



Universidade de Brasília

Faculdade de Tecnologia
Departamento de Engenharia Elétrica

Arbitragem Estatística: Design de Protótipos baseados em Cointegração

Caio Campos Mameri

Monografia apresentada como requisito parcial
para conclusão do Bacharelado em Engenharia Elétrica

Orientador
Prof. Dr. Geovany A. Borges

Brasília
2020



Universidade de Brasília

Faculdade de Tecnologia
Departamento de Engenharia Elétrica

Arbitragem Estatística: Design de Protótipos baseados em Cointegração

Caio Campos Mameri

Monografia apresentada como requisito parcial
para conclusão do Bacharelado em Engenharia Elétrica

Prof. Dr. Geovany A. Borges (Orientador)
Universidade de Brasília

Prof. Dr. Daniel Oliveira Cajueiro Prof. Dr. Marcelino Monteiro de Andrade
Universidade de Brasília Universidade de Brasília

do Bacharelado em Engenharia Elétrica

Brasília, 23 de agosto de 2020

Dedicatória

Ao contribuinte brasileiro, que bancou grande parte dos meus estudos.

Agradecimentos

Primeiramente, agradeço à minha família, que é a base para tudo o que eu fiz e faço. Fica o agradecimento destacado aos meus amigos, sem eles, a vida não teria graça. Em especial, agradeço Bruna e Antonia Cristina pelo carinho e pela ajuda nos momentos mais difíceis. Finalmente, agradeço ao professor Geovany pela orientação na construção desse trabalho.

Resumo

Este trabalho tem como objetivo analisar a performance de estratégias de arbitragem estatística realizadas em cinco protótipos. Há três pilares para o desenvolvimento de cada protótipo: estimação de parâmetros por filtro estocástico, cointegração e sinalização de oportunidade de arbitragem. Os protótipos são realizados e experimentados em um conjunto de 140 ações da Bolsa de Valores de São Paulo, com precificação minuto a minuto e a simulação dos resultados dos protótipos é realizada em um período pós-otimização seguindo uma análise de *walk-forward*. Foi possível perceber que 80% dos protótipos criados ou apresentaram uma performance robusta em termos de rentabilidade, volatilidade e a relação risco-retorno, ou apresentaram o potencial para atingí-la dado que se realizem algumas mudanças apontadas pela análise realizada para cada um.

Palavras-chave: Arbitragem estatística, cointegração, filtragem estocástica

Abstract

This work presents an analysis on the performance of five prototypes that incorporate statistical arbitrage strategies. There are three main building blocks in the development of each prototype: parameter estimation by stochastic filtering, cointegration, and arbitrage opportunities signaling. The prototypes work on a set of 140 stock papers of the São Paulo Stock Exchange priced in a minute based period and its simulation is made according to the walk-forward analysis, which indicates that the results need to be generated in a pos-optimization window. It was possible to perceive that 80% of the prototypes either delivered a robust performance regarding return and volatility or showed the potential to deliver a robust performance once a few changes indicated by the analysis of each prototype were to be made.

Keywords: Statistical arbitrage, cointegration, stochastic filtering

Sumário

1	Introdução	1
1.1	Organização do Texto	1
1.2	Motivação	1
1.3	Contexto	2
1.4	Objetivos	5
1.5	Resultados Obtidos	5
2	Fundamentação	8
2.1	Estratégias Quantitativas	8
2.1.1	Estratégias Direcionais	9
2.1.2	Estratégias Não Direcionais	10
2.2	Econometria de Séries Temporais	11
2.2.1	Fundamentos Estatísticos	12
2.2.2	Testes de Raiz Unitária	14
2.2.3	Cointegração	16
2.3	Filtragem Estocástica	17
2.3.1	Processo Estimado	18
2.3.2	Computação dos Processos	18
2.3.3	Aspectos Probabilísticos	20
2.3.4	Algoritmo do Filtro	20
2.3.5	Parâmetros do Filtro	21
2.4	Tempo de Meia Vida	21
2.5	Cópulas	23
2.5.1	Cópula de Clayton	24
2.5.2	Cópula de Gumbel	25
2.5.3	Cópula de Frank	26
2.6	Análise Walk-Foward	27
2.6.1	Otimização	27
2.6.2	Backtesting	29

2.6.3	Walk-Foward	29
3	Sistema de Trading	31
3.1	Construção do Universo de Operação	32
3.1.1	Construção do Modelo e Estimação de Parâmetros	33
3.1.2	Teste de Cointegração	34
3.2	Entrada e Saída de Posições	35
3.2.1	Z-Score	37
3.2.2	Cópulas	38
3.3	Gestão de Risco e Dimensionamento de Posições	41
3.4	Construção dos Protótipos	44
3.4.1	Protótipo 1	45
3.4.2	Protótipo 2	46
3.4.3	Protótipo 3	47
3.4.4	Protótipo 4	47
3.4.5	Protótipo 5	48
4	Resultados e Análises	50
4.1	Protótipo 1: Pairs Trading - Z-Score	52
4.2	Protótipo 2: Pairs Trading (ON/PN) - Z-Score	55
4.3	Protótipo 3: Pairs Trading - Cópulas	57
4.4	Protótipo 4: Pairs Trading (ON/PN) - Cópulas	59
4.5	Protótipo 5: Trading de Tríades - Z-Score	61
4.6	Benchmarking e Análises Finais	64
5	Conclusão	69
	Referências	73

Lista de Figuras

1.1 Comparativo de rentabilidade.	6
1.2 Comparativo de volatilidade.	6
2.1 Ilustração: série estacionária e série com tendência estacionária.	13
3.1 Parâmetros variantes no tempo.	34
3.2 Sinalização de oportunidades de arbitragem: z-score.	37
3.3 Exemplo de dependência com distribuição não gaussiana.	38
3.4 Ilustração de bandas de confiança: long (esquerda) e short (direita).	40
3.5 Ilustração da comparação entre $CVaR$ e VaR	43
4.1 Rentabilidade: Protótipo 1.	52
4.2 Rentabilidade: Protótipo 2.	55
4.3 Rentabilidade: Protótipo 3.	57
4.4 Rentabilidade: Protótipo 4.	60
4.5 Rentabilidade: Protótipo 5.	62
4.6 Rentabilidade: Protótipo 5 com sistema stop-loss.	64
4.7 Desempenho: BOVA11.SA.	65
4.8 Ilustração da comparação entre desempenhos	66
4.9 Comparativo: retornos diários	67

Lista de Tabelas

4.1	Resultados: Protótipo 1	52
4.2	Resultados complementares: Protótipo 1	53
4.3	Resultados considerando liquidez: Protótipo 1	54
4.4	Correlação entre tempo de meia-vida e duração do trading: Protótipo 1 . .	54
4.5	Resultados: Protótipo 2	56
4.6	Resultados complementares: Protótipo 2	56
4.7	Correlação entre tempo de meia-vida e duração do trading: Protótipo 2 . .	57
4.8	Resultados: Protótipo 3	58
4.9	Resultados complementares: Protótipo 3	58
4.10	Resultados considerando liquidez: Protótipo 3	59
4.11	Correlação entre tempo de meia-vida e duração do trading: Protótipo 3 . .	59
4.12	Resultados: Protótipo 4	60
4.13	Resultados complementares: Protótipo 4	61
4.14	Correlação entre tempo de meia-vida e duração do trading: Protótipo 4 . .	61
4.15	Resultados: Protótipo 5	62
4.16	Resultados complementares: Protótipo 5	63
4.17	Correlação entre tempo de meia-vida e duração do trading: Protótipo 5 . .	63
4.18	Comparação de resultados	65
4.19	Testes de correlação de resultados diários	67
4.20	Testes de correlação de resultados diários em <i>bear market</i>	68

Lista de Abreviaturas e Siglas

CVaR *conditional value at risk.*

EnE Departamento de Engenharia Elétrica.

ETF *Exchange Traded Funds.*

LTCM Long Term Capital Management.

ON ação ordinária.

PN ação preferencial.

VaR *value at risk.*

1 Introdução

1.1 Organização do Texto

O trabalho apresentado é composto por cinco capítulos: Introdução (Capítulo 1), Fundamentação (Capítulo 2), Sistema de Trading (Capítulo 3), Resultados e Análises (Capítulo 4) e a Conclusão (Capítulo 5).

Inicialmente, introduz-se o objetivo e a motivação para a realização do projeto, o leitor pode se contextualizar acerca de estratégias quantitativas no geral e os resultados obtidos são mostrados de forma superficial.

Em um segundo momento, a teoria que embasa o sistema de trading é apresentada. Assim, no segundo capítulo, são abordados os temas como: estratégias quantitativas com ênfase na arbitragem estatística, estacionaridade e cointegração de séries temporais, filtragem estocástica para estimação de parâmetros, cópulas de arquimedes como estruturas de dependência e, finalmente, análise *walk-forward* como simulação utilizada para geração dos resultados.

No terceiro capítulo, o sistema de trading construído é descrito. Assim, cada conceito explicado no Capítulo 2 é relacionado a um aspecto da operação montada. Logo, são percorridos os temas acerca da construção do universo de operação, da entrada e saída de posições, da gestão de risco e dimensionamento de posições e os protótipos construídos para a realização das estratégias são apresentados.

No Capítulo 4, os resultados gerados pelo sistema de trading construído são exibidos e analisados. Dessa forma, nesse capítulo se concentram as análises do desempenho de cada protótipo montado e seus potenciais.

Finalmente, no Capítulo 5, são apresentadas a conclusão do trabalho e as sugestões para trabalhos futuros a serem realizados.

1.2 Motivação

As motivações que levaram à confecção desse trabalho são duas, uma de cunho pessoal e outra de cunho institucional. Pessoalmente, esse trabalho visa ser uma pequena analogia

ao que Billy Beane e Paul DePodesta conseguiram realizar no time de beisebol Oakland Athletics. Institucionalmente, o desejo é que se fomente tanto o estudo da engenharia financeira quanto as realizações interdisciplinares na Universidade de Brasília.

O desejo individual de replicar *Moneyball* em um estudo financeiro não é algo raro. Assim como Billy Beane e Paul DePodesta mudaram o beisebol ao utilizar análises estatísticas para reduzir o *gap* de performance entre seu time e os times de orçamento maior [1], vários investidores procuram e desenvolvem métodos analíticos como forma de obter resultados próximos ou melhores do que os obtidos por "investidores profissionais". Dessa forma, um dos motivos por trás desse trabalho é investigar um modo de investimento baseado em ferramentas da engenharia e da estatística e mesmo assim obter resultados expressivos, dando mais uma ferramenta para o *little guy*.

A motivação institucional decorre da vontade de que a Universidade de Brasília e, em especial, a Faculdade de Tecnologia proporcione aos estudantes de engenharia, e quaisquer outros cursos, a oportunidade de aplicar os conhecimentos adquiridos ao longo da graduação e pós para qualquer área que tenham interesse e que tenha recursos para que os estudantes e professores possam se desenvolver pessoal e profissionalmente nas mais diversas esferas do conhecimento, indiferentemente de sua especialização, fomentando assim a interdisciplinaridade, a pluralidade e a diversidade no ambiente acadêmico e profissional. Dessa forma, esse trabalho é uma contribuição do Departamento de Engenharia Elétrica aos estudos da esfera financeira.

1.3 Contexto

Duas escolas econômicas relevantes sustentam visões distintas acerca do mercado financeiro e decisões de cunho econômico no geral. Enquanto a hipótese do mercado eficiente supõe os agentes tomadores de decisão como indivíduos racionais que buscam maximizar sua função utilidade, a psicologia cognitiva aplicada à economia, ou economia comportamental, sugere que os processos de decisão humanos estão sujeitos à várias ilusões. Alguns dos processos de ilusão cognitiva aos quais estão sujeitos os processos de decisão incluem: ancoragem, falácia do jogador, aversão à perda e viés da confirmação.

Os defensores da economia comportamental sugerem que o viés e os efeitos de enquadramento causados pela heurística fazem com que os preços de mercado se desviem dos seus valores fundamentais [2]. Dessa forma, os processos de decisão tomados no mercado financeiro teriam dois efeitos, um imposto ao tomador de decisão e outro imposto aos valores do próprio mercado. No que diz respeito ao tomador de decisão, posições tomadas influenciadas por alguns dos processos de ilusão cognitiva podem causar prejuízos ao seu

portfólio. No que tange o mercado, haveriam oportunidades que surgem pela ineficiência temporária do mercado, refletida nos preços dos ativos.

Essa posição defendida pela escola da economia comportamental é sustentada por alguns investidores quantitativos. Nesse tópico, George Soros, no painel de regulamentação de hedge funds sediado pelo Congresso Americano após a crise de 2008 disse: “Primeiramente, mercados financeiros não refletem condições subjacentes exatamente. Eles provêm uma imagem que é sempre enviesada ou distorcida de uma forma ou outra. Além disso, as visões distorcidas de participantes do mercado e expressadas nos preços dos ativos podem, sob certas circunstâncias, afetar os chamados fundamentos que os preços do mercado supostamente refletem. Eu chamo essa relação circular de dois rumos, entre preços de mercado e a realidade subjacente, de reflexividade. Eu afirmo que mercados financeiros são sempre reflexivos e em certas ocasiões eles podem desviar consideravelmente do chamado equilíbrio” [3].

É nesse contexto, de um suposto mercado ineficiente associado a um método de tomada de decisão não enviesado, que se pode criar um *framework* de operação no mercado a partir do estudo de séries temporais, associadas à engenharia. A convolução de métodos estatísticos e algoritmos pautados em dados e otimizações fez surgir um novo tipo de jogador no mercado financeiro, os chamados quants.

Os quants são investidores cujas decisões de operação no mercado são pautadas em algoritmos ou estratégias sistematizadas. Em outras palavras, são formadas regras de operação ao invés de tomadas de decisão discricionárias. A limitação do julgamento humano pelo modelo matemático criado se sustenta em três pilares: dados, estatística e sistematização.

Quando se fala dos quants, a história de três fundos (LTCM, D.E. Shaw e Medallion) se destaca por dois motivos: primeiramente seu alto rendimento e impacto nos mercados, além de seu papel como precursor de muitas estratégias desenvolvidas com a atração de muitos talentos da academia e da indústria, estudiosos relacionados principalmente às áreas de matemática, estatística, engenharia e computação.

Long Term Capital Management, LTCM, foi provavelmente o fundo mais famoso da história. Liderado pelo matemático John Meriwether, o *hedge fund* contou com a presença de alguns ganhadores do prêmio Nobel como Robert Scholes e Robert C. Merton no desenvolvimento de suas estratégias e na administração geral do fundo. Contando com uma equipe reconhecida no mercado, o fundo altamente capitalizado acumulou retornos impressionantes baseando suas operações principalmente em arbitragens realizadas com produtos de renda fixa.

Basicamente, o fundo identificava anomalias em investimentos similares e, apostando que essa anomalia desapareceria, o fundo capitalizava na conversão dos investimentos

ao seu comportamento normal. Algumas das operações principais do fundo consistiam em comprar títulos cujos preços haviam caído abaixo de uma média histórica e operar vendido em títulos similares que pareciam supervalorizados, assim, quando os preços dos instrumentos convergiam o fundo capitalizava. Como essas anomalias normalmente eram de poucos pontos percentuais, a capitalização era amplificada pela alavancagem do fundo. Isso fez com que o LTCM ganhasse uma média de 50% nos seus primeiros três anos, administrando algo próximo de sete bilhões de dólares. Devido à sua posição altamente alavancada, eventos de cauda desencadeados pela inadimplência russa e crise asiática fizeram com que o fundo chegasse a perder mais de 500 milhões de dólares em um só dia (contrariando seu gerenciamento de risco baseado em *value at risk*, que previa uma perda máxima de 35 milhões) e, eventualmente, o Federal Reserve teve de formar um consórcio com os maiores bancos de investimento mundiais para resgatar o fundo e eventualmente dissolvê-lo [4].

O grupo D.E. Shaw foi fundado pelo cientista da computação David Shaw com capitalização do investidor Donald Sussman. Crescendo baseado na estratégia de arbitragem estatística no mercado de ações desenvolvida por Shaw quando trabalhava no banco de investimentos Morgan Stanley, o fundo acumulou retornos médios anuais de 18% desde seu lançamento. O portfólio do fundo era neutro com relação ao mercado, ou seja, na teoria, era blindado dos movimentos de valorização e desvalorização do mercado no geral. O grupo de investimentos de Shaw foi descrito pela revista Fortune como “a força mais intrigante e misteriosa de Wall Street... o melhor quant shop, um ninho de matemáticos, cientistas da computação e outros devotos da análise quantitativa” [5].

Finalmente, talvez o *hedge fund* mais bem-sucedido da história, liderado pelo matemático Jim Simons, conhecido como o homem que bateu o mercado, o Medallion e seu time de acadêmicos mantém em segredo suas estratégias guardadas em seus modelos de trading automatizados. Mesmo em segredos, sabe-se que umas das contribuições do fundo para a análise quantitativa do mercado foi o emprego de técnicas de reconhecimento de linguagem natural para previsões dos comportamentos dos ativos. Em essência, cria-se um modelo capaz de digerir quantidades incertas de informação e gerar palpites confiáveis do que ocorrerá em seguida. Esse tipo de estratégias adornadas com seus segredos vem retornando ao fundo algo em torno de 39% ao ano, superando a rentabilidade de investidores como Warren Buffet [5].

O que esses fundos trazem em comum é a análise sistematizada das séries temporais financeiras, baseada em análise estatística, métodos matemáticos e métodos de processamento computacional para que se possa desenvolver estratégias rentáveis sem que seja necessário se basear na intuição e na experiência prévia humana, que, como vem sendo estudado pela escola econômica comportamental, pode sofrer com algumas ilusões cognitivas

e prejudicar o desempenho financeiro dos investidores. Dessa forma, o desenvolvimento de estratégias quantitativas expande a capacidade do investidor de aumentar seus retornos no mercado.

1.4 Objetivos

Percebe-se que tanto a possível existências de ineficiências no mercado quanto o método de tomada de decisão não enviesado podem dar uma vantagem ao investidor individual ou institucional. Dessa forma, o presente projeto busca avaliar os resultados de cinco estratégias embasadas em análises estatísticas de séries temporais, em especial, a cointegração, para operar no mercado de ações brasileiro e, assim:

1. Avaliar a atratividade das estratégias no que diz respeito a risco e retorno;
2. Avaliar qual estratégia apresenta melhor performance no cenário investigado e quais apresentam maior potencial;
3. Avaliar pontos a serem melhorados, rejeitados e aprovados.

O presente projeto considera as seguintes limitações de escopo: de tempo, de mercado, de ativos e de capacidade de processamento computacional.

O estudo tem como alvo a análise de 140 ativos do mercado brasileiro. Os algoritmos são desenvolvidos no período que compreende 18/03/2019 a 16/03/2020. Sendo que é desenvolvida uma análise walk-forward para 10 semanas entre 06/01/2020 e 13/03/2020.

O mercado brasileiro foi escolhido por três motivos: pela disponibilidade de dados minuto a minuto dos ativos em questão, para que se promova um estudo que busque estimular a participação do investidor individual brasileiro no mercado de capitais e para que se contribua para a literatura acerca de estratégias considerando dados nacionais.

A limitação da capacidade computacional se mostra nas restrições de otimização e na pré-seleção dos dados a serem utilizados nas estratégias.

Por fim, as estratégias utilizarão uma avaliação sistemática e automatizada de séries financeiras temporais afim de simular, dentro das limitações de um ambiente sem interação com outros investidores, a atratividade e a validade da estratégia em questão

1.5 Resultados Obtidos

Os resultados obtidos para os cinco protótipos envolvidos se desenvolvem principalmente ao redor de dois parâmetros: rentabilidade e volatilidade, sendo a rentabilidade a medida do tamanho da capitalização derivada da estratégia e a volatilidade um indicador de risco

e incerteza. Dessa forma, quanto maior a rentabilidade da estratégia, maior é a sua atração, mas essa atração deve ser ponderada pela medida de incerteza que ela carrega. O comportamento desses parâmetros como medida de atratividade é inverso, ou seja, o que se quer é uma estratégia de alta rentabilidade e pouca incerteza.

Os resultados relacionados a ambos os parâmetros podem ser vistos nas Figura 1.1 e Figura 1.2.

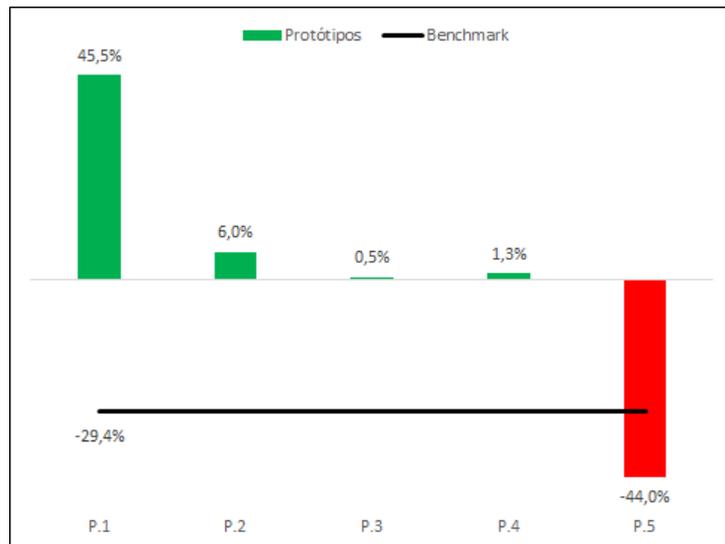


Figura 1.1: Comparativo de rentabilidade.

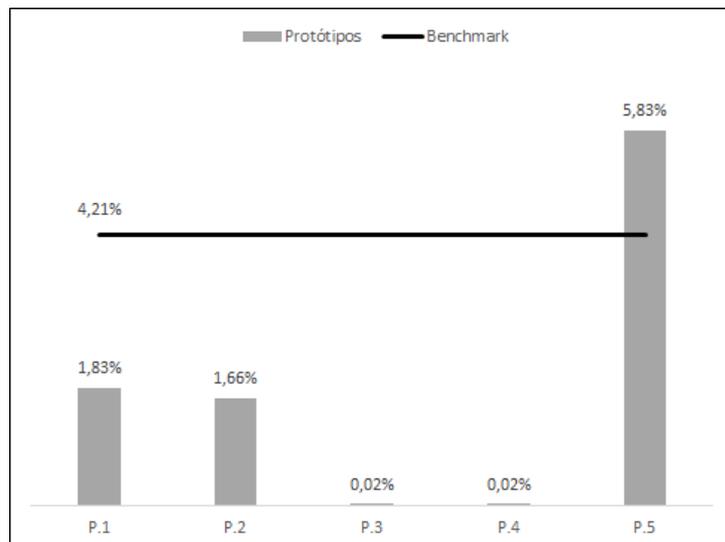


Figura 1.2: Comparativo de volatilidade.

Como se pode observar pelos resultados obtidos, para quatro dos cinco protótipos desenvolvidos, tanto a rentabilidade é maior quanto a volatilidade é menor do que os

parâmetros obtidos para o *benchmark*. A atratividade desses protótipos e uma análise mais aprofundada é feita no Capítulo 4.

2 Fundamentação

Esse capítulo descreve a fundamentação teórica utilizada para a implementação do sistema de trading descrito no Capítulo 3. Nesse contexto, são descritas estratégias quantitativas no geral, de forma a contextualizar o leitor acerca dessa forma de investimento, antes de se detalhar o funcionamento da arbitragem estatística, que é o modelo explorado no projeto.

Descrito o modelo de investimento, pode-se seguir para os conceitos estatísticos e matemáticos utilizados em sua construção. Dessa forma, descrevem-se tópicos em integração de séries temporais, filtragem estocástica para estimação de parâmetros, distribuições bivariadas para entrada e saída de posições e, finalmente, o método de simulação pós-otimização que aproxima os resultados obtidos do trading em tempo real.

2.1 Estratégias Quantitativas

Uma estratégia quantitativa é um processo de trading que se baseia na análise sistemática de dados e modelos para implementar as decisões de compra e venda de ativos ao invés de utilizar processos de decisão discricionários.

Historicamente, há alguns fundos quantitativos que se destacaram pelo retorno proporcionado aos seus investidores como o Medallion, D. E. Shaw e Quantitative Management Associates. Enquanto outros se destacaram pelo impacto geral no mercado, não necessariamente de forma positiva, como o Long Term Capital Management.

Hoje, há motivos para se acreditar que vantagens que *hedge funds* quantitativos possuem só irão aumentar. Isso se deve às previsões dos novos tipos de dados que os modelos computacionais serão capazes de digerir e analisar. A IBM estima que 90% do conjunto de dados presentes no mundo foram criados nos últimos três anos somente, e que 40 zettabytes, ou 40 trilhões de gigabytes, de dados serão criados até o fim de 2020. Tanto a diversidade quanto a quantidade de dados digitalizados e disponibilizados para investidores são uma realidade que fomenta o uso de análises quantitativas [5].

O crescimento exponencial da capacidade de processamento computacional, guiado pela Lei de Moore, e armazenamento tem possibilitado o manejo de todo tipo de dados.

Esse potencial permite que *quants* descubram e testem muito mais modelos preditivos do que jamais fora imaginado. “Ao invés de analisar estratégias *hit-and-miss*, utilizando a criatividade para achar sinais preditivos, agora é possível instituir uma classe de fórmulas em um sistema de *machine learning* e testar milhões de distintas possibilidades” diz um especialista da Renaissance Technologies [5].

Com todo esse entusiasmo, porém, as limitações da abordagem quantitativa são bem claras. Processar e descobrir sinais preditivos dos dados do mercado não é algo simples e fácil: todo o ruído contido nas séries temporais financeiras e a limitação dos métodos de aprendizado de máquina desenvolvidos não são suficientes para tornar esse processo confiável e eficiente sem a supervisão humana. Dessa forma, mesmo que as análises quantitativas possuam um grande apelo, seu estudo é algo em andamento e um campo em expansão, não só quando consideramos o mercado financeiro, mas a indústria como um todo.

As estratégias quantitativas podem ser divididas em dois grupos: as direcionais e as de valor relativo. As estratégias direcionais normalmente se constroem baseadas em tendências de mercado ou outros padrões sugestivos de *momentum*. Já as estratégias de valor relativo tentam capitalizar em relações de preço previsíveis entre vários ativos.

2.1.1 Estratégias Direcionais

Estratégias direcionais tipicamente se baseiam em seguimento de tendências ou outra forma de seguimento de padrões que sugiram um momento de alta ou de baixa no preço de um ativo ou no mercado em geral. Vários métodos podem ser implementados para que a direção possa ser estimada: análise técnica, análises de sentimento, combinações entre elas, entre outras.

Análises técnicas ou *charting* envolvem o estudo dos preços e volumes do ativo no passado recente como um tipo de preditor dos preços no seu futuro próximo. Elas podem envolver o uso de médias móveis, bandas referentes aos níveis de confiança do desvio padrão dos preços, níveis de suporte e resistência e regressões como modelos de previsão.

Além disso, a direção a ser estimada pode ser derivada de análises de sentimento. Esses tipos de estudo normalmente envolvem métricas que influenciam o ativo de uma forma indireta, como medidas de volatilidade implícita em *puts* e *calls*, taxas de juros de curto prazo, posições de grandes fundos e indicadores de volume de posições de compra e venda.

Vários métodos estatísticos e de *machine learning* são implementados para complementar ou aprofundar tanto as análises técnica quanto a de sentimento. Alguns exemplos são o uso de *deep-learning* para prever a formação de resistências, suportes, e outros indi-

cadres utilizados; e também o uso de interpretadores de linguagem natural para estimar métricas de sentimento de redes sociais ou sites de publicação oficial [6].

2.1.2 Estratégias Não Direcionais

Estratégias não direcionais normalmente envolvem algum tipo de arbitragem. Arbitragem é o processo que envolve a capitalização pela precificação relativa ineficiente. É importante ressaltar que a precificação relativa envolve o preço de um ativo com relação a outro ativo, e não seu valor absoluto. Nesse tipo de estratégia, o ativo subvalorizado é comprado e o ativo supervalorizado é vendido e a capitalização ocorre com o retorno dos preços relativos à média.

Dentre as várias formas de arbitragem, se destacam três: arbitragem estatística, arbitragem de risco e arbitragem de renda fixa. Arbitragem de renda fixa compreende: arbitragem de *spread* de *swap*, arbitragem da curva de juros, arbitragem de hipotecas, arbitragem de volatilidade, arbitragem de estrutura de capital, entre outras [7]. Arbitragem de risco envolve a ocorrência eminente de uma fusão, dessa forma, com a divulgação dos preços de compra de uma empresa pela outra, pode haver uma oportunidade de lucro por uma precificação diferente entre a soma dos preços das ações das empresas separadas e do valor final da ação da empresa pós-fusão.

Arbitragem estatística é uma estratégia que se tornou conhecida devido ao seu uso pelo fundo Long Term Capital Management, que acumulou retornos expressivos antes de entrar em uma espiral de perdas iniciada pelas crises financeiras asiática de 1997 e russa de 1998 [4]. Nesse tipo de estratégia, um modelo matemático estabelece uma relação de preços entre dois ou mais ativos e desvios nesse modelo indicam uma precificação ineficiente e uma oportunidade de capitalização.

Para definir em termos simples, arbitragem estatística compreende um conjunto de estratégias quantitativas que buscam explorar a movimentação da precificação relativa para milhares de instrumentos financeiros a partir da análise dos padrões e das diferenças desses preços. Esse tipo de estratégia possui flexibilidade temporal na medida em que pode ser aplicada em trades de alta, média e baixa frequência.

Quanto aos conceitos utilizados nas estratégias de arbitragem estatísticas, eles são empregados principalmente na análise da precificação relativa e na análise da diferença da precificação para abertura de uma posição. Para essas análises, pode-se empregar conceitos de séries temporais, autoregressão, cointegração, correlação, modelagem de volatilidade, PCA (*principal component analysis*), entre outros [8].

Com relação aos tipos de arbitragem estatística, elas podem incluir: arbitragens neutras de mercado, arbitragem de ativos cruzados, arbitragem de mercados cruzados, arbitragem com ETF, entre outras. As estratégias neutras de mercado envolvem uma posição

comprada e uma vendida (*long and short*) em um ativo subvalorizado e um sobrevalorizado, respectivamente. Os ativos são comprados e vendidos de forma a compensar suas volatilidades, dessa forma, uma apreciação do mercado causaria um ganho em uma posição e uma perda na outra, e o resultado da estratégia se daria da diferença entre essas apreciações.

As estratégias de ativos cruzados buscam explorar diferenças de precificações entre um ativo e uma forma de instrumento financeiro que também envolva o ativo, como uma ação e suas opções de compra e venda. Já as estratégias de mercado cruzado buscam explorar diferentes precificações de um mesmo ativo em diferentes mercados. As arbitragens de ETF podem ser vistas como um tipo de arbitragem com ativo cruzado, na medida em que são arbitragens entre um ETF e os ativos que o compõem.

Quando observada sob uma perspectiva evolutiva, arbitragem estatística é uma evolução do chamado *pairs trading*, introduzida pelo *trader* Nunzio Tartaglia [9]. O processo envolve identificar pares de ações em que os preços apresentavam movimentos similares. Quando era identificada uma anomalia nessa relação, montava-se uma posição com a ideia de que a anomalia se corrigiria. Tartaglia e seu grupo empregaram a estratégia de *pairs trading* ao longo de 1987. Apesar do grupo ter se desmontado em 1989, a estratégia de *pairs trading* se tornou popular entre investidores individuais, institucionais e *hedge funds* [10].

Mesmo que esse tipo de estratégia tenha retornado grandes lucros para aqueles que a empregaram, ela possui seus riscos inerentes, em especial, destacam-se dois: a estratégia depende de forma acentuada de relações de reversão à média, o que pode não ocorrer em certos casos e certos períodos do mercado; além disso, o fluxo constante do mercado pode causar mudanças nas relações entre os ativos e lucros não podem ser garantidos a todos os tempos.

2.2 Econometria de Séries Temporais

O estudo econométrico das séries temporais financeiras, representativas dos preços dos ativos, é utilizado nesse projeto para que relações de cointegração entre ativos sejam identificadas. A cointegração entre os ativos é o ponto central da construção do universo de ativos operados pelos protótipos construídos e a característica intrínseca de retorno à média de um resíduo estacionário é o pivô da arbitragem estatística desenvolvida.

2.2.1 Fundamentos Estatísticos

Processos Estocásticos

No escopo limitado pela precificação dos ativos do mercado de ações, as séries temporais compostas por esses dados são séries estocásticas. De forma geral, processos estocásticos podem ser definidos como uma coleção de variáveis aleatórias ordenadas no tempo - $\{x_t \in T\}$ -, em que T é um conjunto ordenado de índices. Um processo estocástico fica caracterizado se definirmos a função de distribuição conjunta das variáveis aleatórias que o compõem para qualquer valor de T . Diz-se que a estrutura probabilística de um processo estocástico é conhecida quando se conhecem as seguintes distribuições: funções de médias, funções de variâncias, função de autocovariâncias e função de autocorrelação [11].

Os processos estocásticos podem ser estacionários ou não. Processos estacionários estocásticos z_t ocorrem quando as propriedades estatísticas de qualquer sequência finita z_1, \dots, z_k de componentes de Z_t são semelhantes às da sequência z_{1+h}, \dots, z_{k+h} para qualquer h inteiro. Já os processos não estacionários ocorrem quando as propriedades estatísticas de qualquer sequência finita finita z_1, \dots, z_k de componentes de Z_t não são semelhantes às da sequência z_{1+h}, \dots, z_{k+h} para qualquer h inteiro. Um exemplo de processo estocástico estacionário é o ruído branco, já o de um processo não estacionário é o consumo elétrico anual de uma região [11].

Processos Estacionários

O conceito de estacionaridade é uma das principais ideias que se deve ter para que se possa estimar uma série temporal. É, fundamentalmente, a constatação de estacionaridade que permitirá proceder inferências estatísticas sobre os parâmetros estimados com base na realização de um processo estocástico [12].

Um processo estocástico pode ser fracamente estacionário ou estritamente estacionário. A série temporal é fracamente estacionária se:

1. $E|y_t|^2 < \infty$;
2. $E(y_t) = \mu$, para todo t ;
3. $E(y_t - \mu)(y_{t-j} - \mu) = \gamma_j$.

Segundo Rodrigo L. S. Bueno em [12], a primeira condição afirma que o segundo momento não centrado deve ser finito. A segunda condição assegura que a média é igual para todo o período, mesmo que a distribuição da variável aleatória se altere ao longo do tempo. A última condição estabelece que a variância é sempre igual para todo período e que a autocovariância não depende do tempo.

O processo estocástico também pode ser estritamente estacionário se a função distribuição conjunta de $\{y_{t_i}\}_{i=1}^k$ for igual à função de distribuição conjunta de $\{y_{t_i+h}\}_{i=1}^k$. Intuitivamente, estacionaridade estrita implica que os gráficos da função distribuição da série em quaisquer dois intervalos de tempo de igual tamanho exibirão propriedades estatísticas similares [12].

Para esse trabalho, a estacionaridade fraca é suficiente para que sejam atendidas as características desejadas de um conjunto de ativos a serem operados. Graficamente, é possível se observar estacionaridade se uma série flutua em torno de uma média fixa. Não obstante, serão necessários testes estatísticos para se verificar ou não a estacionaridade da série.

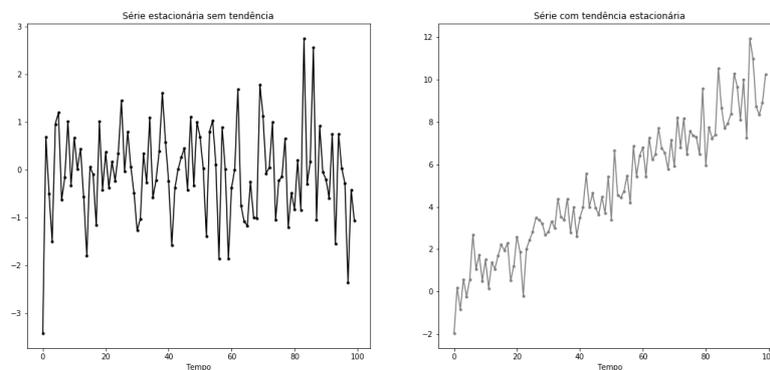


Figura 2.1: Ilustração: série estacionária e série com tendência estacionária.

Tendências

De modo geral, uma série econômica tem três partes distintas: tendência, ruído e um componente estacionário. A equação

$$y_t = y_0 + \delta t + \phi(L)\epsilon_t \quad (2.1)$$

define o que se chama de tendência estacionária, porque a série flutua em torno de uma tendência determinística, sem se distanciar dela. Outra forma de se ver isso é observando que a subtração da tendência determinística gera um componente estacionário [13].

A equação

$$\Delta y_t = \delta + \epsilon_t \rightarrow y_t = y_{t-1} + \delta + \epsilon_t \quad (2.2)$$

modela as variações esperadas de y , que são compostas recursivamente, culminando em

$$y_t = y_0 + \delta t + \sum_{i=1}^t \epsilon_t. \quad (2.3)$$

A variável aleatória y é, assim, dada pela composição de todos os choques ocorridos, representados pelo somatório. Define-se tal série como de tendência estocástica ou diferença estacionária, pois a primeira diferença da série é estacionária. Nesse caso, os choques produzem mudanças permanentes na série y , ainda que aleatórias [13].

Uma série cuja tendência é estocástica é chamada série integrada, $I(d)$. A ordem de integração, d , depende do número de somatórios dos erros. No caso anterior, diz-se que a série é integrada de ordem 1, $I(1)$, porque os erros são todos somados uma única vez, como indica a presença de um somatório. Dessa forma, a série é integrada de ordem 1, $I(1)$, porque precisa ser diferenciada apenas uma vez para ser estacionarizada [13].

2.2.2 Testes de Raiz Unitária

Testes de raízes unitárias são capazes de verificar a estacionaridade ou não de uma série temporal. A condição necessária para a estacionaridade fraca é que as raízes da equação característica devem estar localizadas fora do círculo unitário cartesiano [12].

Dickey-Fuller

A inspeção visual de uma série raramente permite distingui-la como de tendência estocástica ou tendência determinística. Isso levou ao desenvolvimento de testes para a verificação de existência de raízes unitárias.

Para se entender o teste, considere o modelo:

$$y_t = \phi y_{t-1} + \epsilon_t. \quad (2.4)$$

A tendência inicial é estimar esse modelo e usar um teste convencional de t sobre ϕ , tendo como hipótese nula $H_0 : \phi = 1$ [13]. Em geral, os pacotes econométricos reportam os testes nos coeficientes contra a nula de serem iguais a zero. Então, pode-se alterar o teste subtraindo y_{t-1} de ambos os lados:

$$\Delta y_t = (\phi - 1)y_{t-1} + \epsilon_t = \alpha y_{t-1} + \epsilon_t. \quad (2.5)$$

Assim, $H_0 : \phi = 1 \rightarrow \alpha = 0$. O problema é que, sob a nula, a distribuição do teste infelizmente não é convencional, ou seja, não é igual à distribuição estatística t , pois y_t não é estacionário. A distribuição do teste pode ser encontrada analiticamente usando-se o teorema central do limite funcional e pode ser visualizada por meio de experimentos de Monte Carlo [13].

Dessa forma, supondo uma série com $T + 1$ observações, a estimação do parâmetro α é feita por mínimos quadrados:

$$\hat{\alpha} = \frac{\sum_{t=1}^T y_t y_{t-1}}{\sum_{t=1}^T y_{t-1}^2} - 1, \quad (2.6)$$

sua variância amostral é calculada:

$$S^2 = \frac{1}{T-1} \sum_{t=1}^T (\Delta y_t - \hat{\alpha} y_{t-1})^2, \quad (2.7)$$

e depois o desvio padrão do coeficiente estimado é computado:

$$s(\hat{\alpha}) = \frac{S}{\sqrt{\sum_{t=1}^T y_{t-1}^2}}, \quad (2.8)$$

como enunciou Rodrigo L. S. Bueno em [13].

Finalmente, a estatística $\hat{\tau}$ é calculada por:

$$\hat{\tau} = \frac{\hat{\alpha}}{s(\hat{\alpha})}. \quad (2.9)$$

O critério de rejeição é: se $\hat{\tau} < \tau$ (em que τ são os valores críticos obtidos por experimentos de Monte Carlo) rejeita-se a nula e conclui-se que a série não possui raiz unitária. Por outro lado, não rejeitar a nula significa admitir que a série possui uma raiz unitária [13].

Dickey-Fuller Aumentado

Segundo Rodrigo L. S. Bueno em [13], o problema do teste anterior é que Dickey e Fuller consideram o erro um ruído branco. Mas, frequentemente, o erro é um processo estacionário qualquer e, assim, esse problema pode causar distorções no poder do teste.

A ideia por trás do teste de Dickey Fuller Aumentado é estimar o modelo com variáveis autorregressivas. Essa é uma forma de corrigir o desvio do valor correto da estatística, ou seja, intuitivamente trata-se de encontrar os desvios de y_t em relação à sua média, para deslocar a distribuição de α em direção a zero, caso a hipótese nula seja verdadeira.

Do ponto de vista prático, o modelo dado pela Equação 2.4 é reescrito como um processo autorregressivo de ordem p :

$$y_t = \mu + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \epsilon_t, \quad (2.10)$$

$$\Delta y_t = \alpha y_{t-1} + \sum_{i=1}^{p-1} \lambda_i \delta y_{t-1} + \epsilon_t, \quad (2.11)$$

em que $\alpha = -(1 - \sum_{i=1}^p \phi_i)$ e $\lambda_i = -\sum_{j=1}^{p-1} \phi_{j+1}$. O teste então pode ser feito usando-se os mesmos valores críticos encontrados por Dickey e Fuller. Ou seja, desde que se faça a correção do modelo, de forma a considerar as demais variáveis defasadas, em nada se alteram os valores críticos do teste ou a interpretação do modelo [13].

A ordem p para execução do teste de raiz unitária deve ser escolhida de forma a considerar que o número de defasagens incluídas deve ser tal que os resíduos estimados sejam um ruído branco. Assim, um opção é fixar um p_{max} relativamente alto e estimar o modelo por mínimos quadrados ordinários para cada p_i . No caso desse projeto, o valor de p_{max} é dado pelo critério de escolha de Schwert [14]:

$$p_{max} = \text{int}[12 \times (\frac{T}{100})^{\frac{1}{4}}]. \quad (2.12)$$

2.2.3 Cointegração

Engle e Granger em [15] definiram que os elementos de um vetor X_t , $n \times 1$, são ditos cointegrados de ordem (d, b) , denotados por $X_t \sim CI(d, b)$ se:

1. Todos os elementos de X_t são integrados de ordem d , ou seja, são $I(d)$;
2. Existe um vetor não nulo β , tal que:

$$u_t = X_t^T \beta \sim I(d - b), b > 0. \quad (2.13)$$

A primeira condição indica que todas as variáveis constantes de X_t devem estar na mesma ordem de integração para que possam ser cointegradas [16]. A segunda condição indica que existe uma combinação linear entre os elementos do vetor X_t que resulta em um resíduo integrado de ordem menor que a dos elementos combinados.

As séries econômicas e financeiras, em geral, têm ordem de integração menor que dois e séries estacionárias têm ordem de integração igual a zero. Em outras palavras, como a ordem de integração é o número de diferenças que deve ser aplicado à série para torná-la estacionária, a diferença $(d - b)$ para o caso de séries financeiras é, em geral, nula - caracterizando um resíduo estacionário. Economicamente, isso representa uma relação de longo prazo entre as coordenadas desse vetor, pois a ordem de integração da série diminui. Essa relação de longo prazo então é perturbada por choques de curto prazo representados por u_t . Porém, a estacionaridade da série temporal garante que essa perturbação tende a se dissipar ao longo do tempo, de modo que as variáveis originais voltam ao seu equilíbrio de longo prazo [16].

Dessa forma, considerando a necessidade de se obter um resíduo estacionário, a questão que emerge é como testar se os resíduos são de fato estacionários. É natural então pensar no teste de raiz unitária sobre os resíduos. Logo, o modelo de cointegração exige dois passos:

1. Estimar os parâmetros que geram o resíduo dado pela combinação linear as séries temporais;
2. Aplicar o teste de raiz unitária sobre o resíduo gerado para garantir seu comportamento estacionário.

Com relação a esse modelo, há dois testes principais que são utilizados na literatura: o teste de cointegração de Engle-Granger e o teste de Johansen. No caso desse trabalho, faz-se um modelo próprio, utilizando a estimação de parâmetros baseada em modelos de filtragem estocástica e no teste de Dickey Fuller Aumentado.

2.3 Filtragem Estocástica

O problema de se estimar o estado de um sistema dinâmico estocástico de observações ruidosas é de importância central na engenharia. O interesse nesse tipo de problema se remota a Gauss que, interessado em determinar os elementos da órbita de um corpo celestrial a partir de observações, desenvolveu o que é conhecido hoje como mínimos quadrados. Mais recentemente, os nomes de Kalman e Wiener são associados a avanços na teoria de estimação [17].

Um sistema dinâmico estocástico discreto pode ser descrito por

$$x_{k+1} = \phi(x_k, t_{k+1}, t_k) + \Gamma(x_k, t_k)w_{k+1}, \quad (2.14)$$

em que o estado no tempo t_k é x_k , um vetor de dimensão n , ϕ é uma função vetor de dimensão n , Γ é uma matriz $n \times r$ e w_k é um vetor de dimensão r tal que $w_k \sim N(0, Q_k)$. As observações discretas e ruidosas são um vetor de dimensão m :

$$y_k = h(x_k, t_k) + v_k, \quad (2.15)$$

em que h é uma função vetor de dimensão m e v_k é um vetor de dimensão m tal que $v_k \sim N(0, R_k)$ [17].

Dada uma realização da sequência de observações y_1, \dots, y_l , ou seja, dado Y_l , a estimação discreta do problema consiste em computar uma estimativa x_k , baseada em Y_l . Se $k < l$, o problema é chamado de *discrete smoothing*; se $k = l$, o problema é chamado de *discrete filtering*; e caso $k > l$, é chamado de *discrete prediction* [17].

Problemas de filtragem e predição são normalmente associados com operações em tempo real, em que estimativas são necessárias com base em observações e dados disponíveis agora. Para esse projeto, a ideia é utilizar o filtro estocástico, em específico, o filtro de Kalman para se estimar os parâmetros necessários para que se combinem os ativos de forma a gerar um resíduo, para depois submetê-lo ao teste de Dickey-Fuller Aumentado.

O filtro de Kalman é um conjunto de equações matemáticas que fornece um método eficiente, computacional e recursivo de se estimar o estado de um processo de uma forma que minimiza a média do erro quadrático. O filtro é importante devido ao fato de que ele suporta estimativas de estados passados, presentes e até futuros e pode fazê-lo mesmo quando a natureza do sistema modelado é desconhecida [18].

2.3.1 Processo Estimado

Segundo Welch e Bishop em [18], o filtro de Kalman aborda o problema geral de estimar o estado de um processo discreto que é governado por uma equação linear estocástica diferencial:

$$x_k = Ax_{k-1} + Bu_{k-1} + w_{k-1}, \quad (2.16)$$

com medições dadas por:

$$z_k = Hx_k + v_k. \quad (2.17)$$

As variáveis w_k e v_k representam os ruídos do processo e da medição, respectivamente. Elas são consideradas independentes e com distribuição normal de probabilidades:

$$p(w) \sim N(0, Q_k), \quad (2.18)$$

$$p(v) \sim N(0, R_k), \quad (2.19)$$

em que Q e R são as matrizes de covariância do ruído do processo e da medição respectivamente, e podem ser adaptativas com cada período de tempo, ou assumidas constantes.

A matriz A , $n \times n$, relaciona o estado no período anterior $k-1$ com o estado no período atual k na ausência de uma função condutora ou um ruído de processo. Já a matriz B relaciona a entrada ótima de controle u com o estado x . A matriz H , $m \times n$, relaciona o estado à medição z_k [18].

2.3.2 Computação dos Processos

São definidos dois estados do filtro: um *a priori* e outro *a posteriori*. Define-se \hat{x}_k^- como uma estimativa *a priori* do estado no período k , dado o conhecimento do processo anterior ao período k ; e \hat{x}_k é considerada uma estimativa *a posteriori* do estado no período k , dada

a medida z_k [18]. Assim, define-se erros de estimativa *a priori* e *a posteriori* como:

$$\hat{e}_k^- \equiv x_k - \hat{x}_k^-, \quad (2.20)$$

$$\hat{e}_k \equiv x_k - \hat{x}_k. \quad (2.21)$$

Dessa forma, a covariância de erro estimada a priori e a posteriori são, respectivamente:

$$\hat{P}_k^- = E[e_k^- e_k^{-T}], \quad (2.22)$$

$$\hat{P}_k = E[e_k e_k^T]. \quad (2.23)$$

Derivando as equações para o filtro de Kalman, nota-se que o objetivo começa em se encontrar uma equação que compute a estimativa de estado a posteriori x_k como uma combinação linear da estimativa *a priori* e uma diferença ponderada entre a atual medida e a predição da medição Hx_k :

$$\hat{x}_k = \hat{x}_k^- + K(z_k - H\hat{x}_k^-), \quad (2.24)$$

em que diferença $(z_k - H\hat{x}_k^-)$ é chamada inovação de medição ou resíduo. Dessa forma, resíduo reflete a discrepância entre a medição predita e a medição realizada [18].

A matriz K , $n \times m$, é denominada matriz de ganho de Kalman, que minimiza a covariância do erro *a posteriori*. A minimização é realizada primeiro substituindo na Equação 2.21, substituindo-a na equação de estimativa da covariância do erro *a posteriori*, realizando as expectativas indicadas, tomando a derivada do traço do resultado em relação a K , definindo o resultado igual a zero e depois resolvendo para K :

$$K_k = P_k^- H^T (H P_k^- H^T + R)^{-1} = \frac{P_k^- H^T}{H P_k^- H^T + R_k}. \quad (2.25)$$

Nota-se, pela Equação 2.25, que na medida em que a medida da covariância do erro R se aproxima de 0, o ganho K tem maior peso sobre o resíduo:

$$\lim_{R_k \rightarrow 0} K_k = H^{-1}. \quad (2.26)$$

Por outro lado, à medida em que a estimativa da covariância de erro *a priori* se aproxima de 0, o ganho de Kalman tem menor peso sobre o resíduo:

$$\lim_{P_k^- \rightarrow 0} K_k = 0. \quad (2.27)$$

Uma outra forma de abordar a questão do peso do ganho de Kalman é que a medida em

que o erro de covariância R se aproxima de 0, a medida realizada z_k se torna mais confiável enquanto a predição da medida Hx_k perde importância. Por outro lado, à medida em que a estimativa a priori da covariância do erro P_k^- se aproxima de 0, a medida realizada z_k se torna menos confiável e a medição predita Hx_k ganha importância [18]. O ganho de Kalman é a medida que balanceia o *trade off* entre a medida realizada e a predita.

2.3.3 Aspectos Probabilísticos

Como ressaltado por Welch e Bishop em [18], a justificção para a equação de resíduo está relacionada na regra de Bayes. O Filtro de Kalman mantém os dois primeiros momentos do estado da distribuição, ou seja:

$$E[x_k] = \hat{x}_k, \quad (2.28)$$

$$E[(x_k - \hat{x}_k)(x_k - \hat{x}_k)^T] = P_k. \quad (2.29)$$

A estimativa *a posteriori* reflete a média (o primeiro momento) do estado da distribuição. A estimativa da covariância do erro *a posteriori* reflete a variância do estado da distribuição (o segundo momento), como mostra a equação abaixo:

$$p(x_k|z_k) \sim N(E[x_k], E[(x_k - \hat{x}_k)(x_k - \hat{x}_k)^T]) = N(\hat{x}_k, P_k). \quad (2.30)$$

2.3.4 Algoritmo do Filtro

O Filtro de Kalman estima o processo utilizando uma malha fechada de controle: o filtro estima o estado do processo em algum tempo e obtém feedback na forma de medições ruidosas [19]. Dessa forma, as equações do filtro de Kalman se encaixam em dois grupos: atualização temporal e atualização de medições.

As equações de atualização temporal são responsáveis por projetar no tempo o estado atual e as estimativas de erro de covariância para obter as estimativas de estado para o próximo período de tempo. As equações de atualização de medição são responsáveis pelo *feedback*, ou seja, por incorporar uma nova medição na estimativa *a priori* para obter uma estimativa melhorada *a posteriori* [18]. Dessa forma, as equações de atualização temporal podem ser vistas como equações de predição, enquanto as medições de atualização de medição podem ser vistas como equações de correção. De fato, o algoritmo de estimação lembra àquele de um algoritmo preditor-corretor para resolução de problemas numéricos.

Predição:

$$\hat{x}_k^- = Ax_{k-1} + Bu_{k-1}, \quad (2.31)$$

$$P_k^- = AP_{k-1}A^T + Q_k. \quad (2.32)$$

Correção:

$$K_k = P_k^- H^T (HP_k^- H^T + R_k)^{-1}, \quad (2.33)$$

$$\hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-), \quad (2.34)$$

$$P_k = (I - K_k H)P_k^-. \quad (2.35)$$

Assim, pode-se notar que as equações de atualização temporal projetam o estado e as estimativas de covariância de um período $k - 1$ para um período k . Já as equações de atualização de medição funcionam de forma sequencial. A primeira etapa é computar o ganho de Kalman seguida da medição do processo para obter z_k e assim gerar a estimativa de estado *a posteriori* incorporando a medição. O passo final é obter a estimativa de covariância dos erros *a posteriori*.

2.3.5 Parâmetros do Filtro

Na implementação do filtro de Kalman, a covariância do ruído de medição R é geralmente medida antes da operação do filtro. Porém, a medição do erro de covariância é geralmente possível porque as medições do processo permitem a estimação inicial de R e depois essa matriz pode ser ajustada a cada período de tempo para os novos dados de medição.

A determinação da covariância do ruído de processo Q é geralmente mais difícil já que, tipicamente, não se tem a habilidade de observar diretamente o processo que se estima. Essa matriz é normalmente estimada por meio de *cross-testing* com vários parâmetros testados. Para o caso desse trabalho, a matriz Q utilizada foi sugerida em [20] após otimizações por *cross-testing*.

2.4 Tempo de Meia Vida

Processos de difusão têm sido utilizado para modelar processos dinâmicos estocásticos em vários ramos da ciência como física, biologia, entre outros. Talvez o uso mais eminente desses processos contínuos estocásticos nas últimas décadas tenha sido em finanças, seguindo o trabalho de Merton, Black e Scholes que estabeleceu a fundação da teoria de precificação de opções.

Uma aplicação importante dos processos de difusão é a modelagem das taxas de juros de curto prazo, que são quantidades fundamentais em finanças já que elas definem o excesso de retornos e prêmios de risco. Uma família de processos de difusão normalmente utilizados nessas modelagens é:

$$dZ_t = \theta(\mu - Z_t)dt + \sigma(Z_t)^\rho dB_t, \quad (2.36)$$

em que a tendência linear descreve uma reversão de Z_t à média μ com velocidade θ . A função de difusão $\sigma(Z_t)^\rho$ pode acomodar um conjunto de padrões de volatilidade à medida que Z_t aumenta. Membros importantes dessa família são o modelo de Vasicek, com $\rho = 0$ e o modelo CIR, com $\rho = 1/2$. Ambos esses modelos são utilizados em finanças por dois motivos: primeiro, ambos possuem interpretações simples e atrativas em finanças; segundo, ambos admitem soluções de forma fechada [21].

Apesar dos papéis críticos desempenhados por esses processos na modelagem de instrumentos financeiros importantes, é sabido empiricamente que a estimação dos parâmetros θ e μ pode envolver grandes vieses e variações. Esse é o caso para virtualmente todos os processos de estimação incluindo o de máximo verossimilhança [21].

Para esse trabalho, o processo de Vasicek é utilizado já que satisfaz a equação diferencial estocástica:

$$dZ_t = \theta(\mu - Z_t)dt + \sigma dB_t, \quad (2.37)$$

caracterizando assim um processo de Ornstein-Uhlenbeck com distribuição condicional de Z_t dado Z_{t-1} :

$$Z_t|Z_{t-1} \sim N(Z_{t-1}e^{-\theta\delta} + \mu(1 - e^{-\theta\delta}), \frac{1}{2}\sigma^2\theta^{-1}(1 - e^{-2\theta\delta})), \quad (2.38)$$

em que δ é o intervalo de amostragem.

A média e variância condicionais de Z_t dado Z_{t-1} são:

$$E(Z_t|Z_{t-1}) = Z_{t-1}e^{\theta\delta} + \mu(1 - e^{\theta\delta}), \quad (2.39)$$

$$Var(Z_t|Z_{t-1}) = \frac{1}{2}\sigma^2\theta^{-1}(1 - e^{-2\theta\delta}). \quad (2.40)$$

Assim, sendo $\phi(z)$ a função densidade da distribuição normal $N(0, 1)$, então a função de verossimilhança de $\kappa = (\theta, \mu, \sigma^2)$ é

$$L(\kappa) = \phi(\sigma^{-1}\sqrt{2\theta}(Z_0 - \mu)) \prod_{t=1}^n \phi(\sigma^{-1}\sqrt{2\theta(1 - e^{-2\theta\delta})^{-1}}\{Z_t - E(Z_t|Z_{t-1})\}), \quad (2.41)$$

em que n é o número de amostras do processo.

Dessa forma, Song Xi Chen define em [21] os estimadores de máxima verossimilhança como:

$$\hat{\theta} = -\delta^{-1} \log(\hat{\gamma}_1), \quad (2.42)$$

$$\hat{\mu} = \hat{\gamma}_2, \quad (2.43)$$

$$\hat{\sigma}^2 = 2\hat{\theta}\hat{\gamma}_3(1 - \hat{\gamma}_1^2)^{-1}, \quad (2.44)$$

em que

$$\hat{\gamma}_1 = \frac{n^{-1} \sum_{i=1}^n Z_i Z_{i-1} - n^{-2} \sum_{i=1}^n Z_i \sum_{i=1}^n Z_{i-1}}{n^{-1} \sum_{i=1}^n Z_{i-1}^2 - n^{-2} (\sum_{i=1}^n Z_{i-1})^2}, \quad (2.45)$$

$$\hat{\gamma}_2 = \frac{n^{-1} \sum_{i=1}^n (Z_i - \hat{\gamma}_1 Z_{i-1})}{1 - \hat{\gamma}_1}, \quad (2.46)$$

$$\hat{\gamma}_3 = n^{-1} \sum_{i=1}^n \{Z_i - \hat{\gamma}_1 Z_{i-1} - \hat{\gamma}_2 (1 - \hat{\gamma}_1)\}^2. \quad (2.47)$$

A Equação 2.39 e a Equação 2.40 sugerem que as observações discretas dadas por $\{Z_t\}_{t=0}^n$ seguem um processo autorregressivo AR(1) com o coeficiente de autorregressão dado por $\gamma_1 = e^{-\theta\delta}$ [21].

2.5 Cópulas

Cópulas são funções que unem ou “acoplam” distribuições multivariadas às suas funções de distribuição marginal unidimensional. De forma equivalente, cópulas são distribuições multivariadas, cujas marginais são uniformes no intervalo $(0, 1)$. Há duas razões pelas quais cópulas são do interesse de estatísticos: primeiro, como uma maneira de estudar medidas de dependência sem escala; e segundo, como um ponto de partida para a construção de famílias de distribuições bivariadas. Especificamente, as cópulas são uma parte importante do estudo da dependência entre duas variáveis, pois permitem separar o efeito da dependência dos efeitos das distribuições marginais. Esse recurso é análogo à distribuição normal bivariada, na qual os vetores médios são diferentes da matriz de covariância e determinam conjuntamente a distribuição [22].

Três tipos de copulas podem ser diferenciados: cópulas fundamentais, implícitas e explícitas. Cópulas fundamentais são aquelas que representam dependência perfeita positiva, independência e dependência perfeita negativa. Cópulas implícitas são cópulas extraídas de distribuições multivariadas conhecidas e não possuem forma fechada. As cópulas explícitas, que são de interesse desse projeto, são também conhecidas como Cópulas de Arquimedes e são aquelas que possuem uma forma simples de expressão fechada e que seguem regras gerais de construção matemática.

Cópulas de Arquimedes são todas construídas pela especificação de uma função geradora. Dessa forma, sendo ϕ um gerador estrito com ϕ^{-1} completamente monotônico em $[0, \infty)$, uma cópula bivariada de Arquimedes é definida como:

$$C(u, v) = \phi^{-1}(\phi(u) + \phi(v)). \quad (2.48)$$

Uma fonte de funções geradoras importantes de Cópulas de Arquimedes consiste nas inversas da transformada de Laplace de funções de distribuições contínuas [23].

Uma das características atrativas das Cópulas de Arquimedes é que elas são facilmente relacionadas às medidas de dependência. Genest e MacKay em [24] provaram que a relação entre a função geradora de cópula e a medida de dependência conhecida como τ de Kendall é dada por:

$$\tau_k = 1 + 4 \int_0^1 \frac{\phi(x)}{\phi'(x)} dx. \quad (2.49)$$

Dentre as diversas cópulas arquimedianas, três são foco desse trabalho de forma a contemplar e capturar diferentes relações de dependência além de identificar uma precificação ineficiente entre dois ativos, ou seja, derivar o sinal que permite a entrada e a saída de uma posição.

2.5.1 Cópula de Clayton

A função geradora da cópula de Clayton é

$$\phi(u) = \theta^{-1}(u^{-\theta} - 1). \quad (2.50)$$

A cópula de Clayton normalmente é utilizada para correlação de riscos por sua habilidade de capturar dependência de cauda inferior. A forma fechada da cópula de Clayton bivariada é dada por

$$C(u, v; \theta) = (u^\theta + v^\theta - 1)^{\frac{-1}{\theta}}, \quad (2.51)$$

enquanto sua função densidade de cópula é dada por

$$c(u, v; \theta) = (\theta + 1)(u^{-\theta} + v^{-\theta} - 1)^{-2-\frac{1}{\theta}} u^{-\theta-1} v^{-\theta-1}, \quad (2.52)$$

em que θ é o parâmetro da cópula restrito ao intervalo $(0, \infty)$ e é um indicador da força da dependência das distribuições. Se θ for nulo, então as distribuições marginais são independentes. Se θ tende ao infinito, então a cópula de Clayton se aproxima do limite superior de Fréchet-Hoeffding¹. Dessa forma, a cópula de Clayton não é capaz de capturar uma relação de dependência negativa, já que o limite inferior de Fréchet-Hoeffding não pode ser atingido [23].

¹Os limites de Fréchet são definidos como

$$\max \sum_{i=1}^d u_i + 1 - d, 0 \leq C(u) \leq \min u_1, \dots, u_d, \quad (2.53)$$

em que os limites superior e inferior são chamados limites superior de Fréchet-Hoeffding e inferior de Fréchet-Hoeffding, respectivamente, para uma função de cópula $[0, 1]^d \rightarrow [0, 1]$.

A estimativa do parâmetro θ pode ser feita de maneira simples e indireta a partir da computação do parâmetro de dependência, τ de Kendall:

$$\tau_k = \frac{\theta}{\theta + 2}, \quad (2.54)$$

assim, caso o valor de dependência possa ser calculado, o valor de θ é obtido de forma simples.

Finalmente, as funções de cópula condicionais bivariadas são dadas tomando-se as derivadas parciais da cópula:

$$C(v|u) = u^{-\theta-1}(u^{-\theta}v^{-\theta} - 1)^{-\frac{1}{\theta}-1}, \quad (2.55)$$

$$C(u|v) = v^{-\theta-1}(u^{-\theta}v^{-\theta} - 1)^{-\frac{1}{\theta}-1}. \quad (2.56)$$

2.5.2 Cópula de Gumbel

A função geradora da cópula de Gumbel é

$$\phi(u) = -(\ln(u))^\theta. \quad (2.57)$$

A cópula de Gumbel é utilizada principalmente como modelo de dependência assimétrica. Essa cópula é famosa por sua habilidade de capturar dependência forte de cauda superior e dependência fraca de cauda inferior. Se os resultados esperados exibem correlação forte em valores altos, mas correlação fraca em valores baixos, então a Cópula de Gumbel é uma escolha apropriada. A cópula bivariada de Gumbel é dada por

$$C(u, v; \theta) = \exp -[(-\log(u))^\theta + (-\log(v))^\theta]^{\frac{1}{\theta}}, \quad (2.58)$$

enquanto sua função de densidade é dada por

$$c(u, v; \theta) = C(u, v; \theta)(uv)^{-1}A^{-2+\frac{2}{\theta}}[\ln(u)\ln(v)]^{\theta-1}[1 + (\theta - 1)A^{\frac{-1}{\theta}}], \quad (2.59)$$

$$A = (-\ln(u))^\theta + (-\ln(v))^\theta, \quad (2.60)$$

em que θ é o parâmetro de cópula restrito ao intervalo $[1, \infty)$. Quando θ se aproxima de 1, as distribuições marginais são independentes e quando θ tende a infinito, a Cópula de Gumbel se aproxima do limite superior de Fréchet, indicando dependência positiva [23].

A estimação do valor de θ também pode ser dada pelo τ de Kendall:

$$\tau_k = 1 - \frac{1}{\theta}. \quad (2.61)$$

As funções de cópula condicionais bivariadas são dadas por:

$$C(v|u) = C(u, v; \theta)[(-\ln(u))^\theta + (-\ln(v))^\theta]^{\frac{1-\theta}{\theta}} (-\ln(u))^{\theta-1} \frac{1}{u}, \quad (2.62)$$

$$C(u|v) = C(u, v; \theta)[(-\ln(u))^\theta + (-\ln(v))^\theta]^{\frac{1-\theta}{\theta}} (-\ln(v))^{\theta-1} \frac{1}{v}. \quad (2.63)$$

2.5.3 Cópula de Frank

A função geradora da cópula de Frank é

$$\phi(u) = -\ln\left(\frac{e^{-\theta u} - 1}{e^{-\theta} - 1}\right), \quad (2.64)$$

assim, a forma fechada da cópula de Frank é dada por

$$C(u, v; \theta) = -\theta^{-1} \log\left(1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1}\right), \quad (2.65)$$

enquanto sua função densidade é dada por

$$c(u, v; \theta) = \frac{-\theta(e^{-\theta} - 1)(e^{-\theta(u+v)})}{((e^{-\theta u} - 1)(e^{-\theta v} - 1) + (e^{-\theta} - 1))^2}, \quad (2.66)$$

em que θ é o parâmetro de cópula que pode assumir qualquer valor real. Diferentemente da cópula de Clayton e de Gumbel, a cópula de Frank permite o alcance máximo de dependência. Isso significa que o parâmetro de dependência da cópula de Frank permite a aproximação dos limites superior e inferior de Fréchet e, dessa forma, ela permite a modelagem de dependência positiva e negativa. Apesar disso, a cópula de Frank não permite dependência de cauda nem superior nem inferior [23], logo, a Cópula de Frank é adequada para modelagem de dados que possuam baixa dependência de cauda.

A estimação do parâmetro θ também pode ser computada indiretamente a partir do τ de Kendall:

$$\tau_k = 1 + 4 \frac{D_1(\theta) - 1}{\theta}, \quad (2.67)$$

$$D_1(\theta) = \frac{1}{\theta} \int_0^\theta \frac{t}{e^t - 1} dt. \quad (2.68)$$

As distribuições condicionais bivariadas da cópula de Frank são dadas por:

$$C(v|u) = \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1) + (e^{-\theta v} - 1)}{(e^{-\theta u} - 1)(e^{-\theta v} - 1) + (e^{-\theta} - 1)}, \quad (2.69)$$

$$C(u|v) = \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1) + (e^{-\theta u} - 1)}{(e^{-\theta u} - 1)(e^{-\theta v} - 1) + (e^{-\theta} - 1)}. \quad (2.70)$$

2.6 Análise Walk-Foward

A análise *walk-foward* é um modo sistemático e formalizado de se implementar o que é referenciado como otimização periódica. O benefício central dessa análise é determinar o grau de confiança com o qual o implementador da estratégia consegue antecipar a viabilidade ou não da estratégia em tempo real quando comparado à performance em uma amostra histórica.

Uma outra vantagem dessa análise é a capacidade de produzir uma performance maximizada a medida que o mercado, suas tendências e sua volatilidade se modificam. Como a análise *walk-foward* provê uma otimização periódica, isso indica que ela consegue representar uma performance superior àquela provida pela otimização tradicional ou *back-testing* [25].

2.6.1 Otimização

Otimizar uma estratégia é fazer o uso mais efetivo dela. O processo de otimização realiza isso pelo exame e pela avaliação dos potenciais parâmetros utilizados na estratégia. Durante o processo de otimização, uma simulação histórica é calculada para um número suficiente de valores para que os parâmetros chave da estratégia possuam o maior impacto possível.

Há diversas formas de se descrever um processo de otimização: o teste em lotes, teste de execução, varredura de variáveis, processos de estimação, entre outros. A palavra otimização é utilizada como a seleção dos parâmetros mais robustos para a estratégia a ser executada. Além disso, otimizar é identificar e validar os parâmetros capazes de gerar a melhor performance da estratégia em tempo real.

Diferentes parâmetros para uma estratégia de trading podem levar a resultados completamente diferentes em termos de risco e retorno. Idealmente, a performance de uma estratégia robusta terá menor variância entre o seu conjunto de parâmetros. Na prática, porém, parâmetros diferentes levam a performances muito distintas, o que pode ser a diferença em sucesso e fracasso [26].

Framework de Otimização

O framework de um processo de otimização é relativamente simples e direto. Ele exige, porém, atenção aos detalhes. Segundo Pardo em [26], há quatro decisões-chave que devem ser feitas para que o framework possa ser montado:

1. Os parâmetros da estratégia utilizados na otimização devem ser selecionados e suas faixas de teste também;

2. Uma amostra de dados adequada deve ser selecionada;
3. A função objetivo ou as funções objetivo que identificam os parâmetros a serem otimizados devem ser selecionados;
4. Regras para a avaliação do processo de otimização devem ser estabelecidas.

Os Parâmetros: as estratégias de trading podem ter vários parâmetros que podem ser otimizados. Mesmo assim, é sempre melhor ter o menor número de parâmetros otimizáveis possível. Quanto maior o número de parâmetros a serem otimizados, maior é a possibilidade de os resultados sofrerem com *overfitting*² ou que a otimização seja feita de forma incorreta [26].

As Faixas de Teste: há dois princípios na determinação das faixas a serem incorporadas no framework de otimização: o princípio teórico e o prático. Primeiramente, consideram-se os princípios teóricos. As faixas de teste dos parâmetros devem ser apropriadas de forma intuitiva e razoável ao indicador, regra ou fórmula que está sendo utilizada. Além disso, considerando os princípios práticos, o tempo de processamento computacional é o principal fator a ser balanceado. Obviamente, um número menor de simulações históricas consome menos tempo do que um número maior e o número dessas simulações é igual ao produto dos candidatos em cada parâmetro otimizável no framework de otimização [26].

É difícil estabelecer uma regra rígida para a determinação da faixa de testes apropriada. É uma prática adequada, porém, manter as proporções constantes para os parâmetros a serem testados.

A Função Objetivo: a função objetivo tem um propósito único, que é selecionar os parâmetros mais robustos para execução da estratégia. Os parâmetros mais robustos não são, porém, muitas vezes, os mais rentáveis. Por definição, os parâmetros mais robustos são aqueles que são os que proporcionam uma maior chance de produzir resultados sustentáveis e confiáveis na execução da estratégia em tempo real [26].

Na prática, é recomendado que o estrategista selecione um conjunto de funções objetivo e trabalhe com elas em várias seções de otimização. Seria algo semelhante a uma otimização de funções objetivos. Isso é recomendado porque uma vez selecionada uma função objetiva capaz de identificar os parâmetros mais robustos, é improvável que ela seja mudada ao longo do tempo.

²*Overfitting* é um termo utilizado em aprendizado de máquinas que se translada bem para o processo de otimização de estratégias. Ele ocorre quando o modelo se adapta muito bem aos dados fora da amostra (aos dados com os quais está sendo treinado) porém, não generaliza bem para novos dados. Isto é, o modelo só se comporta bem para aquele conjunto de dados específicos e não tem boa aplicabilidade para conjunto de dados diferentes.

2.6.2 Backtesting

Em seu sentido mais estrito, o *backtesting* é uma simulação histórica da performance de uma estratégia em um período passado. Segundo Marcos Lopez de Prado em [25], o *backtesting* é uma das técnicas mais utilizadas e menos compreendidas no arsenal dos investidores quantitativos. Um engano comum é pensar no *backtesting* como uma ferramenta de pesquisa que aponta a robustez de uma estratégia.

Mesmo que um *backtesting* não apresente falhas no sentido de que seus resultados são reproduzíveis por qualquer um, ele provavelmente retorna resultados que não se aproximam da realidade. Isso porque simulações históricas são feitas diversas vezes em uma mesma amostra de dados, então, tem-se que levar em conta a probabilidade de ser uma descoberta falsa, ou seja, um acaso estatístico que inevitavelmente aparece depois que são rodados múltiplos testes em uma amostra de dados. Dessa forma, quanto melhor você se torna no *backtesting*, maior a probabilidade de que descobertas falsas irão emergir [25].

A principal função de um *backtesting* é o descarte de modelos ruins. O ajuste de um modelo em função dos resultados da simulação histórica é algo perigoso já que o *backtesting* provê muito pouco *insight* na razão pela qual uma estratégia é ou não robusta. A falta de *insight* está relacionada ao fato de que toda estratégia simulada historicamente está sujeita ao *overfitting* até certo ponto como resultado do viés de seleção, que advém da monetização de padrões históricos aleatórios. Como esses padrões provavelmente não se repetirão, a estratégia desenvolvida provavelmente falhará [25].

2.6.3 Walk-Foward

Quando uma estratégia de trading mostra que pode ser beneficiada por otimização, o próximo passo natural é passar a estratégia pelo último estágio no processo de desenvolvimento de estratégias: a análise de *walk-foward*. A análise de *walk-foward* julga a performance de um sistema de trading exclusivamente com base no período de pós-otimização ou fora da amostra.

Segundo Prado em [26], a performance de uma estratégia em dados que não foram parte do processo de otimização é uma medida muito mais confiável do que a performance baseada somente em simulação na amostra (*backtesting*). Se uma estratégia de trading apresenta uma boa performance na análise de *walk-foward*, ela mostra ser robusta e capaz de produzir bons resultados quando aplicada em tempo real. Além disso, a análise *walk-foward* é a simulação mais próxima da operação de uma estratégia otimizada em tempo real.

Teste de Walk-Foward

O teste de *walk-foward* é um processo de dois passos. O primeiro passo consiste em uma otimização tradicional. Um espaço de parâmetros é explorado para a otimização na amostra e um parâmetro ótimo ou um conjunto de parâmetros ótimos é selecionado pela função objetivo.

O segundo passo é o que o distingue e é a fonte de sua importância. Nesse passo, a performance dos parâmetros ótimos selecionados é avaliada em uma amostra de dados adjacente, fora da amostra de otimização. Em outras palavras, o modelo já otimizado é testado em uma simulação de trading de tempo real, é uma medida do resultado pós-otimização. Esse segundo passo é chamado duplo-cego [26].

Análise de Walk-Foward

A análise de *walk-foward* é um conjunto de testes individuais e sequenciais de *walk-foward* aplicados em uma amostra representativa de dados. Tipicamente, quanto maior a amostra, maior o número de testes realizados e, assim, maior a confiança estatística da análise realizada. Essa é a forma mais abrangente e exaustiva de se testar uma estratégia. É também o mais próximo que uma simulação chega de uma performance em tempo real. Uma análise de *walk-foward* completa, dessa forma, “anda para frente” nos dados, ou seja, tanto a janela de otimização quanto a janela de *walk-foward* se deslocam no tempo de forma a varrer o conjunto de dados históricos. O tamanho da janela determina o intervalo no qual as amostras são deslocadas no tempo [26].

3 Sistema de Trading

Esse capítulo descreve o sistema de trading desenvolvido e sua aplicação nos respectivos protótipos afim de gerar e avaliar os resultados da estratégia elaborada.

A aplicação do sistema desenvolvido em protótipos se baseia na ideia de que um protótipo levanta uma questão e obtém dados sobre alguma área de interesse. Dessa forma, os protótipos aqui desenvolvidos podem isolar os aspectos de um problema e projetar uma experiência de trading real de forma que seja permitido tentar e falhar logo, sem que se invista demais em um caminho sem ter informações suficientes [27].

A construção dos protótipos é baseada nas componentes principais enumeradas por Robert Pardo em *The Evaluation and Optimization of Trading Strategies* [26] e na adição do universo de operação como limitador.

1. Construção do universo de operação;
2. Entrada e saída de posições;
3. Gestão de risco;
4. Dimensionamento de posições.

O primeiro item, a construção do universo de operação, seleciona, dentre os 140 papeis disponíveis, os conjuntos que serão operados, ou seja, quais conjuntos apresentam as características desejadas para que se possa aplicar o método de abertura e fechamento de posições escolhido.

O segundo item - entrada e saída de posições - é o *driver* do trading, ou seja, é ele que indica a abertura e fechamento das posições, sendo a fundação da estratégia. Como uma estratégia compreende uma série de comandos de compra e venda de um ou mais ativos, é a entrada e saída de posições que dita o ritmo de toda a operação, enquanto os demais componentes se preocupam com a manutenção da mesma.

O terceiro item - gestão de risco - tem como objetivo a limitação das perdas para que seja possível que a estratégia permaneça operável depois de um período estendido de rendimentos negativos.

Finalmente, o dimensionamento de posições conduz o número de contratos ou papéis que serão comprometidos em cada operação. É fácil perceber o impacto desse componente em uma estratégia: se uma posição é muito pequena, o capital negociado não está sendo utilizado em nível ótimo, em contra partida, se uma posição é muito grande pode aumentar o risco de perda irreparável para o investidor.

O projeto em questão utiliza um processo de quatro etapas inspirado pelo desenvolvido por Ray Dalio em Princípios [28] para desenvolver o sistema de trading:

Definir objetivos claros: obter protótipos robustos em termo da relação risco e retorno, ou seja, que ofereçam índices de Sharpe superiores a 1.0;

Identificar os problemas que interferem nesses objetivos: identificar em que consiste o problema que impede que o índice de Sharpe supere o valor unitário (rentabilidade ou volatilidade);

Diagnosticar os problemas para atingir as causas raízes: a partir da identificação da localização do problema, fazer uma análise dos parâmetros envolvidos nesse componente para que se possa superá-lo;

Projetar planos para superá-los: após a identificação da causa raiz, desenvolver um plano para resolvê-lo ou descartar o protótipo.

Dessa forma, por mais que o sistema de trading seja desenvolvido nesse trabalho, outro ponto importante envolve a análise do que pode impedir esse sistema de alcançar o objetivo destacado, o que é melhor desenvolvido no Capítulo 4.

Para o Capítulo 3, o foco é a descrição dos itens relacionados às componentes da estratégia de trading e da realização dessas estratégias em termos de cinco protótipos desenvolvidos. Dessa forma, o Capítulo 3 consiste em quatro partes: construção do universo de operação, descrição dos sistemas de entrada e saída de posições, descrição do sistema de gestão de risco e do dimensionamento das posições e, finalmente, os elementos que compõem cada protótipo e seu universo de operação.

3.1 Construção do Universo de Operação

A construção do universo de operação tem por objetivo selecionar entre os ativos disponíveis aqueles que apresentem uma relação de longo prazo e uma propriedade de retorno à média. Isso garante que, apesar de a relação estar sujeita a perturbações causadas por choques de curto prazo, esse efeito tende a se dissipar de modo que a relação estabelecida volte ao equilíbrio de longo prazo.

O processo de construção desse universo utiliza dados históricos imediatamente anteriores ao dia de operação e é composto por duas partes:

1. Construção do modelo e estimação de parâmetros;
2. Teste de Cointegração.

Dessa forma, para que seja possível selecionar os pares e trios que constituirão o universo de operação de cada um dos cinco protótipos, para cada dia de operação, todas as combinações possíveis dos ativos envolvidos passam por um método de estimação de parâmetros, pela combinação desses ativos na forma de um modelo e pelo teste de cointegração. Caso seja confirmado que o par é cointegrado nos parâmetros estimados, ele pode fazer parte do universo diário para que sejam determinados os momentos de entrada e saída das posições.

3.1.1 Construção do Modelo e Estimação de Parâmetros

Para que o conjunto de papeis possa passar pelo teste de cointegração, previamente, é necessária a construção de um modelo que possa combinar os ativos de forma que seja possível obter um resíduo. Dessa forma, para o projeto em questão, foi determinado que a relação entre os ativos seria linear. Assim, a equação a seguir define o modelo que é utilizado para um conjunto de $N + 1$ ativos:

$$y_t = \beta_1 x_{1,t} + \dots + \beta_N x_{N,t} + \alpha + \epsilon_t. \quad (3.1)$$

Para que esse modelo seja computado, porém, é necessário que se utilize um método de estimação dos parâmetros envolvidos (β_i, α). O método de estimação de parâmetros do projeto é baseado em filtragem estocástica, de forma mais específica, o Filtro de Kalman é utilizado como estimador linear. Dessa forma, as equações abaixo compõem o algoritmo de estimação dos parâmetros.

Modelo de Estimação:

$$\Theta_{k+1} = [I] \Theta_k + w_k, \quad (3.2)$$

$$z_k = [x_{1,k} \quad \dots \quad x_{N,k} \quad 1] \Theta_k + v_k, \quad (3.3)$$

$$\Theta_k = [\beta_{1,k} \quad \dots \quad \beta_{N,k} \quad \alpha_k]^T, \quad (3.4)$$

$$Q = [I]_{N+1 \times N+1} 10^{-5}, \quad (3.5)$$

$$R = \begin{bmatrix} \sigma_y^2 & \sigma_{x_1}\sigma_y & \dots & \sigma_{x_N}\sigma_y \\ \sigma_y\sigma_{x_1} & \sigma_{x_1}^2 & \dots & \sigma_{x_N}\sigma_{x_1} \\ & & \dots & \\ \sigma_y\sigma_{x_N} & \sigma_{x_1}\sigma_{x_N} & \dots & \sigma_{x_N}^2 \end{bmatrix}_{N+1 \times N+1} . \quad (3.6)$$

A principal vantagem de se utilizar o Filtro de Kalman como estimador dos parâmetros do modelo de regressão linear consiste no fato de ele permitir que a relação entre as variáveis sofra choques ao longo do tempo, dessa forma, tem-se parâmetros flutuantes no tempo, como se pode ver na Figura 3.1.

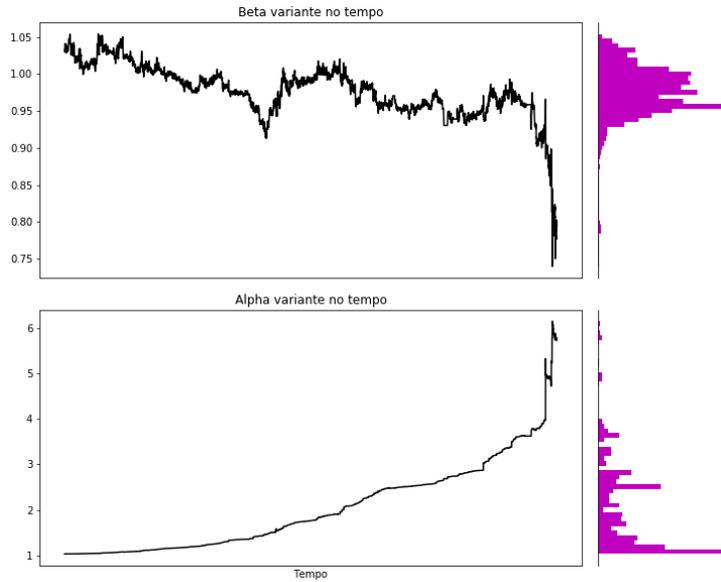


Figura 3.1: Parâmetros variantes no tempo.

Quando se permite que esses parâmetros mudem no tempo, tem-se dois benefícios de um modelo mais representativo da realidade do momento: um *hedging* mais eficiente e um resíduo de cointegração mais adequado, que pode prevenir a operação com pares cuja relação de cointegração esteja comprometida temporariamente.

3.1.2 Teste de Cointegração

Só depois que a estimação dos parâmetros foi feita é que se pode computar o termo de erro do modelo, também chamado de resíduo ou *spread*, para que se possa fazer o teste de cointegração.

O resíduo obtido para o modelo desenvolvido é dado pela equação abaixo:

$$\epsilon_t = y_t - (\beta_1 x_{1,t} + \dots + \beta_N x_{N,t} + \alpha). \quad (3.7)$$

É esse resíduo computado que é testado para garantir a relação de longo prazo e a propriedade de regressão à média do conjunto de ativos. O teste estatístico que garante a propriedade desejada é chamado de teste de raiz unitária e, dentre os testes de raiz unitária, escolheu-se o Teste de Dickey-Fuller Aumentado para estimar a estacionaridade do resíduo calculado.

Dessa forma, após a computação do resíduo, o Teste de Dickey-Fuller Aumentado é aplicado sob os seguintes parâmetros: regressão constante sem tendência e p-valor baseado nos resultados obtidos por MacKinnon em [29]. Como valor limite do p-valor, para garantir uma cointegração forte, utilizou-se 1%.

Assim, caso o conjunto de ativo testado retorne um p-valor menor que 1% no Teste de Dickey-Fuller Aumentado, esse conjunto (par ou trio) é considerado cointegrado e pode compor o universo de ativos operados. Finalmente, toda essa operação é feita diariamente, de forma a limitar os ativos cointegrados do dia para que se possa passar à fase de abertura e fechamento de posições.

É importante ressaltar, porém, que todas essas variáveis estão disponíveis na forma de uma base de dados que compõem as chamadas séries temporais financeiras de cada papel. Logo, o processo é feito com uma parcela da série temporal correspondente aos preços dos ativos em uma janela temporal, de tamanho a ser determinado, que se estende por uma quantidade de dias anteriores ao dia presente. Logo, a cointegração é verificada para essa janela de tempo somente, sendo então um processo que utiliza os preços anteriores ao dia de operação.

3.2 Entrada e Saída de Posições

Após a construção do universo de operação, tem-se o conjunto de ativos cointegrados para os quais serão verificadas as oportunidades de abertura e fechamento de posições para o dia.

Limitado o universo de ativos, são utilizados dois métodos capazes de identificar possibilidades de arbitragem:

1. Z-Score;
2. Cópuas de Arquimedes.

O primeiro - z-score - é um método de precificação relativa baseado na média histórica e na suposição de uma distribuição gaussiana do spread. Já o segundo método, utiliza um processo mais sofisticado para identificar se o *spread* está supervalorizado ou subvalorizado. Ambos os métodos, porém, compartilham o princípio normativo de investimentos de comprar um instrumento que esteja subvalorizado e vender um instrumento que esteja

supervalorizado. Esse princípio é ainda mais importante para esse projeto já que os conjuntos de ativos operados apresentam o comportamento de regressão à média, garantido pela relação de cointegração [8].

Para a operação com pares, caso um par esteja supervalorizado, a propriedade de reversão à média indica que a capitalização do investimento é positiva para uma posição vendida no par. Isso significa que para que se tenha um retorno positivo nesse caso, é necessário abrir uma posição *short* no par. Para o caso de um par subvalorizado, a propriedade de reversão à média indica que a capitalização é positiva para uma posição comprada no par, ou seja, é necessária uma operação *long* no par. A operação com trios é análoga.

Operacionalmente, o significado de *long* e *short* em um *spread* é um pouco delicada.

Long: a posição *long* em um conjunto de ativos relacionados pela Equação 3.1 é definida como:

1. *long* no ativo simbolizado por y_t ;
2. *short* nos ativos simbolizados por $x_{1,t}, \dots, x_{N,t}$.

Short: a posição *short* em um conjunto de ativos relacionados pela Equação 3.1 é operacionalizada como:

1. *long* nos ativos simbolizados por $x_{1,t}, \dots, x_{N,t}$;
2. *short* no ativo simbolizado por y_t .

Vale ressaltar, finalmente, que posições *long* e *short* possuem diferentes formas de operacionalização. Quando monta-se uma posição *long*, o valor do papel é descontado do caixa, enquanto quando se monta uma posição *short*, o valor do papel é adicionado ao caixa. Da mesma forma, quando liquida-se uma posição *long*, o valor do papel é adicionado ao caixa enquanto a liquidação de uma posição *short* implica a subtração do valor do papel ao caixa. Dessa forma, a capitalização de uma posição *long* é positiva quando um papel se valoriza, ou seja, quando o valor do papel no momento da liquidação é maior do que no momento da compra e, a capitalização de uma posição *short* é positiva quando o valor do papel no momento de liquidação é menor do que no momento da compra. Com a operacionalização descrita, é possível prosseguir para a descrição dos métodos utilizados para abertura e fechamento de posições.

3.2.1 Z-Score

O z-score utiliza o resíduo de cointegração para identificar oportunidades de arbitragem. Dessa forma, depois de estimados os parâmetros, o *spread histórico* para a janela temporal definida deve ser calculado pela Equação 3.7.

Duas métricas são calculadas para o *spread histórico*: a média e o desvio padrão, respectivamente, μ_h e σ_h . Computados esses valores, o *spread atual* é calculado para cada minuto com os parâmetros estimados pelo Filtro de Kalman. Os valores do *spread atual*, denominados ϵ_i , são então utilizados para obter o z-score:

$$z_{score} = \frac{\epsilon_i - \mu_h}{\sigma_h}. \quad (3.8)$$

A fórmula do z-score é uma maneira de se precificar o *spread atual* com relação à sua média histórica e colocar esse valor em termos de desvio-padrão. Para esse trabalho, definiu-se os valores de ± 2 desvios-padrão como níveis de confiança para abertura e fechamento de posições. Assim, um valor de ϵ_i acima de $+2$ é considerado supervalorizado e abaixo de -2 é considerado subvalorizado.

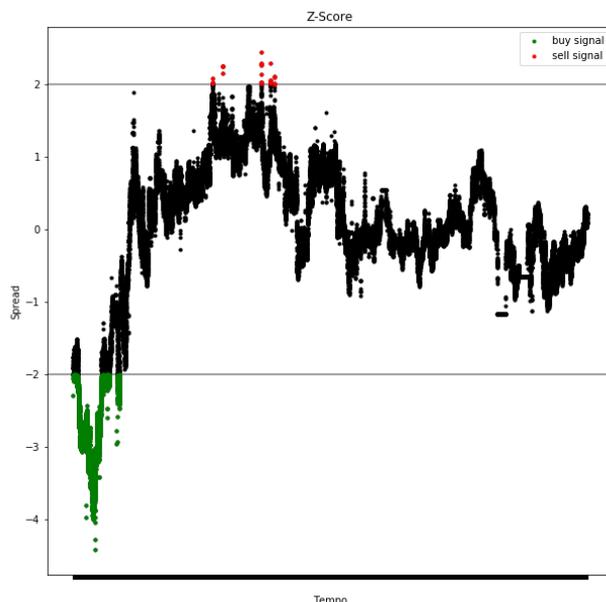


Figura 3.2: Sinalização de oportunidades de arbitragem: z-score.

Como é ilustrado na Figura 3.2, a propriedade de retorno a média reforça a oportunidade que se abre quando as bandas de confiança estabelecidas são rompidas. Esses valores de ± 2 foram escolhidos por representar 2 desvios-padrão, compreendendo mais

de 95% dos valores da distribuição gaussiana, além de ser um valor padrão para outras pesquisas nessa área [30].

3.2.2 Cópulas

Tradicionalmente, estratégias que envolvem pares, ou *pairs-trading*, dependem de alguma forma implícita da suposição de distribuições gaussianas. Uma abordagem diferente pode ser utilizada. Ao contrário do z-score, que, implicitamente, considera uma distribuição gaussiana do spread, o uso de cópulas bivariadas modela a relação entre um par de ativos de forma a capturar a estrutura de dependência entre os pares de ações e permitir que as características individuais de cada ação, como assimetria e distribuição de cauda-larga, sejam preservadas.

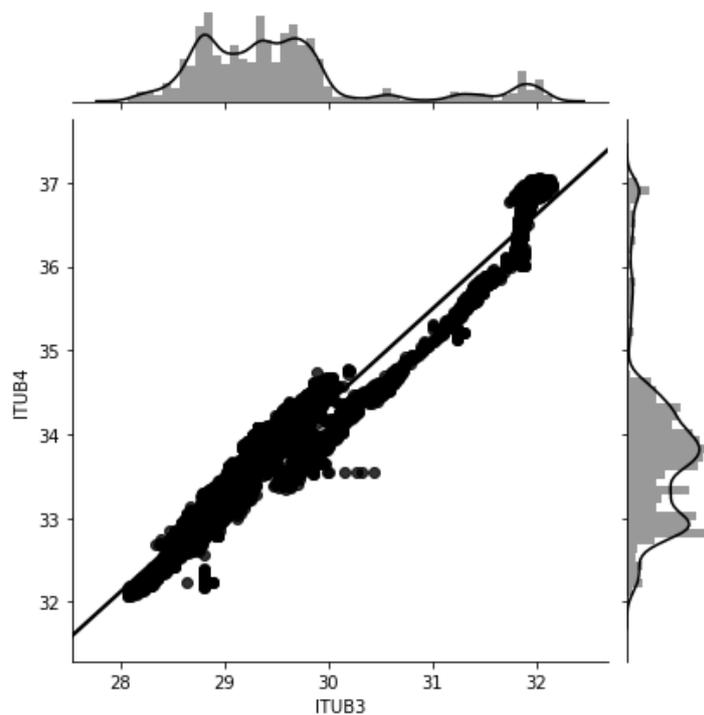


Figura 3.3: Exemplo de dependência com distribuição não gaussiana.

A importância de se permitir que distribuições assimétricas e de cauda-larga se manifestem decorre dos estudos variados que consideram as diferenças de comportamento das principais variáveis de mercado em mercados emergentes e desenvolvidos. Os mercados emergentes mostram mais choques extremos, ou eventos de cauda, do que seus pares desenvolvidos e o maior grau de volatilidade leva as variáveis-chave a apresentarem distribuições de cauda-larga [31].

As cópulas de Arquimedes são de interesse desse projeto por dois motivos: primeiro, elas possuem uma fórmula matemática fechada, o que facilita sua construção e implementação; segundo, os parâmetros importantes possuem uma propriedade interessante e podem ser estimados de forma indireta e mais simples que os parâmetros das outras cópulas.

Para que se possa utilizar as cópulas como sinalizadoras de entrada e saída de posição, passa-se por dois passos:

Fit de cópulas: estabelece qual cópula melhor descreve a estrutura de dependência dos papéis que formam o par a ser operado.

Sinalização de oportunidade: deriva as bandas de confiança para a função de distribuição condicional de V dado U , denotado por $C(V|U = u)$, assim como para a função de distribuição condicional de U dado V , denotado por $C(U|V = v)$.

O primeiro item, fit de cópulas, segue um algoritmo:

1. Sendo P_x e P_y as séries de preços históricos para os papéis X e Y , o logaritmo dos retornos para ambas as ações são derivados e computados em R_x e R_y ;
2. Para que se possa construir as cópulas, R_x e R_y devem ser transformados em variáveis uniformemente distribuídas U e V . Uma função de distribuição empírica é utilizada para fazer essa transformação;
3. O parâmetro θ é estimado a partir da relação entre a cópula e o coeficiente de correlação dado pelo τ de Kendall, para cada uma das três cópulas;
4. Com a estimativa de θ e a formação das cópulas, é possível selecionar a cópula com o melhor ajuste.

Como não há consenso sobre um critério estatístico que selecione a cópula que fornece o melhor ajuste para os dados, após o *fit* de cópulas para as variáveis uniformes U e V , utiliza-se o critério conhecido por AIC (Akaike Information Criterion) como utilizado na tese *High-Frequency copula-based pairs trading on U.S. Goldmine Stocks* [32]:

$$AIC = -2 \times L(\theta) + 2k, \quad (3.9)$$

em que $k = 1$ é o número de parâmetros no modelo de cópulas e

$$L(\theta) = \sum_{t=1}^T \log[c(u_t, v_t; \theta)]. \quad (3.10)$$

Computados esses valores para as cópulas de Gumbel, Frank e Clayton, a cópula que provê o melhor ajuste é aquela com o menor valor dado pela fórmula do AIC [32].

Depois que o *fit* de cópulas é feito, a sinalização de posições para os pares de ações pode ser realizada. Dessa forma, sendo X e Y duas variáveis aleatórias com função de probabilidade cumulativa dada por $F_1(X)$ e $F_2(Y)$ e sendo $U = F_1(X)$ e $V = F_2(Y)$, então a função de cópula será computada por

$$C(u, v) = P(U < u, V < v). \quad (3.11)$$

As derivadas parciais da função de cópula com relação a u e v dão a função de distribuição condicional das variáveis:

$$P(U < u|V = v) = \frac{\partial C(u, v)}{\partial v}, \quad (3.12)$$

$$P(V < v|U = u) = \frac{\partial C(u, v)}{\partial u}. \quad (3.13)$$

Essa estratégia de trading com cópulas utiliza a cópula de melhor ajuste como meio de se derivar bandas de confiança para funções de distribuição condicional, isto é, índices de precificação relativa. Dessa forma, quando as observações se encontram fora dessa banda de confiança, há uma indicação de uma oportunidade de compra ou venda do par. Nesse trabalho, foram escolhidos os valores 95% e 5% como bandas de confiança superior e inferior respectivamente.

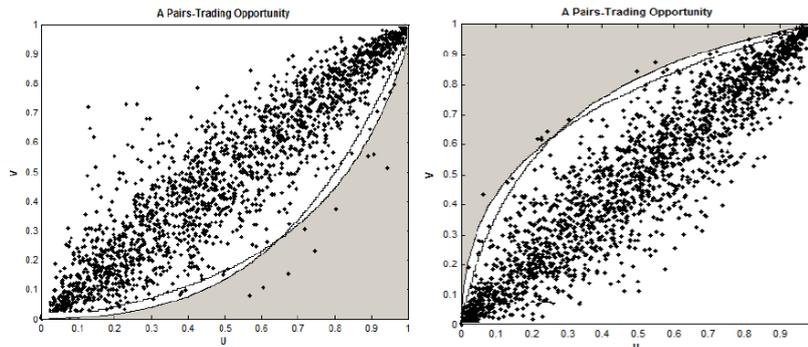


Figura 3.4: Ilustração de bandas de confiança: long (esquerda) e short (direita) (Fonte: [31]).

Assim, dadas as variáveis aleatórias X e Y , os índices de precificação relativas são dados por:

$$M_{X|Y} = P(U < u|V = v) = \frac{\partial C(u, v)}{\partial v}, \quad (3.14)$$

$$M_{Y|X} = P(V < v|U = u) = \frac{\partial C(u, v)}{\partial u}. \quad (3.15)$$

Dessa forma, para cada dia, os retornos logarítmicos das ações x_t e y_t são convertidos em U e V utilizando funções de distribuições empíricas. Depois disso, os índices de precificação relativa são calculados utilizando a cópula que apresente o melhor ajuste estimado [31]. Quanto se tem $MI_{Y|X} < 5\%$ e $MI_{X|Y} > 95\%$, para a relação construída pela Equação 3.1, o papel denotado por y_t está subvalorizado enquanto o papel x_t está supervalorizado, assim, uma posição *long* é construída no par. Analogamente, quando se tem $MI_{Y|X} > 95\%$ e $MI_{X|Y} < 5\%$, o papel y_t se encontra supervalorizado enquanto o ativo x_t se encontra subvalorizado, logo, uma posição *short* é construída no par.

3.3 Gestão de Risco e Dimensionamento de Posições

O processo de gestão de risco passa por duas partes: primeiro pela diversificação e alocação de capital e depois pelo limite de acúmulo de posições. O processo de gestão de risco foi simplificado ao máximo para que pudesse ser feita a análise dos protótipos sem ancoragem em nenhum valor específico e desembasado.

Dessa forma, como o universo de pares cointegrados é calculado diariamente, o limite de acúmulo de posições impõe que sejam no acumuladas no máximo 479 posições diárias, uma por minuto disponível na base de dados diária, exceto o minuto final, que é guardado para uma eventual liquidação de todas as posições. Além disso, o capital disponível para investimento é dividido igualmente entre os pares cointegrados.

Com relação ao dimensionamento de posições, esse processo envolve a quantificação do risco de cada ativo que compõe o par ou trio operado e o *hedging*. O risco é quantificado por um método denominado *expected shortfall* ou *conditional value at risk*, uma extensão do método tradicional VaR.

O entendimento da quantificação de risco por *expected shortfall* é mais intuitivo depois de se compreender o VaR. O *value at risk* é um modo de se estimar a maior perda possível em um portfólio dada uma janela de tempo e um certo nível de confiança. Há duas formas de calculá-lo: uma paramétrica e uma histórica (ou não-paramétrica) [33].

Paramétrico: quando se assume que os retornos de determinado ativo seguem uma certa distribuição, diz-se que o VaR é paramétrico. Normalmente, a distribuição que se assume para o VaR Paramétrico é a gaussiana. A premissa de que os retornos possuem distribuição gaussiana permite que a volatilidade do ativo possa ser descrita em termos do seu desvio-padrão. Dessa forma, o VaR é computado por:

$$VaR = \sigma \times P, \quad (3.16)$$

em que P é o tamanho da posição.

Histórico: quando se observa a distribuição de retornos passados dos ativos para se estimar o VaR, diz-se que ele é uma medida não paramétrica ou histórica. O VaR histórico utiliza um histograma do retorno dos ativos e calcula um percentil k para estimar uma perda esperada com certo nível de confiança, assumindo que os retornos futuros são próximos aos retornos passados. Dessa forma, o cálculo do VaR Histórico com cobertura γ é feito com o percentil $100(1 - \gamma)$ dos menores retornos observados e seu produto com o tamanho da posição e sua computação é a mesma da Equação 3.16, só que com o percentil calculado substituindo o desvio-padrão.

Experiências passadas sugerem que o método do *value at risk* é uma forma de gestão de risco pouco eficiente. Os motivos por trás disso são diversos, por exemplo: os retornos das ações não possuem forma de curva gaussiana; em eventos de cauda, a correlação entre ativos muda e muitos ativos antes não correlacionados, apresentam correlação próxima do valor unitário. O maior exemplo da falha em só se basear no VaR foi a crise gerada pelo desempenho do fundo Long Term Capital Management. Segundo a gestão de risco do fundo (baseada no VaR), o valor máximo a ser perdido em um dia com 95% de confiança eram 35 milhões de dólares. Mesmo assim, o fundo acumulou perdas superiores a 100 milhões por vários dias seguidos e perdeu todo seu capital, tendo de ser socorrido por um consórcio, um evento que teria probabilidade de 1 em 10^{24} de ocorrer [4].

O CVaR responde à pergunta: “Qual a perda média sofrida pela carteira nos α piores cenários históricos?”. Matematicamente, ele considera a forma do histograma dos retornos, pesando assim parâmetros específicos de distribuição como *skewness* e *kurtosis* de forma a calcular a medida de risco [34]:

$$CVaR_{\alpha}(x) = \int_{f(x,y) \geq VaR} f(x,y)p(y)dx dy. \quad (3.17)$$

Dessa forma, pode-se utilizar o *conditional value at risk* para calcular a perda média sofrida pela carteira formada pelo par ou pelo trio de ativos nos α piores cenários históricos. Isso permite que se estabeleça uma relação entre o risco que se está disposto a tomar em cada posição aberta e o risco da carteira. A razão entre o risco em que se está disposto a tomar em cada posição e o CVaR percentual resulta no número de ações totais que podem compor o portfólio.

O *hedge* é uma estratégia que visa proteger o investidor de perdas financeiras. É esse fator que torna a arbitragem estatística uma estratégia neutra de mercado. Sendo uma estratégia neutra de mercado, o *hedge* consiste em balancear a posição *long* com a posição *short* construída. Esse balanceamento pode ser feito de duas formas: as quantidades balanceadas podem ser o número de ações ou a quantidade de dinheiro envolvida. Nesse

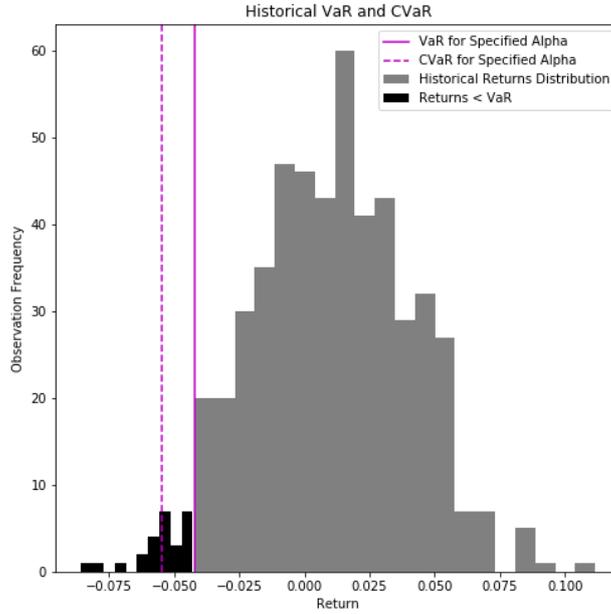


Figura 3.5: Ilustração da comparação entre $CVaR$ e VaR .

trabalho, o número de ações foi escolhido como o parâmetro a ser balanceado pelo *hedge*, justamente porque são os parâmetros realmente operados. Assim, o *hedge* consiste em balancear a posição *long* e *short* com o fator β estimado pelo filtro de Kalman.

Dessa forma, como indicado pela Equação 3.1, para cada ação y_t comprada (ou vendida), vende-se (ou compra-se) β_1 ações $x_{1,t}$, β_2 ações $x_{2,t}$, ..., β_N ações $x_{N,t}$. Esse processo é responsável por eliminar a exposição ao risco direcional do mercado, tal que o retorno obtido não deve apresentar alta correlação com o índice de referência de mercado [30].

Assim, com o número total de ações dado pelo $CVaR$ e a relação entre o número de ações dado pelo *hedge*, calcula-se o tamanho de cada posição em número de ações compradas e vendidas. Vale ressaltar que o número escolhido como parâmetro de risco tomado em cada posição foi de 1%, logo, o número total de ações compradas é dado por:

$$n = \frac{1\%}{CVaR_{95\%}(\Omega)}, \quad (3.18)$$

em que Ω é o portfólio de ações contendo $y_t, x_{1,t}, \dots, x_{N,t}$. E a quantidade de cada papel operado é dado por:

$$\begin{bmatrix} n_y & n_{x_1} & \dots & n_{x_N} \end{bmatrix} = \frac{\begin{bmatrix} 1 & \beta_1 & \dots & \beta_N \end{bmatrix}}{1 + \beta_1 + \dots + \beta_N} \times n. \quad (3.19)$$

Há uma ressalva porém, mesmo sendo um avanço com relação ao VaR, o CVaR ainda baseia seu método de quantificação de risco em medidas de correlação. Isso se mostra algo não muito confiável já que essa é uma variável cujo comportamento se inverte em tempos de crise [35].

3.4 Construção dos Protótipos

A construção de todos os protótipos passa por uma fase comum: a determinação da janela temporal de cointegração para a formação do universo de ativos operados e cálculo dos fatores históricos necessários para os métodos de entrada e saída de posições.

A determinação da janela temporal é feita a partir do método de otimização *walk-forward*. Segundo o *framework* definido por Pardo em *The Evaluation and Optimization of Trading Strategies* [26], há três fatores que precisam ser determinados para que o processo de otimização seja eficiente.

Determinação do parâmetro a ser otimizado: nesse caso, o parâmetro-chave a ser otimizado é a janela temporal;

Amostra de dados selecionada: como a janela temporal é o parâmetro a ser otimizado, a amostra de dados é automaticamente selecionada e, por limitações de capacidade de processamento, a amostra de dados foi tomada de 30 em 30 minutos ao invés de minuto a minuto;

Determinação da função-objetivo a ser otimizada: como a amostra de dados considera a precificação dos ativos de 30 em 30 minutos, a função-objetivo a ser otimizada foi o número de pares cointegrados no período e, caso houvessem dois períodos com o mesmo número de pares cointegrados, utilizou-se o rendimento do último dia de trading como fator de desempate.

A janela temporal é determinada pela otimização segundo o *framework* para cada uma das dez semanas de operação. Assim, dentre um espaço de parâmetros pré-determinado (30, 60, 90, 120, 150 e 200 dias), é realizada a otimização tradicional e selecionada a janela temporal ótima a partir da função-objetivo determinada. Essa seleção, porém, ocorre em uma amostra de dados adjacente, imediatamente anterior à semana de operação, e a janela temporal selecionada é aplicada fora da amostra de otimização. Em outras palavras, o modelo já otimizado é testado em uma simulação de trading pós-otimização, chamado duplo-cego, o que caracteriza uma análise *walk-forward*.

A intenção do trabalho era selecionar os pares cointegrados para cada dia utilizando a janela temporal selecionada para a semana dentre os 9730 pares possíveis. Contudo, a

limitação da capacidade de processamento computacional tornou essa intenção impraticável. Logo, houve uma pré-seleção dos pares a serem testados diariamente para a formação do universo de ativos. Essa pré-seleção resultou em três universos pré-selecionados distintos: um composto por seis pares, um com 95 pares e outro com 61 trios.

A seleção do primeiro conjunto de pares pré-selecionado se baseia na característica intrínseca fundamental da ação ordinária e ação preferencial, que é o fato de elas serem papéis relacionados à mesma empresa, dessa forma, espera-se uma cointegração natural desses ativos. Dentre os 140 ativos disponíveis, há seis pares formados por ações ON e PN.

O segundo conjunto, com 95 pares, foi selecionado de forma empírica a partir dos preços de fechamento diário dos 140 ativos. Essa seleção utilizou os preços de fechamento dos ativos para a estimação de parâmetros do filtro de Kalman e testes de cointegração foram realizados para todas as combinações possíveis para nove períodos diferentes, que compreenderam períodos imediatamente anteriores ao início da data de trading. Isso resultou em 462 pares possíveis, dos quais se selecionaram os 95 com menor tempo de meia-vida (até 50 dias).

Para o último conjunto, o mesmo procedimento foi realizado, mas o teste de cointegração só foi realizado para dois períodos imediatamente anteriores ao início do trading, o que resultou em 638 trios possíveis, dos quais se selecionaram os 61 com menor tempo de meia-vida (até 20 dias).

A estimação do tempo de meia-vida para a pré-seleção dos universos utilizou o modelo de Ornstein-Uhlenbeck [36]. Seja z_t o spread entre um conjunto de ações, com propriedade de reversão à média, então, é possível modelá-lo como [37]:

$$dz_t = \theta(\mu - z_t)dt + \sigma \Delta W_t + \epsilon_t, \quad (3.20)$$

em que μ é a média de longo prazo do spread entre os preços, z_t é o valor do spread em um dado ponto do tempo t , θ a força de reversão à média, σ o desvio padrão, W_t o processo de Wiener e ϵ_t o resíduo da regressão. Matematicamente, se diz que o valor médio de z_t segue um decaimento exponencial para sua média μ , e a meia-vida do decaimento exponencial é igual a $\ln(2)/\theta$, que é tempo esperado que o spread leva para reverter a metade do desvio inicial em relação à média [30].

3.4.1 Protótipo 1

O protótipo 1 utiliza o teste de cointegração diariamente para selecionar os pares operáveis dentro do universo constituído pelos 95 pares pré-selecionados.

Assim, para cada dia de trading, os parâmetros α e β da Equação 3.1 são estimados pelo filtro de Kalman e o teste de cointegração seleciona o universo de pares operáveis dentre os 95 pré-selecionados. O protótipo 1 utiliza o z-score como método de entrada e saída de posições e utiliza o β calculado para o dia como fator de *hedging*.

Para cada dia de trading, o caixa total é dividido igualmente entre os pares cointegrados, independente da sinalização do z-score. Para os pares que apresentarem oportunidades de arbitragem, elas são realizadas e para os que não, o investimento não apresenta rentabilidade para aquele dia. Dessa forma, o primeiro dia de trading supõe um caixa total de 20 mil reais, que é atualizado ao final do dia a depender das rentabilidades de cada par. Para o próximo dia, o caixa é novamente dividido igualmente entre os pares cointegrados e assim por diante.

A rentabilidade de cada dia do trading é determinada a partir do somatório das variações dos sub-caixas de cada par do início ao final do dia. Dessa forma, tem-se para cada par um sub-caixa, e enquanto o *spread* está abaixo de -2 desvios-padrão (acima de +2 desvios-padrão), monta-se uma posição de *long* no par (*short* no par) por minuto. Assim, para cada posição *long* montada (*short*), o sub-caixa é reduzido (aumentado) no valor do *spread*. Quando o *spread* retorna para a banda de confiança estabelecida, o total de papéis operados são liquidados nos valores daquele minuto e tem-se um novo valor para o sub-caixa. As variações do sub-caixa de cada par compõem a variação total do caixa e a rentabilidade do dia.

3.4.2 Protótipo 2

O protótipo 2 utiliza o teste de cointegração diariamente para selecionar os pares operáveis dentre o universo de seis pares compostos por ações ON e PN.

Assim, para cada dia de trading, os parâmetros α e β da Equação 3.1 são estimados pelo filtro de Kalman e o teste de cointegração seleciona o universo de pares operáveis dentre os seis pré-selecionados. O protótipo 2 utiliza o z-score como método de entrada e saída de posições e também utiliza o β calculado para o dia como fator de *hedging*.

Para cada dia de trading, o caixa total é dividido igualmente entre os pares cointegrados, independente da sinalização do z-score. Para os pares que apresentarem oportunidades de arbitragem, elas são realizadas e para os que não, o investimento não apresenta rentabilidade para aquele dia. Dessa forma, o primeiro dia de trading supõe um caixa total de 20 mil reais, que é atualizado ao final do dia a depender das rentabilidades de cada par. Para o próximo dia, o caixa é novamente dividido igualmente entre os pares cointegrados e assim por diante.

A rentabilidade de cada dia do trading é determinada a partir do somatório das variações dos sub-caixas de cada par do início ao final do dia. Dessa forma, tem-se para cada

par um sub-caixa, e enquanto o *spread* está abaixo de -2 desvios-padrão (acima de +2 desvios-padrão), monta-se uma posição de *long* no par (*short* no par) por minuto. Assim, para cada posição *long* montada (*short*), o sub-caixa é reduzido (aumentado) no valor do *spread*. Quando o *spread* retorna para a banda de confiança estabelecida, o total de papéis operados são liquidados nos valores daquele minuto e tem-se um novo valor para o sub-caixa. As variações do sub-caixa de cada par compõem a variação total do caixa e a rentabilidade do dia.

3.4.3 Protótipo 3

O protótipo 3 utiliza o teste de cointegração diariamente para selcionar os pares operáveis dentro do universo constituído pelos 95 pares pré-selecionados.

Assim, para cada dia de trading, os parâmetros α e β da Equação 3.1 são estimados pelo filtro de Kalman e o teste de cointegração seleciona o universo de pares operáveis dentre os 95 pré-selecionados. O protótipo 3 utiliza as cópulas de Arquimedes como método de entrada e saída de posições e utiliza o β calculado para o dia como fator de *hedging*.

Para cada dia de trading, o caixa total é dividido igualmente entre os pares cointegrados, independente da sinalização das cópulas. Para os pares que apresentarem oportunidades de arbitragem, elas são realizadas e para os que não, o investimento não apresenta rentabilidade para aquele dia. Dessa forma, o primeiro dia de trading supõe um caixa total de 20 mil reais, que é atualizado ao final do dia a depender das rentabilidades de cada par. Para o próximo dia, o caixa é novamente dividido igualmente entre os pares cointegrados e assim por diante.

A rentabilidade de cada dia do trading é determinada a partir do somatório das variações dos sub-caixas de cada par do início ao final do dia. Dessa forma, tem-se para cada par um sub-caixa, e enquanto o par é identificado como subvalorizado (supervalorizado), monta-se uma posição de *long* no par (*short* no par) por minuto. Assim, para cada posição *long* montada (*short*), o sub-caixa é reduzido (aumentado) no valor do *spread*. Quando o *spread* não é mais identificado como distorcido, o total de papéis operados são liquidados nos valores daquele minuto e tem-se um novo valor para o sub-caixa. As variações do sub-caixa de cada par compõem a variação total do caixa e a rentabilidade do dia.

3.4.4 Protótipo 4

O protótipo 4 utiliza o teste de cointegração diariamente para selecionar os pares operáveis dentre o universo de seis pares compostos por ações ON e PN.

Assim, para cada dia de trading, os parâmetros α e β da Equação 3.1 são estimados pelo filtro de Kalman e o teste de cointegração seleciona o universo de pares operáveis dentre os seis pré-selecionados. O protótipo 4 utiliza as cópulas de Arquimedes como método de entrada e saída de posições e também utiliza o β calculado para o dia como fator de *hedging*.

Para cada dia de trading, o caixa total é dividido igualmente entre os pares cointegrados, independente da sinalização das cópulas. Para os pares que apresentarem oportunidades de arbitragem, elas são realizadas e para os que não, o investimento não apresenta rentabilidade para aquele dia. Dessa forma, o primeiro dia de trading supõe um caixa total de 20 mil reais, que é atualizado ao final do dia a depender das rentabilidades de cada par. Para o próximo dia, o caixa é novamente dividido igualmente entre os pares cointegrados e assim por diante.

A rentabilidade de cada dia do trading é determinada a partir do somatório das variações dos sub-caixas de cada par do início ao final do dia. Dessa forma, tem-se para cada par um sub-caixa, e enquanto o par é identificado como subvalorizado (supervalorizado), monta-se uma posição de *long* no par (*short* no par) por minuto. Assim, para cada posição *long* montada (*short*), o sub-caixa é reduzido (aumentado) no valor do *spread*. Quando o *spread* não é mais identificado como distorcido, o total de papéis operados são liquidados nos valores daquele minuto e tem-se um novo valor para o sub-caixa. As variações do sub-caixa de cada par compõem a variação total do caixa e a rentabilidade do dia.

3.4.5 Protótipo 5

Finalmente, o protótipo 5 utiliza o teste de cointegração diariamente para selecionar os trios operáveis dentro do universo de 61 trios pré-selecionados.

Assim, para cada dia de trading, os parâmetros α , β_1 e β_2 da Equação 3.1 são estimados pelo filtro de Kalman e o teste de cointegração seleciona o universo de trios operáveis dentre os 61 pré-selecionados. O protótipo 5 utiliza o z-score como método de entrada e saída de posições e utiliza β_1 e β_2 como fatores de *hedging*.

Para cada dia de trading, o caixa total é dividido igualmente entre os trios cointegrados, independente da sinalização do z-score. Para os trios que apresentarem oportunidades de arbitragem, elas são realizadas e para os que não, o investimento não apresenta rentabilidade para aquele dia. Dessa forma, o primeiro dia de trading supõe um caixa total de 20 mil reais, que é atualizado ao final do dia a depender das rentabilidades de cada trio. Para o próximo dia, o caixa é novamente dividido igualmente entre os trios cointegrados e assim por diante.

A rentabilidade de cada dia do trading é determinada a partir do somatório das variações dos sub-caixas de cada trio do início ao final do dia. Dessa forma, tem-se para cada

trio um sub-caixa, e enquanto o *spread* está abaixo de -2 desvios-padrão (acima de +2 desvios-padrão), monta-se uma posição de *long* no trio (*short* no trio) por minuto. Assim, para cada posição *long* montada (*short*), o sub-caixa é reduzido (aumentado) no valor do *spread*. Quando o *spread* retorna para a banda de confiança estabelecida, o total de papéis operados são liquidados nos valores daquele minuto e tem-se um novo valor para o sub-caixa. As variações do sub-caixa de cada trio compõem a variação total do caixa e a rentabilidade do dia.

4 Resultados e Análises

Este capítulo visa apresentar os resultados obtidos a partir da implementação das estratégias de trading e realizar as respectivas análises de forma a responder dois questionamentos principais para cada protótipo:

1. A estratégia aplicada no protótipo é robusta?
2. Há melhoras a serem implementadas no protótipo? Quais?

Além disso, esse capítulo também inclui a comparação da performance das estratégias com um benchmark a fim de obter uma visão holística em termos de risco e retorno.

A avaliação da robustez da estratégia aplicada para cada protótipo passa por três pontos:

1. Rentabilidade diária e acumulada;
2. Volatilidade;
3. Índice de Sharpe.

Além desses parâmetros, a avaliação dos pontos de melhora de cada protótipo envolvem a análise de mais cinco itens:

1. Volume médio de operações diárias;
2. Perda máxima possível diária de pares vencedores;
3. Correlação entre tempo de meia-vida e duração do trading;
4. Percentual de pares vencedores;
5. Percentual de dias vencedores.

Vale adicionar um racional por trás de alguns dos itens enumerados. A perda máxima possível diária é uma métrica que calcula, para os pares vencedores, qual a perda máxima possível de um par em um dia, caso sua posição fosse encerrada em um período diferente

do que foi estabelecido. O percentual de pares vencedores é um indicativo da frequência do acerto da estratégia com relação às escolhas de abertura e fechamento de posições no dia. Por fim, o percentual de dias vencedores é um indicativo da relação entre o tamanho das posições de ganho e perda abertas em um dia, que resultam em dias de ganhos maiores que as perdas.

Esses dados, de certa forma complementares, são indicativos de uma série de características importantes da estratégia. O percentual de pares e dias vencedores são indicativos tanto da performance da estratégia no longo prazo quanto do balanceamento do tamanho das perdas e ganhos. Além disso, o volume de operações diárias indica se a estratégia baseia sua rentabilidade em volume ou em tamanho de posições. Já a perda máxima possível diária de pares vencedores pode funcionar com uma referência para um sistema de *stop-loss* caso seja verificado que há um desbalanceamento entre tamanho de ganhos e perdas. Finalmente, a correlação entre tempo de meia-vida e rentabilidade pode dar mais uma métrica preditiva para ser utilizada em aberturas de posição.

Dessa forma, são desenvolvidas cinco seções para avaliação individual de cada estratégia e uma sexta seção para comparação das performances dos protótipos com o benchmark. Vale também ressaltar que o benchmark escolhido para comparação das estratégias foi o papel BOVA11.SA, representativo da performance das principais empresas da Bovespa. Finalmente, a taxa livre de risco utilizada para computação do índice de Sharpe¹ é a SELIC.

¹O índice de Sharpe (IS) desse trabalho é calculado pela equação:

$$IS = \frac{R_i - 0,007996}{\sigma_i - 4.806 \times 10^{-6}}. \quad (4.1)$$

4.1 Protótipo 1: Pairs Trading - Z-Score

O primeiro protótipo é a aplicação da estratégia de pairs trading envolvendo cointegração e z-score como método de entrada e saída de posições.

Os resultados diário e acumulado para a estratégia podem ser vistos na Figura 4.1:

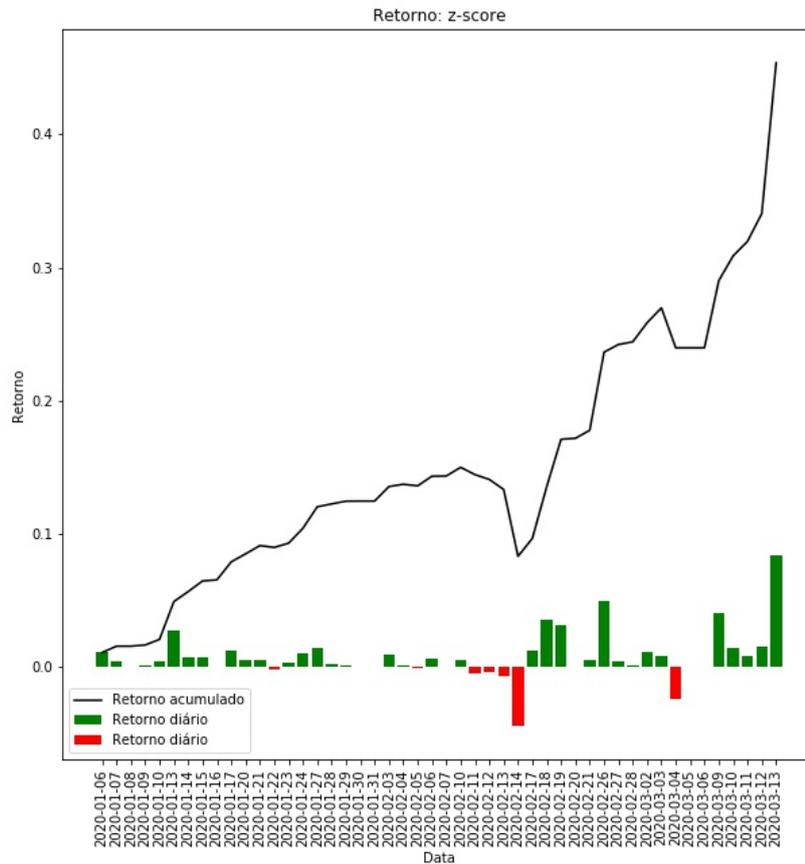


Figura 4.1: Rentabilidade: Protótipo 1.

A Tabela 4.1 também resume os parâmetros restantes para que se possa avaliar a robustez da estratégia:

	Resultados
Rentabilidade Acumulada	45,46%
Rentabilidade Média Diária	0,78%
Volatilidade	1,83%
Índice de Sharpe	24,36

Tabela 4.1: Resultados: Protótipo 1

Para que se tenha uma visão mais completa da estratégia em termos de suas demais características, a Tabela 4.2 apresenta os dados relativos ao volume médio das operações diárias, da perda máxima possível, e dos percentuais vencedores.

	Resultados
Volume de operações diária	227 ± 183
Perda máxima possível diária de um par vencedor	39,43%
Percentual vencedor (Pares)	86,32%
Percentual vencedor (Dias)	81,25%

Tabela 4.2: Resultados complementares: Protótipo 1

Com base tanto nos resultados mostrados na Figura 4.1 quanto na Tabela 4.1, pode-se ver que a rentabilidade da estratégia é elevada. Essa rentabilidade, porém, não vem acompanhada de um indicativo de risco muito elevado, vide a volatilidade, que não alta. Sabe-se, nesse caso, que a relação de risco e retorno da estratégia é extremamente benéfica pelo índice de Sharpe. Índices de Sharpe maiores que o valor unitário são considerados bons, e no caso desse protótipo, o valor unitário é superado em muitas vezes.

Dessa forma, pode-se concluir que a estratégia, nos parâmetros aplicados ao protótipo, é robusta em termos de risco e retorno ao investidor.

A Tabela 4.2 nos dá uma visão mais completa da estratégia. Por meio dela, pode-se inferir que a rentabilidade está associada a um volume grande de posições estabelecidas em um dia. Além disso, é visível que o percentual vencedor relativo aos pares é algo muito favorável, ou seja, o percentual de acerto na abertura de posições é alto, indicando uma rentabilidade robusta no longo prazo também. O percentual de dias vencedores também dá uma visão favorável da estratégia. Por ser um valor elevado e próximo ao percentual de pares vencedores, vê-se que há um balanceamento entre o tamanho das perdas e dos ganhos. Isso é um bom indicativo já que, por mais que se tenha um grande percentual de pares vencedores, na presença de um desbalanceamento, o pequeno número de perdas pode superar em tamanho o alto número de ganhos. A presença desse balanceamento entre ganhos e perdas associada a um alto valor de máxima perda possível de um par vencedor aponta para a inadequação da aplicação de um sistema de *stop-loss* a essa estratégia.

Para o protótipo 1, faz-se necessário um comentário sobre liquidez de ativos. A liquidez é um indicativo da rapidez com que um ativo pode ser revertido em dinheiro. Como não foi feito qualquer tipo de consideração acerca de liquidez para a seleção dos pares operados e a estratégia usa o volume de operações como base de sua rentabilidade, a presença de ativos de baixa liquidez na simulação pode ser uma prerrogativa inadequada na operação em tempo real. Dessa forma, vale indicar a rentabilidade da estratégia que está associada a ativos de pouca liquidez.

Os principais resultados, considerando o aspecto da liquidez seriam dados pela Tabela 4.3.

	Resultados
Rentabilidade Acumulada	32,42%
Rentabilidade Média Diária	0,59%
Volatilidade	1,33%
Índice de sharpe	23,76

Tabela 4.3: Resultados considerando liquidez: Protótipo 1

A análise dos resultados presentes na Tabela 4.3 permite concluir que, mesmo considerando a limitação da liquidez de certos ativos e mesmo que a rentabilidade acumulada da estratégia caia, a relação entre risco e retorno não é tão afetada. Assim, conclui-se que a estratégia nos parâmetros aplicados ao protótipo 1 se mantém robusta apesar da queda de 28,7% da rentabilidade, que é compensada por uma queda de 27,3% da volatilidade.

Finalmente, três métodos foram utilizados afim de calcular a correlação entre o tempo de meia-vida para cada par cointegrado diário e a duração do trading. O esperado seria que se detectasse uma correlação positiva, indicativa de que quanto menor o tempo de meia-vida, mais rápida é a regressão à média e menor a duração do trading. Os métodos utilizados para o cálculo da correlação foram: ρ de Pearson, τ de Kendall e a correlação de Spearman e os valores obtidos são mostrados na tabela Tabela 4.4.

	Resultados
Pearson	-0,12
Kendall	-0,03
Spearman	-0,04

Tabela 4.4: Correlação entre tempo de meia-vida e duração do trading: Protótipo 1

Correlações fortes são indicadas por valores em módulo próximos à unidade e cointegrações fracas são indicadas por valores em módulo próximos ao zero. Como é visível na Tabela 4.4, as correlações existentes entre tempo de meia-vida e a duração do trading são muito fracas, logo o tempo de meia vida não pode ser utilizado como preditor da duração do trading e não auxilia na abertura e fechamento de posições.

4.2 Protótipo 2: Pairs Trading (ON/PN) - Z-Score

O segundo protótipo é uma variação do primeiro. Essa variação consiste em limitar o universo dos pares operados a pares que representem papéis da mesma empresa, no caso, ações ordinárias (ON) e preferenciais (PN). Isso é feito porque já é esperado que esses pares apresentem uma cointegração forte, já que, por serem representativos da mesma empresa, na teoria, suas variações deveriam ser muito similares.

Os resultados diário e acumulado para a estratégia podem ser vistos no gráfico abaixo:

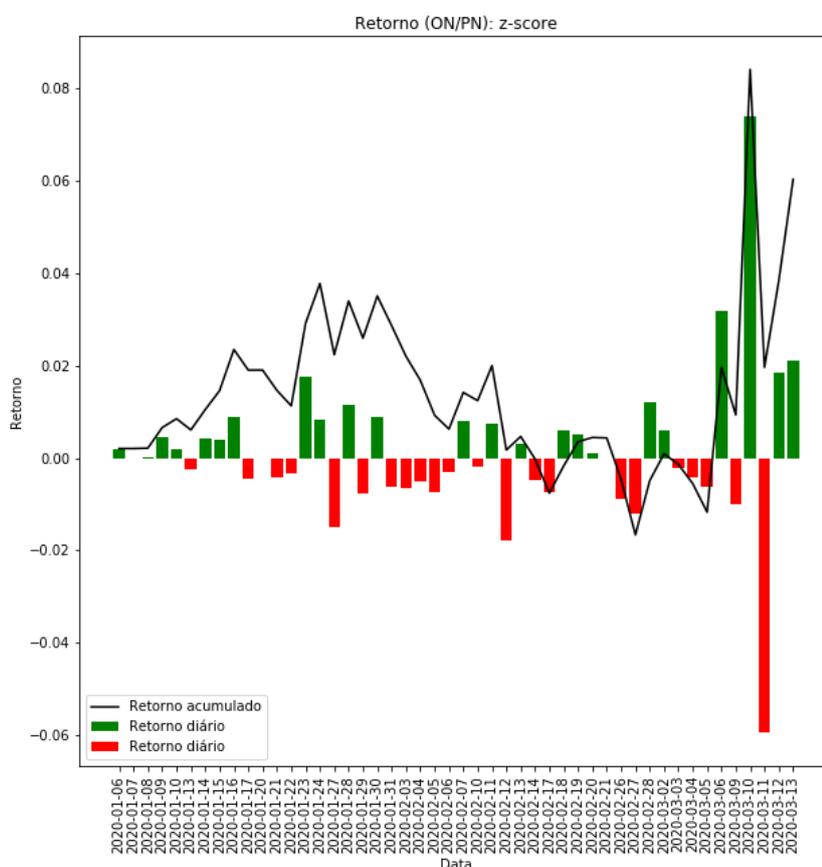


Figura 4.2: Rentabilidade: Protótipo 2.

A Tabela 4.5 também resume os parâmetros restantes para que se possa avaliar a robustez, enquanto a Tabela 4.6 mostra uma visão mais completa da estratégia.

Baseado nos resultados mostrado na Tabela 4.5 e na Figura 4.2, pode-se concluir que a estratégia apresenta um resultado positivo, mas em determinado momento chegou a apresentar pequenas perdas com relação ao valor inicial investido. Mesmo que a rentabilidade acumulada não seja tão expressiva quanto à do protótipo 1, ainda é um resultado

	Resultados
Rentabilidade Acumulada	6,02%
Rentabilidade Média Diária	0,12%
Volatilidade	1,66%
Índice de Sharpe	3,13

Tabela 4.5: Resultados: Protótipo 2

considerável, já que se fala de uma estratégia com rendimento anualizado de quase 36%. O que se nota, por meio da relação entre rentabilidade e volatilidade computada pelo índice de Sharpe, é que a relação risco e retorno da estratégia ainda é maior que o valor unitário, indicando que a estratégia é robusta nos parâmetros do protótipo 2.

	Resultados
Volume de operações diária	275 ± 203
Perda máxima possível diária de um par vencedor	11,27%
Percentual vencedor (Pares)	72,91%
Percentual vencedor (Dias)	50,00%

Tabela 4.6: Resultados complementares: Protótipo 2

Quando se analisam os pontos complementares mostrados na Tabela 4.6, nota-se que a estratégia ainda baseia sua performance no volume de posições, o que é indicado pelo alto número de operações diárias. No que diz respeito à relação mostrada entre o percentual vencedor de pares e dias, nota-se que, ao contrário do protótipo 1, há um desbalanceamento entre ganhos e perdas. Enquanto o percentual de pares vencedores é consideravelmente alto, apontando para um provável sucesso no longo prazo, o percentual de dias vencedores é igual ao de dias perdedores. Logo, mesmo que o número de pares vencedores seja maior que o de pares perdedores, o resultado médio dos pares perdedores é maior que o dos pares vencedores. Isso normalmente indica que a estratégia pode se beneficiar de um sistema de *stop-loss*. Como o valor máximo de perda diária de um par vencedor foi 11,27%, esse seria um valor no qual se poderiam ancorar os testes de *stop-loss*, mas esse processo precisa ser melhor desenvolvido. Dessa forma, a adição desse sistema fica como ponto de melhora, para aumentar a rentabilidade da estratégia e melhorar sua relação risco-retorno.

Como no caso do protótipo 2, todos os pares operados apresentam alta liquidez, nenhum comentário é necessário nessa área. Finalmente, podemos observar a correlação entre o tempo de meia-vida e a duração do trading dos pares na Tabela 4.7.

Como nenhuma correlação significativa é verificada, não se pode utilizar o tempo de meia-vida como parâmetro preditivo da duração do trading para o protótipo 2.

	Resultados
Pearson	0,02
Kendall	-0,03
Spearman	-0,06

Tabela 4.7: Correlação entre tempo de meia-vida e duração do trading: Protótipo 2

4.3 Protótipo 3: Pairs Trading - Cópulas

O terceiro protótipo utiliza a cointegração para formar os pares a serem operados e utiliza as cópulas de Arquimedes como método de abertura e fechamento de posições.

Os resultados diário e acumulado para a estratégia podem ser vistos no gráfico abaixo:

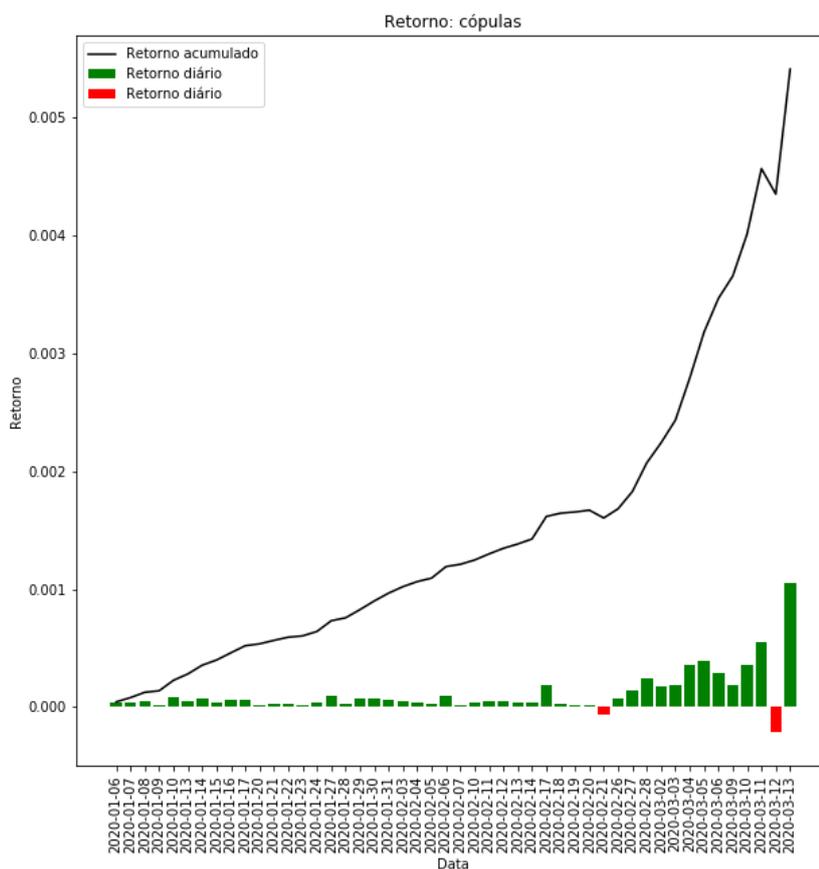


Figura 4.3: Rentabilidade: Protótipo 3.

Já a partir da análise gráfica, pode-se observar que os resultados obtidos por essa estratégia são bem mais suaves quando comparadas às duas primeiras, mas para que a

análise possa ser feita de forma mais sistemática, os demais resultados são mostrados na Tabela 4.8.

	Resultados
Rentabilidade Acumulada	0,54%
Rentabilidade Média Diária	0,01%
Volatilidade	0,02%
Índice de Sharpe	-14,11

Tabela 4.8: Resultados: Protótipo 3

O interessante de se notar nos resultados do protótipo 3 é que, nos parâmetros atuais desse protótipo, embora a curva de rentabilidade seja bem suave, a rentabilidade acumulada da estratégia é menor do que a da SELIC. Assim, a partir da análise dos parâmetros complementares, os pontos de melhora da estratégia podem ser destacados de forma a tornar a estratégia robusta.

	Resultados
Volume de operações diária	3 ± 4
Perda máxima possível diária de um par vencedor	2,5%
Percentual vencedor (Pares)	76,47%
Percentual vencedor (Dias)	95,83%

Tabela 4.9: Resultados complementares: Protótipo 3

Quando se analisam os valores na Tabela 4.9, percebe-se que a questão da rentabilidade da estratégia é algo simples de ser entendido. Pode-se notar que o protótipo 3, em contraste com os protótipos 1 e 2, baseia sua rentabilidade no tamanho das posições ao invés de baseá-la na quantidade de posições. O fato de o volume de operações diárias ser bem baixo mostra que para que a rentabilidade seja aumentada, o tamanho da posição para a estratégia com cópulas deve ser aumentado. Para ver se esse aumento na posição é algo que vale a pena, a análise dos demais parâmetros complementares é necessária. Esse protótipo também apresenta um percentual de pares vencedores consideravelmente alto, o que indica um possível sucesso de longo prazo, mas o ponto mais atrativo da estratégia está no percentual de dias vencedores. Como o percentual de dias vencedores é muito elevado, o desbalanceamento entre os ganhos e perdas da estratégia é algo positivo, ou seja, o ganho médio das posições vencedoras é maior que a perda média das posições perdedoras. Dessa forma, como a relação de pares e dias vencedores é algo extremamente favorável, a modificação no tamanho das posições deve resultar em uma estratégia extremamente robusta e o sistema de *stop-loss* se mostra desnecessário.

O protótipo 3 também opera com pares pouco líquidos. Dessa forma, a análise da estratégia com os pares ilíquidos é algo necessário. Esses dados são mostrados na Tabela 4.10.

	Resultados
Rentabilidade Acumulada	0,41%
Rentabilidade Média Diária	0,008%
Volatilidade	0,01%
Índice de sharpe	-35,62

Tabela 4.10: Resultados considerando liquidez: Protótipo 3

Como se pode notar, as operações com pares pouco líquidos representam pouco menos de 25% da rentabilidade acumulada. É importante notar que, embora a liquidez seja algo a ser estudado, como essa estratégia baseia sua rentabilidade em poucas posições, a falta de liquidez não é um fator tão considerável quanto no protótipo 1.

Finalmente, os resultados da análise de correlação entre tempo de meia-vida e a duração do trading são mostrados na Tabela 4.11.

	Resultados
Pearson	-0,03
Kendall	0,02
Spearman	0,03

Tabela 4.11: Correlação entre tempo de meia-vida e duração do trading: Protótipo 3

Novamente, como os testes de correlação mostram valores próximos ao nulo, o tempo de meia-vida não pode ser utilizado como fator preditivo para duração do trading.

4.4 Protótipo 4: Pairs Trading (ON/PN) - Cópulas

O quarto protótipo também é uma variação do terceiro e a variação consiste na mesma limitação do universo aos pares operados no protótipo 2 e pelos mesmos motivos. Dessa forma, para o quarto protótipo, somente pares compostos por ações ordinárias (ON) e preferenciais (PN) são considerados.

Os resultados diário e acumulado para a estratégia podem ser vistos na Figura 4.4.

Novamente, nota-se que a rentabilidade vinda do protótipo que utiliza as cópulas como método de abertura e fechamento de posição é muito suave quando comparada à rentabilidade dos protótipos 1 e 2, que utilizam o z-score como método de abertura e fechamento de posição. Os resultados para análise da robustez da estratégia são mostrados na Tabela 4.12.

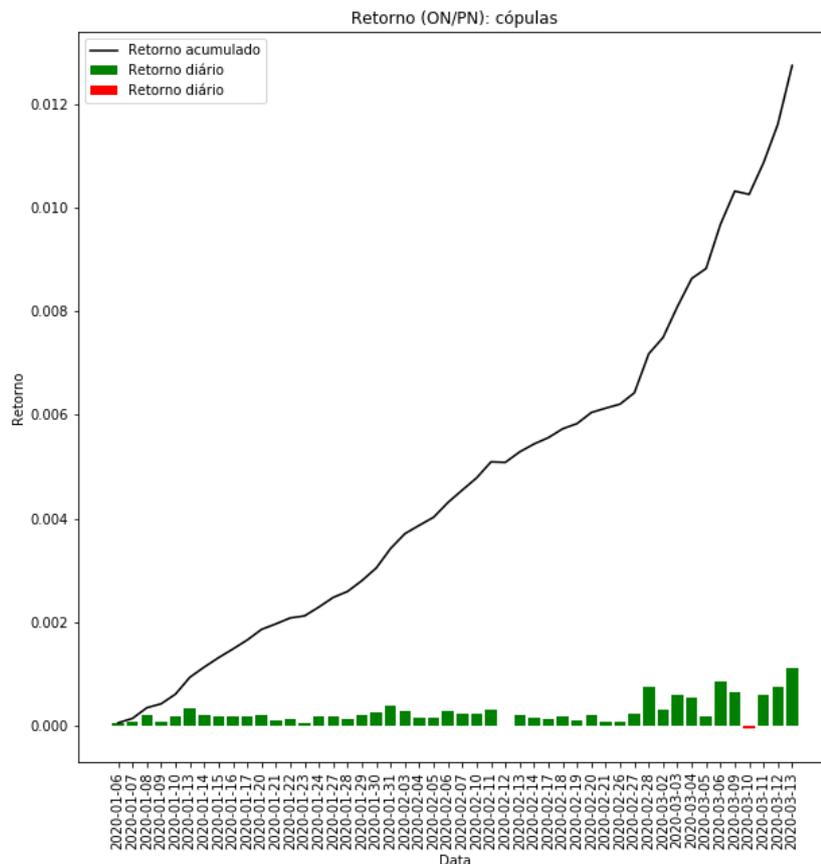


Figura 4.4: Rentabilidade: Protótipo 4.

	Resultados
Rentabilidade Acumulada	1,27%
Rentabilidade Média Diária	0,03%
Volatilidade	0,02%
Índice de Sharpe	20,46

Tabela 4.12: Resultados: Protótipo 4

Como se pode ver, o protótipo 4 apresenta rentabilidade maior que a do protótipo 3 e maior que a SELIC. Mesmo assim, essa estratégia apresenta baixa rentabilidade (rentabilidade anualizada de 6,87%), mas só pelo fato de apresentar rentabilidade pouco maior que a SELIC, o índice de Sharpe do protótipo 4 é altíssimo, o que evidencia a volatilidade extremamente baixa da estratégia. O protótipo 4, mesmo nos parâmetros atuais, é robusto pela relação risco-retorno evidenciado pelo índice de Sharpe, mas mesmo assim, a análise dos parâmetros complementares se faz útil para que se possa buscar um modo de

melhorar a rentabilidade da estratégia.

	Resultados
Volume de operações diária	4 ± 4
Perda máxima possível diária de um par vencedor	2,05%
Percentual vencedor (Pares)	87,85%
Percentual vencedor (Dias)	95,83%

Tabela 4.13: Resultados complementares: Protótipo 4

A análise da Tabela 4.13 permite tirar algumas conclusões. Primeiramente, analogamente ao protótipo 3, a estratégia baseia sua rentabilidade no tamanho das posições e não na quantidade, dessa forma, uma forma de se aumentar a rentabilidade da estratégia é aumentar o tamanho das posições alocadas nos pares. Além disso, percebe-se que o percentual de pares vencedores é maior que o do protótipo 3, essa vantagem provavelmente foi um dos fatores que fez com que a rentabilidade do protótipo 4 fosse maior que a do protótipo 3 e conseguisse superar a SELIC. Finalmente, percebe-se que o desbalançamento também existe no protótipo 4 e ele também é positivo, ou seja, o ganho médio das posições vencedoras supera em módulo a perda média das posições perdedoras. Isso é evidenciado pelo fato de que o percentual de dias vencedores é maior do que o de pares vencedores. Logo, essa estratégia também não necessita de um sistema de *stop-loss*.

Como esse universo de pares também não apresenta papéis pouco líquidos, análises acerca da liquidez são desnecessárias. Finalmente, a análise de correlação entre tempo de meia-vida e rentabilidade evidenciada na tabela Tabela 4.14 confirma novamente que a correlação é muito fraca para que esse parâmetro seja preditivo da duração do trading e não pode ser utilizado para auxiliar abertura e fechamento de posições

	Resultados
Pearson	-0,03
Kendall	0,04
Spearman	0,06

Tabela 4.14: Correlação entre tempo de meia-vida e duração do trading: Protótipo 4

4.5 Protótipo 5: Trading de Tríades - Z-Score

O quinto e último protótipo envolve a extensão natural do *pairs trading*, que é a arbitragem com trio de ativos, ou trading de tríades.

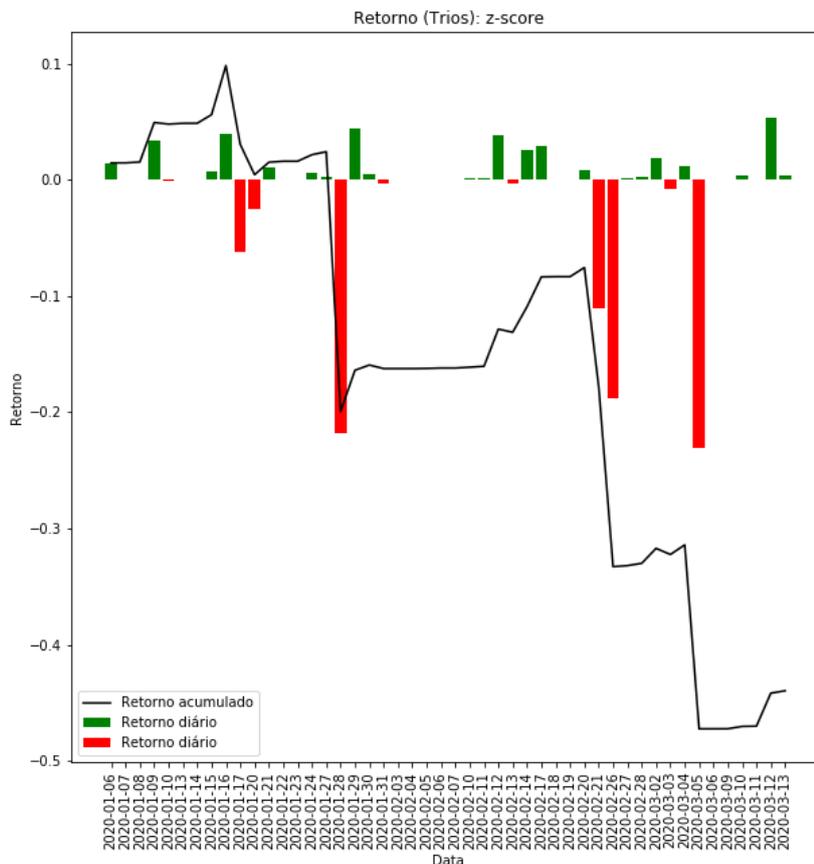


Figura 4.5: Rentabilidade: Protótipo 5.

Somente com a análise gráfica da rentabilidade, já é possível perceber que o protótipo com seus parâmetros atuais não é robusto, mas a análise do protótipo permite também que sejam identificados pontos de melhora afim de tornar a estratégia robusta.

	Resultados
Rentabilidade Acumulada	-43,97%
Rentabilidade Média Diária	-0,76%
Volatilidade	5,83%
Índice de Sharpe	-7,67

Tabela 4.15: Resultados: Protótipo 5

A partir da análise dos resultados da Tabela 4.15, percebe-se não só que a rentabilidade é negativa, como a volatilidade da estratégia é muito superior às dos demais protótipos.

A análise dos dados da Tabela 4.16 mostra que, assim como as outras estratégias que utilizam o z-score, o protótipo 5 também baseia sua rentabilidade no volume de

	Resultados
Volume de operações diária	214 ± 174
Perda máxima possível diária de um par vencedor	32,30%
Percentual vencedor (Pares)	76,19%
Percentual vencedor (Dias)	64,58%

Tabela 4.16: Resultados complementares: Protótipo 5

posições. O que mais chama atenção nos resultados é que o percentual de pares vencedores é alto. Dessa forma, mesmo que a estratégia acerte na maioria da escolha dos pares, a rentabilidade ainda é negativa. Além disso, o percentual de dias vencedores é menor que o percentual de pares vencedores, o que aponta para um desbalanceamento desfavorável entre as perdas e ganhos.

A partir da análise gráfica e das análises dos percentuais vencedores, pode-se perceber que o principal problema da estratégia é que quando ela tem um dia de resultado negativo, esse resultado é muito grande em módulo quando comparado aos ganhos dos dias vencedores. Logo, o protótipo 5 se beneficiaria de um sistema de *stop-loss*. Porém, quando é analisada a perda máxima diária de um par vencedor, percebe-se que esse número é elevado em módulo, mas menor que a perda média dos pares perdedores (41,30%). Assim, percebe-se que essa estratégia pode apresentar resultados melhores com a adição desse sistema.

Estabelecendo um sistema *stop-loss* ancorado no valor de perda máxima possível de um par vencedor, a estratégia apresentaria a rentabilidade mostrada na Figura 4.6, o que indica que mesmo com a implementação desse sistema, a rentabilidade ainda é negativa e a volatilidade ainda é relativamente alta (rentabilidade acumulada de -10,32% e volatilidade de 3,07%). Logo, a estratégia não pode ser considerada robusta. A implementação de um sistema *stop-loss* com parada abaixo do valor de perda máxima possível de um par vencedor interfere nos ganhos assim como nas perdas, portanto, um estudo mais detalhado deve ser feito.

Novamente, como o universo de trios possíveis para essa estratégia não incluiu pares de baixa liquidez, não é necessária uma análise desse tipo. Por fim, pode-se analisar se a correlação entre tempo de meia-vida e a rentabilidade pode ajudar a melhorar os resultados da estratégia.

	Resultados
Pearson	-0,09
Kendall	0,00
Spearman	0,01

Tabela 4.17: Correlação entre tempo de meia-vida e duração do trading: Protótipo 5

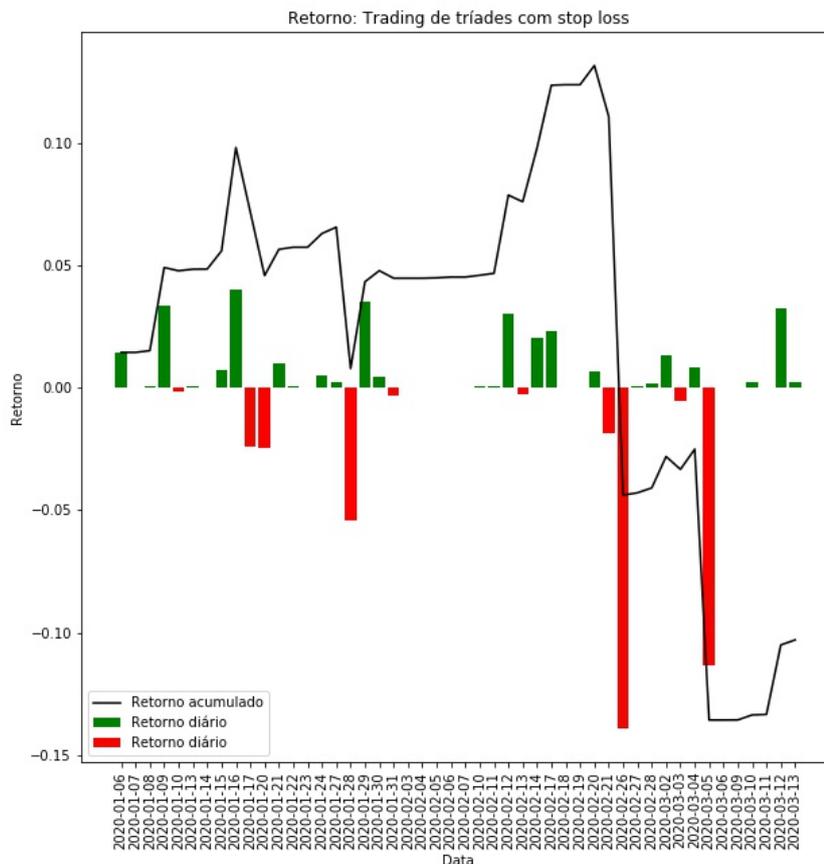


Figura 4.6: Rentabilidade: Protótipo 5 com sistema stop-loss.

Logo, pela análise da Tabela 4.17, percebe-se que o tempo de meia vida não é um bom preditor da duração do trading.

4.6 Benchmarking e Análises Finais

A análise de cada protótipo de forma individual, por mais que possa iluminar diversos pontos de melhora e mostrar a robustez da estratégia em termos de risco e retorno, ainda pode ser complementada pela técnica de *benchmarking*, que é um método derivado dos estudos de marketing, que consiste em avaliar o desempenho de um objeto em relação a outros.

Foi escolhido o BOVA11.SA como principal instrumento comparativo, já que esse papel é um fundo de índice que apresenta rentabilidade semelhante ao Ibovespa, servindo como *proxy* do comportamento geral do mercado de ações brasileiro.

No período que compreende a data de avaliação dos protótipos, o desempenho do BOVA11.SA é mostrado na Figura 4.7. Nela, observam-se dois períodos diferentes de comportamento de mercado: um neutro, em que o mercado apresentou comportamento lateralizado; e um *bear market*, em que o mercado apresentou queda acentuada devido às notícias relacionadas ao COVID-19 principalmente.



Figura 4.7: Desempenho: BOVA11.SA.

Nesse mesmo período, pode-se avaliar os parâmetros de rentabilidade e volatilidade relacionados ao benchmark escolhido. Dessa forma, em termos de comparação, a tabela Tabela 4.18 aponta os principais parâmetros a serem avaliados.

	Rentabilidade acumulada	Volatilidade	Índice de Sharpe
BOVA11.SA	-29,39%	4,21%	-7,17
Protótipo 1	45,46%	1,83%	24,36
Protótipo 2	6,02%	1,66%	3,13
Protótipo 3	0,54%	0,02%	-14,11
Protótipo 4	1,27%	0,02%	20,46
Protótipo 5	-43,97%	5,83%	-7,67

Tabela 4.18: Comparação de resultados

A comparação entre o desempenho fica mais fácil por meio da ilustração mostrada na Figura 4.8. Quanto maior a bolha, representativa do índice de Sharpe, melhor a avaliação de risco-retorno para as bolhas cujo índice é positivo. Além disso, quanto maior a rentabilidade e menor a volatilidade, melhor o desempenho da estratégia, logo,

pode-se perceber que somente o protótipo 5 teve desempenho abaixo do benchmark. Todos os outros protótipos tiveram performance superior em termos de volatilidade e rentabilidade. A questão da volatilidade é algo a ser ressaltado: todas essas quatro estratégias apresentaram volatilidades significativamente menores do que o benchmark, o que é algo interessantíssimo em termos de investimentos em mercados de renda variável, já que aponta para uma incerteza sensivelmente menor.

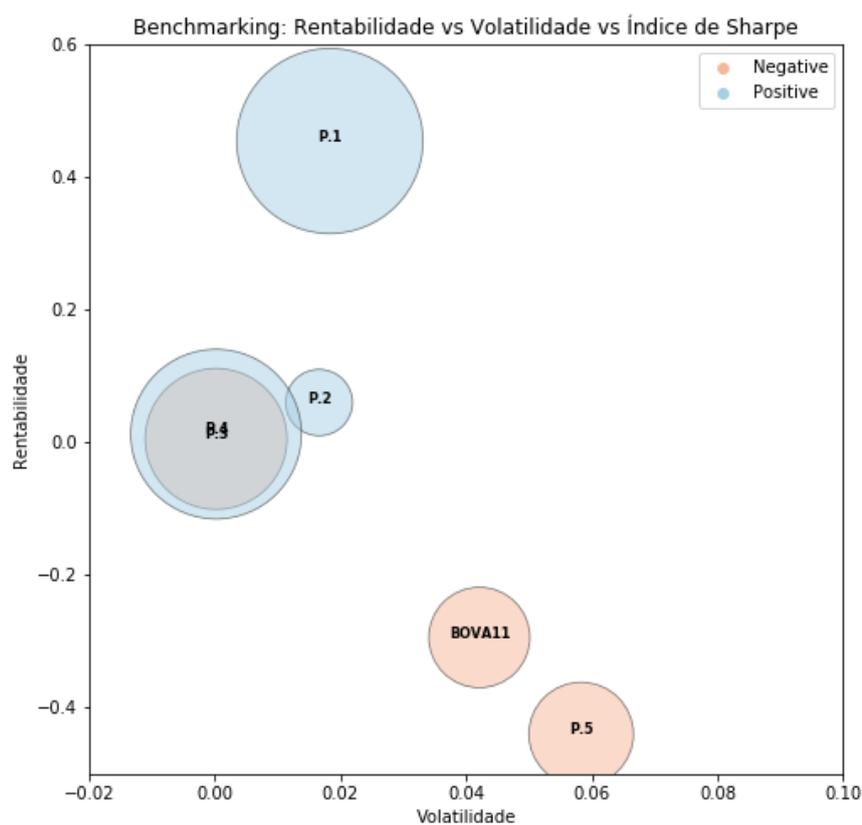


Figura 4.8: Ilustração da comparação entre desempenhos .

Além da comparação do desempenho, por serem estratégias neutras de mercado, é importante avaliar a correlação entre os retornos diários dos protótipos e o retorno diário do benchmark escolhido. Isso pode ser feito por meio dos testes de correlação e por uma inspeção visual.

Como é visível nos resultados mostrados na Tabela 4.19, mesmo que todos os resultados estejam no espectro mais próximo do nulo do que do unitário, temos correlações em alguns protótipos que não são insignificantes. Primeiramente, percebe-se que as estratégias que utilizaram de um universo de pares restritos aos pares preferenciais (PN) e ordinários

Comparativo: Retornos diários

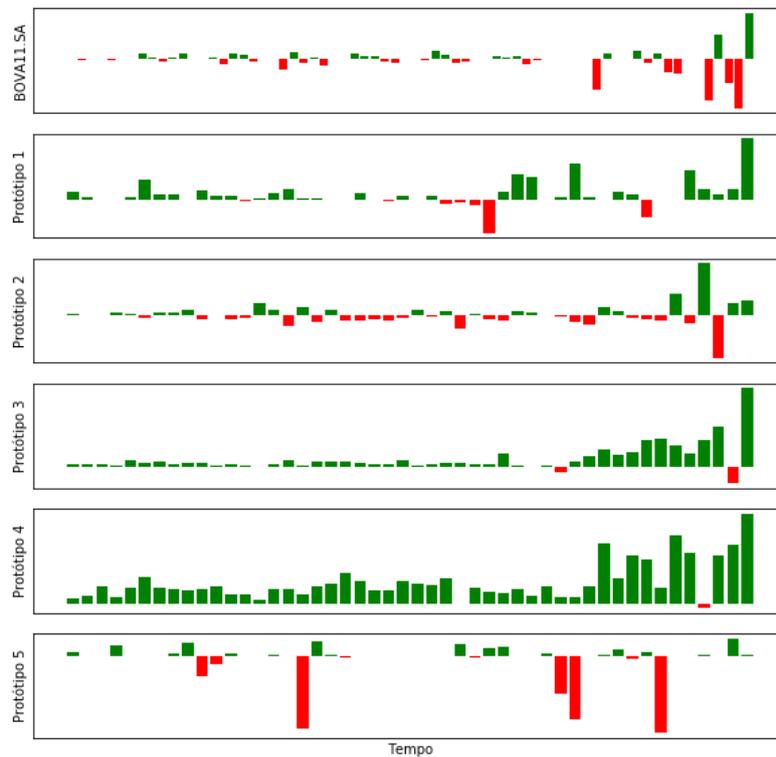


Figura 4.9: Comparativo: retornos diários .

	Pearson	Kendall	Spearman
Protótipo 1	0,19	0,02	-0,03
Protótipo 2	0,38	0,24	0,29
Protótipo 3	0,02	0,02	0,36
Protótipo 4	-0,13	-0,19	-0,21
Protótipo 5	-0,03	0,02	0,03

Tabela 4.19: Testes de correlação de resultados diários

(ON) possuem maior correlação com o retorno de mercado indicado pelo BOVA11.SA do que as demais estratégias. Isso ocorre principalmente em função dos papéis utilizados, que possuem um peso significativamente maior no mercado do que os operados nas demais estratégias. Dessa forma, pode-se confirmar que as estratégias que possuem uma maior diversificação de ativos são menos correlacionadas ao mercado. Finalmente, mesmo que as estratégias do universo ON/PN apresentem ativos de grande influência no mercado, o retorno das estratégias ainda apresenta uma correlação menor que 0,5 em módulo, o

que indica que a correlação é fraca apesar de não ser insignificante. Conclui-se então que os protótipos representam uma estratégia neutra com relação ao mercado, apesar de o universo operado influenciar na força da correlação: enquanto o universo restrito a papéis ON/PN incorre uma correlação mais forte, um universo mais diversificado dilui mais essa correlação.

Como a amostra inclui um pequeno período de *bear market*, é importante avaliar a correlação nesse período, já que em momentos de estresse no mercado, as correlações entre os ativos tendem a mudar e se tornarem próximas ao valor unitário, em módulo. Para o período em questão, as correlações são bem distintas das anteriores.

	Pearson	Kendall	Spearman
Protótipo 1	0,32	0,06	0,05
Protótipo 2	0,46	0,39	0,51
Protótipo 3	0,71	0,36	0,51
Protótipo 4	0,06	-0,12	-0,08
Protótipo 5	0,00	0,12	0,24

Tabela 4.20: Testes de correlação de resultados diários em *bear market*

Como é visível pelos testes de correlação, no período de *bear market* as correlações entre o protótipo 2 e o mercado e entre o protótipo 3 e o mercado se tornam bastante significativas. Isso é preocupante no que tange as possíveis perdas que podem acontecer em tais períodos. Logo, do ponto de vista de gerenciamento de risco, essas estratégias merecem um controle maior em períodos de estresse. Para os demais protótipos, como os valores dos testes variam significativamente e não são tão significativos, temos um indicativo que essas estratégias não são tão afetadas pelas mudanças provocadas no período de queda brusca do mercado.

5 Conclusão

Nesse trabalho, propõe-se a criação de uma gama de estratégias de arbitragem baseadas em cointegração e que exploram a característica de reversão à média do conjunto de ativos operados. As estratégias passam por dois passos: primeiro é feita a identificação diária dos conjuntos de ativos cointegrados, depois é feita a sinalização das oportunidades de entrada e saída das posições, que são de tamanho determinado por um método quantitativo de medida de risco.

A identificação dos conjuntos de ativos cointegrados é constituída por dois fatores: primeiro o filtro estocástico de Kalman estima os parâmetros que relacionam os ativos de um conjunto linearmente, após essa estimação de parâmetros, o resíduo resultante para cada conjunto de ativos é submetido ao teste de Dickey-Fuller Aumentado para que se possa testar para a cointegração, garantindo a estacionaridade do resíduo e sinalizando a cointegração existente entre os ativos do conjunto.

O universo de conjunto de ativos cointegrados é formado diariamente para que seja possível identificar oportunidades de entrada e saída de posição por dois métodos distintos: z-score e Cópulas de Arquimedes. Esses dois modelos permitem que se identifiquem os momentos para a construção de posições *long* e *short* no conjunto de ativos cointegrados. Para cada oportunidade identificada, a posição é dimensionada segundo dois fatores: a razão entre o apetite de risco para cada posição e o risco dimensionado pelo método *expected shortfall*, e o *hedging* necessário para manter a estratégia blindada da direcionalidade do mercado.

Dessa forma, são construídos cinco protótipos caracterizados por três fatores: a quantidade de ativos em cada conjunto cointegrado (par ou trio), o método de sinalização de entrada e saída de posição (z-score ou Cópulas de Arquimedes) e o universo de ativos que podem constituir os conjuntos cointegrados. Assim, o protótipo 1 opera sobre um conjunto de 95 possíveis pares utilizando o z-score, o protótipo 2 opera sobre um conjunto de 6 pares possíveis utilizando o z-score, o protótipo 3 opera sobre um conjunto de 95 pares utilizando as cópulas, o protótipo 4 opera sobre um conjunto de 6 pares utilizando as cópulas e, finalmente, o protótipo 5 opera sobre um conjunto de 61 trios possíveis utilizando o z-score como método de sinalização de oportunidades de arbitragem.

Os resultados obtidos para os protótipos que sintetizam as estratégias são então visualizados sob duas óticas, uma de resultado entregue e uma de potencial de entrega. Para que isso seja possível, é realizada uma macro-análise, que compara três parâmetros da estratégia (rentabilidade, volatilidade e índice de Sharpe) com um *benchmark* e é realizada uma micro-análise da estratégia que tenta estimar seu potencial a partir de quatro parâmetros: volume de operações diárias, perda máxima possível de um par vencedor, percentual de pares vencedores e percentual de dias vencedores.

Os parâmetros da macro-análise são relacionados ao resultado que o protótipo conseguiu entregar em termos de risco e retorno. Esses parâmetros são comparados a um índice de fundo de mercado para que se possa ter noção de como esse protótipo se compara a um simples investimento em um ETF. Cada parâmetro da micro-análise, tem seu racional: o volume de operações diárias é uma forma de inferir a adequação do dimensionamento das posições para cada protótipo; a perda máxima possível de um par vencedor é uma *proxy* para a implementação possível de um sistema de gestão de risco mais incisivo; o percentual de pares vencedores é um indicativo da possibilidade de sucesso de longo prazo da estratégia e sua relação com o percentual de dias vencedores é um indicativo de um possível desbalanceamento entre o módulo dos ganhos e das perdas na estratégia.

O protótipo 1 apresentou tanto a macro-análise quanto a micro-análise favoráveis. Com rentabilidade acumulada de 45,46% e volatilidade de 1,83%, o protótipo 1 apresentou um índice de Sharpe de 24,36, o que expressa uma relação de risco e retorno sensivelmente vantajosa. Além disso, a micro-análise da estratégia indica uma adequação no que diz respeito ao dimensionamento das posições tomadas e tanto ao sucesso de longo prazo, quanto ao balanceamento entre os módulos de ganhos e perdas. Mesmo com a consideração feita com relação à falta de liquidez de alguns ativos operados, a estratégia continuou apresentando ambas as análises favoráveis. Dessa forma, é possível concluir que a estratégia realizada no protótipo 1 é robusta nos parâmetros definidos pelo projeto.

O protótipo 2 apresentou uma macro-análise favorável, mas a micro-análise da estratégia alertou para um possível ponto de melhora. Com rentabilidade acumulada de 6,02% e volatilidade de 1,66%, o protótipo 2 apresentou um índice de Sharpe de 3,13, o que expressa uma boa relação de risco e retorno. Dos elementos da micro-análise que merecem destaque, o desbalanceamento entre o percentual de pares vencedores (72,91%) e de dias vencedores (50,00%) aponta para um ponto ineficiente da estratégia. Esse desbalanceamento é uma sinalização de que o resultado médio dos pares perdedores é, em módulo, maior do que o dos pares vencedores, o que corroi a rentabilidade da estratégia. Um sistema que limite as perdas máximas a um valor ancorado na perda máxima experimentada por um par vencedor é um método que merece atenção. Dessa forma, mesmo que a estratégia realizada no protótipo 2 seja robusta nos parâmetros definidos, essa estratégia

se beneficiaria de um refinamento para que seu sucesso no longo prazo seja possível.

O protótipo 3 apresentou uma macro-análise desfavorável e a micro-análise da estratégia permitiu que o ponto crítico para a melhora dessa estratégia fosse descoberto. Com rentabilidade acumulada de 0,54% e volatilidade de 0,02%, esse protótipo apresentou um índice de Sharpe de -14,11. O índice de Sharpe da estratégia aponta para o ponto de melhora: a rentabilidade, mas é a micro-análise da estratégia que mostra em que ponto há inadequação. O volume de operações pequeno da estratégia aponta para uma inadequação com relação ao dimensionamento das posições de arbitragem. Isso é corroborado pelo fato de que a estratégia apresenta um desbalanceamento positivo entre pares e dias vencedores, ou seja, o resultado médio das posições vencedoras é, em módulo, maior que o das posições vencedoras. Isso é indicado pelo fato de o percentual de dias vencedores (95,83%) ser maior do que o dos pares vencedores (76,47%). Logo, mesmo que o protótipo 3 não seja robusto dos parâmetros definidos, essa estratégia apresenta grande potencial se revisto o dimensionamento das posições.

O protótipo 4 apresentou tanto uma macro-análise favorável quanto uma micro-análise eficiente, vide que ela identificou o mesmo possível ponto de melhora necessário ao protótipo 3. Com uma rentabilidade acumulada de 1,27% e uma volatilidade de 0,02%, o protótipo 4 apresentou índice de Sharpe de 20,46, o que aponta para uma relação de risco e retorno vantajosa. O fato de uma rentabilidade relativamente pequena apresentar um índice de Sharpe tão expressivo aponta para o ponto de melhora da estratégia: a rentabilidade. A micro-análise da estratégia aponta novamente que o dimensionamento das posições pode ser melhor projetado para que a estratégia apresente uma rentabilidade maior. Isso é corroborado também pelo desbalanceamento positivo entre pares vencedores (87,85%) e dias vencedores (95,83%). Dessa forma, mesmo que o protótipo 4 seja robusto nos parâmetros desse projeto, ele pode ser melhorado para retornar uma rentabilidade mais atrativa.

O protótipo 5 apresentou uma macro-análise sensivelmente desfavorável e a micro-análise apontou uma possível deficiência dessa estratégia. Com rentabilidade de -43,97% e volatilidade de 5,83%, o protótipo apresentou índice de Sharpe de -7,67, o que aponta para uma relação de risco-retorno desvantajosa. A micro-análise apontou que há um desbalanceamento negativo entre percentual de pares vencedores (76,19%) e dias vencedores (64,58%), mas como a perda máxima possível de um par vencedor é alta (32,30%), quando realizado um teste com um sistema de *stop-loss*, o protótipo ainda apresentou rentabilidade negativa (-10,3%). Dessa forma, o protótipo 5 não é eficiente nos parâmetros estabelecidos e uma análise mais detalhada em termos de rentabilidade e volatilidade deve ser feita.

Dessa forma, quatro dos cinco protótipos são ou eficientes ou seu ponto de melhora

para o alcance da eficiência é claro. O protótipo 5, porém, foi o único que performou de forma pior que o *benchmark* tanto em rentabilidade quanto em volatilidade e sua análise posterior é mais delicada. Pode-se concluir, assim, que 80% do trabalho realizado tem aplicação quase imediata, o que indica uma adequação entre teoria e simulação, faltando a análise em tempo real. A análise em tempo real, porém, possui suas peculiaridades que não são reproduzidas em simulação, principalmente no que diz respeito à reflexividade do mercado, ou seja, no impacto que uma posição de compra ou venda pode ter na precificação do ativo em geral. Como a simulação é um ambiente passivo, em que um comando de posição não implica reprecificação do ativo, esse limitante pode impactar a estratégia e a dimensão desse impacto é um fator importante.

Finalmente, ficam como recomendações para perspectivas de trabalhos futuros dois conjuntos de construção distinta: um relacionado aos pontos chave da estratégia como estão e outro relacionado às modificações no corpo da estratégia. As construções que estendem a estratégia incluem tanto os pontos de melhora indicados para cada protótipo quanto a análise do impacto que o trading em tempo real tem sobre a estratégia. Além disso, elas incluem a ampliação do conjunto de períodos de cointegração a serem analisados e funções objetivo analisadas, para que a análise *walk-forward* seja mais abrangente. Podem também incluir a aplicação da estratégia em outros mercados e outros tipos de ativos e, finalmente, a operação da estratégia com o conjunto completo de ativos a serem formados, sem que o processamento computacional limite sua quantidade de forma tão contundente. Com relação às construções que modificam o corpo da estratégia, fica como recomendação a exploração de outros modelos utilizados para a estimação de parâmetros do filtro de Kalman e a consideração de métodos de construção de portfólio para a alocação de capital de maneira mais inteligente.

Referências

- [1] Lewis, Michael: *Moneyball. O homem que mudou o jogo*. Intrínseca, 2003. 2
- [2] Singh, Sudhir: *Investor irrationality and self-defeating behavior: Insights from behavioral finance*. Journal of Global Business Management, 8:116–122, 2012. 2
- [3] Soros, George: *Hedge fund regulation, fund managers panel*. <https://www.c-span.org/video/?282391-2/hedge-fund-regulation-fund-managers-panel>. 3
- [4] Lowenstein, Roger: *When Genius Failed: The Rise and Fall of Long-Term Capital Management*. Random House Trade, 2001. 4, 10, 42
- [5] Zuckerman, Gregory: *The Man Who Solved the Market: How Jim Simons Launched the Quant Revolution*. Portfolio, 2019. 4, 8, 9
- [6] Adam Atkins, Mahesan Niranjan, Enrico Gerding: *Financial news predicts stock market volatility better than close price*. The Journal of Finance and Data Science, 4:120–137, 2018. 10
- [7] Jefferson Duarte, Francis A. Longstaff, Fan Yu: *Risk and return in fixed income arbitrage: Nickels in front of a steamroller?* California Digital Library, 2005. 10
- [8] Marco Avellaneda, Jeong Hyun Lee: *Statistical arbitrage in the u.s. equities market*, 2008. 10, 36
- [9] Panariello, André: *Trading por arbitragem estatística*, 2016. 11
- [10] Vidyamurthy, Ganapathy: *Pairs Trading: quantitative methods and analysis*. John Wiley Sons, 2004. 11
- [11] Davila, Victor Hugo L.: *Introdução às séries temporais*. <https://www.ime.unicamp.br/hlachs/MaterialSeries.pdf>. 12
- [12] Bueno, Rodrigo L. S.: *Econometria de Séries Temporais*, capítulo 2. Cengage Learning, 2ª edição, 2011. 12, 13, 14
- [13] Bueno, Rodrigo L. S.: *Econometria de Séries Temporais*, capítulo 4. Cengage Learning, 2ª edição, 2011. 13, 14, 15, 16
- [14] Schwert, G. William: *Tests for unit roots: A monte carlo investigation*. Journal of Business and Economic Statistics, 1989. 16

- [15] Robert F. Engle, C. W. J. Granger: *Co-integration and error correction: Representation, estimation, and testing*. *Econometrica*, 1987. 16
- [16] Bueno, Rodrigo L. S.: *Econometria de Séries Temporais*, capítulo 7. Cengage Learning, 2ª edição, 2011. 16
- [17] Jazwinski, Andrew H.: *Stochastic Processes and Filtering Theory*, capítulo 5. Academic Press, 1970. 17
- [18] Greg Welch, Gary Bishop: *An introduction to the kalman filter*. University of North Carolina at Chapel Hill, 1997. 18, 19, 20
- [19] KALMAN, R. E.: *A new approach to linear filtering and prediction problems*. *Journal of Basic Engineering*, 82:35–45, 1960. 20
- [20] Jack Farmer, Ram Seshadri: *Using machine learning in trading and finance - kalman filter trading applications*. <https://www.coursera.org/lecture/machine-learning-trading-finance/kalman-filter-trading-applications-jEt1T>. 21
- [21] Chen, Song Xi: *Parameters estimation and bias correction for diffusion processes*. Iowa State University Digital Repository, 2006. 22, 23
- [22] N. Balakrishnan, Chin Diew Lai: *Continuous Bivariate Distributions*, capítulo 2, 3. Springer, New York, 2ª edição, 2009. 23
- [23] Mahfoud, Maria: *Bivariate archimedean copulas: an application to two stock market indices*, 2012. 24, 25, 26
- [24] Christian Genest, Jock MacKay: *The joy of copulas: Bivariate distributions with uniform marginals*. *The American Statistician*, 40:280–283, 1986. 24
- [25] Prado, Marcos Lopez de: *Advances in Financial Machine Learning*, capítulo 12. Wiley, 2018. 27, 29
- [26] Pardo, Robert: *The Evaluation and Optimization of Trading Strategies*. John Wiley Sons, 2008. 27, 28, 29, 30, 31, 44
- [27] Bill Burnett, Dave Evans: *O Design da Sua Vida*. Rocco, 2016. 31
- [28] Dalio, Ray: *Princípios*. Intrínseca, 2018. 32
- [29] MacKinnon, James G.: *Critical values for cointegration tests*, 2010. 35
- [30] Caldeira, João F.: *Arbitragem estatística, estratégia long-short pairs trading abordagem com cointegração aplicada ao mercado de ações brasileiro*. *Revista EconomiA*, páginas 521—546, 2013. 38, 43, 45
- [31] Yolanda Stander, Daniël Marais, Ilse Botha: *Trading strategies with copulas*. *Journal of Economic and Financial Sciences*, páginas 83–108, 2013. 38, 40, 41
- [32] Landgraf, Nikolaus: *High-frequency copula-based pairs trading on u.s. goldmine stocks*, 2016. 39, 40

- [33] Best, Philip: *Implementing Value at Risk*. Wiley, 1999. 41
- [34] Rocha, José Eduardo N. Da: *Sistema Inteligente de Diagnósticos Energéticos e de Análise de Investimentos em Projetos de Eficiência Energética Gerenciados pelo Lado da Demanda*. Tese de Doutorado, Pontifícia Universidade Católica do Rio de Janeiro, 2013. 42
- [35] Taleb, Nassim N.: *Against value-at-risk: Nassim taleb replies to philippe jorion*. <https://www.fooledbyrandomness.com/jorion.html>, 1997. 44
- [36] Gillespie, Daniel T.: *The mathematics of brownian motion and johnson noise*. American Journal of Physics, 64:225–240, 1995. 45
- [37] G. E. Uhlenbeck, L. S. Ornstein: *On the theory of brownian motion*. Physical Review, 36:823–841, 1930. 45