir ordens de venda para o mesmo ativo no tempo  $t + \delta$ . Baseado nesta estratégia, a tarefa de predição de preços pode ser simplificada para uma tarefa de detecção de oportunidades em janelas de negociação  $[t, t + \delta]$ . Cada oportunidade corresponde a um aumento do preço superior a um limiar  $\epsilon$ , que representa um limite de segurança para uma previsão de uma alta de preços. Quanto maior o valor de  $\epsilon$ , mais conservadora é a estratégia, pois apenas variações de preço mais elevadas são consideradas. O limiar  $\epsilon$  ajuda no controle da agressividade da estratégia.

Seja  $p_{s,t}$  o preço de um ativo  $s$  em um instante de tempo  $t$ ; e  $p_{s,t+\delta}$  o preço do mesmo ativo  $s$  em um instante  $t + \delta$  posterior. Formalmente definimos que uma *oportunidade* ocorre quando

$$\frac{p_{s,t+\delta} - p_{s,t}}{p_{s,t}} > \epsilon \quad (4-1)$$

Fixada a largura  $\delta$  das janelas de negociação, o problema de detecção de oportunidades pode ser abordado como um problema de classificação binária tal que para cada instante de mercado  $t$ , podemos ter duas possibilidades de classe: oportunidade (+1) ou não oportunidade (0).

$$\text{classe} = \begin{cases} 1 & \text{Se oportunidade} \\ 0 & \text{Caso contrário.} \end{cases}$$

Como temos para cada instante  $t$  uma classe associada, o problema de classificação binária pode ser resolvido por um algoritmo de Aprendizado de Máquina Supervisionado. Em vista disto, propomos uma metodologia de anotação dos exemplos, que são informações de mercado até o instante  $t$ . Nossa metodologia consiste em dividir o dia de mercado em intervalos de comprimentos distintos e gerar subconjuntos de dados agrupados por comprimento de intervalo e ativo. Desta forma, para cada conjunto de dados, verificamos se existem oportunidades, a partir de cada instante  $t$  do dia de mercado, e associamos um detector de oportunidades independente.

Na figura 4.1, ilustramos um exemplo hipotético do comportamento de um ativo ao longo do dia de mercado e um subconjunto das diversas oportunidades possíveis. Na figura 4.1, o dia de mercado é discretizado e dividido em intervalos de comprimentos distintos. As oportunidades são: (a,i), (e,i), (a,c), (e,g), (a,b), (c,d), (e,f) e (g,h). Neste exemplo, teríamos um detector de oportunidades para cada intervalo de mesmo comprimento. Neste caso, teríamos quatro detectores de oportunidades.

Para os detectores de oportunidades utilizamos dois algoritmos: um classificador denominado de Máquina de Vetores de Suporte (SVM) (Boser et al., 1992) e um regressor denominado de Mínimos Quadrados Parciais (Wold et al., 1983). Os preditores gerados por estes algoritmos de Aprendizado de Máquina são reconhecidos pelo seu bom desempenho em diferentes domínios.

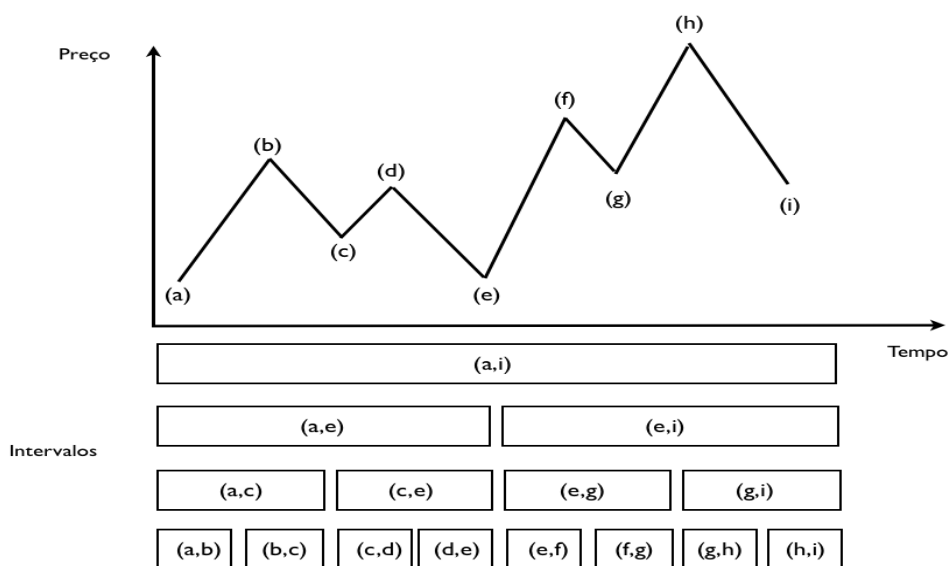


Figura 4.1: Exemplo de oportunidades possíveis.

## 4.2

### Engenharia de Atributos

Nesta seção, descrevemos os atributos utilizadas pelos detectores de oportunidade e como utilizamos as informações desses atributos para algoritmos de Aprendizado de Máquina Supervisionado.

#### 4.2.1

##### Atributos Básicos

Análise Técnica é uma ferramenta muito utilizada por investidores do mercado financeiro. A partir desta, investidores tentam encontrar, com base em um histórico de mercado, padrões que auxiliem na tomada de decisão.

Munehisa Homma é um dos investidores mais famosos a empregar a Análise Técnica, tendo ganho uma grande fortuna no mercado de arroz do Japão em 1700 (Nison, 2010). As técnicas de Munehisa Homma introduziram o que é conhecido hoje como Candlestick Patterns. Em 1800, Charles Dow e Willian Peter Hamilton desenvolveram um estudo teórico, a Teoria de Dow, que mostrava que o mercado de ações apresentava, em certos momentos, padrões previsíveis. Apesar da ampla utilização desta ferramenta por investidores, apenas em 1933 (Cowles 3rd, 1933) ocorreu a primeira publicação, em uma revista acadêmica de finanças, sobre um estudo empírico da real efetividade da Análise Técnica. O estudo indicou que as previsões propostas por Hamilton, baseadas na Teoria Down, para o período de 1904 a 1929, foram eficazes em 55% das vezes. De 1929 até 1960, não houve quase produção acadêmica nesse tema. Entre 1960 e 1970 foram publicados dois estudos, (Fama e Blume, 1966, Jensen e Benington, 1970), que indicaram a ineficácia de certos indicadores para a previsão do mercado. Tal conclusão levou a conhecida Hipótese do Mercado Eficiente (Fama, 1970), que sugere que os preços correntes de mercado já refletem toda a informação histórica de mercado e que não é possível realizar previsões de mercado baseadas em um histórico de mercado. No entanto, tais estudos foram feitos para um conjunto reduzido de ativos e para um período curto de tempo. Apesar da Hipótese do Mercado Eficiente, existem trabalhos publicados em veículos relevantes ao mercado financeiro, relacionados a Análise Técnica, onde a lucratividade propiciada por esta estratégia é evidenciada (Brock et al., 1992, Lo et al., 2000).

Na modelagem proposta nesta tese, além dos indicadores de análise técnica, utilizamos as informações de preço e volume de fechamento e abertura; preço e volume médio, máximo e mínimo para intervalos de cinco minutos.

Seja  $t$  um instante de tempo e  $n$  o número de instantes anteriores, definimos os seguintes atributos simples de mercado: preço de abertura  $p_t$ ,



preço de fechamento  $c_t$ , quantidade total  $v_t$ , preço mínimo  $min_{t,n}$ , preço máximo  $max_{t,n}$ , volume mínimo  $minv_{t,n}$  e volume máximo  $maxv_{t,n}$ . Esses atributos simples de mercado, além de utilizados individualmente no modelo, são utilizados pelos indicadores de Análise Técnica descritos a seguir.

Começaremos por um indicador responsável por capturar tendências de mercado, denominado Média Móvel Simples ( $SMA_t$ ), que pode ser descrito como

$$SMA_t = \frac{1}{n} \cdot \sum_{t=1}^n p_t.$$

Os próximos indicadores, Taxa de Mudança ( $ROC_{t,n}$ ), Williams ( $WR_{t,n}$ ) e Indicador de Força Relativa ( $RSI_{t,n}$ ), pertencem a categoria de indicadores de Momentum. Estes indicadores são responsáveis por capturar variações de preço relativas a instantes anteriores e são muito utilizados para ajudar na antecipação de variações futuras de preço.

Taxa de Mudança,

$$ROC_{t,n} = \begin{cases} 100 \cdot \frac{p_t - p_{t-n}}{p_{t-n}} & \text{se } t > 0, \\ 0 & \text{caso contrário.} \end{cases}$$

indica condições de compra em excesso e venda em excesso do mercado.

Tal como a Taxa de Mudança, Williams também mede condições de compra em excesso e venda em excesso do mercado. No entanto, Williams mede a variação do preço de fechamento em função dos preços mínimo e máximo. Segue, abaixo, a descrição do indicador Williams:

$$WR_{t,n} = \begin{cases} 100 \cdot \frac{max_{t,n} - c_t}{max_{t,n} - min_{t,n}} & \text{se } max_{t,n} - min_{t,n} > 0, \\ 100 & \text{caso contrário.} \end{cases}$$

O último indicador de momentum é o Indicador de Força Relativa, responsável por medir a velocidade das variações de preços. Indicador de Força Relativa pode ser descrito como

$$RSI_{t,n} = 100 - \frac{100}{1 + \frac{SMA_{t,n}^{up}}{SMA_{t,n}^{dw}}}$$

em que  $SMA_{t,n}^{up}$  e  $SMA_{t,n}^{dw}$  correspondem á Médias Móveis das variações de alta de preços e variações de queda de preços, respectivamente.

A partir daqui descrevemos os indicadores da categoria Quantidade que são: Balanço de Quantidade ( $OBV_t$ ), Tendência de Preços e Quantidade ( $PVT_t$ ), Indicador de Quantidade Positiva ( $PVI_t$ ), Indicador de Quantidade

Negativa ( $NVI_t$ ) e Fluxo de Dinheiro ( $MFI_t$ ).

Balanco de Quantidade,

$$OBV_t = \begin{cases} v_0 & \text{se } t = 0 \\ OBV_{t-1} + v_t & \text{se } p_t > p_{t-1} \text{ e } t > 0, \\ OBV_{t-1} - v_t & \text{caso contrário.} \end{cases}$$

mede o fluxo positivo e negativo de quantidades considerando as variações positivas e negativas dos preços. Este indicador tem por base a hipótese de que a informação prévia da quantidade auxilia na previsão do comportamento de preços futuros.

Tendência de Quantidade e Preço,

$$PVT_t = \begin{cases} v_0 \cdot ROC_{0,1} & \text{se } t = 0 \\ PVT_{t-1} + v_t \cdot ROC_{t,1} & \text{se } t > 0, \\ 0 & \text{caso contrário.} \end{cases}$$

pode ser interpretado da mesma forma que Balanco de Quantidades. No entanto, Tendência de Quantidade e Preço pondera incrementos de volume pela magnitude das variações de preço. Relacionado a este indicador, aplicamos mais dois outros indicadores muito utilizados.

Um é o Indicador de Quantidades Positivas,

$$PVI_t = \begin{cases} 0 & \text{se } t = 0 \\ PVI_{t-1} + v_t \cdot ROC_{t,1} & \text{se } v_t > v_{t-1} \text{ e } t > 0, \\ 0 & \text{caso contrário.} \end{cases}$$

que pondera o volume apenas para as variações de preço positivas. Este indicador representa a força compradora do mercado.

Outro indicador é o Indicador de Quantidades Negativas,

$$NVI_t = \begin{cases} 0 & \text{se } t = 0 \\ NVI_{t-1} + v_t \cdot ROC_{t,1} & \text{se } v_t < v_{t-1} \text{ e } t > 0, \\ 0 & \text{caso contrário.} \end{cases}$$

que é exatamente o oposto do indicador anterior e representa as forças vendedoras do mercado.

O último indicador de quantidades é o indicador de Fluxo de Dinheiro,

$$MFI_t = \frac{PMF_t}{PMF_t + NMF_t}$$

em que,  $PMF_t$  representa o fluxo de dinheiro positivo,

$$PMF_t = \begin{cases} 0 & \text{se } t = 0, \\ PMF_{t-1} + p_t \cdot v_t & \text{se } p_t > p_{t-1} \text{ e } t > 0, \end{cases}$$

e  $NMF_t$  representa o fluxo de dinheiro negativo,

$$NMF_t = \begin{cases} 0 & \text{se } t = 0, \\ NMF_{t-1} + p_t \cdot v_t & \text{se } p_t < p_{t-1} \text{ e } t > 0. \end{cases}$$

Fluxo de Dinheiro representa a razão entre as forças compradores e vendedoras que são representadas pelo fluxo de dinheiro positivo e fluxo de dinheiro negativo. Aqui, o dinheiro é representado pelas quantidades transacionadas multiplicado pelo preços do ativo num determinado período de tempo. Fluxo de dinheiro positivo representa variações de preço positivas multiplicadas pela quantidade transacionada. Fluxo de dinheiro negativo representa variações de preço negativas multiplicadas pela quantidade transacionada.

Aqui, descrevemos a última categoria, de indicadores Volatilidade, de indicadores de análise técnica. Indicadores de Volatilidade são divididos em duas categorias: os que medem o risco de mercado e os que medem o lucro por unidade de risco. Os indicadores que iremos descrever aqui são: Variância  $VAR_t$ , índice de Ulceração  $UI_t$ , Draw Down  $DD_t$ , Draw Down Máximo  $MDD_t$ , que são indicadores de risco, e Razão de Sharpe  $SHP_t$ , Razão de Martin  $MR_t$ , que são indicadores de ganho por unidade de risco.

Variância,

$$VAR_t = \frac{1}{N} \cdot \sum_{i=t}^N (p_t - SMA_t)^2$$

é a medida de risco mais simples. Este indicador considera que ativo é arriscado se este apresenta uma média elevada de variações de queda e alta de preços. Apesar de muito utilizado, alguns analistas criticam a variância como medida de risco, pois esta contabiliza variações positivas de preço como parte do risco.

Em vista disto, temos outro indicador denominado de índice de Ulceração,

$$D_t = 100 \cdot \frac{p_t - \max_{t,n}}{\max_{t,n}}$$

$$UI_t = \sqrt{\frac{\sum_{i=t}^n D_i}{n}}$$

que contabiliza apenas variações de preço em função de um valor máximo, o que é conveniente, pois variações positivas não representam risco.

Outro indicador, um dos mais utilizados para medir o risco de estratégias,

é o Draw Down. O Draw Down mede o declínio máximo de preços para um dado período de tempo. O Draw Down pode ser descrito como

$$DD_t = \begin{cases} 0 & \text{se } t = 0, \\ \max[DD_{t-1}, \frac{p_t - p_t}{p_t}] & \text{se } t > 0. \end{cases}$$

em que  $t$  é uma constante que representa o tempo corrente.

Outro indicador derivado deste é o Draw Down Máximo,

$$MDD_t = \begin{cases} 0 & \text{se } t = 0, \\ \max[MDD_{t-1}, DD_t] & \text{se } t > 0. \end{cases}$$

que representa o maior dos máximos declínios para um dado período de tempo.

Por fim descrevemos os dois últimos indicadores que representam ganho por unidade de risco.

Razão de Sharpe,

$$SHP_t = \frac{R_t - R_f}{S_t}$$

mede o retorno excedente por unidade de risco, tal que  $S_t$  representa o desvio padrão do retorno, que como a variância também representa o risco,  $R_t$  representa o retorno médio da série de preços e  $R_f$  representa o retorno de um investimento livre de risco.

Por fim, nosso último indicador, descrevemos a Razão de Martin,

$$MR_t = \frac{R_t - R_f}{UI_t}$$

este indicador utiliza como risco o índice de Ulceração.

#### 4.2.2

##### Janela Deslizante

Nesta subseção, descrevemos Janela Deslizante (JD) (Dietterich, 2002), cujo objetivo é converter as informações de um problema com dados sequenciais para a abordagem de Aprendizado de Máquina Supervisionado.

Formalmente, podemos descrever o JD da seguinte forma: Seja um conjunto de dados sequenciais  $D = \{(x_t, y_t)\}_{t=1}^n$  composto por  $n$  exemplos, tal que cada par  $(x_t, y_t)$  contém um vetor de atributos  $x_t$  e uma variável dependente  $y_t$  para o tempo  $t$ . A abordagem JD consiste em mapear a uma saída  $y_t$  o conjunto dos últimos  $w$  atributos  $x_{t-w}, \dots, x_{t-1}, x_t$  até o instante  $t$ . O tamanho da janela  $w$ , corresponde ao número de instantes passados que se deseja utilizar.

Com o emprego do JD, algoritmos de Aprendizado de Máquina Supervisionado podem achar combinações lineares ou não-lineares entre atributos de tempos distintos, o que contribui para a melhora do desempenho dos algoritmos para a tarefa (Dietterich, 2002).

### 4.3 Métricas

Como cada detector de oportunidade possui um algoritmo de Aprendizado de Máquina e este apresenta parâmetros, devemos ajustar esses de forma a melhorar a qualidade da detecção de oportunidades. Não há um consenso de qual métrica utilizar para ajustar os parâmetros de algoritmos quando o objetivo final é o lucro da negociação de mercado (Atsalakis e Valavanis, 2009b). No entanto, a abordagem mais comum consiste em utilizar uma métrica próxima da função objetivo de negociação que, em nosso caso, é o lucro. Também não há consenso quanto à heurística adequada para o ajuste de parâmetros. Nesta tese, utilizamos a heurística Grid Search (Hsu et al., 2000), muito empregada para o algoritmo SVM. Esta heurística é guiada pela métrica de lucro médio, descrita a seguir como

$$r = \frac{1}{n} \sum_{i=1}^n c_i r_i, \quad (4-2)$$

em que

$$r_i = \frac{p_{i+1} - p_i}{p_i}$$

e

$$c_i = \begin{cases} 1 & \text{se oportunidade estimada} \\ 0 & \text{caso contrário.} \end{cases}$$

tal que  $n$  é o número de exemplos;  $p_{i+1}$  e  $p_i$  são, respectivamente, os preços dos ativos para os instantes  $i + 1$  e  $i$ ;  $c_{i+1}$  e  $c_i$  são, respectivamente, classes de *oportunidade* e *não-oportunidade* estimadas pelo detector. Quando o detector estima a classe *oportunidade* e acerta a classificação, este é premiado por  $\frac{r_i}{n}$ , cujo valor é positivo. Quando o detector estima a classe *oportunidade* e erra a classificação, este é penalizado por  $\frac{r_i}{n}$ , cujo valor é negativo. No entanto, quando o detector estima a classe *não-oportunidade*, independente se acerta ou erra a classe, não contabiliza na métrica. Desta forma, enfatizamos que não estamos preocupados em penalizar ou premiar a classe *não-oportunidade*, pois o que nos dá lucro é acertar a classe *oportunidade*.

#### 4.4

#### Considerações Finais

Neste capítulo, apresentamos a tarefa de detecção de oportunidades de negociação, a engenharia de atributos e métricas dos detectores de oportunidade. Na primeira seção, vimos que esta tarefa pode ser facilmente transformada em um problema de classificação binária e este pode ser resolvido por um algoritmo de Aprendizado de Máquina Supervisionado. Na seção Engenharia de Atributos, podemos concluir que indicadores de Análise Técnica foram e ainda são amplamente utilizados por acadêmicos e investidores de mercado para o apoio à tomada de decisão. Muitos dos estudos empíricos enfatizam sua eficácia e contradizem a Hipótese do Mercado Eficiente (Fama, 1970). Nesta mesma seção, apresentamos a técnica de Janela Deslizante que contribui para melhoria da qualidade dos algoritmos de Aprendizado de Máquina clássicos em problemas de natureza sequencial. Por fim, na última seção, apresentamos a métrica de avaliação dos detectores de oportunidades, que corresponde ao lucro médio. Descartamos as métricas de acurácia, precisão e demais devido a estar distante do objetivo final das estratégias, que é o lucro..

## 5

### Construção Eficiente de Estratégias de Negociação

Neste capítulo, apresentamos duas formulações: uma para o problema de maximização do retorno acumulado e outra para o problema de maximização do retorno acumulado com o controle do risco. Para isto, descrevemos, primeiramente, o conceito de Trader Atômico, utilizado nestes dois problemas. Na seção seguinte, abordamos o problema de maximização de lucro e propomos a estratégia de Composição de um Time de Traders. Na última seção, abordamos o problema de maximização de lucro com controle do risco e propomos a estratégia de Composição de Carteiras por Intervalo.

#### 5.1

##### Trader Atômico

Primeiramente, definimos o conceito de Trader Atômico  $\tau$ . Dividindo o dia de mercado em intervalos discretos de tempo  $[b, e]$ , podemos representar um Trader Atômico  $\tau$  por uma tripla  $(s, b, e)$ , tal que  $s$  representa um ativo,  $b$  é o tempo inicial do intervalo de negociação,  $e$  o tempo final do intervalo de negociação e  $b < e$ . Um Trader Atômico  $\tau = (s, b, e)$  satisfaz as seguintes condições:

- (a) opera um único ativo  $s$ .
- (b) está associado apenas ao intervalo  $[b, e]$ .
- (c) apresenta apenas dois estados: *operante*,  $t_{s,b} = 1$  em que uma compra é feita no instante  $b$  e uma venda é feita no instante  $e$ , ou *não operante*  $t_{s,b} = 0$ .

De forma a computar o lucro de um Trader Atômico  $\tau$  para um dia  $d$ , associamos um prêmio  $r_{\tau,d}$  que representa a razão entre preços para um intervalo, descrita por

$$r_{\tau,d} = \left[ \frac{p_{s,e,d}}{p_{s,b,d}} \right]^{t_{s,b,d}}$$

tal que  $p_{s,b,d}$  e  $p_{s,e,d}$  representam, respectivamente, os preços do ativo  $s$  para os instantes  $b$  e  $e$  no dia  $d$  e  $t_{s,b,d}$  corresponde ao estado do Trader Atômico para o dia  $d$ .

Adicionalmente, denominamos de retorno  $l_{\tau,d}$  para um Trader Atômico como

$$l_{\tau,d} = \left[ \frac{p_{s,e,d} - p_{s,b,d}}{p_{s,b,d}} \right]^{t_{s,b,d}}$$

## 5.2

### Composição de um Time de Traders

Com um único Trader Atômico dentro do dia, tal que a duração do intervalo é menor que a duração do dia de mercado, ainda há espaço para a adição de mais Traders Atômicos. Desta forma, consideramos que dois Traders Atômicos distintos  $\tau = (s, b, e)$  e  $\tau' = (s', b', e')$  não possuem intervalos sobrepostos, se e somente se,

$$(b' - e) > 0 \text{ ou } (b - e') > 0. \quad (5-1)$$

Podemos então representar uma estratégia de negociação que possui  $k$  Traders Atômicos não-sobrepostos para o dia como

$$T = \{(s_i, b_i, e_i)\}_{i=1}^k.$$

Desta forma, denominamos  $T$  como uma estratégia bem formada.

Vamos formular o problema para a estratégia de Composição de um Time de Traders (CTT). Primeiramente, assumiremos que nosso objetivo é operar com  $n$  ativos pertencentes a um conjunto  $S$  de ativos. Seja  $A = \{(s_i, b_i, e_i)\}_{i=1}^n$  um conjunto disponível de Traders Atômicos tal que todos os ativos  $s_i$  pertencem a  $S$ .

Agora, assumimos que nos é fornecido um conjunto de dados com  $m$  dias de mercado consecutivos tal que cada dia possui  $H$  horas de operação. Então, para cada dia de operação  $d$ , instante de operação  $j$  e ativo  $s \in S$ , o conjunto de dados provê o preço correspondente a  $p_{s,j,d}$ .

A partir dos preços fornecidos pelo conjunto de dados, podemos calcular as operações de negociação  $t_{s,b,d}$  e recompensas  $r_{\tau,d}$  e  $r_{\tau}$  para todos  $\tau = (s, b, e) \in A$ .

Dado o conjunto de dados de entrada, devemos escolher os Traders Atômicos que farão parte da estratégia  $T$ , nossa estratégia bem formada. Então, para cada Trader Atômico  $\tau \in A$ , definimos uma correspondente variável de decisão  $x_{\tau}$  como



$$x_\tau = \begin{cases} 1 & \text{se } \tau \in T \\ 0 & \text{caso contrário.} \end{cases}$$

Assim, o prêmio final para a estratégia  $T$ , considerando os  $m$  dias do conjunto de dados, é

$$\prod_{d=1}^m \prod_{\tau \in A} r_{\tau,d}^{x_\tau}$$

Seja  $r_\tau^{x_\tau}$  o prêmio total de um Trader Atômico  $\tau$  para  $m$  dias, descrito como

$$r_\tau^{x_\tau} = \prod_{d=1}^m r_{\tau,d}^{x_\tau}$$

Podemos representar, de forma equivalente, o prêmio final para a estratégia  $T$ , considerando os  $m$  dias do conjunto de dados, como

$$\prod_{\tau \in A} r_\tau^{x_\tau}.$$

Formalmente, o problema para a estratégia  $CTT$  é representado como

$$\text{Maximizar } \prod_{\tau \in A} r_\tau^{x_\tau} \quad (5-2)$$

tal que

$$T = \{\tau \in A | x_\tau = 1\} \text{ é uma estratégia bem formada.}$$

Como maximizar o prêmio acumulado equivale a maximizar a soma do  $\log$  de cada prêmio, então obtemos

$$\sum_{\tau \in A} l_\tau x_\tau$$

tal que  $l_\tau = \log(r_\tau)$ .

Portanto,  $CTT$  pode ser transformado no problema

$$\text{Maximizar } \sum_{\tau \in A} l_\tau x_\tau \quad (5-3)$$

tal que

$$T = \{\tau \in A | x_\tau = 1\} \text{ é uma estratégia bem formada.}$$

Como podemos observar, este último problema é equivalente ao problema de Agendamento de Intervalos Ponderados visto no capítulo 3. Fazendo uma

analogia ao problema de Agendamento, o retorno de uma tarefa corresponde ao logaritmo do produto de todas as razões de preços de Traders Atômicos que atuam num mesmo intervalo do dia, ou seja, a soma de todos os retornos das tarefas em um mesmo intervalo. Queremos encontrar a combinação ótima de tarefas que forneça a maior soma dos retornos.

Na figura 5.1, ilustramos um exemplo com dois ativos e doze Traders Atômicos distribuídos em dois dias de mercado. Cada linha corresponde a um Trader Atômico tal que o valor acima de cada linha indica o prêmio retirado pela decisão de compra e venda. Note que dois Traders Atômicos possuem prêmios iguais a zero. Isto ocorre quando não é tomada uma decisão de compra e venda. Podemos observar também que existem dois valores de prêmio inferiores a 1, o que significa que houve uma decisão errônea de compra e venda.

Na figura 5.2, apresentamos os prêmios gerados aplicando-se a função *log* nos produtos das operações dos Traders Atômicos de mesmo ativo e intervalo. Com este novo cenário, resolvemos o problema de Agendamento de Intervalos Ponderados. A solução ótima para o problema está ilustrada na parte inferior da linha do tempo.

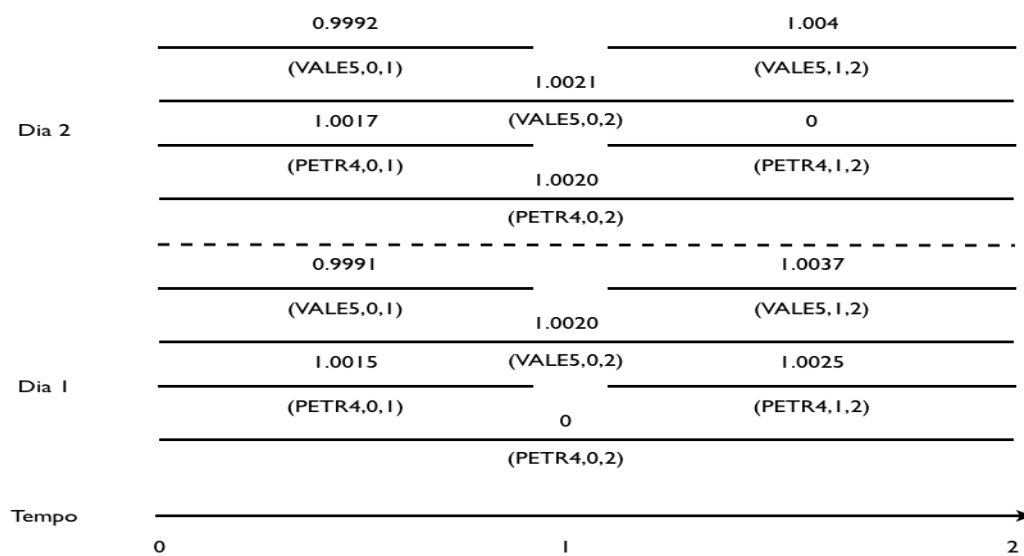


Figura 5.1: Exemplo de distribuição de operações para dois dias de mercado.

### 5.3

#### Composição de Carteiras por Intervalos

A estratégia CTT maximiza o prêmio total independente do risco. Naquela, investe-se todo o montante de dinheiro em um ativo por vez, mesmo havendo uma diversificação natural de ativos ao longo do tempo, o que pode

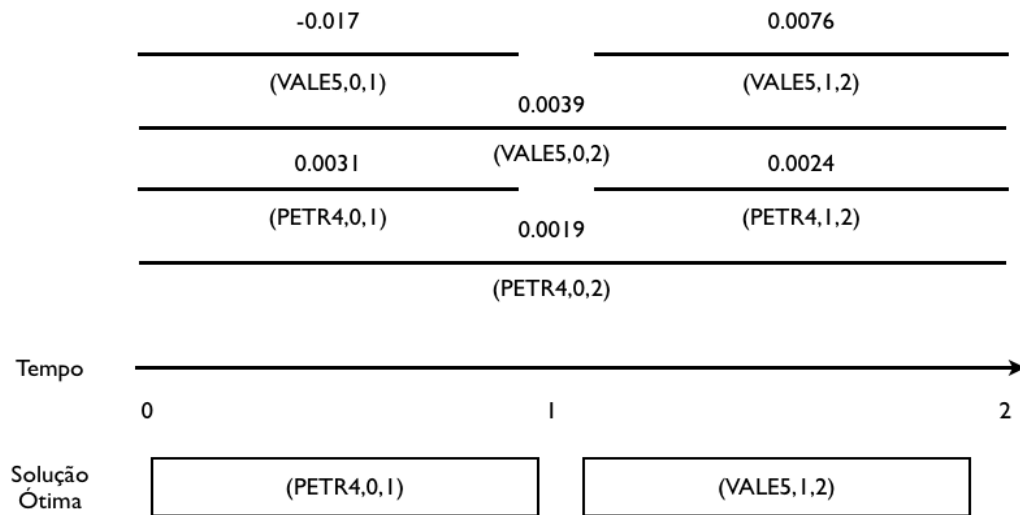


Figura 5.2: Exemplo de solução para o problema de Agendamento de Intervalos Ponderados.

ser uma estratégia inadequada para certos perfis investidores. No entanto, determinados investidores necessitam ter um maior controle do risco de forma a medir prováveis perdas ao longo do tempo.

Dado este cenário, desejamos controlar o risco por intervalo através da diversificação de Traders Atômicos para o mesmo intervalo. Para isso, permitimos que um conjunto de Traders Atômicos compartilhem um mesmo intervalo. Desta forma, introduzimos o conceito de carteira de Traders Atômicos para cada intervalo. Ao invés de compor uma carteira a partir do desempenho de cada ativo, compomos uma carteira a partir dos desempenhos dos traders que estão associados a um ativo cada.

Seja  $i = (b, e)$  um intervalo e  $I$  um conjunto com todos os intervalos do dia de mercado, tal que  $i \in I$ , consideramos que dois intervalos  $i = (b, e)$  e  $i' = (b', e')$  não são sobrepostos, se e somente se, a equação (5-1) é satisfeita. Podemos então representar uma estratégia de negociação bem formada com múltiplas carteiras não-sobrepostas como

$$T = \{C_i\}_{i \in I} .$$

tal que,  $C_i = \{(i, s_k, w_k, p_k)\}_{k=1}^n$  representa uma carteira em que:  $k$  é um ativo da carteira,  $n$  é o número total de ativos da carteira,  $i \in I$  é o intervalo de operação de uma carteira tal que  $i = (b, e)$ ,  $s_k$  um trader que opera em um ativo individual  $k$  da carteira,  $w_k \in [0, 1]$  o percentual de investimento associado ao ativo da carteira e  $p_k \in [0, 1]$  o retorno estimado para um trader da carteira. O retorno  $p_k$  de cada trader em  $i$  pode ser estimado a partir do

histórico das operações para o intervalo  $i$  utilizando-se uma média ou produto dos rendimentos de cada operação. Representamos o retorno esperado de um trader  $s_k$  para uma carteira  $C_i$  por  $r_{i,s_k} = w_k p_k$ . Representamos, também, o retorno total  $R_i$  de uma carteira como  $R_i = \sum_{k=1}^n w_k p_k$ . Representaremos o risco a partir da métrica de correlação entre o histórico dos desempenhos, descrito na seção 3.2 do capítulo 3, de uma carteira pela variável  $Z_i$ .

Vamos formular o problema da estratégia de Composição de Carteiras por Intervalos (CCI). Para isso, devemos escolher as carteiras que farão parte da estratégia  $T$ , nossa estratégia bem formada. Então, para cada intervalo  $i \in I$ , definimos uma correspondente variável de decisão  $x_i$  como

$$x_i = \begin{cases} 1 & \text{se } i \in T \\ 0 & \text{caso contrário.} \end{cases}$$

Dado que temos as informações de retorno e risco de cada carteira e a variável de decisão, podemos descrever o problema de Composição de Carteiras por Intervalos (CCI) como

$$\text{Maximizar } \prod_{i \in I} \left( \frac{R_i}{Z_i} \right)^{x_i} \tag{5-4}$$

tal que o retorno  $R_i$  é penalizado pelo risco  $Z_i$ .

Como queremos maximizar o prêmio acumulado considerando o risco, aplicamos a função *log* na equação (5-4) e transformamos no problema de Agendamento de Intervalos Ponderados,

$$\text{Maximizar } \sum_{i \in I} l_i x_i \tag{5-5}$$

tal que  $l_i = \log\left(\frac{R_i}{Z_i}\right) = \log(R_i) - \log(Z_i)$  e que

$T = \{i \in I | x_i = 1\}$  é uma estratégia bem formada.

A solução ótima apresenta a maior diferença entre os logaritmos dos retornos das carteiras e os logaritmos dos riscos. Caso o investidor deseje fixar um mesmo risco para todas as carteiras, o problema pode ser simplificado para

$$\text{Maximizar } \prod_{i \in I} R_i^{x_i} \tag{5-6}$$

o que nos leva ao problema de Agendamento de Intervalos Ponderados

$$\text{Maximizar } \sum_{i \in I} l_i x_i \tag{5-7}$$

tal que  $l_i = \log(R_i)$  e que

$$T = \{i \in I | x_\tau = 1\} \text{ é uma estratégia bem formada.}$$

Na figura 5.3, apresentamos um exemplo com três intervalos possíveis e apenas dois Traders Atômicos que operam em um ativo cada. Para cada intervalo do dia, temos uma carteira de traders  $C_i$  com pesos para cada trader  $w_i$ , retorno total  $R_i$  e risco total  $Z_i$ . Criamos um cenário que indica os prêmios  $a$ ,  $b$  e  $c$  associados para cada carteira tal que a solução ótima respeita a condição  $a + b > c$ .

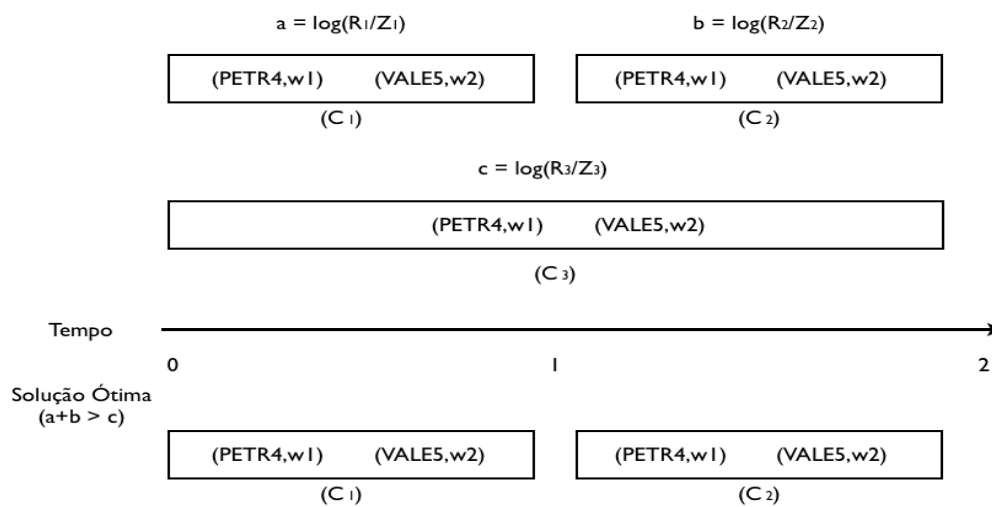


Figura 5.3: Exemplo de carteiras de traders por intervalo.

## 5.4 Considerações Finais

Neste capítulo, apresentamos nossas contribuições teóricas para dois problemas: maximização do lucro e maximização do lucro com controle do risco. Para a formulação de maximização do lucro, vimos que o produtório da razão do preços pode ser transformado em um somatório dos lucros a partir do emprego da função log. Para a formulação de maximização do lucro com o controle do risco, vimos que o produtório da razão entre lucro e risco, pode ser transformado em um somatório das diferenças entre os logs do lucro e do risco. Para as duas formulações descritas neste capítulo, mostramos que as mesmas podem ser transformadas na formulação do problema de Agendamento por Intervalos Ponderados. Este, que poder resolvido de forma eficiente computacionalmente.

## 6 Sistema de Trading

Um sistema de trading possui um papel fundamental para a realização de experimentos de qualidade. Neste capítulo descrevemos as premissas de mercado adotadas para a simulação, custos operacionais, métricas de desempenho para avaliação das estratégias propostas e por fim a arquitetura de nosso sistema de Trading.

### 6.1 Premissas de Mercado

A seguir, esclarecemos certas premissas do cenário de simulação (Boyarshinov, 2005). As premissas são:

- (a) *Tudo ou nada* significa que ou o montante de dinheiro é inteiramente investido no mercado ou o montante está guardado na carteira.
- (b) *Mercado fracionário* significa que quantidades fracionadas de ações podem ser compradas ou vendidas a qualquer momento.
- (c) *Impacto de mercado* corresponde ao efeito causado pela compra ou venda de um ativo por um participante de mercado. A cada compra o preço do ativo aumenta, prejudicando o próprio comprador. Para a venda de ativos, este efeito ocorre. Como a maioria dos trabalhos, nesta tese não consideramos o impacto mercado. Isto porque não é possível mensurar com exatidão este efeito em simulação.

O *impacto de mercado* não é significativo para ações de moderada à alta liquidez, que é o nosso caso.

- (d) *Estratégia longa* significa que só pode ser executada uma ordem de venda de um ativo, se houver previamente uma ordem de compra do mesmo ativo.

## 6.2

### Custos Operacionais

Com o objetivo de reproduzir um cenário mais realístico, nós consideramos os seguintes custos operacionais (Gomide e Milidiú, 2010): imposto de renda e taxa de corretagem.

A taxa de corretagem é a taxa cobrada pelas corretoras aos seus clientes investidores. Esta taxa pode ser cobrada com base em desconto percentual no valor da operação ou desconto constante por operação. A forma de desconto é estipulada pela corretora. Para o nosso caso, consideramos o desconto de US\$11 por operação (Gomide e Milidiú, 2010), taxa comum empregada em grandes corretoras.

Com relação ao imposto de renda, deduzimos 20% dos ganhos. Esta dedução está de acordo com as regras atuais de mercado do Brasil.

## 6.3

### Métricas de Desempenho

Para avaliar o desempenho dos traders, nós empregamos as métricas usuais encontradas na literatura. Nós dividimos as métricas em dois módulos:

- *Módulo de trading* contém indicadores de lucro anual, lucro médio, lucro mínimo e máximo para o dia de mercado. Adicionalmente, medimos o número de operações de negociação que lucram, o número de operações de negociação que dão prejuízo, o número de médio de operações de negociação e o ganho médio por operação de negociação.
- *Módulo de risco* contém os seguintes indicadores: Valor em Risco e Valor em Risco Condicional, descritos no capítulo 3, para avaliar a estratégia CCI; Draw Down Máximo e Índice de Ulceração para a estratégia CTT. Adicionalmente, calculamos o ganho por unidade de risco para a estratégia CTT, a partir dos indicadores de Sharpe, Calmar, Sterling, Martin e Burke. Estes indicadores são comumente utilizados para avaliar o risco de estratégias de negociação. Estes indicadores de risco estão detalhados no capítulo 4.

## 6.4

### Arquitetura do Sistema de Trading

A arquitetura de trading proposta neste tese contém três módulos: aprendizado dos detectores de oportunidade, construção da estratégia de trading e execução de ordens.

Primeiramente, no aprendizado dos detectores de oportunidade, treinamos um algoritmo de Aprendizado de Máquina por ativo e por comprimento de intervalo. Durante o treinamento, salvamos os modelos de cada algoritmo e aplicamos em um novo conjunto de dados, que corresponde ao conjunto de validação. Então, agrupamos os resultados por Trader Atômico, que opera em um ativo e por intervalo do dia, e computamos a métrica de retorno.

Na figura 6.1, ilustramos um exemplo para  $n$  dias  $D_n$  de mercado que possui quatro detectores de oportunidades  $C_i$  e seis Traders Atômicos  $T_i$  que produzem seis resultados de lucro  $R_i$ . Cada retângulo representa um intervalo do dia. Para este exemplo, nossos detectores de oportunidades correspondem a classificadores SVM. Nesta etapa, ajustamos os parâmetros de cada detector em um conjunto de validação que é distinto do conjunto de treino. Neste processo utilizamos o Grid Search e resgatamos o modelo cuja classificação ou regressão acarreta em maior lucro para os Traders Atômicos que atuam nos intervalos correspondentes.

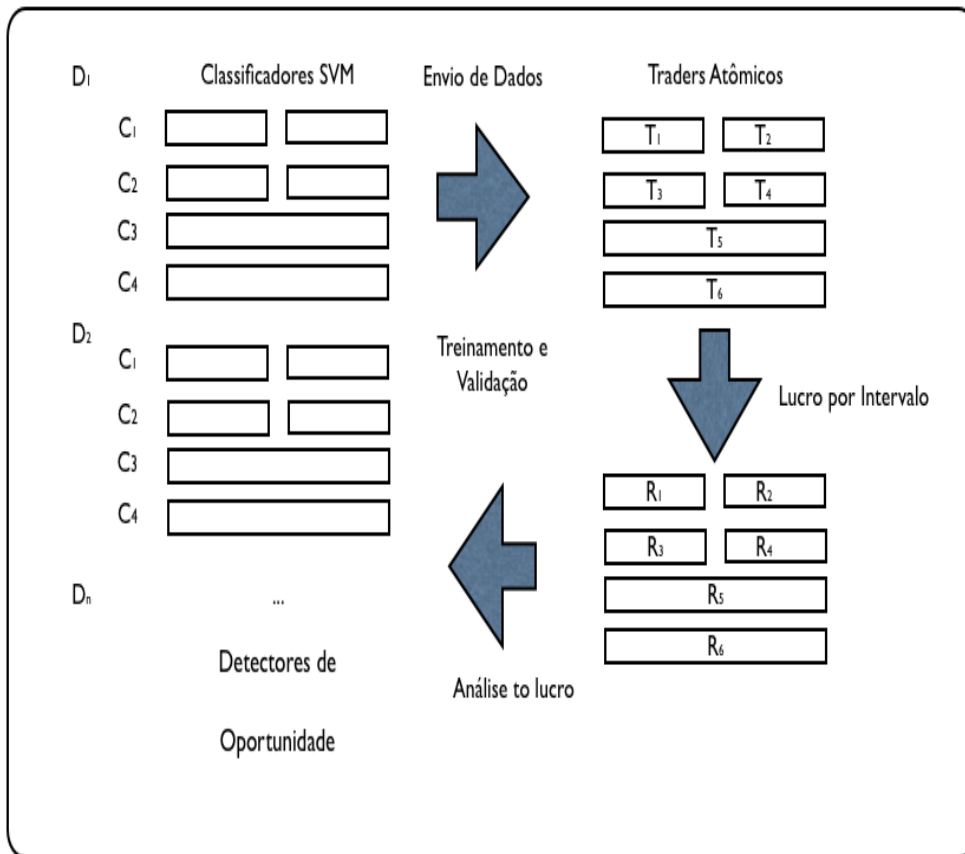


Figura 6.1: Etapa de detecção de oportunidades.

A segunda etapa, que corresponde à construção da estratégia de nego-



ciação, depende do tipo de problema a ser resolvido. Para a maximização do retorno, como temos os retornos associados a cada Trader Atômico, basta resolver o problema de Agendamento de Intervalos Ponderados por programação dinâmica, que corresponde à estratégia de Composição de um Time de Traders descrita no capítulo 5. Para a maximização do retorno com controle do risco, devemos antes aplicar, para cada intervalo do dia, o método de seleção de carteira de Markowitz. Após a execução de cada seletor de carteiras, temos o retorno e o risco por intervalo. Com isto, associamos, para cada intervalo, o retorno dividido pelo risco e resolvemos o problema de Agendamento de Intervalos Ponderados por Programação Dinâmica. A etapa de seleção de carteiras por intervalos e posterior resolução do problema de Agendamentos por Intervalos Ponderados corresponde à estratégia de Composição de Carteiras por intervalos. Ao final desta etapa, armazenamos a solução do problema de Agendamento por Intervalos Ponderados. Para o caso da estratégia de Composição de Carteiras por Intervalos, devemos armazenar, além de cada carteira selecionada para o intervalo, as informações das proporções de investimento para cada Trader Atômico da carteira.

Na figura 6.2, ilustramos esta etapa de construção das estratégias de negociação, que dependendo do objetivo do investidor, maximização do retorno ou maximização do retorno com controle do risco, o sistema opera de forma distinta.  $R_i$  é o prêmio de cada intervalo do problema de maximização de retorno;  $M_i$  corresponde ao retorno estimado por intervalo que é utilizado para cada carteira; e  $S_i$  representa o risco. Dadas as informações, em ambos os casos, resolve-se o problema de Agendamento de Intervalos Ponderados. A solução para o problema de maximização de retorno é um conjunto de traders com o melhor desempenho. Para a maximização com controle do risco, a solução corresponde a um conjunto de carteiras com o melhor desempenho.

Por fim, na etapa de execução de ordens, aplicamos a solução encontrada no conjunto de validação, em um conjunto de teste independente. Para o problema da maximização do retorno, cada Trader Atômico da solução investirá todo montante de dinheiro em seu intervalo e ativo de operação no caso de uma sinalização do detector de oportunidade. Caso não seja sinalizada uma oportunidade pelo detector, então o dinheiro não é investido. Ao final de cada intervalo, o montante total de dinheiro é repassado para o próximo Trader Atômico. Cabe ressaltar que não há rateio do montante entre os traders. Para o problema da maximização do retorno com controle do risco, o montante de dinheiro é fracionado e alocado pelos traders de cada carteira que opera em um determinado intervalo. Cabe ressaltar que os percentuais de investimento  $w_i$  de cada ativo não necessariamente são iguais. Ao final de cada intervalo,

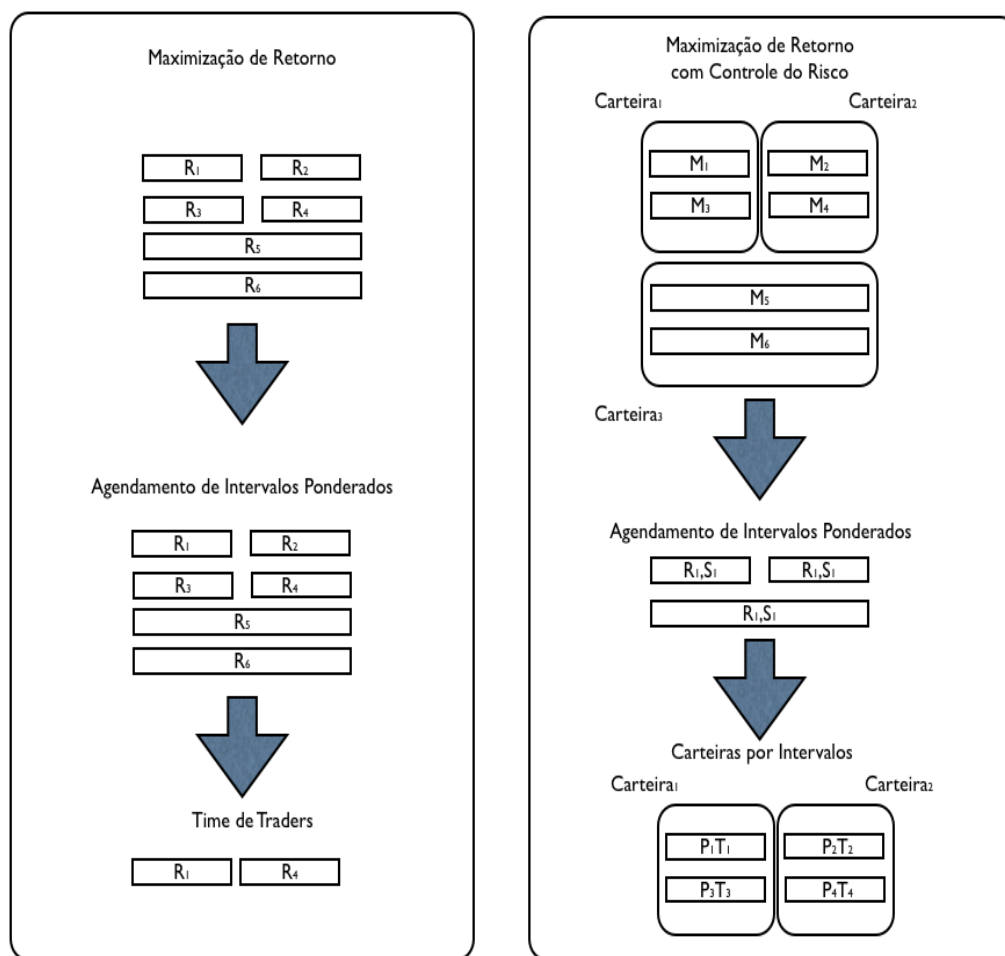


Figura 6.2: Etapa de construção das estratégias de negociação.

o montante de dinheiro é repassado à próxima carteira. Nesta etapa, cada carteira possui um conjunto de Traders Atômicos que podem ou não operar. Isto porque cada Trader Atômico possui um detector de oportunidade, que pode ou não detectar uma nova oportunidade para o intervalo. Mesmo que um Trader Atômico não opere, as proporções de investimento continuam as mesmas para os demais Traders Atômicos da carteira.

### 6.5 Considerações Finais

Neste capítulo vimos o sistema de trading proposto para operacionalizar e avaliar as estratégias de trading propostas. Vimos que o sistema considera os custos operacionais para operacionalização das estratégias, o que é fundamental para uma simulação mais próxima da realidade de mercado. Adicionalmente, vimos que o sistema de trading considera como métricas de desempenho indicadores de lucro, risco e de trading. Estas métricas são de fundamental importância para a análise final das estratégias. Além disto, tais métricas são

amplamente implementadas no mercado e na academia. Por fim, vimos que o sistema de trading atende às duas estratégias propostas nesta tese.

descrevemos as premissas de mercado adotadas para a simulação, custos operacionais, métricas de desempenho para avaliação das estratégias propostas e por fim a arquitetura de nosso sistema de Trading.

## 7

### Avaliação Empírica

Neste capítulo, descrevemos a avaliação empírica das estratégias CTT e CCI propostas nesta tese. Na primeira seção, descrevemos nosso conjunto de dados de ações utilizado para os experimentos. Para este conjunto de dados, descrevemos o total de ativos, os setores de cada ativo e estatísticas de lucro por ativo. Por fim, indicamos nossa metodologia de validação, e as divisões percentuais do conjunto de dados por período. Na seção seguinte, apresentamos informações dos nossos detectores de oportunidades. Primeiramente indicamos as implementações dos algoritmos básicos aqui adotados e aplicados à cada estratégia. Após isto, descrevemos o total de detectores de oportunidades e as resoluções de negociação utilizadas. Por fim, descrevemos os parâmetros dos detectores de oportunidades. Estes que possuem atributos de análise técnica com diferentes janelas de tempo. Na seção posterior, denominada baselines, justificamos a escolha dos baselines adotados para comparação das estratégias propostas. Fazemos isto referenciando estudos da literatura acadêmica. Em seguida, descrevemos dois baselines, um para cada estratégia proposta. Na seção seguinte, apresentamos os experimentos relacionados à estratégia CTT. Primeiramente, apresentamos os resultados do baseline sugerido na seção anterior. Posteriormente, apresentamos os resultados de negociação da estratégia CTT em função de métricas adotadas de desempenho de negociação. Ao final da seção, descrevemos o impacto dos custos operacionais para a estratégia CTT. Na seção seguinte, apresentamos os experimentos relacionados à estratégia CCI. Esta seção consiste de metodologia similar a da seção anterior, em que apresentamos os resultados do baseline, métricas de desempenho de negociação da estratégia CCI e impacto dos custos operacionais. No entanto, devido aos métodos de seleção de carteiras e seus respectivos parâmetros, também descrevemos o ajuste destes parâmetros. Por fim, na última seção, apresentamos um resumo dos resultados das duas estratégias propostas.

#### 7.1

## Conjunto de Dados

Para os experimentos, utilizamos os dez ativos de mais alta liquidez da BMF&Bovespa no ano de 2009. Os ativos foram fornecidos por uma corretora. Para escolher os ativos, selecionamos dezesseis ativos de maior liquidez. Desses dezesseis ativos, seis foram removidos devido a estarem em períodos distintos.

Na Tabela 7.1, listamos os dez ativos e seus respectivos setores. Os ativos estão distribuídos em quatro setores: mineração, petróleo, bebidas, telecomunicações e finanças. Esta diversificação dos ativos em setores distintos é de fundamental importância para o bom desempenho das estratégias propostas. Isto porque a diversificação acarreta em maiores possibilidades de ganhos para estratégias de maximização do lucro e menor risco para estratégias que consideram o risco.

Aqui apresentamos o método de divisão do conjunto de dados para os experimentos, que está relacionado com a metodologia empírica de validação dos experimentos. Dividimos o conjunto de dados em treino, validação e teste com as proporções 60%, 20% e 20%. Esta técnica de validação é denominada de Hold-Out. Aqui, não definimos a melhor partição para este tipo de técnica. No entanto, as proporções descritas anteriormente, são as mais comuns. Não utilizamos a validação cruzada, pois nesta, dados de partições com datas futuras seriam utilizadas para validar dados passados. Processo este que não faz sentido para problemas que abordam séries temporais.

Utilizamos o conjunto de validação para ajustar parâmetros dos detectores de oportunidades e dos traders. Utilizamos o conjunto de teste para aplicar a melhor configuração obtida no conjunto de validação. Ressaltamos que o critério para a escolha dos traders é a sua lucratividade, independente do número de operações de compra e venda necessárias. Isto porque estamos considerando a taxa de corretagem constante e, desta forma, os ganhos compensam esse custo constante quanto maior for o montante investido.

<b>Ativo</b>	<b>Empresa</b>	<b>Setor</b>
ITSA4	Itausa Investimentos Itau	Finanças
PETR4	Petrobras	Petróleo
USIM5	Usiminas	Mineração
VALE5	Vale do Rio Doce	Mineração
BBDC4	Banco Bradesco	Finanças
CSNA3	Cia. Siderúrgica Nacional	Mineração
AMBV4	Ambev	Bebidas
BBAS3	Banco do Brasil	Finanças
VIVO4	Vivo	Telecomunicações
TNLP4	Telemar	Telecomunicações

Tabela 7.1: Informações do conjunto de dados.

Na Tabela 7.2, listamos a média dos retornos diários de cada ativo para os conjuntos de dados de validação e de teste. Aqui, o retorno corresponde a diferença relativa entre os preços de abertura e de fechamento no dia de mercado. Podemos observar na Tabela 7.2 que, no período de validação, a maioria dos ativos tem retorno médio positivo. Para o conjunto de teste, a maioria dos ativos apresenta retorno negativo. Esta variação do retorno médio é bastante comum no mercado acionário, principalmente em períodos de crise. A desvantagem de um conjunto de dados com comportamentos distintos no treino e teste é a possível queda de desempenho das estratégias. Isto porque, para nosso enfoque, o número de oportunidades é menor em períodos de queda. No entanto, este é um cenário mais próximo à realidade de mercado e serve como um limite inferior.

Ativo	Retorno	Médio (%)
	Validação	Teste
ITSA4	0,065	0,001
PETR4	0,028	-0,198
USIM5	0,213	-0,101
VALE5	0,150	-0,191
BBDC4	0,057	-0,006
CSNA3	0,195	-0,206
AMBV4	0,164	-0,012
BBAS3	0,174	-0,003
VIVO4	0,132	-0,258
TNLP4	-0,037	-0,112

Tabela 7.2: Retorno médio por ativo.

## 7.2

### Detectores de Oportunidade

Para os detectores de oportunidades, utilizamos dois algoritmos, um para cada estratégia. Para a estratégia CTT empregamos o algoritmo regressor PLS (Wold et al., 1983). Para a estratégia CCI empregamos o algoritmo classificador SVM (Boser et al., 1992). Para o SVM utilizamos a implementação Liblinear (Fan et al., 2008). Cabe ressaltar que, os dois algoritmos apresentam apenas um parâmetro cada, sendo os mesmos ajustados no conjunto de validação.

Para os experimentos, utilizamos detectores de oportunidades com as seguintes resoluções de negociação: 15, 30, 45, 60, 75, 90 e 105 minutos. Não incluímos resoluções menores que 15 minutos, pois a dificuldade de previsão do comportamento do mercado aumenta demasiadamente. Não fizemos experimentos para resoluções maiores que 105 minutos. As informações de mercado

são atualizadas de cinco em cinco minutos. No caso, para um dia de mercado com cinco horas temos 84 atualizações de cinco minutos por resolução menos o número de intervalos que extrapolam o final do dia de mercado. Dadas as resoluções, utilizamos 70 detectores de oportunidades, sendo um por resolução e ativo, e um Trader Atômico por intervalo do dia e ativo.

Com relação aos parâmetros dos detectores de oportunidade, utilizamos uma Janela Deslizante de tamanho 28. Para os atributos de análise técnica utilizamos as janelas de tempo de 15, 25 e 35 minutos passados. Cabe ressaltar que não há na literatura um padrão com relação ao ajuste de parâmetros dos indicadores de análise técnica e Janela Deslizante. Para ajustar descobrir estes parâmetros fizemos alguns testes para diferentes valores no conjunto de validação.

### 7.3

#### Baselines

Primeiramente, ressaltamos que não há um padrão ou uma sugestão na literatura para baselines (Krollner et al., 2010). De acordo com (Krollner et al., 2010), os baselines mais comuns são: estratégia Buy and Hold; estratégia Random Walk; o desempenho de indicadores de mercado, como por exemplo desempenho do BMF&Bovespa; ou outra estratégia com detectores de oportunidades que utilizam Aprendizado de Máquina.

Para enfatizar as vantagens da utilização de estratégias multirresolução, adotamos como primeiro baseline (BLS1) uma estratégia de única resolução e ativo, cujo os resultados estão descritos na subseção 7.4.1. Para os resultados do baseline, utilizamos a metodologia padrão de Aprendizado de Máquina em que são feitos experimentos individuais para cada resolução fixa de operação e ativo. Ao final, seleciona-se a estratégia que produz o melhor resultado. Este baseline é utilizado para comparar com a estratégia CTT.

Adicionalmente, utilizamos um outro baseline (BLS2), que corresponde ao método de seleção de carteiras de Markowitz. Os resultados do baseline BLS2 estão descritos na subseção 7.5.2. Aplicamos o baseline BLS2 para o mesmo conjunto de ativos, como na abordagem tradicional (da Silva e Laber, 2011). Para computar a matriz de correlação necessária para o risco, utilizamos, para cada ativo, a série histórica de retornos. Cada retorno da série histórica corresponde à diferença relativa entre os preços de abertura e de fechamento do dia de mercado. Para estimar os retornos para cada ativo, calculamos a média aritmética da série histórica de retornos para o conjunto de validação.

## 7.4

### Composição de um Time de Traders

Nesta seção, reproduzimos os resultados da estratégia CTT para maximização de retorno. Primeiramente, apresentamos os resultados de todos os traders de única resolução e ativo, e indicamos o trader mais lucrativo como o baseline BLS1, de forma a comparar com a estratégia CTT. Em seguida, apresentamos os resultados da estratégia CTT e comparamos com os resultados do baseline BLS1, sugerido na seção anterior. Por fim, apresentamos um comparativo dos custos operacionais das estratégias CTT e BLS1.

#### 7.4.1

##### Resultados do Baseline BLS1

Na Tabela 7.3, apresentamos os lucros dos traders por resolução e ativo para o conjunto de validação. Como podemos observar, o conjunto de traders mais lucrativo, com lucro de 0,68% ao dia, opera no ativo USIM5 e em uma resolução fixa de 30 minutos. Desta forma, escolhemos este trader como a estratégia baseline BLS1. Para o conjunto de teste, o lucro médio diário para o ativo USIM5 na resolução de 30 minutos corresponde a 0,06%. Cabe ressaltar que a redução de desempenho desses traders se deve às diferenças de comportamento entre o conjunto de validação, que estava em alta, e o conjunto de teste, que estavam em queda. Na prática, um conjunto de dados em alta apresenta um maior número de oportunidades e a consequência é que os traders lucram mais.

Stock	15 min(%)	30 min(%)	45 min(%)	60 min(%)	75 min(%)	90 min(%)	105 min(%)
ITSA4	0,33	0,14	0,22	0,17	0,14	0,1	0,09
PETR4	0,06	0,17	0,14	0,24	0,06	0,15	0,08
USIM5	0,33	<b>0,68</b>	0,59	0,55	0,45	0,48	0,56
VALE5	0,18	0,07	0,02	0,04	0,05	0,03	0,02
BBDC4	0,29	0,29	0,33	0,23	0,14	0,17	0,15
CSNA3	0,41	0,37	0,33	0,2	0,26	0,28	0,12
AMBV4	0,53	0,34	0,21	0,22	0,16	0,08	0,11
BBAS3	0,36	0,4	0,42	0,25	0,17	0,24	0,32
VIVO4	0,53	0,43	0,24	0,32	0,28	0,18	0,17
TNLP4	0,34	0,53	0,37	0,3	0,15	0,03	0,17

Tabela 7.3: BLS1 - Média dos lucros diários.

#### 7.4.2

##### Resultados de Negociação

Na Tabela 7.4, comparamos BLS1 e CTT de acordo com as métricas de lucro anual, médio, máximo e perda máxima diárias. Nossos resultados indicam que a estratégia CTT supera em 0,24% o BLS1 para a métrica de lucro médio diário. Com relação ao desempenho anual, as estratégias CTT e



<b>Indicadores de Lucro</b>	<b>BLS1 (%)</b>	<b>CTT (%)</b>
Lucro Anual	16,07	<b>77,26</b>
Lucro Médio Diário	0,06	<b>0,24</b>
Perda Máxima Diária	3,80	<b>1,30</b>
Lucro Máximo Diário	2,75	<b>3,89</b>

Tabela 7.4: CTT - Resultados de lucro.

BLS1 apresentam lucro de 77,26% e 16,07%, respectivamente. Cabe ressaltar que desempenho anual do Índice Bovespa para o período corresponde a um prejuízo de -9,38%.

<b>Indicadores de Risco</b>	<b>BLS1 (%)</b>	<b>CTT (%)</b>
Draw Down Máximo	<b>0,51</b>	0,62
Índice de Ulceração	<b>0,29</b>	<b>0,29</b>

Tabela 7.5: CTT - Risco.

<b>Lucro/Risco</b>	<b>BLS1 (%)</b>	<b>CTT (%)</b>
Sharpe	0,35	<b>1,33</b>
Martin	0,16	<b>0,75</b>
Calmar	0,09	<b>0,35</b>
Sterling	0,13	<b>0,50</b>
Burke	0,08	<b>0,39</b>

Tabela 7.6: CTT - Lucro por unidade risco.

Na Tabela 7.5, comparamos o risco das estratégias CTT e BLS1. CTT apresenta um risco ligeiramente superior para o indicador de Draw Down Máximo com relação à estratégia BLS1 e um risco similar para o índice de Ulceração. No entanto, uma análise dos resultados de lucro por unidade de risco, apresentados na Tabela 7.6, nos indica que o lucro da estratégia CTT compensa um risco mais elevado.

<b>Operações</b>	<b>BLS1</b>	<b>CTT</b>
Ganhadoras	1,19	2,72
Perdedoras	1,02	2,59
Total	2,21	5,31
Ganhadoras/Perdedoras	1,03	<b>1,20</b>

Tabela 7.7: CTT - Número médio de operações ao dia.

Na Tabela 7.7, apresentamos o número de operações ganhadoras, perdedoras e o total de operações ao dia. A estratégia CTT executa um número maior operações ao dia com relação à estratégia BLS1. No entanto, apesar de um número maior de operações, CTT apresenta um desempenho de lucro sobre

prejuízo 16% maior. Tais resultados indicam que CTT gera um custo de taxa de corretagem maior, tal que para compensar este custo, deve-se investir um montante maior de capital.

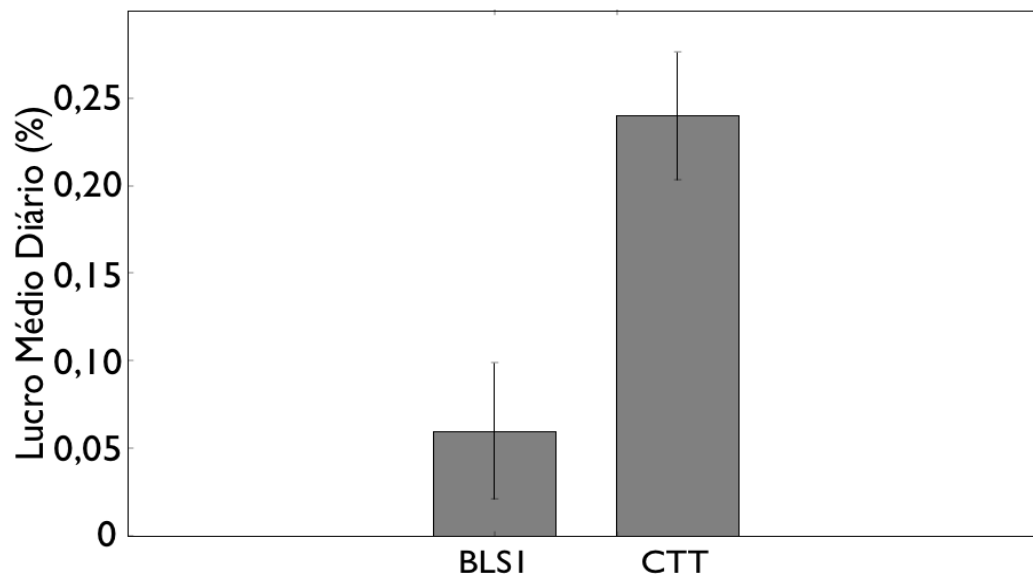


Figura 7.1: CTT - Comparativo.

Na figura 7.1, comparamos os lucros diários médios, representados pelas barras verticais, e os erros padrões, representados pelas linhas verticais, para as duas estratégias CTT e BLS1. Podemos notar que há um ganho significativo de 300% do lucro médio da estratégia CTT com relação à estratégia BLS1. Adicionalmente, podemos observar que não há sobreposição dos erros padrões dos lucros médios das duas estratégias abordadas.

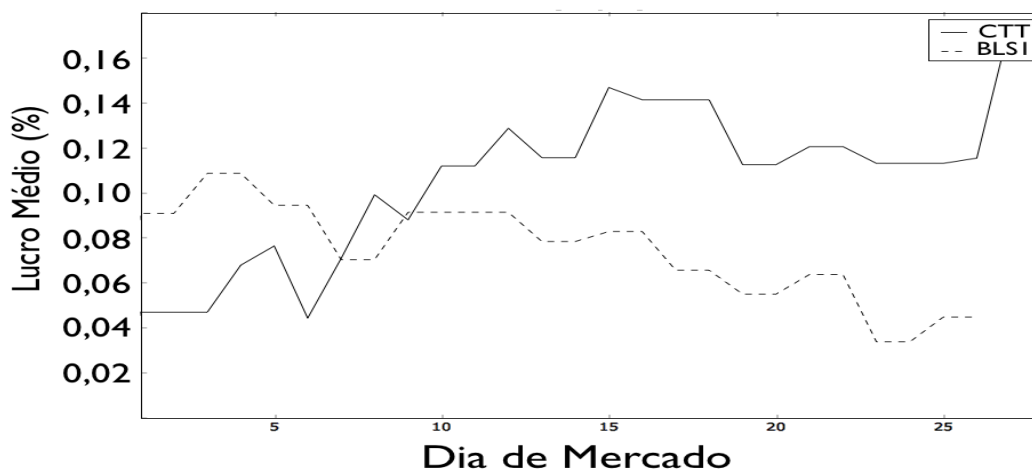


Figura 7.2: CTT - Desempenho das estratégias BLS1 e CTT ao longo do dia.

Na figura 7.1, ressaltamos os lucros médios acumulados, ao longo do dia de mercado, para as estratégias CTT e BLS1. O eixo-y representa o lucro médio e o eixo-x representa o dia de mercado que está dividido em vinte e oito intervalos de quinze minutos cada. Como podemos observar, na figura 7.2, CTT apresenta uma curva de lucro crescente e mais consistente em comparação à curva de lucro da estratégia BLS1.

### 7.4.3

#### Impacto dos Custos Operacionais

Na figura 7.3, nós ilustramos os lucros das estratégias BLS1 (BLS) e CTT considerando a taxa de corretagem e o desconto do imposto de renda. Adicionalmente, ilustramos o índice inflacionário para o período.

Como podemos observar, na figura 7.3, a estratégia CTT é superior à estratégia BLS1 em todos os cenários avaliados e que o lucro da estratégia CTT supera o índice inflacionário.

Na figura 7.4, nós ilustramos os custos operacionais com relação ao capital de investimento inicial. O eixo x representa o capital de investimento inicial e o eixo y representa o custo percentual associado à taxa de corretagem. Podemos observar que os custos com as taxas de corretagem da estratégia CTT são superiores aos custos de taxa de corretagem da estratégia BLS1. Isto se deve ao

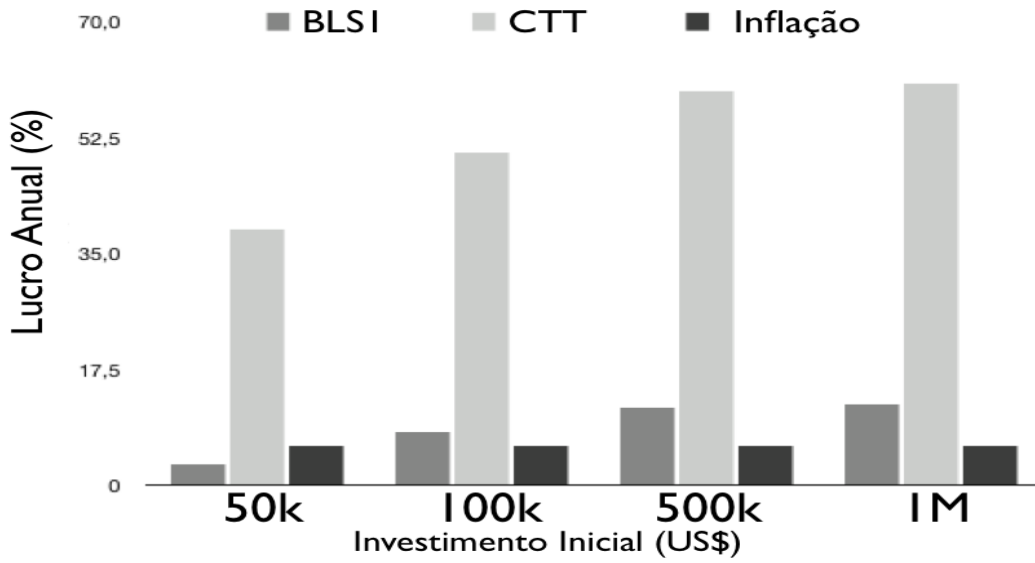


Figura 7.3: Simulação das estratégias CTT e BLS1 com custos operacionais.

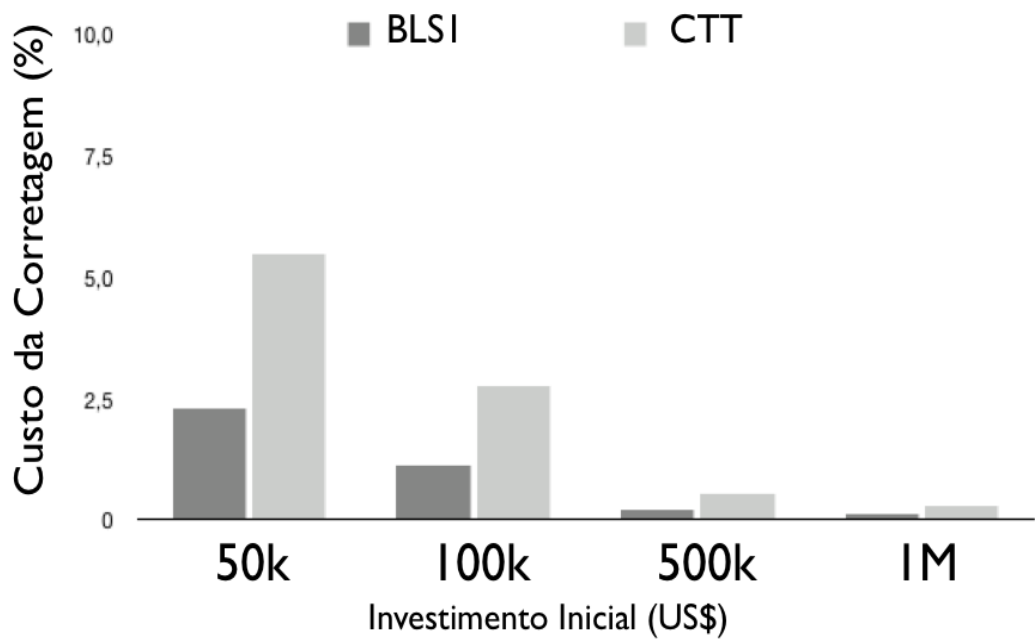


Figura 7.4: CTT - Custos operacionais.

maior número de operações ao dia da estratégia CTT com relação à estratégia BLS1. No entanto, o lucro da estratégia CTT compensa o seu maior custo comparado a estratégia BLS1. Devido ao custo de corretagem ser um valor constante, o custo operacional tende a reduzir quando se aumenta o capital de investimento inicial.

## 7.5

### Composição de Carteiras por Intervalos

Nesta seção apresentamos os resultados da estratégia CCI para maximização do retorno em função do risco. Primeiramente, apresentamos o ajuste de parâmetros das carteiras, de acordo com a formulação de Markowitz descrita no capítulo 3. Em seguida, apresentamos os resultados do baseline BLS2, em função dos parâmetros sugeridos. Posteriormente, descrevemos os resultados da estratégia CCI, também em função dos parâmetros definidos. Por fim, apresentamos o impacto dos custos operacionais para a estratégia CCI.

#### 7.5.1

##### Seleção de Parâmetros

A formulação de cada carteira, como descrito no capítulo 3, possui as restrições  $\theta$ ,  $\alpha$  e  $\beta$ . O valor  $\beta$  corresponde ao limite máximo que pode ser investido em um único trader. Para  $\beta$  escolhemos os valores: 1, 0,75 e 0,5. Quanto menor o valor, maior a diversificação do investimento por carteira. Por exemplo, com valor 1, podemos ter carteiras com 100% de investimento em um único trader. Para um valor 0,5, o investimento máximo por trader é de 50%, o que força uma diversificação e consequente possibilidade de redução do risco. O valor  $\alpha$  corresponde ao limite mínimo que pode ser investido em um único trader. Para  $\alpha$ , escolhemos o valor zero. Um valor maior que zero significa forçar a utilização de um trader mesmo que este venha a dar prejuízo, o que não é bom para o desempenho de qualquer estratégia.

Com relação ao parâmetro de risco, devemos ter cautela na sua escolha. Um valor de risco muito baixo pode não ser atendido no processo de otimização de carteira. Isso é ruim para se ter o controle do risco, pois não queremos uma estratégia com risco variado. Desta forma, para encontrar os valores adequados do  $\theta$ , testamos alguns valores e analisamos, no conjunto de validação, o número de soluções viáveis encontradas no processo de otimização de cada carteira.

Na Tabela 7.8, indicamos a variação do parâmetro de risco e o número de soluções viáveis encontradas. Quando uma solução é inviável, não existe carteira para um determinado intervalo e esta não é considerada para o Agendamento de Intervalos Ponderados. Note que há uma relação linear entre risco e o número de soluções viáveis. Quanto menor o risco, maior o número de soluções inviáveis e menor o número de carteiras a serem utilizadas pela estratégia CCI. Desta forma, se desejamos considerar todas as carteiras, basta escolher o valor 1 que está associado a uma estratégia de risco elevado. Para valores abaixo de 0,25, o número de carteiras não é significativo. Neste ponto,

escolhemos os valores 1, 0,75 e 0,5. Com os parâmetros escolhidos, efetuamos experimentos relacionados ao risco, lucro e número de operações.

$\theta$	Carteiras (%)
1	100
0,75	63
0,5	30
0,25	0,5

Tabela 7.8: CCI - Soluções viáveis com relação ao total de carteiras.

### 7.5.2

#### Resultados do Baseline BLS2

Na Tabela 7.9, apresentamos os resultados para o baseline BLS2, abordagem padrão da literatura. Como podemos notar, o baseline BLS2 gera prejuízo. O motivo é que o baseline BLS2 seleciona os pesos da carteira de acordo com o lucro médio estimado de cada ativo e não em função do desempenho de Traders Atômicos. Como todos os ativos descritos na Tabela 7.1 apresentam retorno médio positivo para o período de validação e retorno médio negativo para o período de teste, então BLS2 apresenta prejuízo independente da variação dos parâmetros.

$\beta$	$\theta (\leq)$	Lucro Anual (%)	Lucro Diário (%)
1,00	1,00	-34,87	-0,1412
1,00	0,75	-39,81	-0,1612
1,00	0,50	-40,08	-0,1623
0,75	1,00	-34,00	-0,1377
0,75	0,75	-34,66	-0,1403
0,75	0,50	-34,77	-0,1408
0,50	1,00	-30,29	-0,1226
0,50	0,75	-30,29	-0,1226
0,50	0,50	-30,29	-0,1226

Tabela 7.9: BLS2 - Desempenho.

### 7.5.3

#### Resultados de Negociação

Na Tabela 7.10, indicamos nossos resultados associados ao risco. Nossos resultados indicam que ao reduzir o parâmetro de risco de cada carteira, o risco global da estratégia CCI é reduzido para as métricas CVar e Var. Adicionalmente, ao forçar a diversificação a partir do parâmetro  $\beta$ , para um mesmo valor de risco, o risco global da estratégia CCI também é reduzido para

as métricas CVar e Var. Desta forma, podemos concluir que basta controlar o risco por intervalo para que se tenha o controle do risco geral da estratégia.

$\beta$	$\theta (\leq)$	CVar (5%)	Var (95%)
1,00	1,00	0,0051	0,0029
0,75	1,00	0,0047	0,0030
0,50	1,00	0,0041	0,0027
1,00	0,75	0,0047	0,0029
0,75	0,75	0,0045	0,0030
0,50	0,75	0,0041	0,0027
1,00	0,50	0,0043	0,0027
0,75	0,50	0,0042	0,0026
0,50	0,50	0,0039	0,0026

Tabela 7.10: CCI - Risco

$\beta$	$\theta (\leq)$	Lucro Anual (%)	Lucro Diário (%)
1,00	1,00	62,94	0,1978
0,75	1,00	63,38	0,1989
0,50	1,00	57,37	0,1837
1,00	0,75	50,7	0,1661
0,75	0,75	70,83	0,2170
0,50	0,75	36,62	0,1264
1,00	0,50	60,63	0,1920
0,75	0,50	48,16	0,1592
0,50	0,50	52,06	0,1698

Tabela 7.11: CCI - Lucro.

Os resultados da Tabela 7.11 indicam lucro para todos os parâmetros de risco. Estes resultados de lucro superam o desempenho do baseline BLS2, que apresentou prejuízo. É interessante notar que o melhor resultado da estratégia CCI, que corresponde a risco alto, não ultrapassa o melhor resultado de da estratégia CTT, que ignora o conceito de risco. Adicionalmente, o pior resultado da estratégia CCI ainda é superior ao resultado da estratégia padrão BLS1 que apresenta lucro anual de 16,07%.

De acordo com os resultados da Tabela 7.11 podemos observar que ao reduzir o risco, a partir do parâmetro  $\theta$  ou do parâmetro de limite superior  $\beta$ , o lucro tende a ser menor. No entanto, também ocorre, para um risco menor, um lucro maior. Isto porque o risco está associado, indiretamente, à possibilidade de perdas.

Na Tabela 7.12, apresentamos resultados do número médio de operações ganhadoras, perdedoras e neutras para o dia de mercado. Primeiramente, podemos observar que a estratégia CCI apresenta um número maior de operações ao dia com relação à estratégia CTT. Isto se deve à diversificação

$\beta$	$\theta (\leq)$	Ordens Ganhadoras	Ordens Perdedoras	Ordens Neutras
1,00	1,00	87,27	85,07	15,66
0,75	1,00	104,93	100,54	20,63
0,50	1,00	107,85	105,32	20,1
1,00	0,75	97,22	93,02	17,56
0,75	0,75	103,02	102,39	19,76
0,50	0,75	93,66	88,00	17,17
1,00	0,50	121,80	119,51	23,56
0,75	0,50	111,51	106,34	21,61
0,50	0,50	124,05	120,63	23,95

Tabela 7.12: CCI - Operações.

do capital investido em vários traders que operam simultaneamente. Podemos observar também que há alteração significativa no número de operações quando se reduz o risco. Quanto menor o risco, maior a diversificação do capital investido.

#### 7.5.4 Impacto dos Custos Operacionais

Nesta subseção, aplicamos a estratégia CCI para diferentes montantes de investimento de forma a analisar o impacto dos custos operacionais. Na Tabela 7.13, indicamos os lucros ao ano da estratégia CCI com dedução de imposto de renda e taxa de corretagem para diferentes montantes de investimento. Como podemos observar a estratégia CCI necessita de um capital de investimento maior que a estratégia CTT. Os valores de capital apresentados são elevados para um investidor individual. No entanto, para fundos de investimento, tais valores são comumente praticados.

$\beta$	$\theta (\leq)$	US\$ 2M	US\$ 5M	US\$ 10M	US\$ 20M
1,00	1,00	24,81	40,13	45,24	47,79
0,75	1,00	19,98	38,41	44,56	47,63
0,50	1,00	14,20	33,22	39,55	42,72
1,00	0,75	12,33	29,26	34,91	37,73
0,75	0,75	26,07	44,42	50,54	53,60
0,50	0,75	2,28	18,49	23,89	26,59
1,00	0,50	12,52	34,11	41,30	44,90
0,75	0,50	5,99	25,51	32,02	35,27
0,50	0,50	5,15	27,05	34,34	37,99

Tabela 7.13: CCI - Lucro.

Na Tabela 7.14, indicamos o custo percentual total da taxa de corretagem com relação à quantia de investimento inicial. Podemos observar que o custo percentual passa a ser baixo para valores na casa das dezenas de milhões.



Como estamos utilizando um custo operacional constante, o custo tende a reduzir para quantias de valores maiores.

Aqui, cabe ressaltar que, para reduzir estes custos, o investidor pode buscar taxas de corretagem menores com a corretora. O que é possível para fundos em que se aplicam grandes quantias. Além da negociação dos custos de corretagem com a corretora, pode-se adotar duas medidas para a estratégia CCI de forma a limitar o número de operações ao dia. A primeira medida seria limitar o número de operações por intervalo. Isto poderia ser feito a partir de uma restrição da quantidade de pesos de cada carteira. Outra forma seria, ao invés de utilizar resoluções pequenas como de 15 e 30 minutos, utilizar apenas resoluções mais altas como de 60 minutos para cima.

$\beta$	$\theta (\leq)$	US\$ 2M	US\$ 5M	US\$ 10M	US\$ 20M
1,00	1,00	25,53	10,21	5,10	2,55
0,75	1,00	30,71	12,28	6,14	3,07
0,50	1,00	31,68	12,67	6,33	3,16
1,00	0,75	28,22	11,29	5,64	2,82
0,75	0,75	30,58	12,23	6,11	3,05
0,50	0,75	27,01	10,80	5,40	2,70
1,00	0,50	35,98	14,39	7,19	3,59
0,75	0,50	32,53	13,01	6,50	3,25
0,50	0,50	36,49	14,59	7,29	3,64

Tabela 7.14: CCI - Corretagem.

## 7.6

### Consolidação dos Resultados

Aqui, apresentamos um resumo dos nossos resultados para os dois problemas abordados: maximização de retorno acumulado e maximização de retorno acumulado com risco controlado.

Para o problema de maximização de retorno acumulado, nossa estratégia de Composição de um Time de Traders apresenta 0,24% de lucro médio diário e 77,24% de lucro anual, superando em 300% e 380%, respectivamente, a estratégia de Aprendizado de Máquina padrão (BLS1). Para os custos adotados, a estratégia CTT é viável a partir de US\$ 50.000,00.

Para o problema de maximização do retorno acumulado com risco controlado, a estratégia de Composição de Carteiras por Intervalos apresenta em média 0,179% de lucro diário e 55,85% de lucro anual superando o baseline BLS2, que corresponde ao método de Média-Variância de Markowitz. Para os custos adotados, a estratégia CCI é viável a partir de US\$ 2.000.000,00. Isto porque, para reduzir o risco, a estratégia CCI diversifica seu capital ao longo do dia.

## 8

### Conclusões

Neste capítulo apresentamos nossas considerações finais. Na primeira seção, descrevemos os dois problemas abordados nesta tese, maximização de retorno acumulado e maximização com controle do risco, e estratégias multirresolução para estes problemas. Na seção seguinte, destacamos nossas principais contribuições e o principal artigo derivado desta tese. Por fim, na seção Trabalhos Futuros, sugerimos quatro melhorias para as estratégias propostas nesta tese.

#### 8.1

##### Visão Geral

Nesta tese abordamos dois problemas do dia de mercado: maximização de retorno acumulado e maximização de retorno acumulado com controle do risco. Para estes dois problemas, propomos a construção eficiente de estratégias multirresolução e multiativo, a partir da resolução do problema de Agendamento de Intervalos Ponderados. Para o problema de maximização de retorno acumulado, propomos a estratégia Composição de um Time de Traders. Estratégia adequada a investidores que não são avessos ao risco. Para o problema de maximização de retorno acumulado com controle do risco, propomos a estratégia Composição de Carteiras por Intervalos. Ao contrário da estratégia anterior, esta é adequada ao investidor avesso ao risco. Nesta estratégia, o investidor pode ter um controle do risco global, a partir da parametrização do risco local por intervalos.

Nossa metodologia consiste em dividir o dia de mercado em intervalos, especializar traders por intervalo e associar um prêmio a cada trader. Para o problema de maximização de retorno acumulado, o prêmio de cada trader corresponde ao retorno acumulado entre dias para o intervalo de operação associado. Para o problema de maximização de retorno acumulado com controle do risco, o prêmio de cada trader corresponde ao retorno acumulado dividido pelo risco para o intervalo de operação associado. Diferentemente do problema anterior, empregamos conjunto de traders por intervalo e utilizamos o método de Média-Variância, de Markowitz, para encontrar pesos ótimos para conjunto

de traders de forma a controlar o risco.

Para a sinalização das ordens de compra e venda, nossos traders utilizam detectores de oportunidades. Estes detectores correspondem a algoritmos de Aprendizado de Máquina que processam informações de indicadores de análise técnica e dados de preço e volume.

## 8.2

### Contribuições

Nesta seção apresentamos nossas quatro principais contribuições. A primeira contribuição desta tese está na construção eficiente de estratégias multiativo e multirresolução para os problemas de maximização de retorno acumulado e maximização do retorno acumulado com risco controlado. Isto é feito a partir da resolução do problema de Agendamento de Intervalos Ponderados. Ressaltamos que a abordagem por Agendamento de Intervalos Ponderados para o problema de maximização do retorno acumulado está publicada (Alvim e Milidiú, 2013) na prestigiosa *Decision Support Systems*<sup>1</sup>.

A nossa segunda contribuição corresponde ao emprego de métodos de seleção de carteira utilizando-se das informações de desempenho das decisões de compra e venda de Traders Atômicos. Enquanto a abordagem padrão, para métodos de seleção de carteira, é estimar o retorno e risco de uma carteira a partir dos desempenhos de ativos, nesta tese estimamos o retorno e risco em função do ganho médio e risco de nossos Traders Atômicos. Desta forma, a estratégia fica menos dependente do desempenho de um ativo e mais dependente dos desempenhos dos traders. Como é mais fácil controlar o desempenho de um trader do que de um ativo, podemos ajustar e selecionar melhor nossos traders.

A terceira e última contribuição corresponde à proposta de uma arquitetura do sistema de trading que é dividida em seis etapas: treinamento e validação dos detectores de oportunidades, agrupamento das decisões por intervalos do dia, aplicação de métodos de seleção de carteira, se necessário o risco, avaliação dos desempenhos por intervalos do dia, resolução do problema de Agendamento de Intervalos Ponderados, por programação dinâmica, em um conjunto de validação, aplicação da solução em um conjunto de teste, reinício das etapas anteriores. Nosso sistema combina métodos de otimização e de Aprendizado de Máquina.

<sup>1</sup>periódico internacional Qualis A1

### 8.3

#### Trabalhos Futuros

Como trabalhos futuros, pode-se incorporar um número maior de algoritmos de Aprendizado de Máquina com o intuito de melhorar a qualidade das detecções de oportunidades. Podem-se incorporar estes algoritmos de duas formas. A primeira é a criação de um comitê de algoritmos, de preferência heterogêneos, o que geralmente melhora o desempenho geral das detecções de oportunidades e por consequência reduz o risco de operação. A segunda é adicionar um número maior de algoritmos por intervalo, de forma que concorram entre si.

Adicionalmente, pode-se introduzir um maior número de resoluções de operação dentro do dia de mercado. Com isto, pode-se ampliar as chances de lucro das estratégias.

Um outra direção interessante estaria na redução dos custos operacionais da estratégia CCI. Pode-se adotar duas medidas na estratégia CCI de forma a limitar o número de operações ao dia. A primeira medida seria limitar o número de operações por intervalo. Isto poderia ser feito a partir de uma restrição da quantidade de pesos de cada carteira. Outra medida seria, ao invés de utilizar resoluções pequenas como de 15,30 minutos, utilizar apenas resoluções mais altas como de 60 minutos para cima.

Com relação ao risco processo de otimização de carteiras de Markowitz, experimentos poderiam ser feitos com a restrição Valor em Risco Condicional ao invés da correlação do modelo de média variância de Markowitz. Valor em Risco Condicional está associado diretamente a perdas médias das estratégias e por isso é mais intuitivo que a métrica de correlação.

Por fim, a inclusão de informações textuais como notícias de portais financeiros, Twitter e opiniões de investidores poderiam ser utilizadas nos detectores de oportunidades. Informações desta natureza são de fundamental importância, uma vez que parte dos investidores utilizam-se destas fontes para tomar suas decisões. Ressaltamos aqui que não há abordagem multirresolução que utilize informações desta natureza.

## Referências Bibliográficas

- [Abraham et al., 2003] Abraham, A., Philip, N. S., e Saratchandran, P. (2003). Modeling chaotic behavior of stock indices using intelligent paradigms. In *International Journal of Neural, Parallel & Scientific Computations, USA, Volume 11, Issue*, pages 143–160. 2.1, 3.4
- [Allen e Powell, 2007] Allen, D. E. e Powell, R. J. (2007). Thoughts on var and cvar. In *Proceedings of International Congress on Modelling and Simulation.*, pages 1843–1850. 3.2.1
- [Alvim et al., 2010] Alvim, L. G. M., dos Santos, C. N., e Milidiú, R. L. (2010). Daily volume forecasting using high frequency predictors. In *IASTED International Conference on Artificial Intelligence and Applications. AIA2010*. ACTA-PRESS. track = 674-047. 3.5
- [Alvim e Milidiú, 2013] Alvim, L. G. M. e Milidiú, R. L. (2013). Trading team composition for the intraday multistock market. *Decision Support Systems*, 54:838–845. doi: 10.1016/j.dss.2012.09.009. 1.4, 8.2
- [Atsalakis e Valavanis, 2009a] Atsalakis, G. S. e Valavanis, K. P. (2009a). Surveying stock market forecasting techniques - part i: Conventional methods. *Journal of Computational Optimization in Economics and Finance*, 2(1, Part 1):45–92. 2.1
- [Atsalakis e Valavanis, 2009b] Atsalakis, G. S. e Valavanis, K. P. (2009b). Surveying stock market forecasting techniques - part ii: Soft computing methods. *Expert Systems with Applications*, 36(3, Part 2):5932 – 5941. 2.1, 4.3
- [Boser et al., 1992] Boser, B. E., Guyon, I. M., e Vapnik, V. N. (1992). A training algorithm for optimal margin classifiers. In *Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory*, pages 144–152. ACM Press. 3.4, 4.1, 7.2
- [Boyarshinov, 2005] Boyarshinov, V. (2005). *Machine learning in computational finance*. PhD thesis, Troy, NY, USA. AAI3173253. 1, 2.2, 6.1

- [Brock et al., 1992] Brock, W., Lakonishok, J., e LeBaron, B. (1992). Simple Technical Trading Rules and the Stochastic Properties of Stock Returns. *The Journal of Finance*, 47(5):1731–1764. 4.2.1
- [Chawla, 2009] Chawla, S. (2009). Advanced algorithms. 3.3, 3.3, 3.3
- [Cortes e Vladimir, 1995] Cortes, C. e Vladimir, V. (1995). Support vector networks. In *Machine Learning*, pages 273–297. 3.4
- [Cowles 3rd, 1933] Cowles 3rd, A. (1933). Can stock market forecasters forecast? *Econometrica*, 1:309–324. 4.2.1
- [Cuoco et al., 2008] Cuoco, D., He, H., e Isaenko, S. (2008). Optimal dynamic trading strategies with risk limits. *Operations Research*, 56(2):358–368. 2.3
- [da Silva e Laber, 2011] da Silva, T. A. e Laber, E. S. (2011). Estudo experimental para técnicas de otimização de carteiras. Master's thesis, Pontifícia Universidade Católica do Rio de Janeiro - PUC-Rio, Brasil. 1, 2.3, 3.2, 7.3
- [Dietterich, 2002] Dietterich, T. G. (2002). Machine learning for sequential data: A review. In *Structural, Syntactic, and Statistical Pattern Recognition*, pages 15–30. Springer-Verlag. 4.2.2
- [Dostl, 2009] Dostl, Z. (2009). *Optimal Quadratic Programming Algorithms: With Applications to Variational Inequalities*. Springer Publishing Company, Incorporated, 1st edition. 3.4
- [Drucker et al., 1999] Drucker, H., Wu, D., e Vapnik, V. (1999). Support vector machines for spam categorization. *Neural Networks, IEEE Transactions on*, 10(5):1048–1054. 3.4
- [Fama, 1970] Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, 25(2):383–417. 4.2.1, 4.4
- [Fama e Blume, 1966] Fama, E. F. e Blume, M. E. (1966). Filter Rules and Stock-Market Trading. *The Journal of Business*, 39(1):226–241. 4.2.1
- [Fan et al., 2008] Fan, R.-E., Chang, K.-W., Hsieh, C.-J., Wang, X.-R., e Lin, C.-J. (2008). LIBLINEAR: A library for large linear classification. *Journal of Machine Learning Research*, 9:1871–1874. 7.2
- [Golub e Van Loan, 1996] Golub, G. H. e Van Loan, C. F. (1996). *Matrix computations (3rd ed.)*. Johns Hopkins University Press, Baltimore, MD, USA. 3.5.1

- [Gomide e Milidiú, 2010] Gomide, P. e Milidiú, R. L. (2010). Assessing stock market time series predictors quality through a pairs trading system. In *Proceedings of the 2010 Eleventh Brazilian Symposium on Neural Networks, SBRN '10*, pages 133–139, Washington, DC, USA. IEEE Computer Society. 1, 6.2
- [Guyon et al., 2002] Guyon, I., Weston, J., Barnhill, S., e Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. *Mach. Learn.*, 46(1-3):389–422. 3.4
- [Haaland e Thomas, 1988a] Haaland, D. e Thomas, E. (1988a). Partial least-squares methods for spectral analyses. 1. Relation to other quantitative calibration methods and the extraction of qualitative information. *Analytical Chemistry*, 60(11):1193–1202. 3.5
- [Haaland e Thomas, 1988b] Haaland, D. e Thomas, E. (1988b). Partial least-squares methods for spectral analyses. 2. application to simulated and glass spectral data. *Analytical Chemistry*, 60:1202–1208. 3.5
- [Hellström, 2001] Hellström, T. (2001). Optimizing the sharpe ratio for a rank based trading system. In *Proceedings of the 10th Portuguese Conference on Artificial Intelligence on Progress in Artificial Intelligence, Knowledge Extraction, Multi-agent Systems, Logic Programming and Constraint Solving, EPIA '01*, pages 130–141, London, UK, UK. Springer-Verlag. 2.3
- [Hsu et al., 2000] Hsu, C.-W., Chang, C.-C., e Lin, C.-J. (2000). A Practical Guide to Support Vector Classification. 4.3
- [Huang et al., 2005] Huang, W., Nakamori, Y., e Wang, S.-Y. (2005). Forecasting stock market movement direction with support vector machine. *Computers & Operations Research*, 32(10):2513 – 2522. 2.1, 3.4
- [Jensen e Benington, 1970] Jensen, M. e Benington, G. (1970). Random Walks and Technical Theories: Some Additional Evidence. *The Journal of Finance*, 25(2):469–482. 4.2.1
- [Kleinberg e Tardos, 2005] Kleinberg, J. e Tardos, E. (2005). *Algorithm Design*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA. 3.3
- [Krollner et al., 2010] Krollner, B., Vanstone, B., e Finnie, G. (2010). Financial time series forecasting with machine learning techniques: A survey. In *Paper presented at the European symposium on artificial neural networks: Computational and machine learning*. 7.3

- [Lo et al., 2000] Lo, A. W., Mamaysky, H., e Wang, J. (2000). Foundations of technical analysis: Computational algorithms, statistical inference, and empirical implementation. *Journal of Finance*, 40:1705–1765. 4.2.1
- [Markowitz, 1952] Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, 7(1):77–91. 2.3, 3.1, 3.2
- [Nison, 2010] Nison, S. (2010). *Japanese candlestick charting techniques: a contemporary guide to the ancient investment techniques of the Far East*. New York Institute of Finance, New York, 2nd edition edition. 4.2.1
- [Platt, 1999] Platt, J. C. (1999). Advances in kernel methods. chapter Fast training of support vector machines using sequential minimal optimization, pages 185–208. MIT Press, Cambridge, MA, USA. 3.4
- [Renteria e Milidiú, 2003] Renteria, R. e Milidiú, R. L. (2003). *Algoritmos para Regressão por Mínimos Quadrados Parciais*. PhD thesis, Pontifícia Universidade Católica do Rio de Janeiro. 3.5.1
- [Rockafellar e Uryasev, 2000] Rockafellar, R. T. e Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of Risk*, 2:21–41. 3.2.1
- [Tenenhaus, 1998] Tenenhaus, M. (1998). *La régression PLS : théorie et pratique*. Éd. Technip, 1998 (05-Gap, Paris. 3.5
- [Wold et al., 1983] Wold, S., Martens, H., e Wold, H. (1983). The multivariate calibration problem in chemistry solved by the pls method. In *Proceedings Conf. Matrix Pencils*, pages 286–293. Springer Verlag. 3.5, 4.1, 7.2
- [Zhang et al., 2001] Zhang, B., Coggins, R., Jabri, M. A., Dersch, D. R., e Flower, B. (2001). Multiresolution forecasting for futures trading using wavelet decompositions. *IEEE Transactions on Neural Networks*, 12(4):765–775. 1, 2.2