



FACULDADE DE ECONOMIA E FINANÇAS IBMEC  
PROGRAMA DE PÓS-GRADUAÇÃO E PESQUISA EM  
ADMINISTRAÇÃO E ECONOMIA

DISSERTAÇÃO DE MESTRADO  
PROFISSIONALIZANTE EM ECONOMIA

**PROPOSTA DE UM MODELO  
ESTATÍSTICO PARA CLASSIFICAÇÃO DE  
FUNDOS DE INVESTIMENTO**

**CIRO MAGALHÃES DE MELO JORGE**

ORIENTADOR: OSMANI GUILLÉN

**Rio de Janeiro, 15 de agosto de 2008**

# **Livros Grátis**

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

**PROPOSTA DE UM MODELO ESTATÍSTICO PARA CLASSIFICAÇÃO DE  
FUNDOS DE INVESTIMENTO**

CIRO MAGALHÃES DE MELO JORGE

Dissertação apresentada ao curso de  
Mestrado Profissionalizante em Economia  
como requisito parcial para obtenção do  
Grau de Mestre em Economia.  
Concentração: Finanças e Controladoria

ORIENTADOR: OSMANI GUILLÉN

Rio de Janeiro, 15 de agosto de 2008.

**PROPOSTA DE UM MODELO ESTATÍSTICO PARA CLASSIFICAÇÃO DE  
FUNDOS DE INVESTIMENTO**

**CIRO MAGALHÃES DE MELO JORGE**

Dissertação apresentada ao curso de  
Mestrado Profissionalizante em Economia  
como requisito parcial para obtenção do  
Grau de Mestre em Economia.  
Concentração: Finanças e Controladoria

Avaliação:

BANCA EXAMINADORA:

---

Professor OSMANI TEIXEIRA DE CARVALHO GUILLÉN, D.Sc. (Orientador)  
Instituição: IBMEC-RJ

---

Professor ALEXANDRE BARROS DA CUNHA, Ph.D.  
Instituição: IBMEC-RJ

---

Professor EURILTON ALVES ARAÚJO JÚNIOR, Ph.D.  
Instituição: IBMEC-SP

Rio de Janeiro, 15 de agosto de 2008.

332.63  
J82

Jorge, Ciro Magalhães de Melo.

Proposta de um modelo estatístico para classificação de fundos de investimento / Ciro Magalhães de Melo Jorge. - Rio de Janeiro: Faculdades Ibmecc, 2008.

Dissertação de Mestrado Profissionalizante apresentada ao Programa de Pós-Graduação em Economia das Faculdades Ibmecc, como requisito parcial necessário para a obtenção do título de Mestre em Economia.

Área de concentração: Finanças e Controladoria.

1. Classificação de fundos de investimento. 2. Análise de Estilo. 3. Análise de *Clusters*. 4. Análise Discriminante

Dedico este trabalho à minha avó  
Áurea, meus pais, meu irmão e à  
Taissa, minha namorada.

## AGRADECIMENTOS

Em primeiro lugar, à minha querida avó Áurea! Sem seu apoio muitas coisas não seriam possíveis ao longo da vida, muito menos o início e a conclusão deste mestrado.

A meu pai Adílio, minha mãe Adélia e meu irmão Vitor, pelo interminável carinho e suporte. Por se preocuparem com meu futuro e estarem presentes para ajudar em todo o necessário e da forma mais natural do mundo. Qualquer sucesso que eu tenha na vida é devido a vocês!

À minha namorada linda, a Inhinha, pelo companheirismo incondicional e pelos enormes sacrifícios ao longo destes quase três anos de estudos praticamente diários. Te amo demais! Agora é hora de nós curtirmos muito juntos!

Ao Fábio Chung, o Dioin Brother. Vários debates emocionantes sobre o tema com esse camarada para o resto da vida. Quase uma banca antecipada de tão bons que eram os questionamentos! Obrigado também pela intermediação na obtenção de dados e pelas opiniões especializadas de um profissional da área!

Aos amigos todos, por manterem os convites para os programas, mesmo eu estando um chato que não fazia nada a não ser estudar. Agora estou de volta com força total!

Ao Professor Osmani Guillén. Digo com toda a franqueza que, além de professor e orientador, o tenho como um amigo!

Ao Professor Alexandre Cunha pela tremenda seriedade e profissionalismo na condução da Coordenação do Mestrado em Economia do Ibmec.

À banca como um todo, que reservou com dedicação sua atenção para ler e ouvir a obra de pesquisa de um humilde candidato a mestre.

À equipe de professores do Ibmec, composta 100% por profissionais de altíssima qualidade e competência!

Aos colegas do mestrado. Só nós sabemos como a carga é pesada! Para trabalhar e estudar ao mesmo tempo tem que ter garra. Nada que um chope no intervalo de aulas não resolva.

À toda a equipe do Ibmec (e em especial à Rita de Cássia), que confere a essa Escola nível internacional de ensino e pesquisa sem sombra de dúvida!

À ANBID, à ANDIMA e à Quantum Fundos, pela generosa disponibilização de dados para a realização do trabalho e, quando aplicável, pelas entrevistas concedidas por seus profissionais. Seu apoio à pesquisa no Brasil é de importância inestimável. Parabéns pela postura pró-ativa!

Ao Carlos Gutierrez, pelos preciosos ensinamentos em Matlab.

A ignorância humana tende ao infinito.

Fato permanente e imutável.

A ciência, incrédula, se recusa a aceitar...

Que sua teimosia seja eterna!

## RESUMO

A classificação adequada de fundos de investimento é relevante porque organiza a informação disponível de forma que investidores possam tomar melhores decisões de aplicação de recursos e a custos de procura menores. No Brasil, há dois sistemas com esse intuito, o da CVM e o da ANBID, porém ambos possuem categorias escolhidas subjetiva e arbitrariamente. Isso implica em má definição de fronteiras entre categorias, o que pode prejudicar a diversificação de carteiras. Adicionalmente, tornam complexa a decisão dos investidores por apresentarem listas artificialmente extensas de categorias. A presente dissertação traz uma nova proposta com metodologia objetiva. Para gerar a grade de classificação, se baseia no próprio comportamento dos fundos, o qual é definido pelo conceito de estilo introduzido por Sharpe (1992). A partir dos estilos computados para uma amostra de fundos, foram formadas, através de Análise de Clusters, categorias compostas por fundos com comportamentos semelhantes. Análise Discriminante foi então aplicada para avaliar a consistência das fronteiras entre essas categorias. A eficácia do novo modelo foi confrontada com o da ANBID. Constatou-se que seu poder de explicar retornos futuros é maior, mesmo possuindo um número bem inferior de categorias.

**Palavras-chave:** classificação de fundos de investimento, análise de estilo, análise de *clusters*, análise discriminante.

## ABSTRACT

The existence of proper models for mutual fund classification is desirable because efficient organization of available information can help investors make better financial decisions while at the same time reducing search costs. In Brazil, there are two distinct ones: CVM's and ANBID's, but both define categories via arbitrary and subjective ways, thus resulting in inadequate classification frontiers that can cause damage to portfolio diversification. Additionally, investor decision making turns out to be more complex than it should be because of the extensive list of categories provided by these models. The present study brings a new theoretically based approach with objective methodology. In order to create the classification matrix, the proposed model takes into account the actual fund managers' investment behaviors, which are defined by the concept of style introduced by Sharpe (1992). Once computed the styles for a selected sample, Cluster Analysis was applied to discover natural groupings among the funds. Discriminant functions were then estimated to test the consistency of frontiers between fund categories. The ANBID classification system was chosen as a benchmark to allow model evaluation. The results showed that the latter explains future returns better than the former even though its number of categories is significantly less.

**Keywords:** mutual fund classification, style analysis, cluster analysis, discriminant analysis.

## SUMÁRIO

1. INTRODUÇÃO .....	1
2. REVISÃO DA LITERATURA.....	9
2.1. Testes sobre sistemas tradicionais/oficiais de classificação de fundos .....	9
2.2. Propostas de modelos para classificação de fundos de investimento .....	11
3. METODOLOGIA DE PESQUISA .....	14
3.1. Análise de Estilo .....	14
3.2. Análise de <i>Clusters</i> .....	17
3.3. Análise Discriminante Múltipla.....	21
3.4. Seqüência lógica da metodologia.....	23
4. APRESENTAÇÃO DOS DADOS.....	26
4.1. A variável dependente.....	26
4.2. As variáveis independentes .....	32
5. MODELAGEM E RESULTADOS .....	42
5.1. Análise de Estilo .....	42
5.2. Análise de <i>Clusters</i> .....	49
5.3. Análise Discriminante.....	56
5.4. Avaliação do desempenho prático do modelo .....	64
6. CONCLUSÃO .....	67
6.1. Limitações do estudo .....	71
6.2. Direcionamentos futuros.....	73
7. REFERÊNCIAS BIBLIOGRÁFICAS .....	74
APÊNDICE A – Fundos pesquisados no sistema de consulta a carteiras da CVM .....	78
APÊNDICE B – Código do programa em Matlab para Análise de Estilo .....	80
APÊNDICE C – Histogramas das variáveis independentes da análise discriminante .....	82
APÊNDICE D – Gráficos de dispersão das variáveis discriminantes por categoria.....	85

## LISTA DE GRÁFICOS

Gráfico 1 – Evolução na oferta de fundos de investimento .....	1
Gráfico 2 – Ofertas públicas de recebíveis imobiliários e de crédito privado.....	37
Gráfico 3 - Histograma da posição a descoberto individual mais relevante .....	45
Gráfico 4 - Histograma da alavancagem total .....	46
Gráfico 5 - Histograma da concentração máxima em uma única classe de ativos.....	46
Gráfico 6 - Histograma dos $R^2$ .....	47
Gráfico 7 - <i>Boxplots</i> dos fundos por categoria .....	52

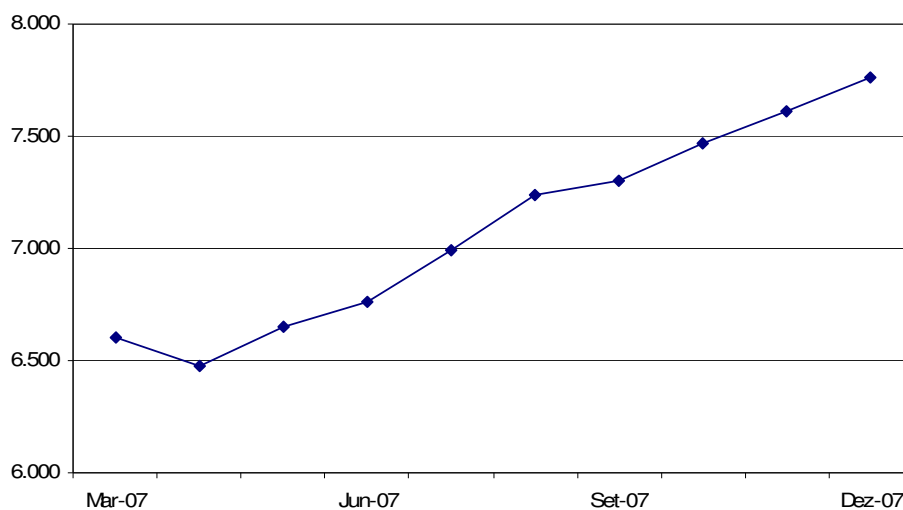
## LISTA DE TABELAS

Tabela 1 – Quantidade de fundos existentes no Brasil.....	31
Tabela 2 – Índices utilizados em outros trabalhos envolvendo Análise de Estilo.....	32
Tabela 3 - Ações integrantes dos índices setoriais construídos.....	34
Tabela 4 - Sumário dos índices selecionados.....	36
Tabela 5 - Participação de debêntures (sem <i>leasing</i> ) no estoque de investimentos brasileiro ....	38
Tabela 6 - Componentes resultantes da análise iterativa.....	40
Tabela 7 - Variáveis independentes da Análise de Estilo e respectivas características .....	41
Tabela 8 - Matriz de correlação entre as classes de ativos da Análise de Estilo .....	42
Tabela 9 - Centróides do <i>cluster</i> hierárquico inseridos como entrada na análise de <i>k</i> -médias ...	50
Tabela 10 - Número de fundos de acordo com os métodos hierárquico e <i>k</i> -médias .....	51
Tabela 11 - Centróides das categorias da nova classificação de fundos proposta.....	51
Tabela 12 - Matriz de transição da Classificação ANBID para a nova classificação proposta...	53
Tabela 13 - Estatísticas do teste de normalidade.....	57
Tabela 14 - Teste Box's M .....	57
Tabela 15 - Matriz de correlação entre os pesos das 5 primeiras classes de ativos.....	59
Tabela 16 - Separação entre amostra de estimação e amostra de validação.....	59
Tabela 17 - Probabilidades prévias das categorias na amostra inicial e grupo de estimação.....	60
Tabela 18 - Parâmetros estimados para as funções classificatórias .....	61
Tabela 19 - Matriz de previsão por tipo de amostra.....	62
Tabela 20 - Percentuais de erro e acerto de previsão por tipo de amostra .....	62
Tabela 21 - Resultados dos Testes de Huberty.....	63
Tabela 22 - Testes-t sobre a diferença nos $R^2$ médios entre o modelo proposto e o da ANBID.	67

## 1. INTRODUÇÃO

A indústria de fundos mútuos tem papel relevante como canal de investimento em vários países ao redor do mundo (Xue-jun e Xiao-lan, 2004; Moreno et al., 2006; Malhotra et al., 2007). No Brasil, desde a década de 90, um expressivo crescimento no seu volume de poupança vem sendo experimentado. Em dezembro de 2007, segundo a Associação Nacional dos Bancos de Investimento (ANBID), as aplicações em fundos atingiram valor de mercado igual a R\$1,1 trilhão.<sup>1</sup>

A indústria envolve também um elevado número de participantes. Entre pessoas físicas e jurídicas, existiam 10,9 milhões de cotistas no fim do ano de 2007 (fonte: ANBID). A oferta de fundos de investimento disponíveis no mercado vem aumentando consistentemente visando suprir essa demanda. Tal fato pode ser claramente observado através do Gráfico 1, o qual demonstra sua evolução no período compreendido entre março e dezembro de 2007.



Fonte: ANBID

**Gráfico 1 – Evolução na oferta de fundos de investimento**

<sup>1</sup> Inclui fundos de ações, multimercado, previdência (entidades abertas), referenciado DI, curto prazo, renda fixa, privatização, câmbio, dívida externa, índice e de investimentos em direitos creditórios. Considera fundos de investimento em cotas de outros fundos. Isto é válido para as demais estatísticas sobre a indústria de fundos brasileira apresentadas ao longo do capítulo corrente.

O alcance tomado pela indústria de fundos mútuos traz forte evidência de que ela oferece – ao menos para os cotistas participantes – benefícios que prevalecem sobre seus custos.

Do lado dos custos, tem-se:

- Despesas operacionais incorridas pelo fundo e que não recairiam sobre investidores individuais, por exemplo, taxa de fiscalização da Comissão de Valores Mobiliários (CVM)<sup>2</sup>, despesas com auditoria, preparação e publicação de relatórios financeiros e organização de assembleias gerais, etc.<sup>3</sup>
- Dependendo do fundo, taxa de performance<sup>4</sup> (cobrada somente quando o retorno pós-despesas ultrapassa um determinado parâmetro de referência),
- Dependendo do fundo, taxa de ingresso e/ou de saída,<sup>5</sup>
- *Bid-ask spread* nas operações de compra e venda de cotas,<sup>6</sup>
- Taxa de administração,
- Custos de agência.

Os dois últimos tópicos merecem discussão mais aprofundada. No Brasil, a taxa de administração é utilizada para remunerar, segundo o Art. 61 da Instrução CVM nº 409/04, os serviços de gestão, processamento e controle de operações, além da distribuição de cotas, ou seja, a venda. Afora a venda de cotas, os demais custos também seriam incorridos por um investidor individual. Conseqüentemente, a taxa de administração não deve ser considerada um peso adicional ao investidor, a menos que este realmente possua uma estrutura mais eficiente para a execução das atividades pertinentes à gestão de sua carteira.

---

<sup>2</sup> Órgão responsável pela normatização e fiscalização do mercado de valores mobiliários brasileiro.

<sup>3</sup> Despesas atribuídas diretamente ao fundo, de acordo com o Artigo 99 da Instrução CVM nº 409/2004.

<sup>4</sup> Tais taxas podem atuar positivamente. Elton et al. (2003) encontraram que fundos que prevêm taxas de performance (*incentive fees*) em seus regulamentos apresentam melhor seleção de ativos e menores despesas relativas que os demais. Por outro lado, assumem maiores riscos. A taxa de performance busca alinhar os interesses de gestores e cotistas. Porém, devido à sua natureza assimétrica (não há divisão de prejuízos com o gestor), a tomada de risco indiscriminada é um potencial gerador de conflitos.

<sup>5</sup> Sua cobrança é permitida, de acordo com o Art. 61 da Instrução CVM nº 409/04. Tais taxas são remuneração do administrador no Brasil, diferentemente do que se observa em países como Estados Unidos, Inglaterra e Austrália, onde são receitas do fundo, visando proteger investidores de longo prazo dos custos de transação gerados ao fundo pela saída precoce de cotistas de curto prazo (Parwada, 2003).

<sup>6</sup> Não existente no Brasil, porém comum em países como Inglaterra e Austrália (Parwada, 2003). Seu propósito é o mesmo das taxas de ingresso e saída.

A questão é que alguns autores documentaram a existência de economias de escala em despesas administrativas quando associadas ao tamanho do fundo, em termos de patrimônio líquido (Dermine e Roller, 1992; Mahoney, 2004 e Malhotra et al., 2007).<sup>7</sup> Isso tenderia a tornar o investimento via fundos mais eficiente do que a atuação individual.

O fato de, no Brasil, em respeito ao Art. 41 da Instrução CVM nº 409/04, a taxa de administração ser cobrada como um percentual sobre o patrimônio líquido do fundo, logo, linearmente em relação ao seu tamanho, não impede a realização destas economias. Conforme Mahoney (2004) lembra, a premissa econômica básica é que a grande competitividade presenciada na indústria de fundos – fenômeno também válido para o ambiente brasileiro – tenda a trazer a cobrança aos cotistas para o custo marginal de prestação dos serviços. Sendo verdadeira esta premissa, se experimentaria redução natural nos percentuais cobrados como taxa de administração.

Por outro lado, é importante ter em mente que em virtude da racionalidade dos consumidores ser limitada, os preços poderiam se desviar do custo marginal mesmo em mercados de intensa concorrência (Gabaix e Laibson *apud* Mahoney, 2004). Adicionalmente, a presença de custos de procura, estudada no contexto da indústria de fundos por Sirri e Tufano (1998), é fator relevante na formação da taxa de administração e, por conseguinte, pode afetá-la de modo prejudicial ao investidor.

No tocante aos custos de agência, decerto que ambientes competitivos contribuem para sua minimização (Mahoney, 2004). Havendo concorrência, um limitador de atitudes do gestor danosas aos cotistas seria a possível fuga de recursos que ocorreria como reação a um desempenho ruim do fundo (Chevalier e Ellison, 1999). Ainda assim, a literatura traz evidências de que conflitos persistem afetando a relação entre essas pessoas.

Mahoney (2004) argumenta que o mecanismo de remuneração do gestor não é desenhado adequadamente e gera desalinhamento de objetivos. Uma vez que a utilidade

---

<sup>7</sup> Dermine e Roller (1992), no entanto, encontraram que fundos com patrimônio muito grande enfrentam deseconomias de escala. Uma possível explicação para este achado é a restrição relativa que tais fundos exibem à mobilidade de recursos. Em função de apresentarem posições em carteira muito elevadas, geralmente não encontram liquidez no mercado que absorva instantaneamente seus planos de investimento e desinvestimento sem forte impacto nos preços.

do gestor é uma função crescente dos ativos do fundo, há incentivo para que ele aumente as despesas com distribuição (comissões de venda) e promoção do fundo acima do nível ótimo sob o ponto de vista do investidor. As economias de escala presentes na indústria podem manter alinhados os objetivos de ambos até onde o benefício marginal decorrente da redução no custo médio seja igual ao gasto marginal com *marketing*. A partir daí, não existe interesse do cotista em aumentar despesas. Todavia, o mesmo pode não ser válido para o gestor.

Parwada (2003), por exemplo, sugere que as diferenças positivas encontradas entre taxas de *bid-ask spread* de fundos e seus efetivos custos de transação poderiam representar pagamentos na modalidade de “*soft-dollar commissions*”. Tais pagamentos se refeririam a comissões direcionadas acima do valor de mercado a corretores que recomendassem tendenciosamente as cotas do fundo para clientes.

Os fenômenos de custos de agência na indústria de fundos não se esgotam aí. Outros já documentados e que podem ser citados são: “*late trading*” (Mahoney, 2004) e preocupações de gestores com a carreira (Chevalier e Ellison, 1999). No primeiro, gestores permitiriam arbitragem a alguns privilegiados, em troca de aplicações no fundo, através da aceitação de ordens de negociação da cota após o momento estipulado para cálculo e fechamento da carteira, porém ao valor apurado para a cota nesse fechamento. No segundo, gestores, principalmente os novos, tenderiam a evitar risco não-sistemático agindo em “manada” para se igualarem aos índices referenciais de mercado e, assim, aumentar as chances de manterem seus empregos.

Do lado dos benefícios do investimento em fundos, podem ser citados:

- Acesso a emissões exclusivas de títulos e valores mobiliários, disponibilizadas somente a investidores qualificados e institucionais,
- Diversificação automática dos investimentos (Mahoney, 2004), o que permite redução ou até mesmo eliminação do risco não-sistemático através da compra de um ou poucos ativos,
- Utilização de representantes que, por atuarem no mercado em tempo integral, podem reagir mais rapidamente a novas informações (Mamaysky e Spiegel, 2001),

- Delegação das decisões de aplicação de recursos a profissionais potencialmente mais preparados (Mahoney, 2004),
- Redução do custo médio de transações, como, por exemplo, o *bid-ask spread* em operações de balcão e a corretagem (Parwada, 2003), em função de maior poder de barganha em negociações com intermediários,<sup>8</sup>
- Menores custos de transação em bolsa: emolumentos, taxa de liquidação e taxa de registro (esta última válida apenas para operações com derivativos),<sup>9</sup>

Foi mencionado que a aplicação de recursos em fundos traz o benefício da diversificação automática, pois o titular de uma cota é na verdade proprietário de uma fração de cada um dos ativos detidos na carteira do fundo. Foi mencionado também o benefício de oferecer representação a indivíduos que não podem (ou não querem) acompanhar o mercado incessantemente.

Um benefício que aí está implícito é a simplificação da procura ao investidor. Ao invés de diariamente analisar ativos individuais e montar combinações entre os mesmos, ele pode, em intervalos de tempo maiores, comparar características de carteiras já prontas. Para que este trabalho seja eficiente, contudo, são necessários instrumentos que facilitem o processo de coleta e análise destas informações e, conseqüentemente, a decisão do investidor.

Com relação aos custos de procura, vale destacar a conclusão de Sirri e Tufano (1998) de que fatores que os reduzem são determinantes na definição dos fluxos de recursos entre fundos de investimento. Os fatores estudados pelos autores foram: exposição espontânea do fundo na mídia, sua despesa com promoção e tamanho do gestor em termos de *asset under management* (sendo que tamanho seria uma *proxy* para reconhecimento).

Modelos de classificação de fundos de investimento são outro artifício muito comumente utilizado com o intuito de reduzir custos de procura nesta indústria. Tais modelos buscam sumarizar em algumas categorias a informação contida no universo de

---

<sup>8</sup> Parwada (2003) encontrou que este benefício se pronuncia mais em fundos maiores, consistente com economias de escala. O autor não agrega a seu modelo, contudo, outra variável com impacto possivelmente relevante: o giro empreendido pelo fundo aos ativos da carteira.

<sup>9</sup> Conforme tabela de custos operacionais da Bolsa de Valores de São Paulo (BOVESPA).

fundos disponível. Conforme apresentado no Gráfico 1, a oferta brasileira de fundos é bastante significativa e vem crescendo consistentemente. Diante de uma diversidade de opções como essa, a escolha do investidor se torna uma tarefa demasiadamente complexa e custosa. Sendo assim, modelos de classificação de fundos são extremamente bem-vindos.

No Brasil, até pouco tempo atrás, a classificação de fundos existia de um ponto de vista estritamente regulatório e não havia preocupação em refletir a relação risco-retorno do investimento. Sendo estas as duas dimensões que exercem influência sobre a utilidade do investidor, seu conteúdo informacional era, portanto, débil. Varga e Valli (1998) chegaram inclusive a fazer uma crítica neste sentido. Para estes autores, tal informação é imprescindível para que os agentes de mercado possam tomar decisões estratégicas em investimentos.

A lacuna foi suprida em parte pela criação da classificação ANBID e posteriormente também pelo surgimento da Instrução CVM nº 409/04. Não obstante, tais modelos de classificação, além de apresentarem uma quantidade de categorias superdimensionada artificialmente – o que complica desnecessariamente a decisão do investidor – também não distinguem de maneira adequada as categorias por se basearem em critérios definidos subjetiva e arbitrariamente.

Na ANBID, conforme contato com um de seus funcionários, a formulação (ou reformulação) da grade de classificação de fundos envolve a criação de grupos de trabalho com participantes da instituição e do mercado que, através de discussões, tentam mapear a indústria e projetar tendências a fim de gerar subsídios para a definição do número de categorias que melhor a representaria. Tais categorias são caracterizadas pela imposição de limites à composição da carteira, os quais, em última instância, servem como fronteiras entre as categorias. Na CVM, também de acordo com informações prestadas por pessoal interno, o processo não é muito diferente.

Embora as análises e decisões ocorram via órgãos colegiados, o elemento de subjetividade persiste e, conseqüentemente, deixa margem a viés. Moreno et al. (2006) argumentam que categorias mal estabelecidas podem causar danos ao processo de diversificação de investidores pelo fato de não exibirem, apesar da enganosa aparência,

diferenças significativas entre si em termos de risco e retorno. Brown e Goetzmann (1997) defendem que, a fim de atingir consistência na classificação, é imprescindível que o modelo seja determinado objetiva e empiricamente.

Outra fraqueza dos modelos ANBID e CVM é que os mesmos se baseiam na carteira informada pelos próprios administradores do fundo<sup>10</sup>. Inúmeros autores documentaram que classificações baseadas em auto-declarações estão suscetíveis ao problema de desvio moral (*moral hazard*), visto que há incentivo à prestação de informações deliberadamente falsas visando a classificação do fundo numa categoria onde fosse atingida uma posição melhor nos *rankings* de rentabilidade divulgados pela imprensa ou instituições de análise (Brown e Goetzmann, 1997; Kim et al., 2000 e Pattarin et al., 2003).

Tanto a ANBID como a CVM possuem sistemas que filtram estatisticamente as carteiras informadas pelos administradores. Isso certamente reduz erros de classificação, pois caso a carteira do fundo se desvie dos parâmetros da categoria, o filtro acusaria. Entretanto, este procedimento de controle assume que a informação original é prestada corretamente. Nada garante que, na fonte, o administrador a manipule indevidamente.<sup>11</sup> Pattarin et al. (2003) recomendam, portanto, que os procedimentos de classificação se baseiem em retornos passados, pois esta é uma informação que não pode ser distorcida por longos períodos.

Sintetizando, os argumentos colocados acima permitem afirmar que modelos adequados de classificação são de extrema importância para:

- reduzir custos de procura e análise, além de facilitar o processo decisório,
- evitar custos de agência derivados de desvio moral na auto-declaração de objetivos de investimento,
- viabilizar diversificação eficiente de carteiras.

---

<sup>10</sup> No Brasil, o administrador é definido como aquele responsável pelo processamento e registro dos negócios efetuados, bem como pela precificação diária da cota e preparação de relatórios aos investidores, não sendo, necessariamente, responsável pela gestão da carteira, isto é, a decisão de onde aplicar os recursos.

<sup>11</sup> Com relação a isso, convém dizer que a separação das pessoas do gestor e do administrador minimizaria o problema. Todavia, atualmente não há nenhuma obrigação legal nesse sentido no Brasil.

Além de serem importantes para:

- definir corretamente referenciais de desempenho e realizar comparações consistentes entre fundos (Brown e Goetzmann, 1997 e Moreno et al., 2006),
- subsidiar estudos sobre persistência de performance (Moreno et al., 2006),
- auxiliar investidores a explicar melhor os retornos de suas carteiras (Brown e Goetzmann, 1997).

A presente pesquisa, ciente dos problemas apresentados pelos atuais modelos de classificação de fundos, sugere a criação de um novo modelo que busque incorporar as diversas preocupações enunciadas ao longo deste capítulo. Espera-se assim contribuir para o aperfeiçoamento da tomada de decisões financeiras e para o próprio crescimento da indústria. Afinal, é fato estilizado em economia que melhor informação implica em maior eficiência. Considerando a proporção que a indústria de fundos tomou em vários países, fica ainda mais evidente a pertinência de propostas de melhoria neste sentido.

Sendo bem específico na declaração dos objetivos de pesquisa: sugere-se criar um modelo de classificação claro, conciso, teoricamente consistente, e que apresente metodologia objetiva com foco no comportamento realizado dos fundos. O modelo deve identificar fronteiras que efetivamente distingam as categorias em termos de risco-retorno e, adicionalmente, deve servir aos investidores como um eficaz instrumento para explicação de rendimentos futuros. O modelo será aplicado ao caso brasileiro como forma de demonstração da metodologia e também visando testar seu desempenho.

O Capítulo 2 trata da revisão bibliográfica e está dividido em duas etapas. Na primeira, são descritos trabalhos que avaliam e criticam modelos tradicionais de classificação existentes em alguns países. Na segunda, buscou-se identificar modelos alternativos de classificação de fundos já propostos por outros autores. O objetivo com a revisão foi obter subsídios que auxiliassem a produção de um modelo dentro dos critérios estabelecidos nessa dissertação.

O Capítulo 3 apresenta a metodologia empregada na elaboração do modelo e descreve as técnicas quantitativas utilizadas ao longo da pesquisa, enfatizando o papel de cada uma para o alcance dos objetivos propostos.

O Capítulo 4 define o escopo e detalha todo o processo de coleta e preparação de dados para as análises empíricas.

O Capítulo 5 expõe, passo a passo, os resultados obtidos com a aplicação do modelo. Ao fim do capítulo, compara-se o desempenho do modelo proposto com o modelo da ANBID.

O capítulo 6 apresenta as conclusões do trabalho e debate suas limitações. Por fim, são apontadas algumas oportunidades de desenvolvimento que podem ser abordadas em pesquisas futuras.

## **2. REVISÃO DA LITERATURA**

### **2.1. Testes sobre sistemas tradicionais/oficiais de classificação de fundos**

Mayes et al. (2000) desenvolveram um estudo que confrontava se fundos estavam corretamente classificados tomando por base suas exposições a um conjunto de índices de mercado. Aplicando Análise de Estilo<sup>12</sup>, estimaram parâmetros de exposição que foram então utilizados como dados de entrada para prever a que categoria pertencia cada fundo.

A técnica empregada para a etapa de previsão foi a Análise Discriminante<sup>13</sup>. A amostra envolveu 414 fundos americanos, sendo que 240 foram utilizados para estimação e 174 para validação. Observou-se que, na média, houve consonância entre a categoria prevista e os objetivos declarados do fundo em 60,83% dos casos na amostra de estimação, ao passo que na amostra de validação esta medida caiu para 45,98%. Para praticamente todas as categorias, a previsão se saiu melhor do que se esperaria a partir de um procedimento totalmente aleatório. Os autores atribuíram os erros de previsão a três possíveis causas: classificação inicial equivocada do fundo, desvio de comportamento do gestor pós-categorização e/ou má especificação do modelo.

---

<sup>12</sup> No Capítulo 3, são apresentados detalhes sobre esta técnica.

<sup>13</sup> Idem nota anterior.

Kim et al. (2000), em trabalho de natureza similar ao comentado imediatamente acima, conferiram se os objetivos declarados por fundos de investimento americanos estavam em conformidade com uma série de atributos, tais como: retorno médio, desvio-padrão, beta,  $R^2$  (como medida da diversificação do fundo), percentual de ações na carteira, o múltiplo Preço/Lucro das ações, etc.

Os autores operacionalizaram o estudo através de uma Análise Discriminante, a qual foi estimada de maneira iterativa. A cada iteração, os grupos eram reformulados de acordo com a previsão realizada. O procedimento foi repetido até que o percentual geral de acerto na classificação atingisse 99%. A disposição dos fundos nos grupos encontrados pela Discriminante foi então comparada com a disposição da classificação tradicional. Os testes estatísticos constataram que havia diferença significativa entre as duas, sugerindo a necessidade de se melhorar o sistema de classificação vigente à época. No geral, descobriram que 54% dos fundos estavam classificados erroneamente, sendo que 34% estavam severamente mal classificados.

Kim et al. (2000) avaliaram também se havia evidência de desvio moral por parte de gestores. Identificaram que na maior parte dos casos, gestores assumiram menos riscos do que efetivamente declararam. Apesar disso, observaram que a maioria dos fundos que declararam uma política de investimento menos arriscada do que a efetivamente exercida galgaram posições superiores num *ranking* de rentabilidades.

Xue-jun e Xian-lao (2004) foram outros autores que estudaram o posicionamento real de fundos em relação aos objetivos declarados. O ambiente investigado foi a China. Os atributos utilizados para definir objetivos reais de investimento foram: percentual de ações na carteira, retorno médio, desvio-padrão, beta contra a média aritmética dos índices das bolsas chinesas, etc. A metodologia envolveu agrupá-los conforme seus objetivos e testar se a Distância de Mahalanobis entre os centróides dos grupos apresentava significância estatística. A partir de uma matriz de distâncias por pares, concluíram que, a um p-valor de 5%, apenas um, dentre seis pares existentes, era realmente diferente.

Em seguida, utilizaram Análise Discriminante para prever a classificação adequada dos fundos entre as categorias e concluíram que, em média, 50% dos fundos estavam

classificados incorretamente. Os autores não comentam, mas a classificação incorreta pode advir justamente da má distinção entre os grupos com base nos atributos selecionados, fato que tornaria a própria utilização da Análise Discriminante questionável. Há de se destacar também que o trabalho da dupla apresenta uma vulnerabilidade muito grande em relação ao tamanho da amostra, a qual continha apenas 22 fundos.

## **2.2. Propostas de modelos para classificação de fundos de investimento**

Gallo e Lockwood (1997) sugeriram um método de classificação de fundos baseado em Análise de Estilo. Em seu modelo, foi estimada, para 195 fundos mútuos americanos, a exposição do retorno do fundo a quatro diferentes índices de ações, os quais sofriam influência de fatores como tamanho da firma e relação preço/lucro. Quatro classes foram criadas, uma para cada índice, e os fundos foram então destinados à classe referente ao índice ao qual estavam mais expostos.

A fim de avaliar o desempenho do sistema de classificação proposto, Gallo e Lockwood (1997) construíram 500 carteiras fictícias formadas por investimentos igualmente ponderados em quatro fundos. Cada fundo da carteira era proveniente de sorteios aleatórios realizados para cada uma das classes criadas. O mesmo procedimento de construção de carteiras foi realizado tomando por base dois outros sistemas de classificação de fundos difundidos nos Estados Unidos. Adicionalmente, foram geradas carteiras compostas por quatro fundos selecionados aleatoriamente na população. Cada método foi avaliado tendo como critério o Índice de Sharpe para uma amostra de retornos mensais durante o período de 1986 a 1993. Constatou-se que os sistemas de classificação tradicionais apresentaram desempenho equivalente do ponto de vista estatístico ao sistema de seleção aleatória de fundos. Por sua vez, o sistema sugerido pelos autores superou os demais, evidenciando que suas categorias ofereciam maior potencial de diversificação de carteiras.

Por sua vez, Brown e Goetzmann (1997) propuseram um esquema de classificação que aglomera fundos de acordo com as séries temporais de seus retornos. No modelo da dupla, os retornos período a período são pensados como sendo as observações que definem o objeto a ser aglomerado, isto é, o fundo. A aglomeração é feita através de

uma técnica de *cluster* não-hierárquico (*k-means*), onde o número de grupos *k* deve ser definido *a priori*.<sup>14</sup> Para a definição do número de fundos, utilizam como critério de escolha uma medida de verossimilhança com distribuição qui-quadrada.

Os grupos encontrados precisam de análises adicionais para serem interpretados. Uma alternativa levantada pelos autores é a utilização do modelo de Análise de Estilo de Sharpe (1992). A seqüência de etapas sugerida inicia pelo agrupamento de fundos com base nas séries de retornos. De posse do número de grupos e de seus integrantes, estimam-se as exposições de cada fundo a um conjunto de índices aplicando-se Análise de Estilo. Por fim, calculam-se as médias de cada grupo por índice.

Os exercícios empíricos de Brown e Goetzmann (1997) incluíram a comparação do esquema acima com outros métodos de classificação, os quais foram derrotados na tentativa de explicar retornos futuros. Os outros métodos incluem: (a) um tradicional modelo criado por uma empresa americana com base em critérios não-objetivos, (b) o próprio modelo sugerido pelos autores, porém focando na aglomeração de séries de retornos sumarizadas através de análise de componentes principais, (c) a aglomeração de fundos com base nos parâmetros estimados via Análise de Estilo.

Pattarin et al. (2004) replicaram, a fundos italianos, a idéia de Brown e Goetzmann (1997) com algumas alterações. Primeiro, reduziram o tamanho da série utilizando componentes principais. Em seguida, utilizaram um algoritmo distinto de *cluster* não-hierárquico associado a outro critério estatístico de informação para definição do número de grupos. Foi calculada então, para cada grupo, a série temporal contendo a média dos retornos. Cada série foi regredida em função de índices num procedimento de Análise de Estilo. Ao final, os retornos estimados para os fundos pela Análise de Estilo para anos subseqüentes foram comparados com os retornos estimados por um tradicional modelo de classificação de fundos mútuos da Itália, evidenciando melhor desempenho do modelo proposto por Pattarin et al. (2004).

Em um trabalho com resultados de muito boa qualidade, Moreno et al. (2006) apresentaram um novo modelo de classificação de fundos e o aplicaram ao caso

---

<sup>14</sup> Idem nota anterior.

espanhol. As variáveis utilizadas para comparação da similaridade entre fundos foram diversas e buscaram captar efeitos não-lineares nos comportamentos de investimento de gestores. Delas, destaca-se: retorno médio, desvio-padrão, simetria da distribuição de retornos, curtose, as maiores perdas e ganhos, a razão recompensa-semivariabilidade, betas contra dois índices distintos, etc.

A abordagem sugerida por esses autores é feita em dois níveis. Primeiro, se implementa um algoritmo de redes neurais denominado SOM (*Self-Organizing Maps*). Com os dados do SOM, se prossegue ao segundo nível, no qual é efetuada uma Análise de *Clusters* do tipo *k-means*. O número de partições ideal foi definido através do Índice de Davies-Bouldin. Uma vez que os valores de inicialização foram escolhidos de forma randômica, Moreno et al. (2006) tiveram o cuidado de repetir a análise 50 vezes para reduzir o risco de encontrarem pontos de otimização local muito distantes do ótimo global. As categorias encontradas foram confrontadas por um algoritmo de aprendizado artificial que mostrou que o sistema proposto era bem mais preciso em identificar a real categoria do fundo do que o sistema de classificação vigente na Espanha à época.

Em adição ao teste mencionado acima, os autores investigaram se o modelo proporcionava melhores meios de diversificação de carteiras aos investidores, um experimento com lógica semelhante ao realizado por Gallo e Lockwood (1997). O experimento envolveu a simulação de 10.000 carteiras para cada um dos sistemas de classificação. As carteiras foram formadas por investimentos de igual peso em fundos selecionados aleatoriamente de cada uma das categorias do sistema em consideração. Realizou-se este experimento tanto dentro como fora da amostra, isto é, no período da própria estimação e também no período subsequente. Nos dois casos, o modelo de Moreno et al. (2006) obteve desempenho superior, atingindo melhor retorno a níveis mais baixos de risco. Este resultado foi possível mesmo existindo menos fundos nas carteiras formadas para o modelo proposto, fato que evidencia elevado potencial de diversificação quando comparado ao sistema oficial espanhol.

### 3. METODOLOGIA DE PESQUISA

Neste capítulo, se descreve as três técnicas aplicadas ao longo da pesquisa: a Análise de Estilo, a Análise de *Clusters* e a Análise Discriminante Múltipla. Todas são razoavelmente bem difundidas na literatura, e, por conseguinte, dispensam detalhes muito minuciosos. A explanação sobre as mesmas visa deixar explícito apenas o conhecimento fundamental necessário ao entendimento da metodologia empregada. Ao fim do capítulo, é apresentado o papel de cada uma dentro do trabalho e o relacionamento entre elas.

#### 3.1. Análise de Estilo

Segundo Sharpe (1992), é de notório conhecimento que grande parte da variabilidade no retorno de uma carteira de investimentos se deve à alocação dos recursos em classes de ativos<sup>15</sup>. Por classe de ativos, subentende-se uma combinação de vários ativos financeiros com características comuns em termos de correlação com o núcleo da respectiva classe (risco sistemático).

O fundamento teórico da posição do autor parte de um modelo de fatores vastamente difundido na literatura sobre finanças, cuja expressão genérica é a seguinte:

$$R_i = [b_{i1}F_1 + b_{i2}F_2 + \dots b_{in}F_n] + e_i, \quad (1)$$

onde:

$R_i$  é o retorno do ativo  $i$ ,

$F_j$  é o prêmio do fator  $j$ ,

$b_{ij}$  é a sensibilidade de  $R_i$  em relação a  $F_j$ ,

$e_i$  é um componente idiossincrático do retorno de  $i$ , o qual não é associado aos fatores.

Neste modelo, assume-se que o resíduo de um ativo  $i$  é não correlacionado com o de qualquer outro ativo  $m$ .

---

<sup>15</sup> Varga e Valli (1998) são outros autores que suportam essa visão.

A proposta inovadora trazida por Sharpe (1992) baseava-se num caso especial do modelo de fatores e tinha como objetivo aperfeiçoar os métodos de avaliação de desempenho de carteiras. Denominada comumente de Análise de Estilo (*Style Analysis*), a técnica representava os fatores acima através de séries temporais de retornos de um conjunto de classes de ativos e restringia as sensibilidades para somarem 100%.

O termo entre parênteses na Equação 1 foi então interpretado como sendo a parcela do retorno proveniente do estilo da carteira, enquanto que o componente idiossincrático, a parcela atribuível à seleção específica de ativos dentro de cada uma das classes consideradas. Numa carteira composta por investimentos em vários fundos mútuos, o estilo se torna, na visão de Sharpe (1992), ainda mais expressivo no sentido de explicar seu retorno, pois, através da diversificação, seria possível reduzir fortemente – ou até mesmo eliminar – o risco trazido pelo termo residual, *vis a vis* a referida premissa sobre ausência de correlação.

Ante o exposto, é de extrema relevância que investidores conheçam o estilo dos gestores para os quais se está delegando a decisão de aplicação de recursos. Afinal, este será o principal determinante das duas dimensões de interesse: retorno e risco. Existem duas possíveis maneiras de se descobrir o estilo de um fundo específico: através do exame detalhado de sua carteira ou através da análise do seu retorno.

Na primeira, de acordo com Le Sourd (2006), todos os títulos integrantes da carteira são estudados individualmente e categorizados de acordo com diferentes atributos. Os resultados são então agregados para se chegar ao estilo. O método requer, portanto, a lista completa dos ativos mantidos pelo fundo, assim como seus respectivos pesos. A análise deve ser efetuada regularmente, de modo a acompanhar modificações nas características dos títulos. Logo, sua adoção exige informação e tempo numa quantidade muitas vezes indisponível. Adicionalmente, defendem Brown e Goetzmann (1997), a análise de estilo baseada na carteira está sujeita ao problema de “*window-dressing*”, fenômeno no qual gestores deliberadamente desfariam posições em ativos com performance histórica recente ruim ao término de períodos considerados chave (por exemplo: ao fim de cada mês, quando há apuração de balancetes), uma vez que tais análises são geralmente efetuadas tomando por base estas datas.

A segunda alternativa trata-se do modelo proposto por Sharpe (1992), onde o ativo  $i$  passa a ser representado pela cota de um fundo.<sup>16</sup> Neste modelo, o estilo é definido exclusivamente pelos  $b_{ij}$ . Tais parâmetros refletem a exposição do retorno do fundo ao conjunto de classes de ativos. O que lhe confere eficiência é que não há necessidade de se esmiuçar a carteira para determinar o estilo, mas sim colocar em prática um simples procedimento de estimação. Para que a estimação seja eficaz, todavia, deve-se observar algumas diretrizes teóricas durante a etapa de seleção de classes de ativos. Segundo Sharpe (1992), é desejável, porém não necessariamente mandatório, que as classes:

- sejam mutuamente exclusivas, o que significa que nenhum título aparece em mais de uma,
- sejam exaustivas e, portanto, representem todos os ativos financeiros existentes,
- possuam baixa correlação ou, nos casos em que a correlação é alta, tenham desvios-padrão razoavelmente diferentes.

Merece atenção também o fato de que o domínio dos parâmetros deve refletir a realidade dos fundos. Por este motivo, é de praxe em trabalhos aplicados de Análise de Estilo o estabelecimento de restrições de desigualdade sobre os  $b_{ij}$  a fim de manter seus valores em intervalos condizentes com as práticas de investimento observadas na indústria. Alguns exemplos são: Sharpe (1992), Varga e Valli (1998) e Mayes et al. (2000).

Ao fim da estimação, é recomendável que se avalie o potencial explicativo do modelo e, para isso, Sharpe (1992) sugere a utilização do  $R^2$  como medida do ajuste. Fundos com alto  $R^2$  teriam seu retorno devido majoritariamente ao estilo. Logo, representariam fundos com estratégia de gestão predominantemente passiva, referenciadas geralmente por um ou alguns índices de mercado. No outro extremo, fundos com  $R^2$  baixo seriam fundos com retorno ligados principalmente à seleção de ativos específicos dentro das classes e, por isso, empreendedores de uma gestão ativa na definição da carteira. Esta

---

<sup>16</sup> Faz-se questão de frisar que modelos de fatores e o de Sharpe (1992) são – até onde se sabe – as únicas alternativas existente para estimação do estilo de um fundo tomado individualmente. A técnica de avaliação de estilo proposta por Brown e Goetzmann (1997), por exemplo, é capaz de identificar estilos de grupos, não de um fundo específico. Ademais, não permite a interpretação do estilo do grupo sem análises adicionais.

conclusão é, contudo, intimamente dependente de uma seleção de classes razoavelmente em linha com as diretrizes teóricas listadas acima. Caso contrário, o componente idiossincrático absorveria os problemas derivados da má especificação do modelo e, conseqüentemente, apresentaria uma parcela de retorno atribuível à seleção equivocadamente superavaliada (Mayes et al., 2000).

Por fim, cumpre destacar que o estilo estimado para um fundo através da Análise de Estilo é, na verdade, a média referente ao período sob estudo (Sharpe, 1992). Naturalmente, este estilo pode variar ao longo do tempo.

### **3.2. Análise de *Clusters***

Segundo Sharma (1996), a Análise de *Clusters* é uma técnica usada para combinar objetos em grupos ou conglomerados tal que:

- cada grupo possua objetos homogêneos em relação a determinada(s) característica(s) e
- os objetos de outros grupos sejam diferentes quando tomada(s) esta(s) mesma(s) característica(s) como referência.

A qualidade da Análise de *Clusters*, apontam Hair et al. (2005), está em sua capacidade de classificar objetos. Isto é útil para diversas finalidades, por exemplo: reduzir um grande número de objetos desorganizados a algumas categorias compreensíveis e gerenciáveis.

Vários são os métodos disponíveis para a Análise de *Clusters*, porém, Hair et al. (2005) afirmam que as abordagens seguem, em geral, dois tipos de estratégia: a de hierarquização ou a de partição prévia, esta última também denominada freqüentemente de *k*-médias (*k-means clustering*). Os algoritmos utilizados em cada um destes métodos também são inúmeros. Vão desde procedimentos heurísticos até modelos mais complexos e formais, fundamentados em distribuições estatísticas (Fraley e Raftery, 1998).

A presente dissertação utiliza dois métodos distintos, o primeiro baseado em hierarquização e o segundo, em partição prévia, ambos heurísticos. De acordo com Sharma (1996) e Hair et al. (2005), as análises de *cluster* hierárquico e do tipo *k*-médias podem ser vistas como complementares. A primeira seria utilizada para definir o número natural de grupos e também para fornecer os centróides iniciais para a análise de *k*-médias, cuja função seria refinar o arranjo dos objetos entre os grupos. Aplicando-se este procedimento, flexibiliza-se a movimentação dos objetos entre grupos (impossibilitada pelo desenho em árvore dos processos hierárquicos), ao mesmo tempo em que se reduz a vulnerabilidade que o método de *k*-médias possui à seleção dos centróides iniciais (inclusive a eventual fragilidade, apresentada por alguns algoritmos, ao ordenamento dos dados no arquivo de entrada).

Hair et al. (2005) lembram que os métodos hierárquicos possuem duas lógicas distintas: a aglomerante ou a divisora. Enquanto a primeira inicia dos objetos individuais e os aglutina passo a passo em grupos até chegar a um conglomerado único, a segunda parte de um grupo geral e o divide até que todos os objetos estejam segregados. Segundo estes autores, porém, não há diferenças relevantes entre um e outro. Simplesmente percorrem caminhos inversos.

O algoritmo de aglomeração hierárquica utilizado foi o de ligação completa entre grupos (*complete linkage*). Nele, a similaridade entre dois *clusters* é definida como a máxima diferença entre todos os possíveis pares de objetos. Ao leitor interessado em conhecer detalhes sobre o algoritmo, recomenda-se consultar Sharma (1996). Este autor menciona que estudos de simulação revelaram que o método de ligação completa é menos afetado do que o método de ligação simples (*single linkage*) na presença de distúrbios nos dados, além de ser menos suscetível ao fenômeno de encadeamento, isto é: a propriedade de associar indevidamente objetos a grupos existentes ao invés de formar novos grupos. Há de se mencionar, contudo, que existem diversos outros algoritmos. A escolha entre eles é, em última instância, uma decisão do pesquisador.

Com relação ao método de *k*-médias, o algoritmo empregado foi o de *nearest centroid sorting*, único disponível no *software* SPSS (Statistical Package for the Social Sciences) versão 13.0, ferramenta que serviu de base para esta etapa do trabalho. Esse algoritmo designa indivíduos ao grupo cujo centróide se situa mais próximo (Norusis, 1988). A

fim de atingir maior precisão na classificação dos objetos, o SPSS oferece a opção de classificar e iterar. A cada rodada deste processo, os centróides são recalculados e os objetos são reclassificados até alcançar um pré-determinado critério de convergência ou um número máximo de iterações também estabelecido pelo usuário. Os centróides podem ser reavaliados a cada novo objeto agrupado (*running means*), ou somente após todos terem sido classificados. Nesta dissertação, adotou-se a segunda alternativa visando eliminar por completo problemas derivados de ordenamento dos objetos no arquivo de entrada<sup>17</sup>.

É importante observar que antes de desempenhar a Análise de *Clusters* pelos métodos aqui descritos, deve-se definir um critério para medir a similaridade entre objetos. Conforme defende Sharma (1996), este critério depende exclusivamente dos propósitos do estudo.

De acordo com Hair et al. (2005), existem três tipos fundamentais de medidas: as de correlação, as de distância e as de associação. O primeiro tipo é utilizado para averiguar correspondência entre objetos através de padrões de movimento em suas variáveis. O segundo, para averiguar proximidade nas magnitudes destas variáveis. O terceiro e último, por sua vez, é um caso especial de medida que deve ser empregado quando os objetos forem descritos apenas por variáveis independentes não-métricas (nominais ou ordinais).

Por fim, cumpre mencionar que a revisão da literatura indica que os métodos hierárquicos não suportam bem arquivos com grande quantidade de dados ou por questão de capacidade e tempo de processamento ou por capacidade de armazenagem (Fraley e Raftery, 1998; Hair et al., 2005 e Norusis, 2008). À medida que cresce o número de objetos, cresce exponencialmente a matriz de distâncias que precisa ser analisada. Sem contar que o número de passos no processo de aglomeração também aumenta. Norusis (2008) afirma que um arquivo com 1.000 objetos já pode ser considerado extenso para os procedimentos hierárquicos de Análise de *Clusters*

---

<sup>17</sup> Não se deve confundir o problema de ordenamento dos objetos com o problema de ordenamento da série contendo os centróides iniciais. Este último pode afetar os resultados somente se houver casos de empate na medida de similaridade entre dois objetos distintos e um determinado centróide, algo muito improvável de ocorrer dado o nível de precisão dos números no SPSS, os quais passam de 15 casas decimais.

disponíveis no SPSS. Conforme apresentado mais adiante na seção 4.1, a população sob estudo monta a 5.989 objetos.

A autora recomenda, portanto, que seja utilizado o método de dois estágios (*two-step cluster*) também disponível no SPSS. Este método, todavia, assume a premissa de que a distribuição de cada uma das variáveis características dos objetos seja Normal. Outra condição à obtenção de melhores resultados é que as variáveis sejam independentes entre si. Por fim, o método é vulnerável ao ordenamento dos objetos no arquivo de entrada, fato que pode ser reduzido, porém não eliminado, através de ordenamento randômico.

Num trabalho exploratório onde o número natural de grupos é desconhecido (tal qual é o caso presente), o não atendimento às premissas é um perigo, pois não existe meio de se avaliar os resultados da análise *a posteriori*. Há que simplesmente se acreditar que o número de grupos encontrado é o mais apropriado, o que configura uma vulnerabilidade relevante.

Uma segunda alternativa seria a sugerida por Hair et al. (2005): coletar uma amostra, desde que representativa da população estudada, e manter a estratégia de empregar o método hierárquico. É exatamente a abordagem adotada aqui. Preferiu-se esta alternativa antes mesmo de testar as premissas do *cluster* de dois estágios, porque insistir na proposta de Norusis (2008), visando usar toda a base disponível, representava um risco de desperdício de tempo significativo. O esforço despendido para estimar os parâmetros de todos os 5.989 fundos na Análise de Estilo seria em vão caso fossem rejeitadas as hipóteses do *cluster* de dois estágios. Considerando que é razoavelmente raro encontrar suporte nos dados para as mesmas, este risco se tornava bastante concreto.

De fato, a decisão foi acertada. Conforme demonstram os testes de normalidade realizados na seção 5.3 para a amostra coletada, a hipótese foi rejeitada para absolutamente todas as variáveis e com boa margem de confiança. Adicionalmente, pela Tabela 15, verificou-se que existe correlação entre as mesmas, evidência que atesta a não independência.

### 3.3. Análise Discriminante Múltipla

A Análise Discriminante Múltipla é uma técnica utilizada com ao menos um dos objetivos abaixo listados (Sharma, 1996; Hair et al., 2005):

- identificar se três ou mais grupos podem ser distinguidos a partir de uma determinada combinação de variáveis independentes;
- identificar quais variáveis são significantes para fazer a distinção entre cada um dos grupos;
- construir dimensões que sumariam as variáveis independentes e as substituam na função de discriminar os grupos;
- desenvolver uma regra para prever a qual dos grupos um determinado objeto pertence, isto é, prever sua classificação.

No presente trabalho, é suficiente ater-se apenas ao último tópico.

Segundo Sharma (1996), um problema de classificação se resume a dividir um espaço  $n$ -dimensional, onde  $n$  representa o número de variáveis relevantes à discriminação dos objetos, em regiões mutuamente exclusivas e coletivamente exaustivas. Cada região refletirá um grupo, implicando, portanto, que um determinado objeto será classificado no grupo correspondente à região em que estiver localizado.

O ponto crítico nesta estratégia está em como dividir o espaço  $n$ -dimensional. Sharma (1996) afirma que uma regra deve ser criada tal que um dado critério seja atendido, por exemplo, minimizar o número de objetos classificados erroneamente ou o custo das classificações incorretas (caso haja custos assimétricos entre os grupos). Ele acrescenta que vários são os métodos disponíveis para atingir este objetivo e, dependendo das condições, podem ser matematicamente equivalentes.

A classificação via Análise Discriminante é um destes métodos. Nela, a regra é baseada no chamado valor discriminante (*discriminant score*) e é extremamente simples de ser implementada. Uma vez computadas as funções que determinam estes valores para cada grupo, deve-se classificar um objeto no grupo cujo valor discriminante seja o maior. De

acordo com Sharma (1996), o valor discriminante para um dado objeto  $i$  num grupo  $j$  qualquer ( $d_{ij}$ ) é descrito por:

$$d_{ij} = \mu_j' \Sigma_j^{-1} y_i - 1/2 \mu_j' \Sigma_j^{-1} \mu_j + \ln p_j, \quad (2)$$

onde:

$\mu_j$  é um vetor  $n \times 1$  das médias (centróides) do grupo  $j$  para as  $n$  variáveis discriminantes,

$\Sigma_j$  é a matriz de variância-covariância entre as  $n$  variáveis discriminantes,

$y_i$  é um vetor  $n \times 1$  referente às realizações do objeto  $i$ ,

$p_j$  é a probabilidade prévia de um objeto qualquer na população pertencer ao grupo  $j$ .

O termo  $\mu_j' \Sigma_j^{-1}$  da Equação 2 representa os coeficientes da função classificatória para o grupo  $j$ . Os termos restantes do lado direito, em conjunto, fornecem a constante.

A derivação desta regra é baseada na teoria de decisão estatística e assume que  $\Sigma_j = \Sigma$ , ou seja, as matrizes de variância-covariância são homogêneas entre grupos. Adicionalmente, trata a distribuição de  $y$  como Normal multivariada. O desrespeito às premissas pode afetar negativamente o desempenho classificatório da Análise Discriminante, porém não necessariamente impede sua adoção. Suporte a esta alegação é oferecido por Sharma (1996) quando afirma que os resultados da técnica são bastante robustos a casos de violação.

O fato é que a qualidade da previsão deve ser sempre avaliada e, de preferência, externamente. Isto significa que, após concluído, o modelo deve ser testado em objetos não contemplados durante o processo de estimação. Neste sentido, separar a amostra disponível em duas, uma para efetuar a análise e outra para validá-la, é uma prática bastante comum, sugere a literatura (Sharma, 1996 e Hair et al., 2005). O objetivo com este procedimento é eliminar qualquer viés da medida empregada à avaliação do desempenho.

A viabilidade de se desafiar o modelo *a posteriori* e sem distorções de medida é, logo, um aspecto de grande valia, pois permite que o pesquisador aplique a Análise Discriminante com fins classificatórios mesmo diante de dados não aderentes às premissas. O que realmente definirá sua utilização é o resultado final. Caso seja acurado, se aceita o modelo. Caso contrário, deve ser descartado. Raciocínio simples e pragmático. Afinal, em previsões, o que se deseja é errar o mínimo possível. Isto impõe, inclusive, que mesmo modelos atendendo às premissas devem ser rejeitados quando o desempenho preditivo for insatisfatório.

### **3.4. Seqüência lógica da metodologia**

Apresentadas as técnicas utilizadas, cabe explicar ao leitor como elas se interconectam e contribuem para o alcance dos objetivos da dissertação. Recapitulando, deseja-se classificar fundos de investimento em categorias facilmente compreensíveis e de acordo com seus perfis de risco e retorno.

Antes de continuar com a explanação da metodologia, vale neste momento refletir sobre o fato de que o perfil do fundo pode variar no decorrer do tempo. Isto poderia representar uma ameaça ao êxito da proposta. Afinal, caso fosse muito volátil, seria completamente absurdo atribuir ao fundo uma categoria, pois a mesma poderia não ser válida no dia seguinte. Conseqüentemente, não teria conteúdo informacional nenhum ao investidor.

Todavia, algumas evidências indicam que o comportamento de gestores deve ser razoavelmente estável. O primeiro é a própria existência de inúmeros sistemas de classificação em diversos países. Uma vez que perduraram por tanto tempo, é improvável que sejam completamente inúteis.

Ademais, deve ser considerado que não há incentivo aos gestores para mudarem de perfil muito freqüentemente. Primeiro, porque os custos de transação envolvidos prejudicariam o retorno do fundo (Malhotra et al., 2007). Segundo, porque estes devem ter o que ofertar no mercado a fim de atrair aplicações. Como quaisquer consumidores, investidores desejam informações sobre o produto que comprem, desejam conhecer suas características a fim de formarem um juízo sobre seu valor. Um fundo de personalidade

volátil, por não possuir objetivos definidos e consistentes, não tem como responder a tais questionamentos durante seu processo de captação de clientes e termina não atendendo nenhum segmento de mercado. Com relação a isto, cumpre citar a afirmação de Sharpe (1966) de que gestores de fundos mútuos devem selecionar uma atitude em termos de risco e retorno esperado para então convidar investidores com preferências similares a participar como cotistas do fundo.

A própria regulamentação brasileira também obriga um posicionamento claro por parte dos gestores a respeito da política de investimento. A Instrução CVM nº 409/04 exige, através de seus Artigos 29 e 40, a divulgação de informações em linguagem acessível sobre as metas e objetivos do fundo, o público-alvo, a política de investimento, os riscos inerentes, etc.

Decerto que a promessa de um gestor não é garantia de cumprimento, nem tampouco exigências de regulamentação, pois os sistemas de controle são imperfeitos. Inclusive a própria revisão de literatura trouxe evidência disso. Todavia, os desvios entre comportamento declarado e efetivo apresentados no Capítulo 2 não implicam necessariamente em volatilidade ao longo do tempo. O que foi mostrado foram situações estáticas de não-conformidade. A bibliografia contendo análises dinâmicas é controversa.

Os trabalhos empíricos de Sharpe (1992), por exemplo, suportam a idéia de estabilidade no estilo médio de fundos. No ambiente brasileiro, Varga e Valli (1998) também apresentam resultado com evidência neste sentido. Muito embora Brown e Goetzmann (1997) e Kim et al. (2000) tenham encontrado variações dinâmicas nas atitudes de fundos, a conceituação de estilo nesses trabalhos é diferente da aqui praticada. Por outro lado, Pattarin (2004), utilizando o mesmo conceito de Brown e Goetzmann (1997), obtiveram resultados que demonstraram certa estabilidade.

Tomados coletivamente, os argumentos expostos permitem assumir a premissa de que gestores mantêm alguma estabilidade de comportamento de investimento. A premissa é bastante razoável. Contudo, se reconhece que não é uma verdade absoluta. A melhor alternativa seria confrontá-la empiricamente, procedimento que não é adotado na presente dissertação. Para tanto, o comportamento dos fundos deveria ser avaliado no

decorrer do tempo e testado. Na seção 6.2, sugere-se uma maneira inovadora de atingir tal objetivo em pesquisas futuras.

Retornando à descrição da parte metodológica, a representação do perfil foi feita com base no estilo do fundo. Para sua estimação, foi obviamente útil a técnica de Análise de Estilo. Conhecidos os estilos, foi possível identificar similaridades entre os mesmos e, com base nestas similaridades, definir um número de categorias ideal para agrupar os fundos (método hierárquico de *clusters*) para, em seguida, alocá-los da melhor forma entre estas categorias (método de *k*-médias).

A Análise Discriminante teve como propósito contornar a debilidade que o método hierárquico de aglomeração apresenta, em termos de eficiência computacional, face a uma grande quantidade de dados. Acrescentando a Discriminante como etapa seguinte à Análise de *Clusters*, foi possível criar um modelo geral de classificação trabalhando apenas com uma amostra da população. Neste contexto, seu papel foi identificar padrões subjacentes aos grupos encontrados pela Análise de *Clusters* e, de posse desta informação, prever a qual destes grupos pertencia um fundo qualquer não contemplado na amostra. Ao mesmo tempo em que permitiu economizar tempo e recursos, a Análise Discriminante serviu também como um instrumento de verificação da consistência das fronteiras encontradas para as categorias.

Percebe-se, portanto, que a estratégia de classificação aqui proposta difere das de Brown e Goetzmann (1997) e de Pattarin et al. (2004) na ordem de utilização das técnicas de Análise de Estilo e Análise de *Clusters*. Embora os primeiros autores tenham testado também a classificação via os parâmetros estimados pela Análise de Estilo e tenham verificado que a aglomeração a partir das séries de retorno intocadas se mostrava melhor, alguns pontos sobre sua metodologia representam fraquezas e devem ser discutidos.

Em primeiro lugar, a definição do número de grupos ideal para separação dos fundos se baseou em um critério sujeito à premissa de normalidade, a qual não foi atendida em sua integralidade.

Segundo, o teste sobre a consistência dos grupos formados pelo método de  $k$ -médias, feito através de recursivas estimações com base em *bootstrapping*, evidenciou um erro estatístico médio relevante. Com relação a isto, vale citar inclusive Pattarin et al. (2004), os quais reconhecem que o método de  $k$ -médias está especialmente sujeito a encontrar mínimos locais dependendo dos centróides de entrada. Sharma (1996) e Hair et al. (2005) são outros que também apontam a fragilidade geral dos métodos não-hierárquicos de *cluster* à definição dos centróides iniciais. Sharma (1996) informa ainda que a seleção aleatória costuma ter desempenho bastante ruim. Como alternativa, sugere a utilização dos centróides estimados pelo método hierárquico como valores de inicialização, o que, segundo ele, torna a qualidade dos resultados muito superior. Vale lembrar que este é o procedimento aplicado na presente dissertação.

Por último (e mais importante), Brown e Goetzmann (1997) são negligentes na definição das classes de ativos para a execução da Análise de Estilo. Nada mencionam sobre o processo de escolha dos índices que evidencie cuidado nesta etapa essencial à boa aplicação da técnica. Muito pelo contrário, ao fim do artigo, comentam que uma desvantagem de se fazer a classificação com base nos parâmetros da Análise de Estilo é que geralmente existem colinearidades entre as séries de retornos das classes de ativos. Este fato implicaria em uma estimação imprecisa. Na seção 4.2, o leitor observará que é possível adotar estratégias de redução destas colinearidades, logo, tornando inócua a crítica destes autores.

## **4. APRESENTAÇÃO DOS DADOS**

Neste capítulo, é apresentado todo o processo de coleta e organização de dados para a execução da Análise de Estilo. São definidos o escopo e a amostra e prestadas informações pertinentes às variáveis dependente e explicativas.

### **4.1. A variável dependente**

A variável dependente na Análise de Estilo de um determinado fundo é representada, neste estudo, por dados diários de retornos brutos (antes de imposto de renda) sobre a

aplicação em sua cota.<sup>18</sup> Para a realização dos trabalhos, colheu-se uma série correspondente a exatamente quatro meses em dias úteis, com início em 02 de abril de 2007 e fim em 31 de julho de 2007. Este intervalo é doravante denominado de “período-base do estudo”.<sup>19</sup> São 84 observações no total, o que confere bom número de graus de liberdade para o cômputo do estilo.

Os dados de retornos sobre as cotas dos fundos foram obtidos com a ANBID, através de sua base denominada SI-ANBID. Para constar nesta base, o fundo deve cumprir certas exigências estabelecidas pelo Código de Auto-Regulação da ANBID para Fundos de Investimento. Exigências estas ligadas à adoção das melhores práticas correntes na indústria, por vezes mais rígidas do que a própria regulamentação vigente. Tornar-se assinante deste Código é definitivamente mais custoso para o fundo. Todavia, lhe concede o Selo ANBID, um atestado de qualidade e padronização do processamento e registro de transações, assim como da transparência na divulgação de informações, conferindo-lhe atratividade perante o investidor, razão pela qual a quase totalidade dos fundos existentes no Brasil está contida nesta base. Destaca-se que, mais a frente, são apresentadas estatísticas que corroboram esta afirmação sobre a abrangência da SI-ANBID.

A base, além de fornecer os dados sobre retornos, também fornece a categoria do fundo (de acordo com a Classificação ANBID vigente à época<sup>20</sup>). Merece atenção o fato de que grande parte das categorias ANBID deve sofrer classificação adicional para se adequar à legislação tributária. No Brasil, a Lei nº 11.033 de 21 de dezembro de 2004 estabeleceu a diferenciação entre fundos de curto e longo prazo. As carteiras dos primeiros devem ser formadas por títulos com prazo médio igual ou inferior a 365 dias. O tratamento entre os dois se distingue pela alíquota do “come-cotas” (antecipação do imposto de renda recolhida semestralmente de acordo com a Lei nº 10.892/04) e pela

---

<sup>18</sup> Estes retornos estão obviamente deduzidos de taxa de administração e outros custos operacionais, os quais são apropriados por regime de competência ao resultado do fundo.

<sup>19</sup> A série temporal disponível ao autor continuava até o dia 17 de agosto de 2007. No entanto, os dias restantes foram utilizados para validação prática do modelo proposto, etapa discutida na seção 5.4.

<sup>20</sup> Até 16 de maio de 2007, era válida a Deliberação ANBID nº 29/06. Após esta data, entrou em vigor a classificação aprovada pela Deliberação ANBID nº 32/07, cuja grade está disponível no endereço [http://www.anbid.com.br/documentos/download/auto\\_regulacao/fundos/deliberacoes\\_e\\_pareceres/Deliberao32-anexo1.pdf](http://www.anbid.com.br/documentos/download/auto_regulacao/fundos/deliberacoes_e_pareceres/Deliberao32-anexo1.pdf). As poucas alterações foram a exclusão de três categorias ligadas ao câmbio e a mudança de nome de uma outra categoria. Os dados utilizados na presente dissertação estão em conformidade com a transição imposta pela Deliberação nº 32/07.

tributação de acordo com o intervalo de permanência no fundo. Obviamente, fundos compostos exclusiva ou predominantemente por ações são exemplos de fundos não atingidos por esta legislação, simplesmente porque ações são títulos sem vencimento.

Cabe neste momento salientar que a utilização de retornos brutos como base para a Análise de Estilo é ideal justamente por permitir tratamento estatístico isonômico a todos os fundos durante o processo de categorização. Isso não significa, porém, ignorar o ambiente legal brasileiro. Nada impede que, após executada a classificação pelo método proposto, seja feita a segmentação de acordo com o perfil tributário do fundo.

A informação contida na base sobre a que categoria ANBID pertence o fundo é obviamente útil para efeitos de comparação, visto que se pretende criar uma nova classificação. Adicionalmente, foi útil para eliminar alguns tipos de fundos que não fazem parte do universo sob estudo. Este foi o caso dos classificados como “*Off-shore*”. Uma vez que delimitamos o escopo para considerar somente fundos brasileiros, os “*Off-shore*” não se enquadraram, pois, segundo a própria ANBID, são “constituídos fora do território brasileiro”.

Foram eliminados também os fundos classificados como “Fechados”, isto é, que não estão disponíveis para a adesão de um indivíduo qualquer, ou indivíduo que atenda determinada condição de capital mínimo, mas apenas para um grupo restrito de fundadores e convidados. A razão para este procedimento está ligada à capacidade de intervenção e controle que o cotista de um fundo fechado geralmente tem na sua gestão. Não raro, ele está presente inclusive na sua concepção, momento em que tem a oportunidade de definir a política de aplicação dos recursos. Por este motivo, tal investidor não necessita de uma classificação que lhe auxilie na avaliação do perfil do fundo. Ele próprio pode escolher ou ao menos influenciar o seu destino.

Nos fundos abertos, ofertados ao público geral como produtos de instituições credenciadas pela CVM a operá-los, ao contrário, o poder do cotista pessoa física e até mesmo pessoa jurídica é frequentemente irrelevante em assembleias devido à insignificância de sua participação em relação ao patrimônio total do fundo. Portanto, a adesão só deve ocorrer caso a proposta atenda às expectativas de investimento. Conforme exposto na Introdução do trabalho, em função da oferta de fundos ser muito grande, é socialmente eficiente a prestação de informações padronizadas que facilitem a

decisão de aplicação dos recursos para o investidor. Sendo assim, justifica-se incluir na amostra os fundos abertos, ao passo que o mesmo não é válido para os fundos fechados.

Por último, foram eliminados os fundos denominados como “Investimento no Exterior”. A justificativa para esta exclusão foi tornar o modelo de Análise de Estilo mais parcimonioso, ou seja, dependente da estimação de um número menor de parâmetros. Embora estes não fossem, durante o “período-base do estudo”, os únicos fundos autorizados pela CVM a aplicar recursos em títulos emitidos no exterior, eram os únicos que efetivamente agiam desta maneira num montante relevante da carteira. Logo, excluindo-os da Análise de Estilo, foi possível descartar classes de ativos representativas de investimentos disponíveis fora do Brasil e assim reduzir o número de variáveis explicativas.

Vale enfatizar que as afirmações acima têm fundamento na regulamentação vigente à época e também em discussões com profissionais do setor. No que tange à regulamentação, deve ser notado que os fundos de “Investimento no Exterior” não só estavam autorizados a investir externamente como eram obrigados a fazê-lo. De acordo com a ANBID, a categoria em referência segue o Artigo 96 da Instrução CVM nº 409/04, cujo conteúdo sempre determinou a aplicação de ao menos 80% do patrimônio líquido em títulos representativos da dívida externa de responsabilidade da União, permitindo a aplicação dos eventuais 20% restantes em outros títulos de crédito transacionados no mercado internacional.

Todas as demais categorias de fundos estavam, na prática, até março de 2007, impedidas legalmente de investirem fora do Brasil pela redação originalmente dada ao *caput* do Artigo 86 presente na Instrução CVM nº 409/04:

*“Os títulos e valores mobiliários, bem como outros ativos financeiros integrantes da carteira do fundo, devem estar devidamente custodiados, registrados em contas de depósitos específicas, abertas diretamente em nome do fundo, em sistemas de registro e de liquidação financeira de ativos autorizados pelo Banco Central do Brasil ou em*

*instituições autorizadas à prestação de serviços de custódia pela CVM.”*

A Instrução CVM nº 450/07, publicada em 30 de março de 2007 e vigente a partir daí, alterou o Artigo 86 da Instrução CVM nº 409/04 para tratar de outro assunto completamente diferente e, paralelamente, permitiu, através de mudança no Artigo 85 da mesma Instrução, a aplicação no exterior, pelos fundos de investimento classificados como “Multimercado” (segundo a classificação CVM), de até 20% do patrimônio líquido, enquanto que para os restantes limitou em até 10%.

Não obstante, conforme informações de profissionais atuantes em instituições prestadoras de serviços de administração e/ou gestão de fundos, os fundos brasileiros preferiram predominantemente manter seus recursos em ativos negociados no mercado nacional mesmo após o anúncio da Instrução CVM nº 450/07. Possíveis explicações não exaustivas para o fenômeno seriam: (a) o receio de gestores de aplicar em mercados desconhecidos ou (b) capacidade de replicar internamente qualquer classe de ativo existente externamente por haver fortes correlações entre estes mercados.

Este comportamento razoavelmente homogêneo entre gestores de fundos classificados em outras categorias que não a “Investimento no Exterior” permite, portanto, a exclusão de classes de ativos internacionais da análise. Fique claro, porém, que o modelo aqui proposto deve sofrer atualização e expansão das variáveis explicativas, caso se verifique mudança na postura destes gestores.

Concluída a delimitação do escopo, partiu-se para a quantificação do universo estudado. Foram identificados 6.138 fundos de investimento ativos durante todo o “período-base do estudo”, inclusive de previdência<sup>21</sup>. Descontando os “Fechados”, os de “Investimento no Exterior” e os “*Off-shore*”, os quais contabilizavam, respectivamente: 36, 46 e 67, chegou-se a um universo de 5.989 fundos.

---

<sup>21</sup> Contempla fundos de previdência privada aberta, além dos Fundos de Aposentadoria Programada Individual (FAPIs). Não considera os fundos de pensão, tais como Previ, Petros ou quaisquer outras entidades fechadas de previdência privada.

A efeito de comparação, a Tabela 1 apresenta dados coletados junto à CVM e que mostram o total de fundos existentes no Brasil nos meses em foco. Somente os fundos “*Off-shore*” não são considerados nesses números. Observando a tabela, percebe-se que a base SI-ANBID é realmente muito abrangente, fato que oferece confiança de que os resultados do modelo podem ser generalizados para o universo analisado caso sejam respeitados procedimentos aleatórios na obtenção de amostras.

Mês	Quantidade de fundos
abril/07	6.578
maio/07	6.590
junho/07	6.899
julho/07	7.130

Fonte: CVM

**Tabela 1 – Quantidade de fundos existentes no Brasil**

O total de categorias ANBID distintas atingia um número igual a 41, sendo que destas, apenas 10 logravam classificar 84,4% dos 5.989 fundos, ao passo que 20 classificavam 95,4% e 30 classificavam 99,1%. Absolutamente todos os fundos apresentaram, do início ao fim do período de estimação, a mesma classificação. Vale destacar também que este número de categorias não considera a identificação do fundo entre curto e longo prazo, instituída pela legislação tributária.

A amostra inicial selecionou 600 fundos (aproximadamente 10% do universo), a partir de método aleatório sem reposição executado com o auxílio do pacote estatístico SPSS. Porém, alguns fundos selecionados apresentaram dados em branco ao longo do período e foram desprezados. Sendo assim, permanecemos com uma amostra final de 574 fundos (aproximadamente 9,6%), as quais sofriam 30 classificações ANBID distintas, fato que demonstra representatividade (caso seja tomado este critério como relevante) em relação ao conjunto de dados.

É importante frisar mais uma vez que não foi utilizada toda a amostra visando reduzir o tempo e o custo despendidos com a estimação de parâmetros na Análise de Estilo e com os procedimentos hierárquicos empregados na Análise de *Clusters*, os quais são sensíveis a uma quantidade muito grande de objetos.

## 4.2. As variáveis independentes

Uma vez finalizada a coleta de dados para a variável dependente da Análise de Estilo, procedeu-se à busca por dados para as variáveis independentes, ou seja, as classes de ativos de mercado. Relembrando, Sharpe (1992) explica que, embora não seja estritamente obrigatório, é desejável que as classes de ativos selecionadas:

- sejam mutuamente exclusivas, o que significa que nenhum título aparece em mais de uma classe de ativos,
- sejam exaustivas e, portanto, representem todos os ativos financeiros existentes,
- possuam baixa correlação.

Devem, portanto, segundo o autor, representar carteiras de ativos similares ponderados pelos seus pesos apurados com base no valor de mercado. Tal objetivo é pragmaticamente atingido pela utilização de índices. A fim de tomar ciência da prática presente na literatura, foram consultados trabalhos com aplicação de Análise de Estilo a fundos brasileiros, entre os quais podemos citar Varga e Valli (1998) e Pizzinga (2004). Ambos têm, porém, abrangência bastante reduzida, pois tratam cada um de apenas 2 fundos. As classes de ativos utilizadas nesses estudos estão apresentadas na Tabela 2.

Varga e Valli (1998)	Pizzinga (2004)
<ul style="list-style-type: none"><li>• Ibovespa</li><li>• CDI</li><li>• Índice de Aplicação Cambial (IVC)</li><li>• Índice de Renda Fixa Pré-Fixada (IV6)</li></ul>	<ul style="list-style-type: none"><li>• Ibovespa</li><li>• CDI</li><li>• Quantum Cambial</li><li>• Quantum Pré-Fixado</li><li>• PTAXV</li><li>• IGPM</li></ul>

Tabela 2 – Índices utilizados em outros trabalhos envolvendo Análise de Estilo

Cabe aqui discutir com cuidado o histórico de alguns desses índices. Tanto o IVC como o IV6 foram criados por Varga e Valli (1998) para suprir a ausência de parâmetros brasileiros que servissem como referência dos rendimentos de, respectivamente: aplicações atreladas ao câmbio e aplicações atreladas a taxas pré-fixadas. O primeiro se baseia nas taxas referenciais dos *swaps* DI x Dólar de 1 ano divulgadas pela Bolsa de Mercadorias e Futuros (BM&F) enquanto que o segundo nos *swaps* DI x Pré de 6

meses. Mais tarde, inclusive, tais índices foram objeto de uma resenha específica que os descreveu com maiores detalhes (Varga, 1999).

O Quantum Cambial e o Quantum Pré-Fixado são, na verdade, idênticos aos índices acima. Só trocam de nome, mas têm a mesma metodologia e objetivo. Suas séries são registradas diariamente pela empresa Quantum, especializada em serviços para fundos de investimento. Maiores informações sobre esses índices podem ser obtidas tanto através da consulta aos artigos mencionados como a partir de contato com a empresa.<sup>22</sup>

Percebe-se, pela observação da Tabela 2, que o trabalho de Pizzinga (2004) simplesmente expandiu o número de classes de ativos utilizadas por Varga e Valli (1998), acrescentando a variação na cotação do dólar à vista para venda (PTAXV) e a variação no Índice Geral de Preços do Mercado (IGPM). Entretanto, ressalta-se que nenhum dos dois trabalhos deixa explícita a preocupação em compatibilizar as classes de ativos utilizadas à teoria da Análise de Estilo. Varga e Valli (1998) ensaiam muito brevemente o porquê da seleção daqueles índices. Pizzinga (2004), por sua vez, ignora por completo esta etapa.

Ainda assim, verifica-se que ambos inseriram no modelo somente classes de ativos negociados no Brasil, fato que deve possuir algum embasamento oculto. Reforçando o já comentado, na presente dissertação, ocorre o mesmo: somente índices brasileiros precisam ser selecionados. A justificativa para isto é que a premissa da Análise de Estilo de representar todos os ativos financeiros existentes pôde ser adaptada para refletir um universo reduzido de fundos, uma vez que se eliminou a categoria “Investimento no Exterior” do escopo do estudo.

Diferentemente dos trabalhos citados, considerou-se separar em mais de uma classe os ativos representados por ações. A razão é que poderia haver grupos de ações com características distintas. Colocou-se em prática este objetivo selecionando-se índices setoriais. A Bolsa de Valores de São Paulo (Bovespa) calcula índices para os setores de telecomunicações (ITEL), energia elétrica (IEE) e indústria (INDX11).<sup>23</sup> Para os

---

<sup>22</sup> O endereço da empresa Quantum na *internet* é [www.quantumfundos.com.br](http://www.quantumfundos.com.br).

<sup>23</sup> A metodologia para cálculo destes índices pode ser encontrada em [www.bovespa.com.br](http://www.bovespa.com.br) através da opção “Índices” no menu “Mercado”. As listagens das ações integrantes dos índices no período de maio a

mercados considerados relevantes no cenário brasileiro, mas com ausência de índices, foram construídos índices próprios. São eles: o de mineração (IMIN), o de petróleo e gás (IPG) e o financeiro (IFIN), neste último incluídos os serviços relativos a bancos, seguradoras e gestão de recursos.

A seleção de ações para composição destes índices observou critérios de liquidez diária e significância no volume financeiro negociado no ambiente da Bovespa para o respectivo setor. Os códigos das ações integrantes de cada índice estão apresentados na Tabela 3. Tais códigos são utilizados para identificação das mesmas na Bovespa.

IMIN	IPG	IFIN
<ul style="list-style-type: none"> <li>• VALE3</li> <li>• VALE5</li> <li>• MMXM3</li> </ul>	<ul style="list-style-type: none"> <li>• PETR3</li> <li>• PETR4</li> </ul>	<ul style="list-style-type: none"> <li>• BBAS3</li> <li>• ITAU3</li> <li>• ITAU4</li> <li>• ITSA4</li> <li>• BNCA3</li> <li>• BBDC3</li> <li>• BBDC4</li> <li>• PSSA3</li> <li>• GPIV11</li> </ul>

**Tabela 3 - Ações integrantes dos índices setoriais construídos**

As séries temporais contendo as cotações foram extraídas da base de dados Económica e foram ajustadas para proventos (dividendos ou bonificações) e eventos de *split* ou *inplit*. Foi utilizada a cotação média ao invés da cotação de fechamento com o intuito de imitar o padrão adotado para marcação a mercado de ativos de renda variável nas carteiras de fundos de investimento. Ressalta-se que este padrão vigia no “período-base do estudo”, conforme definido pelo Plano Contábil dos Fundos de Investimento (COFI), aprovado pela Instrução CVM nº 438/06. Entretanto, foi modificado posteriormente pela Instrução CVM nº 465/08, a qual passou a exigir a precificação pela cotação de fechamento.

Os poucos dias sem negociação de algumas ações foram preenchidos com a última cotação disponível. Os retornos de cada ação foram ponderados dia a dia pelo volume financeiro negociado, ou seja, por um critério de fluxo, ao invés do procedimento usual

---

agosto de 2007 foram obtidas em contato direto com a Bovespa. As séries temporais das cotações destes índices foram coletadas a partir da base de dados Económica.

de utilizar o estoque em *free float* (ações em poder de não-controladores). O objetivo com isto foi simplificar o cálculo dos índices.

A respeito dos índices de renda fixa, há semelhança com os trabalhos de Varga e Valli (1998) e Pizzinga (2004) no que concerne à utilização do Quantum Cambial e do Quantum Pré-Fixado. Entretanto, a presente dissertação utiliza também outros índices, conseguidos a partir da Associação Nacional das Instituições do Mercado Financeiro (ANDIMA). São eles, o IRFM, o IMAS, o IMAC e o IMAB. Uma breve descrição destes índices vem a seguir. Ao leitor interessado em detalhes, recomenda-se a leitura da Cartilha publicada pela ANDIMA contendo a metodologia de cálculo.<sup>24</sup>

O IRFM representa a valorização dos títulos públicos federais pré-fixados, ou seja, as Letras do Tesouro Nacional (LTNs) – sem *coupons* – e as Notas do Tesouro Nacional série F (NTNs-F) – com *coupons*. Este índice se diferencia do Quantum Pré-Fixado por estar ligado a títulos de *duration* mais longa. Segundo o Relatório Anual da Dívida Pública Federal de 2007 (RDP 2007) publicado pelo Tesouro Nacional, foram emitidas, no ano em referência, LTNs com prazos de 6, 12 e 24 meses. Já as Notas do Tesouro Nacional série F (NTNs-F) apresentaram prazos de 3, 5 e 10 anos. O prazo médio das emissões de títulos pré-fixados no primeiro semestre foi de 25,6 meses. Merece lembrança, a efeito de comparação, que o Quantum Pré-Fixado simula uma aplicação de prazo de apenas 6 meses

O IMAS é composto por títulos públicos federais pós-fixados indexados à taxa Selic. Isto significa que segue o rendimento das Letras Financeiras do Tesouro (LFTs), eliminando-se daí as séries A e B, as quais são, por força da Lei nº 9.496 de 11 de setembro de 1997, destinadas ao cumprimento dos contratos de assunção pela União das dívidas de responsabilidade dos Estados e do Distrito Federal. A Selic, conforme é de notório conhecimento, além de servir como base para as LFTs, é referência indireta para as aplicações em Certificados de Depósito Interfinanceiro (CDIs) e até mesmo outros títulos privados, como é o caso dos Certificados de Depósito Bancário (CDBs) emitidos por instituições financeiras e das operações compromissadas baseadas em debêntures emitidas por empresas de *leasing* pertencentes a conglomerados financeiros.

---

<sup>24</sup> A Cartilha pode ser acessada no endereço [http://www.andima.com.br/ima/arqs/ima\\_cartilha.pdf](http://www.andima.com.br/ima/arqs/ima_cartilha.pdf).

O IMAC envolve os títulos públicos federais atrelados ao IGPM, os quais são representados pelas Notas do Tesouro Nacional série C (NTNs-C). O IGPM, segundo a Fundação Getúlio Vargas – organização responsável pela sua apuração – procura registrar a inflação de preços desde matérias-primas agrícolas e industriais até bens e serviços finais. Além da atualização monetária a partir do IGPM, as NTNs-C possuem rendimento pré-fixado, a taxa de juros determinada quando da emissão do título.

O IMAB é associado aos títulos públicos federais atrelados ao Índice de Preços ao Consumidor Amplo (IPCA). Tais títulos são as Notas do Tesouro Nacional série B (NTNs-B). O IPCA é o índice que mede oficialmente a inflação no Brasil. Em acréscimo ao IPCA, as NTNs-B são remuneradas a uma taxa de juros pré-fixada definida nos mesmos moldes das NTNs-C.

A última classe de ativos selecionada foi PTAXV (fonte: Economatica). Muito embora o Quantum Cambial já represente investimentos ligados à variação cambial, ele possui horizonte de um ano. Uma vez que a cotação do dólar é muito volátil, aplicações com *durations* diferentes podem apresentar retornos bastante distintos entre si. Por este motivo, decidiu-se por inserir no modelo também a variação do dólar à vista, a fim de representar investimentos atrelados ao câmbio com prazos médios mais curtos.

A Tabela 4 traz um sumário dos índices selecionados. Há, portanto, índices para a maioria dos ativos financeiros negociados no Brasil, fazendo diferenciação, inclusive, em matéria de *duration* dos títulos.

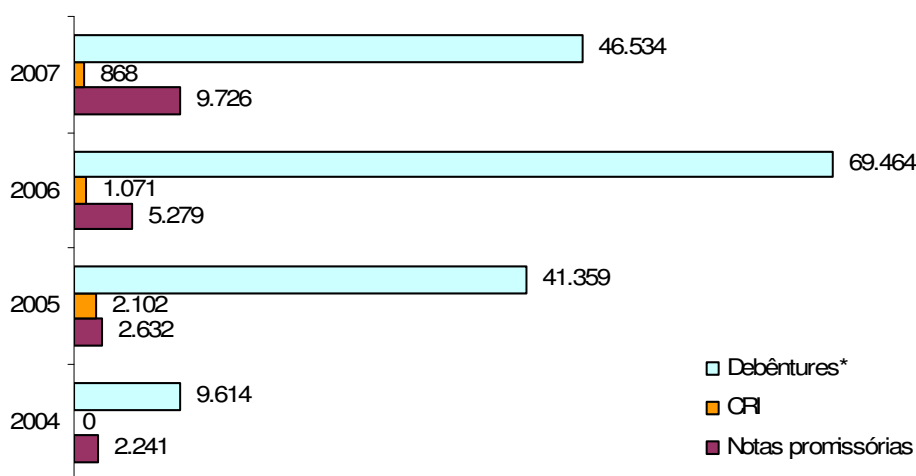
Índice	Referência
ITEL	Ações do setor de telecomunicações
IEE	Ações do setor de energia elétrica
INDX11	Ações do setor industrial
IMIN	Ações do setor de mineração
IPG	Ações do setor de petróleo e gás
IFIN	Ações do setor financeiro
IRFM	Títulos pré-fixados de prazo longo
IMAS	Títulos pós-fixados atrelados a Selic
IMAC	Títulos atrelados ao IGPM
IMAB	Títulos atrelados ao IPCA
Quantum Cambial (QC)	Aplicações atreladas à variação cambial de prazo longo
Quantum Pré-Fixado (QP)	Aplicações pré-fixadas de prazo curto
PTAXV	Aplicações atreladas à variação cambial de prazo curto

**Tabela 4 - Sumário dos índices selecionados**

Todavia, ficaram sem representação as aplicações financeiras ligadas a empreendimentos imobiliários (exs: certificados de recebíveis imobiliários - CRIs) e ao crédito privado (exs: notas promissórias e debêntures ofertadas publicamente). O problema, porém, não é uma deficiência específica do presente trabalho, mas sim uma deficiência do mercado nacional. Simplesmente inexistem *benchmarks* para estes investimentos.

A ausência é muito provavelmente decorrente da menor escala e liquidez observada para os mesmos. O Gráfico 2 mostra a evolução nas emissões públicas destes títulos no período de 2004 a 2007, onde se vê que as debêntures possuem larga vantagem em termos de volume financeiro. Apesar do crescimento expressivo recente, sua participação no estoque de investimentos brasileiros ainda é bastante acanhada. Esta avaliação está detalhada na Tabela 5. Como é de praxe em análises sobre este mercado, desconsiderou-se o estoque de debêntures emitidas por empresas de *leasing* ligadas a conglomerados financeiros, em virtude de tais títulos possuem características e objetivos bastante peculiares, discrimina-se: servir aos bancos destes conglomerados como instrumento alternativo aos CDBs, de forma a evitar os depósitos compulsórios exigidos sobre depósitos a prazo.

Atestada a baixa representatividade dos títulos imobiliários e de crédito privado, é razoável concluir que sua ausência no modelo gere efeitos desprezíveis.



\* Desconsidera emissões de empresas de leasing e do setor financeiro

Fonte: CVM

**Gráfico 2 – Ofertas públicas de recebíveis imobiliários e de crédito privado**

Em R\$ trilhões		Ações Bovespa		Debêntures SND		Estoque (D) = (A) + (B) + (C)	Participação (E) = (C) / (D)
Data	DPMFI <sup>1</sup> (A)	Total <sup>2</sup>	Free float <sup>3</sup> (B)	Total <sup>4</sup>	Sem <i>leasing</i> <sup>5</sup> (C)		
dez/2006	1,09	1,54	0,46	0,16	0,05	1,60	3%
dez/2007	1,22	2,48	0,74	0,22	0,07	2,03	3%

<sup>1</sup> Estoque da Dívida Pública Mobiliária Federal Interna. Fonte: Relatório Anual da Dívida Pública Federal de 2007.

<sup>2</sup> Valor de mercado. Fonte: Boletins Informativos Bovespa Ano 6 nº 96 e Ano 7 nº 114, respectivamente.

<sup>3</sup> Estimativa baseada na premissa de que, em média, 30% do total das ações estariam em poder de não controladores.

<sup>4</sup> Fonte: Boechat (2008).

<sup>5</sup> Estimativa com base na premissa de que 68,7% do total referiam-se a emissões de empresas de *leasing* (utilizou-se a posição de maio/07 divulgada no Boletim Técnico Edição Especial ANDIMA – Ano XIII nº 56, único dado conseguido).

**Tabela 5 - Participação de debêntures (sem *leasing*) no estoque de investimentos brasileiro**

Outros ativos que não possuem índices próprios são as aplicações em volatilidade (opções). De fato, as próprias características deste mercado dificultam seriamente a construção de índices confiáveis. De acordo com Neubert *apud* Varga (1999), a formulação adequada de um índice demanda:

- (a) Conjunto bem definido de títulos nos quais se possa investir;
- (b) Cotações diárias;
- (c) Metodologia de cálculo disponível;
- (d) Classe de ativos bem definida;
- (e) Títulos e classes de ativos especificadas com antecedência;
- (f) Baixo giro de títulos.

Para opções, que são criadas segundo critérios conjunturais e têm horizonte de tempo curto, não são válidos os pontos (a), (e) e (f). Dependendo da distância até o vencimento da opção e até mesmo da natureza do ativo-objeto (opções de dólar e de índice podem apresentar menor liquidez do que a de ações), cotações diárias tampouco estarão disponíveis.

Mais uma vez, entretanto, o montante de investimentos é pouco significativo dentro do mercado nacional. O volume financeiro negociado ao longo de 2007 inteiro foi de R\$44.024 milhões, somando opções de compra e venda, segundo a Bovespa. Ressalta-se que esta é uma medida de fluxo no ano. O estoque médio diário é certamente muito menor do que este valor. Acrescente-se a isto que muitas operações de compra de opções são associadas a uma de venda com preço de exercício diferente visando constituir barreiras a perdas. Sendo assim, o saldo líquido investido reduz-se ainda mais. Por último, vale notar que a instrução CVM nº 409/04 geralmente impõe limites rígidos

à concentração da carteira de fundos abertos em opções, isto quando não veda completamente este tipo de operação. Tais fatores, tomados em conjunto, tornam esta classe de ativos negligenciável.

Retomando, a Tabela 4 consolidou os índices escolhidos para representar ativos negociados no mercado nacional. Contudo, não é possível dizer por ora que estes são realmente os índices definitivos para pesquisa. Isto porque ainda não foram comparados empiricamente visando atestar o atendimento à premissa de baixa correlação imposta pela teoria da Análise de Estilo.

Procedeu-se então à esta avaliação. A matriz de correlação entre os 13 índices selecionados mostrou altos coeficientes entre alguns pares de variáveis, chegando, por vezes, a valores próximos de 0,9. Deve ser notado que níveis tão intensos de correlação geram redundância nos dados e prejudicam, inclusive, a estimação de parâmetros. No intuito de reduzi-los a patamares aceitáveis (definidos nesta dissertação como inferiores, em módulo, ao ponto de corte 0,666), adotou-se a seguinte metodologia.

Selecionou-se o par de variáveis de maior coeficiente e, com a ajuda do pacote SAS (Statistical Analysis Software) versão 9.0, foi elaborada uma análise de componentes principais com base na matriz de covariâncias não corrigida pela média, formando assim uma única variável representativa de ambas. Através desta técnica quantitativa, é geralmente possível sumarizar grande parte da variação de um conjunto de dados, formado por duas ou mais variáveis, em um número bem menor de variáveis, chamadas componentes, as quais são, em última instância, funções lineares das variáveis originais (Sharma, 1996).

Foi utilizada a matriz de covariâncias e não a de correlação com o propósito de se manter o componente na mesma escala das variáveis iniciais – ou seja, numa escala não padronizada – e com isto não causar distorções na Análise de Estilo desempenhada na seqüência. Já o fato da média não ter sido subtraída ocorreu em função das variáveis dependentes da Análise de Estilo (os retornos dos fundos) tampouco terem sido corrigidas.

O par de variáveis identificado como sendo o de maior correlação foi IMAB e IRFM. Após implementada a análise de componentes principais, eliminamos estas variáveis do conjunto de dados e incluímos o componente formado pelas mesmas. Obteve-se em seguida a matriz de correlação dos índices restantes e novamente foi selecionado o par de maior coeficiente, desta vez: QC e PTAXV. Repetiu-se este processo iterativo até não existirem mais, na matriz de correlação, coeficientes maiores ou iguais ao ponto de corte.<sup>25</sup>

Os componentes construídos estão dispostos na Tabela 6, pela qual se percebe que ITEL foi o único dos índices de ações a não ser integrado a um componente. Restaram deste procedimento, portanto, duas variáveis representativas do mercado de ações, uma correspondente ao setor de telecomunicações e outra correspondente aos demais setores da economia. Novamente, um problema de clareza ao investidor poderia daí se derivar, pois os mesmos poderiam indagar o porquê desta separação. Um dos objetivos aqui, conforme já colocado, é trazer praticidade ao investidor e não gerar dúvidas e questionamentos. Logo, prezando mais uma vez pela facilidade de interpretação das categorias na nova classificação de fundos, decidiu-se por alterar a seleção de índices iniciais, substituindo ITEL, IEE, IFIN, INDX, IPG e IMIN, ou seja, todos aqueles associados a ações, pelo Índice Brasil (IBRX), mais geral e que congrega informações sobre todos simultaneamente.

Componente	Formado por
<i>Comp1</i>	IMAB e IRFM
<i>Comp2</i>	QC e PTAXV
<i>Comp3*</i>	IEE e INDX
<i>Comp4*</i>	IFIN e <i>Comp3</i>
<i>Comp5*</i>	IMIN e <i>Comp4</i>
<i>Comp6</i>	IPG e <i>Comp5</i>

\* Eliminado durante o processo iterativo. Representado dentro de *Comp6*

**Tabela 6 - Componentes resultantes da análise iterativa**

<sup>25</sup> Optou-se pelo processo iterativo descrito ao invés de aplicar a análise de componentes principais a todos os índices simultaneamente porque através desta última opção geraríamos componentes de difícil interpretação, o que prejudicaria uma das propostas do trabalho, qual seja: trazer um novo sistema de classificação de fundos que ofereça clareza ao investidor. Caso adotássemos esta alternativa, os primeiros componentes responderiam por quase toda a variação nos dados brutos e carregariam muitas dimensões com forte peso cada uma, logo, tornando impossível distingui-los por uma característica singular e marcante.

Cabe frisar que uma qualidade do IBRX é que o mesmo envolve ainda mais ativos do que o conjunto de todos os índices de ações inicialmente selecionados, os quais atingiam, no período de maio a agosto de 2007, 85 ações, ao passo que o IBRX contempla 100 das ações mais negociadas na Bovespa. Relembrando, isto é positivo pois, segundo Sharpe (1992), as classes de ativos selecionadas para a Análise de Estilo deveriam, sob o rigor teórico, representar todos os ativos existentes no mercado ponderados por seus respectivos volumes financeiros<sup>26</sup>.

A alteração no conjunto de índices gerou a necessidade de se reestimar a matriz de correlações e implementar desde o início o procedimento de junção de variáveis altamente correlacionadas. Após passar pelo mesmo processo iterativo descrito anteriormente, obteve-se, finalmente, as classes de ativos nas quais se baseou a Análise de Estilo (vide Tabela 7).

Classe de ativos	Formada por	Driver
$CL_1$	IMAC	IGPM
$CL_2$	IMAS	Selic
$CL_3$	QP	Juros pré-fixado de curto prazo
$CL_4$	IBRX	Rendimento de ações
$CL_5$	IMAB e IRFM	Inflação + pré de longo prazo
$CL_6$	QC e PTAXV	Variação cambial

**Tabela 7 - Variáveis independentes da Análise de Estilo e respectivas características**

As classes 5 e 6 são fruto do trabalho de sumarização executado.  $CL_5$  explica 97,58% da variação em seus índices formadores, ao passo que  $CL_6$  explica 93,23%, medidas consideradas bastante razoáveis. As equações que descrevem a construção destes componentes são:

$$CL_5 = 0,9367 * IMAB + 0,3501 * IRFM ,$$

$$CL_6 = 0,7948 * QC + 0,6069 * PTAXV .$$

O esforço de redução dos níveis de correlação entre os índices foi efetivo e pode ser visualizado pela matriz de correlação exposta na Tabela 8.

<sup>26</sup> O IBRX é definitivamente a referência mais representativa do mercado de ações brasileiro. Segundo dados da Bovespa (opção “Índices” no Menu “Mercado” de seu *site*), em dezembro de 2006, a capitalização das ações integrantes do índice atingiu R\$1,34 trilhão (87% do total das ações lá negociadas), ao passo que em 2007, somou R\$1,88 trilhão (75%). A título de comparação, o Ibovespa alcançou proporções iguais a 77% e 70% nas mesmas datas.

	$CL_1$	$CL_2$	$CL_3$	$CL_4$	$CL_5$	$CL_6$
$CL_1$	1					
$CL_2$	0,43	1				
$CL_3$	0,14	0,08	1			
$CL_4$	0,28	0,15	0,23	1		
$CL_5$	0,57	0,36	0,37	0,41	1	
$CL_6$	-0,25	-0,05	-0,34	-0,52	-0,62	1

Tabela 8 - Matriz de correlação entre as classes de ativos da Análise de Estilo

## 5. MODELAGEM E RESULTADOS

### 5.1. Análise de Estilo

O modelo de Análise de Estilo utilizado na dissertação está descrito matematicamente abaixo:

$$R_{it} = \sum_j (P_{ij} \cdot CL_{jt}) + e_{it}, \quad \begin{aligned} i &= 1, 2, \dots, 574 \\ t &= 1, 2, \dots, 84 \\ j &= 1, 2, \dots, 6 \end{aligned}$$

$$\text{Min}_{P_{ij}} \sum_t e_{it}^2 \text{ para todo } i,$$

restrito a:

$$\sum_j P_{ij} = 1,$$

$$-0,05 \leq P_{ij} \leq 1,05 \quad \forall j,$$

$$P_{ij} + P_{ik} \geq -0,07 \text{ para todo } j \neq k,$$

$$P_{ij} + P_{ik} + P_{il} \geq -0,08 \text{ para todo } j \neq k \neq l,$$

onde:

$R_{it}$  é o retorno do fundo  $i$  no tempo  $t$ ,

$CL_{jt}$  é o retorno da classe de ativos  $j$  no tempo  $t$ ,

$P_{ij}$  é o peso médio atribuível à classe  $j$  no retorno do fundo  $i$  durante o período amostral,

$e_{it}$  é o componente idiossincrático do retorno do fundo  $i$  no tempo  $t$ .

A restrição de soma 1 é clássica em trabalhos aplicados de Análise de Estilo. Ela facilita o entendimento sobre os  $P_{ij}$ , possibilitando interpretá-los como proporções. Muito embora os parâmetros do modelo não representem efetivamente percentuais de composição da carteira, e sim as sensibilidades do retorno do fundo em respeito aos índices selecionados, é possível pensar neles como *proxies* destes percentuais (Brown e Goetzmann, 1997).

As restrições posteriores são de desigualdade e procuram limitar – a níveis realistas – a alavancagem e a concentração do estilo. Os limites foram definidos a partir de consultas ao sistema de composição de carteiras disponível na página eletrônica da CVM<sup>27</sup>. Foram pesquisados, quando encontrados pela ferramenta de busca do sistema, todos os fundos com as seguintes palavras ou expressões em seus nomes: “agressivo”, “concentrado”, “alavancado”, “alavancagem”, “arriscado”, “alto risco” e “arrojado”. O mês no qual se baseou as consultas foi o de julho de 2007.

O fundo mais alavancado dentre os contemplados foi o “Opus Hedge Agressivo Fundo de Investimento Multimercado”, que apresentou posições a descoberto equivalentes a 9,83% de seu patrimônio líquido, entre empréstimos de ações e operações de venda de opções. Pode ser verificado pelas equações de restrição que, matematicamente, não há impedimento nenhum ao alcance do nível de alavancagem do fundo citado. Isto é inclusive validado empiricamente pela estatística apresentada logo adiante no Gráfico 4.

No modelo, são impostas restrições marginalmente mais rígidas a níveis crescentes de alavancagem. Isto encontra sustentação e justifica-se pela própria realidade do mercado brasileiro, onde, para cada unidade monetária adicional operada sem lastro, são exigidos depósitos de margem de garantia exponencialmente crescentes pelas bolsas de valores a

---

<sup>27</sup> <http://cvmweb.cvm.gov.br/SWB/Sistemas/SCW/CPublica/FormBuscaPartic.aspx?TpConsulta=5>.

A CVM ressalva que as informações são prestadas pelos administradores dos fundos e que não há garantia de que são verdadeiras. Nesta consulta específica, assumiu-se que são.

fim de cobrir os riscos envolvidos, os quais não respeitam um processo linear. Esta exigência torna inviável, em função de recursos escassos, a assunção de níveis de alavancagem muito altos.

Já pelo lado da concentração da carteira, não se observou nenhum fundo com alocações em uma única classe de ativos superiores ao limite de 105% estabelecido para o estudo. O que pode ser assumido destes fatos é que as restrições impostas no modelo são abrangentes o suficiente para o cálculo adequado dos parâmetros. A relação completa dos fundos consultados no sistema da CVM está disponível no Apêndice A. São, ao todo, 38 fundos.

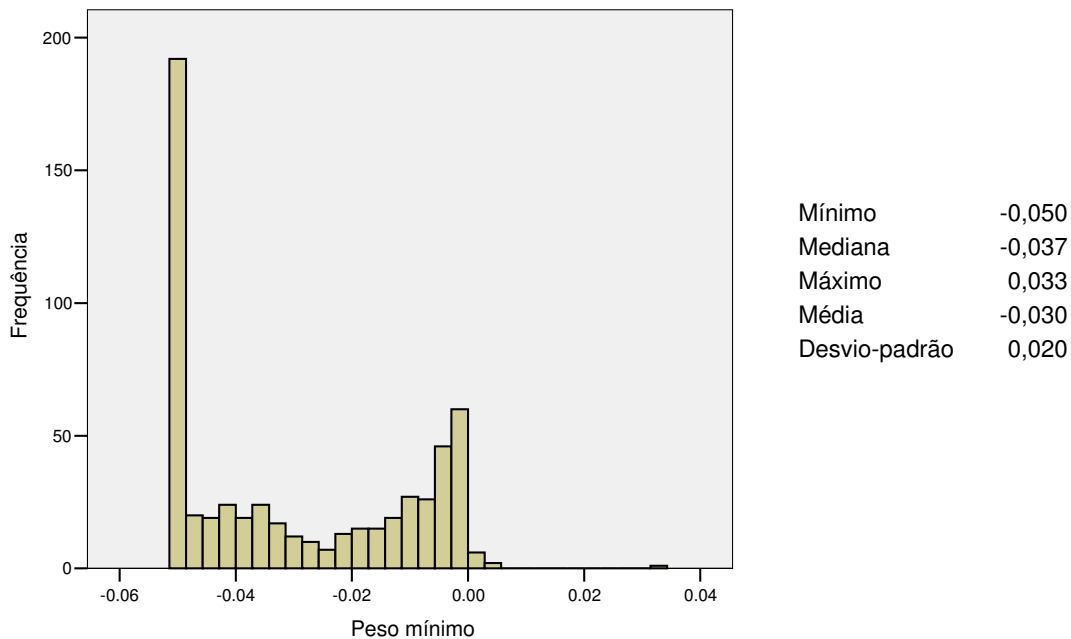
Conforme Sharpe (1992) bem observou, não é viável desempenhar a Análise de Estilo pelo método de regressão em virtude de haver restrições de desigualdade sobre os parâmetros. Uma forma de contornar este problema é utilizar programação quadrática.

Na presente dissertação, a Análise de Estilo foi implementada através da função “fmincon” do pacote de otimização do *software* Matlab em sua versão 7.0. Segundo a ferramenta de ajuda (*help*) deste programa, “fmincon” fornece o mínimo de uma função não linear multivariada sujeita a restrições de igualdade e/ou desigualdade. Adicionalmente, fornece os valores das incógnitas que minimizam a função-objetivo. Por conseguinte, é apropriada ao problema aqui enfrentado. No Apêndice B, está disponível o código de programação escrito para atravessar esta etapa.

Os resultados estão sumarizados por uma seqüência de gráficos que mostram a distribuição de frequência de algumas medidas descritivas com importante conteúdo informacional sobre os parâmetros. O Gráfico 3 é o histograma construído a partir da seleção, para cada fundo  $i$ , do  $P_{ij}$  mínimo. Esta medida representa a posição a descoberto mais intensa de um determinado fundo quando comparados os pesos das classes de ativos individualmente.

Por este gráfico, percebe-se que muitos fundos atingiram a barreira imposta de -0,05. Percebe-se também que quase todos os fundos tem algum grau de alavancagem, por menor que seja. Este resultado talvez não pareça condizente com a realidade à primeira impressão. Porém, deve ser lembrado que, em geral, se não sempre, fundos de

investimento apresentam alguma provisão em seus passivos para honrar obrigações ainda não vencidas, tal qual despesas com a taxa de administração, auditoria, a taxa de fiscalização da CVM, a contribuição à ANBID, etc. Sendo assim, em última instância, há alavancagem, não no sentido de venda a descoberto, mas no sentido de possuírem ativos totais que superam o valor do patrimônio líquido. O modelo parece captar este efeito.

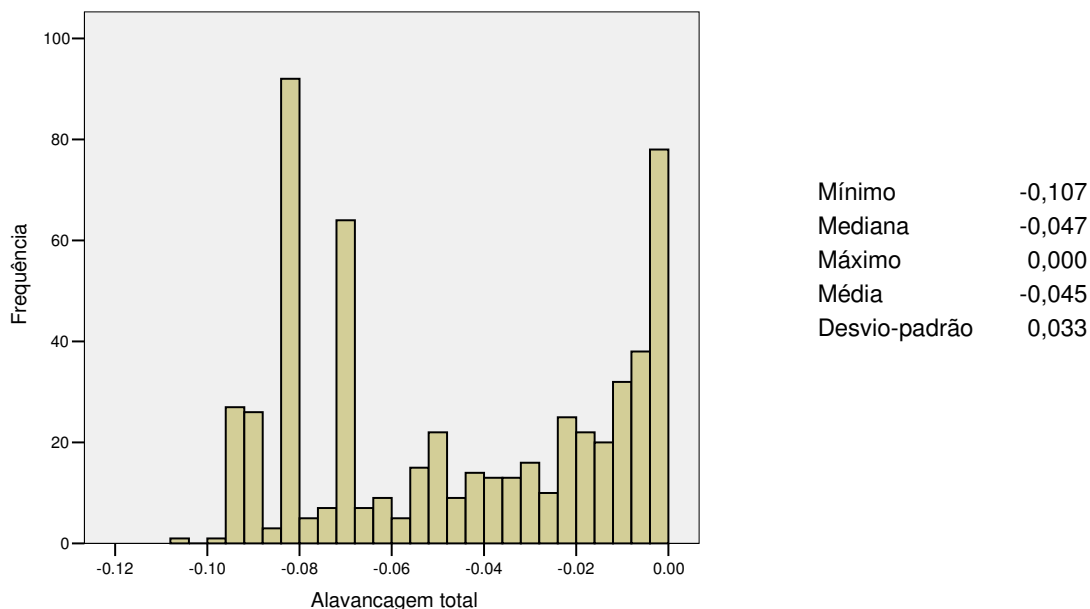


**Gráfico 3 - Histograma da posição a descoberto individual mais relevante**

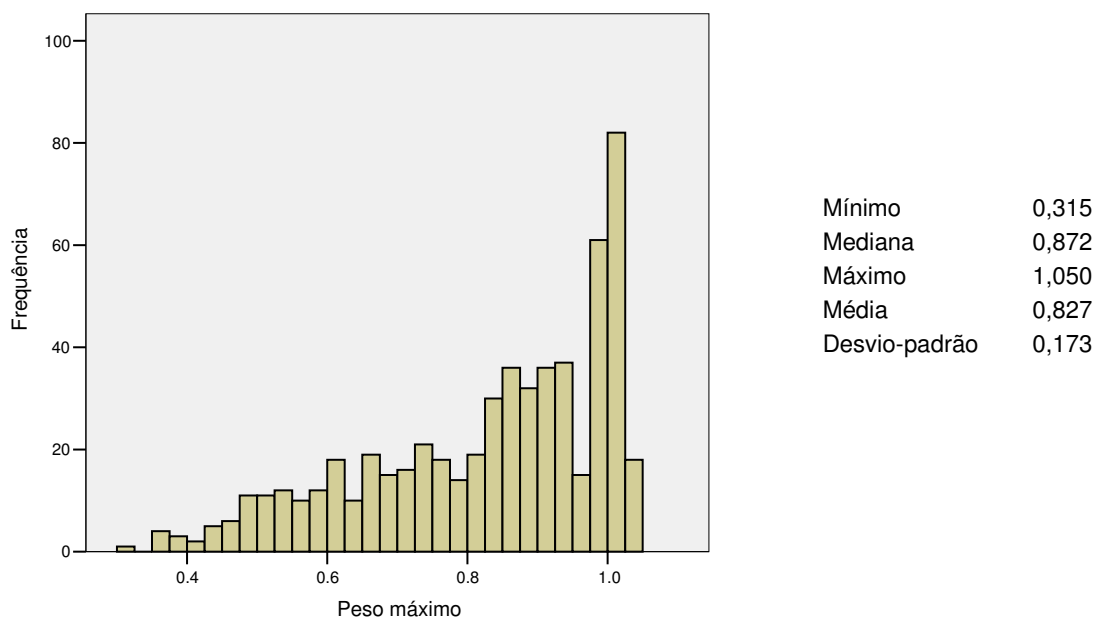
O Gráfico 4 é o histograma do total da alavancagem total computada para cada fundo  $i$ . Analisando-o, corrobora-se a afirmação feita anteriormente de que as restrições impostas ao modelo são flexíveis o bastante para refletir estilos agressivos de investimento dentro do domínio encontrado durante as consultas ao sistema da CVM.

O Gráfico 5, por sua vez, é o histograma construído a partir da seleção, para cada fundo  $i$ , do  $P_{ij}$  máximo. Com este gráfico, procurou-se ilustrar a concentração das carteiras dos fundos na amostra, ao que se vê que existem fundos realmente focados no retorno de um único índice. Para ser mais preciso, a maioria deles atua assim. Dos 574 fundos, exatamente 100 possuem concentração maior do que 100% em alguma classe de

ativos.<sup>28</sup> Outros tantos se aproximam da unidade. Em contraposição, poucos fundos apresentam balanceamento na exposição de seus investimentos. Do total, apenas 32 apresentaram concentração máxima em uma única classe inferior a 50%.



**Gráfico 4 - Histograma da alavancagem total**

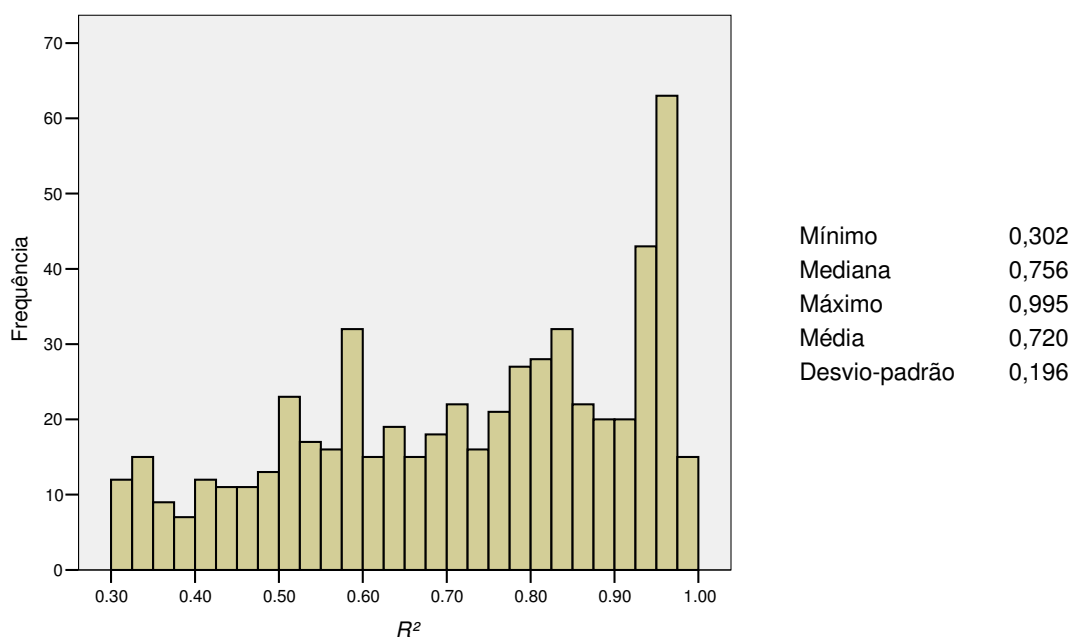


**Gráfico 5 - Histograma da concentração máxima em uma única classe de ativos**

<sup>28</sup> Novamente, isso pode ser reflexo do fato que ativos totais são comumente maiores do que o patrimônio líquido de fundos.

Por fim, o Gráfico 6 mostra a distribuição de freqüência dos  $R^2$  alcançados na reprodução dos retornos de cada fundo  $i$ . Para a maior parte dos fundos, obteve-se um grau de ajuste alto, chegando a um máximo de 99,5%. Entretanto, para outros, a variação residual prevaleceu sobre a explicada.

As possíveis justificativas para o baixo poder explicativo, as quais podem ocorrer isoladas ou em conjunto, são<sup>29</sup>: (a) o fundo possui uma estratégia fortemente ativa de gestão, implicando que seu retorno provém mais da seleção de ativos específicos do que de estilo, (b) o fundo apresentou mudanças acentuadas no estilo durante o período amostral, (c) o fundo possui um estilo predominantemente ligado a uma classe de ativos não representada no modelo, inclusive as internacionais, e esta classe é pouco correlacionada com as classes contempladas, (d) não adequação do fundo aos limites estipulados para as restrições de desigualdade.



**Gráfico 6 - Histograma dos  $R^2$**

As três últimas causas levantadas configuram fragilidades do modelo. No entanto, imagina-se que sejam as de menor efeito. As variações dinâmicas no perfil do fundo,

<sup>29</sup> Não se considerou, na listagem de justificativas, a possível existência de erros nas bases de dados, por serem estes erros um problema potencial em qualquer trabalho empírico.

assunto tratado no item (b), estão certamente presentes. Porém, tais mudanças são muito provavelmente marginais devido aos motivos identificados na seção 3.4.

Quando se trata das duas últimas, vale reforçar novamente os cuidados tomados durante o trabalho para mitigá-las. Com relação ao item (c), a seção 4.2 mostrou que os possíveis investimentos financeiros existentes no país estão bem representados. Muito embora haja classes não incluídas (recebíveis imobiliários, títulos de dívida corporativos, etc.), o estoque de poupança atribuível a elas é bastante baixo em termos relativos. A ausência de classes internacionais, por sua vez, deve gerar efeito mínimo, tendo em vista que o escopo do trabalho foi bem delineado visando eliminar fundos com investimento no exterior.

No tocante ao item (d), a metodologia adotada para definição de limites obviamente não é perfeita, porém oferece razoável segurança de que o comportamento dos fundos não recai com frequência fora do intervalo permitido pela estimação.

O item (a) não representa efetivamente uma fragilidade do modelo, muito pelo contrário. Decerto, a categorização de um fundo de gestão ativa com base num modelo de Análise de Estilo traz menos informação sobre seu retorno do que para os demais. No entanto, alguma informação é melhor do que nada. Ademais, a capacidade de identificar estes fundos pode ser considerada, na verdade, um ponto forte. Afinal, esta é mais uma informação que pode ser fornecida ao investidor para auxiliar seu processo decisório. Ela é relevante, inclusive, por permitir comparação entre as taxas de administração cobradas pelos fundos, as quais costumam ser maiores para os de gestão ativa.

A questão polêmica é certamente a definição de um ponto de corte. Qual é o valor, no intervalo de 0 a 1 do domínio da medida  $R^2$ , mais adequado para diferenciar fundos de gestão passiva e ativa? A resposta é que se trata de algo subjetivo. Pensando como *policy-maker*, talvez a melhor abordagem seja classificar o fundo de acordo com o estilo estimado, mesmo que este estilo seja pouco explicativo, e obrigar o fundo a informar em seu material de divulgação corrente a aderência de seu retorno aos movimentos sistemáticos de mercado. Uma mensagem do tipo: “O retorno deste fundo pôde, historicamente, ser explicado em aproximadamente  $\lambda\%$  por movimentos sistemáticos

em índices de mercado, primordialmente aqueles ligados à categoria na qual se insere. A parcela restante do retorno deveu-se a peculiaridades no estilo de gestão empreendido pelo fundo. Não há garantia de que este percentual se manterá no futuro”.

$\lambda$ , obviamente, seria  $R^2 \times 100$ , e deveria ser calculado por alguma instituição isenta (a CVM ou a ANBID, por exemplo). Tal abordagem, além de conferir transparência ao modelo, deixaria a decisão sobre o que configura gestão passiva ou ativa para o investidor. Logo, sua subjetividade afetaria única e exclusivamente ele próprio.

## 5.2. Análise de *Clusters*

Não há teoria na qual se basear para definir o número mais adequado de categorias para classificação de fundos. Logo, um método exploratório teve que ser empregado. Conforme argumentado na seção 3.2, uma alternativa possibilitada pelo *software* SPSS – a ferramenta utilizada como suporte para esta análise – seria o *cluster* de dois estágios. Entretanto, este método assume algumas premissas (normalidade e independência entre as variáveis características dos objetos) raramente respeitadas na realidade. Sendo assim, foi preterida em favor do método hierárquico.

Tendo em vista que todas as variáveis (pesos das classes de ativos) são medidas na mesma escala (percentual), não houve necessidade de padronizá-las antes de iniciar os procedimentos. Para manter a aderência ao objetivo do trabalho, a medida de similaridade se balizou no critério de distância entre os objetos. Este critério atende aos anseios, pois agrupa fundos que são similares na magnitude de suas alocações relativas na carteira, justamente a característica que dita a exposição de um fundo aos riscos de mercado.

A medida adotada nesta dissertação foi a Distância Euclidiana, a mais popular nas pesquisas com aplicação de *cluster* (Sharma, 1996 e Hair et al. 2005). A utilização dessa medida permite também atingir consistência entre os métodos hierárquico e de partição prévia aqui implementados, visto que essa medida é a única disponível para o algoritmo de *k*-médias do SPSS. Embora provavelmente mais apropriada, por levar em consideração a correlação entre as variáveis independentes, a Distância de Mahalanobis não pôde ser utilizada por esse motivo.

A Distância Euclidiana ( $D$ ) entre realizações ( $X$ ) dos objetos  $i$  e  $j$  em  $p$  dimensões é calculada da seguinte maneira (Sharma, 1996):

$$D_{ij} = \sqrt{\sum_{k=1}^p (X_{ik} - X_{jk})^2}$$

Os algoritmos de *cluster* hierárquico do SPSS se baseiam na lógica aglomerante. Conforme mencionado na seção 3.2, o algoritmo escolhido foi o de ligação completa (*complete linkage*). A medida empregada para definição do número de *clusters* foi a maior distância absoluta entre os coeficientes de aglomeração entre dois estágios subsequentes, a qual indicou uma repartição ótima da amostra em 5 grupos. Logo, este é o número de categorias que, de acordo com o retorno bruto, melhor distingue os fundos no universo estudado. Cabe esclarecer que o coeficiente de aglomeração de um determinado estágio do processo de aglomeração é igual à distância entre os grupos sendo combinados naquele estágio (Norusis, 1988).

Em seguida, foram calculadas as médias das variáveis correspondentes às sensibilidades dos fundos às classes de ativos ( $\bar{P}_{.j}$ ) para cada uma das categorias encontradas e, estes valores, dispostos para consulta na Tabela 9, foram inseridos como entrada numa Análise de *Clusters* do tipo  $k$ -médias.

Categoria	$\bar{P}_{.1}$	$\bar{P}_{.2}$	$\bar{P}_{.3}$	$\bar{P}_{.4}$	$\bar{P}_{.5}$	$\bar{P}_{.6}$
1	-2%	2%	0%	90%	12%	-2%
2	1%	83%	8%	5%	5%	-1%
3	-1%	13%	19%	2%	-2%	69%
4	0%	13%	2%	6%	80%	-1%
5	6%	9%	60%	14%	12%	-1%

As linhas podem não somar 100% por questão de arredondamento.

**Tabela 9 - Centróides do *cluster* hierárquico inseridos como entrada na análise de  $k$ -médias**

Os resultados convergiram na nona iteração. O rígido critério de convergência estabelecido foi a existência de diferença nula entre os centróides de uma determinada rodada e aqueles da rodada anterior. Por sua vez, a medida de distância entre centróides foi novamente a Euclidiana, padrão e única medida disponível no SPSS para *clusters* do tipo  $k$ -médias.

Na Tabela 10, é apresentada uma comparação entre o número de fundos em cada categoria encontrado pelos diferentes métodos empregados, onde fica claro o rearranjo efetuado pelo método de  $k$ -médias.

Categoria	Número de fundos	
	Método Hierárquico	Método $k$ -médias
1	108	108
2	410	344
3	10	10
4	7	22
5	39	90

Tabela 10 - Número de fundos de acordo com os métodos hierárquico e  $k$ -médias

Os centróides definitivos estimados, que formam as categorias da nova classificação de fundos, estão apresentados na Tabela 11. As classes mais relevantes em cada categoria estão destacadas em negrito. Por elas, é possível interpretar a natureza dos fundos em cada categoria e assim dar-lhes identificações.

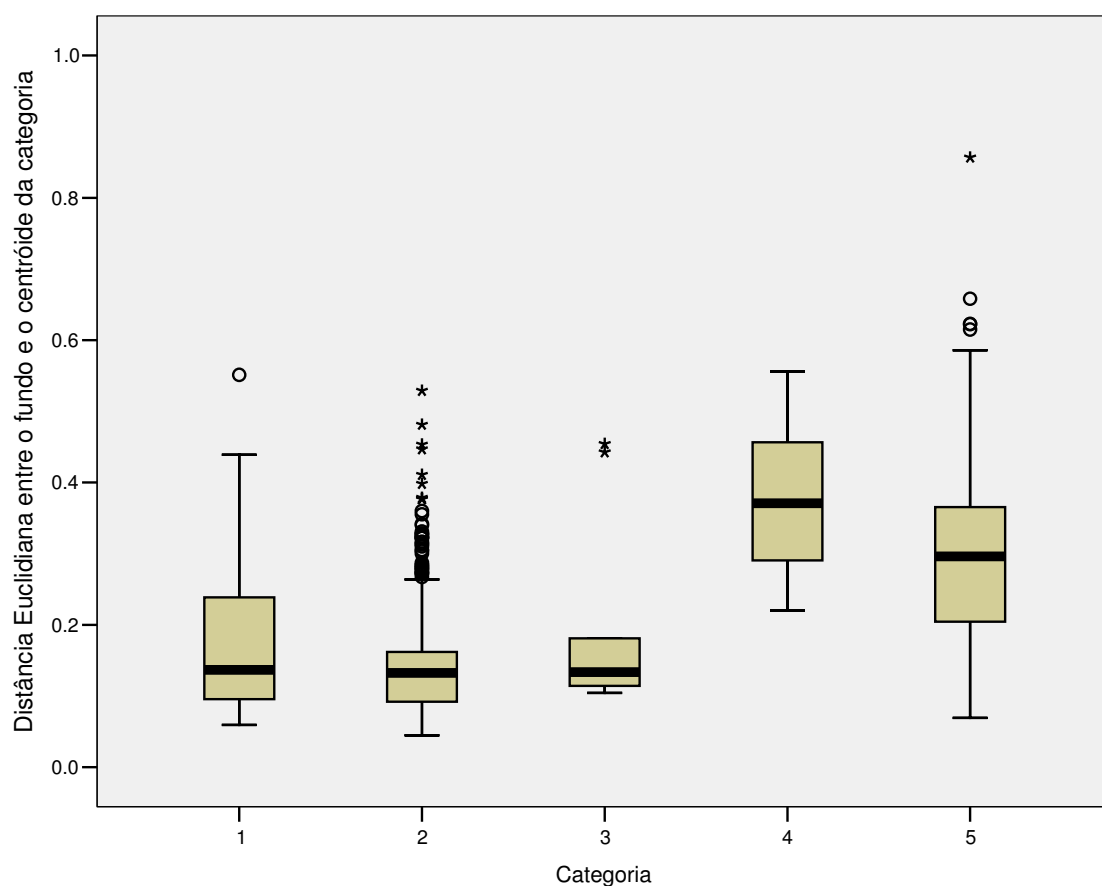
A primeira categoria é composta por fundos identificados como sendo “Fundos de Ações” em virtude desses fundos investirem, em média, 90% de suas carteiras na classe de ativos 4, a qual, conforme visto na Tabela 7, representa o IBRX. Os fundos da segunda categoria foram denominados “Fundos de Referência em Selic” pelo fato de acompanharem em grande parte o IMAS, cujo *driver* é a Selic. Já a terceira categoria envolve os “Fundos Cambiais”, caracterizados por concentrarem seus ativos na variável  $CL_6$ , classe representativa dos índices QC e PTAXV. Os fundos integrantes da quarta categoria foram chamados “Fundos de Capital Preservado” por terem seu foco em retorno atrelado à inflação mais uma taxa de juros pré-fixada. A última categoria, por sua vez, tem sua carteira mais bem distribuída entre as diversas classes de ativos, com destaque para as classes 2 e 3. Seus fundos foram então classificados como “Multimercado”.

Categoria	Identificação	$\bar{P}_{.1}$	$\bar{P}_{.2}$	$\bar{P}_{.3}$	$\bar{P}_{.4}$	$\bar{P}_{.5}$	$\bar{P}_{.6}$
1	Ações	-2%	2%	0%	<b>90%</b>	12%	-2%
2	Referência em Selic	0%	<b>89%</b>	5%	4%	4%	-1%
3	Cambial	0%	6%	25%	3%	-3%	<b>69%</b>
4	Capital Preservado	2%	18%	19%	8%	<b>57%</b>	-3%
5	Multimercado	6%	<b>38%</b>	<b>42%</b>	12%	5%	-2%

As linhas podem não somar 100% por questão de arredondamento.

Tabela 11 - Centróides das categorias da nova classificação de fundos proposta

Com o intuito de avaliar a homogeneidade dos fundos dentro das categorias, foram construídos *boxplots* com base na Distância Euclidiana entre cada fundo e o centróide de sua respectiva categoria. Pelo Gráfico 7, constata-se que as categorias 2, 3 e 5 possuem observações extremas<sup>30</sup>. Tais ocorrências tendem a prejudicar o desempenho preditivo da Análise Discriminante a ser implementada adiante.



**Gráfico 7 - Boxplots dos fundos por categoria**

Por fim, desenvolveu-se, a título de comparação, uma matriz de transição da classificação ANBID para a classificação proposta (apresentada na Tabela 12), onde é possível perceber que, em geral, a categorização se comporta conforme esperado. Alguns fundos, contudo, foram destinados para categorias não condizentes com sua classificação inicial. A origem deste fato está em ao menos uma das seguintes situações: (a) categorização errada na base SI-ANBID devido a desvios de comportamento do gestor do fundo ou erros de registro, (b) fragilidade na estimação da carteira do fundo ou (c) falha na identificação precisa das fronteiras dos *clusters*.

<sup>30</sup> As observações extremas estão ilustradas por estrelas no *boxplot*.

Classificação ANBID	Nova classificação proposta					
	Ações	Referência em Selic	Cambial	Capital Preservado	Multimercado	Total
Ações IBOVESPA Ativo	21	-	-	-	-	21
Ações IBOVESPA Ativo com Alavancagem	5	-	-	-	-	5
Ações IBOVESPA Indexado	7	-	-	-	-	7
Ações IBX Ativo	9	-	-	-	-	9
Ações IBX Ativo com Alavancagem	3	-	-	-	-	3
Ações IBX Indexado	1	-	-	-	-	1
Ações Outros	26	-	-	-	3	29
Ações Outros com Alavancagem	3	2	-	-	1	6
Ações Setoriais Telecomunicações	1	-	-	-	-	1
Balanceados	-	4	-	3	1	8
Cambial Dólar sem Alavancagem	-	1	7	-	-	8
Curto Prazo	-	6	-	-	-	6
Fundos de Privatização CVRD - FGTS	10	-	-	-	-	10
Fundos de Privatização CVRD - Recursos Próprios	3	-	-	-	-	3
Fundos de Privatização Petrobras - FGTS	8	-	-	-	-	8
Fundos de Privatização Petrobras - Recursos Próprios	3	-	-	-	-	3
Multimercados com Renda Variável	1	40	1	4	5	51
Multimercados com Renda Variável com Alavancagem	-	123	1	11	63	198
Multimercados sem Renda Variável	-	30	-	-	-	30
Multimercados sem Renda Variável com Alavancagem	-	4	1	-	4	9
PIBB (Fundos de Índice)	5	-	-	-	-	5
Previdência Balanceados	1	11	-	-	4	16
Previdência Multimercados com Renda Variável	-	2	-	1	3	6
Previdência Renda Fixa	-	4	-	1	-	5
Previdência Renda Fixa Crédito	-	1	-	-	-	1
Privatização FGTS – Carteira Livre	1	-	-	-	-	1
Referenciado DI	-	44	-	-	-	44
Renda Fixa	-	65	-	2	5	72
Renda Fixa com Alavancagem	-	1	-	-	1	2
Renda Fixa Médio e Alto Risco	-	6	-	-	-	6
<b>Total</b>	<b>108</b>	<b>344</b>	<b>10</b>	<b>22</b>	<b>90</b>	<b>574</b>

**Tabela 12 - Matriz de transição da Classificação ANBID para a nova classificação proposta**

Cabe lembrar que, concluída esta etapa, deve-se, a fim de atender a legislação, classificar os fundos segundo seu perfil tributário, quando aplicável. Adicionalmente, é importante notar que a amostra coletada incluiu fundos de previdência. Tais fundos seguem legislação tributária específica, nos termos da Lei nº 11.053 de 29 de dezembro de 2005, e devem possuir identificação própria. Conseqüentemente, o número de categorias encontradas deveria dobrar, pois só assim se estaria atendendo a Lei. A título de exemplo, a categoria “Ações” teria que ser desmembrada em duas: “Ações” e “Ações Previdenciário”, implicando que os fundos de previdência com predominância de investimentos em ações deveriam ser alocados nesta última. O mesmo raciocínio é válido para as demais categorias.

Os resultados da Análise de *Clusters* mostram razoável semelhança das categorias encontradas com aquelas especificadas pela Instrução CVM nº 409/04. Esta Instrução define 7 classes distintas de fundos de investimento, fora os tipos regidos por regulamentação especial, de acordo com o disposto em seu Artigo Primeiro. Convém mencionar que alguns destes fundos “especiais” constam na base SI-ANBID e foram contemplados no universo estudado, tendo sido, inclusive, selecionados durante a amostragem (ex: Fundos Mútuos de Privatização FGTS e Fundos de Índice).

As classes previstas são: I – Fundo de Curto Prazo; II – Fundo Referenciado; III – Fundo de Renda Fixa; IV – Fundo de Ações; V – Fundo Cambial; VI – Fundo de Dívida Externa; e VII – Fundo Multimercado.

Não só alguns nomes são parecidos (ou até iguais), como algumas regras e parâmetros definidos para a composição dos fundos são por vezes compatíveis com os perfis médios estimados. Isso sem contar que a quantidade de categorias se situou razoavelmente próxima uma da outra, sobretudo quando se recorda que os fundos de Dívida Externa foram destacados da análise corrente.

De qualquer forma, um exemplo de ineficácia da classificação CVM concerne à categoria denominada “Curto Prazo”. Fique claro que esta categoria não se confunde com aquela para fins de imposto de renda. Ela se baseia no *caput* do Art. 93 da Instrução CVM nº 409/04, cuja redação é bastante explícita na sua definição:

*“Os fundos classificados como "Curto Prazo" deverão aplicar seus recursos exclusivamente em títulos públicos federais ou privados pré-fixados ou indexados à taxa SELIC ou a outra taxa de juros, ou títulos indexados a índices de preços, com prazo máximo a decorrer de 375 (trezentos e setenta e cinco) dias, e prazo médio da carteira do fundo inferior a 60 (sessenta) dias (...).”*

A definição da ANBID, neste caso, é exatamente igual à da CVM, inclusive para manter aderência legal. Verifica-se, através da Tabela 12, que todos estes fundos foram destinados à categoria “Referência em Selic”. Este resultado, além de bastante coerente, deixa evidente que é absolutamente arbitrária e sem propósito a imposição da CVM, pois os mesmos são similares em seu estilo a diversos outros fundos não classificados como “Curto Prazo”. A existência desta categoria traz distorções no conjunto de informações que podem gerar alocações de recursos sub-otimizadas por parte de investidores. Logo, fica aqui a sugestão de eliminação da categoria.

A arbitrariedade da classificação CVM não é prejudicial somente em termos de definição do número de categorias e sua natureza, mas também na definição de limites de composição das carteiras. Estabelecer o quanto um fundo pode investir em determinada classe de ativos para estar classificado numa categoria específica é ineficiente porque não garante similaridade de estilo entre fundos dentro da mesma faixa. Tampouco garante diferenças significativas de fundos fora dos limites pré-estabelecidos. Somente a partir da avaliação do comportamento efetivo dos fundos é que se pode concluir a respeito de características comuns ou singulares. Esta avaliação deve ser realizada caso a caso. Para isto, é necessário que seja utilizado um método imparcial e objetivo, como o ora empreendido.

Encerrada a comparação à classificação CVM, se passa então ao modelo ANBID, para o qual se percebe diferença gritante quanto ao número de categorias. Conforme já dito, uma lista muito extensa de categorias torna complexo o entendimento por parte dos usuários e aumenta as chances de confundir não só os investidores como possivelmente até mesmo os gestores e administradores de fundos quanto às características de cada uma.

Tem-se ciência de que as determinações da ANBID nunca vão contra as regras impostas pela CVM. Procuram apenas aperfeiçoá-las, exigindo condutas adicionais dos fundos aderentes a seu Código. Aparentemente, porém, não devem estar surtindo o efeito desejado. Uma vez que se afastam mais dos resultados da Análise de *Cluster* do que a classificação CVM, provavelmente atrapalham o mercado de fundos mais do que o desenvolvem.

### **5.3. Análise Discriminante**

Conforme exposto na seção 3.3, a classificação de objetos através da Análise Discriminante assume algumas premissas. Logo, a aderência dos dados às mesmas foi avaliada antes de se iniciar o processo de estimação das funções classificatórias.

A primeira premissa testada foi a de normalidade multivariada das variáveis independentes. Bem verdade, apenas pelo fato de se ter restringido, durante a Análise de Estilo, o domínio das mesmas a um intervalo compreendido entre -0,05 e 1,05, já seria suficiente para saber que não são normalmente distribuídas. Contudo, embora a hipótese não se sustente em teoria, uma vez que a Normal possui 95% de sua massa num raio de 1,96 desvio-padrão a partir da média, havia a possibilidade de exibir equivalência estatística. Sendo assim, procedeu-se aos testes pertinentes.

Tem-se ciência que condição *sine qua non*, porém não garantidora, de normalidade multivariada, é que todas as variáveis sejam Normais individualmente (Sharma, 1996 e Hair et al., 2005). Sendo testes univariados capazes de concluir sobre a falsidade da hipótese multivariada, optou-se pela sua elaboração preliminar. Tal decisão foi guiada pelas seguintes razões apontadas na literatura: (a) menor complexidade envolvida (Hair et al., 2005); e (b) freqüente indisponibilidade do teste multivariado em *softwares* estatísticos (Sharma, 1996). De fato, o SPSS 13.0 e o SAS 9.0 não contêm esta funcionalidade.

O primeiro passo antes dos testes foi corrigir as variáveis independentes de cada fundo pelas médias dos respectivos grupos aos quais pertenciam. Em seguida, a distribuição de freqüência de cada uma das variáveis foi avaliada visualmente. O Apêndice C apresenta gráficos que comparam curvas Normais construídas com valores de média e

desvio-padrão iguais aos dos correspondentes histogramas. Disparidades foram trazidas à tona: picos de moda muito acentuados e, em alguns casos, sobrepeso nas caudas. A rejeição à normalidade foi corroborada pelo teste Jarque-Bera<sup>31</sup>, cujas medidas e respectivas significâncias estão demonstradas na Tabela 13, assim como valores de simetria e curtose para cada variável.

Estatística	$P_{i1}$	$P_{i2}$	$P_{i3}$	$P_{i4}$	$P_{i5}$	$P_{i6}$
Simetria	6,51	-0,64	0,86	0,77	1,10	3,71
Curtose	65,31	3,64	6,74	6,14	6,36	45,78
Jarque-Bera	96.907,69	48,49	404,94	292,36	384,91	45.082,69
Significância	0,00	0,00	0,00	0,00	0,00	0,00

**Tabela 13 - Estatísticas do teste de normalidade**

A Análise Discriminante assume também a premissa de matrizes de variância-covariância homogêneas entre grupos (categorias). A ingênua investigação visual dos gráficos de dispersão consolidados no Apêndice D traz evidência de descumprimento da premissa, suspeita confirmada pelo teste Box's M desempenhado no SPSS, o qual é apropriado a este objetivo (Sharma, 1996 e Hair et al., 2005). Os resultados do referido teste estão dispostos na Tabela 14.

Descrição	Valor
Box's M	1753,24
Estatística-F	26,57
Graus de liberdade 1	60
Graus de liberdade 2	5703,33
Significância	0,00

**Tabela 14 - Teste Box's M**

Uma saída para a não-homogeneidade das matrizes de variância-covariância seria a utilização de uma função discriminante quadrática (Sharma, 1996). Todavia, a estimação desta função demandaria um número de graus de liberdade maior do que o disponível. Sendo assim, foi automaticamente descartada.

Cumprir destacar, conforme argumentado na seção 3.3, que a violação das premissas não necessariamente invalida a utilização da Análise Discriminante como técnica de previsão. Tampouco o atendimento às mesmas a qualifica. O critério realmente decisivo à sua adoção ou não é o resultado final do modelo. Se a previsão for acurada, justifica-

<sup>31</sup> Efetuado com o auxílio do *software* Eviews.

se sua utilização. Logo, é imprescindível que se aprecie ao fim da análise o desempenho preditivo das funções classificatórias.

Isso não significa dizer que o esforço despendido na avaliação das premissas foi em vão. Para todos os efeitos, neste trabalho, tomar ciência do que estava ou não sendo desrespeitado servia mais como um alerta do que como um obstáculo ao prosseguimento. Realizando os testes pertinentes, o investigador pode interpretar melhor os resultados preditivos da Análise Discriminante. Na hipótese de desempenho insatisfatório, o não atendimento às premissas é uma provável explicação do porquê,<sup>32</sup> e permite subsidiar propostas alternativas de ação. Uma alternativa viável, por exemplo, é a utilização de um modelo *logit* multinomial, cuja modelagem é mais flexível por não depender de normalidade (Sharma, 1996), porém mais complexa de se implementar.

A última preocupação que surge antes da implementação da Discriminante é a existência de multicolinearidade nos dados (Sharma, 1996 e Hair et al., 2005). Uma vez que a derivação matemática da função classificatória envolve a inversão das matrizes de variância-covariância, a precisão dos resultados pode ser impactada se houver relações lineares muito fortes entre as variáveis independentes.

O SPSS tem especificada uma medida de tolerância que controla o nível de multicolinearidade e o mantém em níveis aceitáveis protegendo a análise. Esta medida é, de acordo com Sharma (1996), dada por  $1 - R^2$ , onde  $R^2$  é o quadrado da correlação múltipla entre uma determinada variável e as demais. Sendo assim, quanto maior é a correlação e a possibilidade de se representar a variável linearmente por outras, menor é esta medida. O nível de tolerância mínimo padrão do SPSS é igual a 0,001. Quando uma variável ultrapassa este valor, é automaticamente excluída do processo.

Vale lembrar que quando foi efetuada a Análise de Estilo, estabeleceu-se a restrição de que os pesos deveriam somar 1. Portanto, havia multicolinearidade intrínseca a estas variáveis. Para contornar este problema e adiantar o que o próprio SPSS já faria, eliminou-se a variável  $P_{i6}$ . Isto não ocasiona prejuízo à análise, pois toda a informação

---

<sup>32</sup> Outra possível explicação para desempenhos preditivos eventualmente ruins é a existência de fronteiras entre grupos mal definidas.

disponível está contida nas variáveis restantes. A matriz de correlação entre elas está apresentada na Tabela 15.

	$P_{i1}$	$P_{i2}$	$P_{i3}$	$P_{i4}$	$P_{i5}$
$P_{i1}$	1				
$P_{i2}$	-0,07	1			
$P_{i3}$	0,24	-0,32	1		
$P_{i4}$	-0,18	-0,77	-0,22	1	
$P_{i5}$	-0,13	-0,39	-0,01	0,13	1

**Tabela 15 - Matriz de correlação entre os pesos das 5 primeiras classes de ativos**

A preparação dos dados envolveu a separação dos fundos de cada categoria em dois grupos distintos, um para estimação das funções classificatórias e outro para sua validação. O critério utilizado na divisão deve ser enfatizado. Hair et al. (2005) comenta que o número de objetos dentro de cada categoria deve ser superior em no mínimo uma unidade ao número de variáveis independentes. Por este motivo, não foi possível dividir a amostra inicial em duas de igual tamanho, uma vez que a categoria nº 3 apresentou apenas 10 fundos e havia 5 variáveis independentes distintas. A solução foi selecionar 60% dos fundos de cada categoria para estimação enquanto que os 40% restantes ficaram para validação. As proporções realmente atingidas na separação entre amostra de estimação e amostra de validação não foram exatamente estas (60/40), mas sim proporções aproximadas devido à necessidade de arredondamento, conforme a Tabela 16 deixa claro.

Categoria	Identificação	Amostra inicial (I)	Amostra de validação	Amostra de estimação (E)	E / I
1	Ações	108	43	65	60,2%
2	Referência em Selic	344	138	206	59,9%
3	Cambial	10	4	6	60,0%
4	Capital Preservado	22	9	13	59,1%
5	Multimercado	90	36	54	60,0%

**Tabela 16 - Separação entre amostra de estimação e amostra de validação**

A proporção de 60/40 foi utilizada para dividir todas as categorias a fim de se manter as probabilidades prévias (*prior probabilities*) na amostra de estimação equivalentes às da amostra inicial. A amostra inicial fornece probabilidades prévias para cada categoria que são estimadores não viesados das probabilidades universais em virtude de ter sido colhida de forma completamente randômica a partir da base SI-ANBID. De acordo com

a Equação 2 apresentada na seção 3.3, as probabilidades prévias afetam as constantes das funções de previsão.

A Tabela 17 expõe estas probabilidades e as compara com aquelas resultantes da divisão da amostra para formação do grupo de estimação. Percebe-se que a diferença entre as duas é muito pequena, o que torna desprezíveis eventuais distorções nos resultados.

Categoria	Amostra inicial		Amostra de estimação		Diferença
	Número de fundos	Probabilidades prévias	Número de fundos	Probabilidades prévias	
1	108	18,8%	65	18,9%	0,1%
2	344	59,9%	206	59,9%	-0,0%
3	10	1,7%	6	1,7%	0,0%
4	22	3,8%	13	3,8%	-0,0%
5	90	15,7%	54	15,7%	0,0%
Geral	574	100,0%	344	100,0%	N/A

**Tabela 17 - Probabilidades prévias das categorias na amostra inicial e grupo de estimação**

O processo de divisão da amostra entre os grupos de estimação e validação foi totalmente aleatório, à exceção do procedimento implementado para a categoria nº 3, pois, conforme notado no Gráfico 7 (*boxplot*), dois de seus dez fundos apresentaram características peculiares a ponto de serem identificados como observações extremas. Para garantir equilíbrio entre as etapas de estimação e validação, obrigou-se que um deles compusesse o primeiro grupo enquanto que o restante compusesse o segundo. A seleção do fundo destinado ao primeiro grupo foi aleatória.

Não houve necessidade de se padronizar as variáveis independentes pelo fato das mesmas possuírem a mesma natureza e, conseqüentemente, estarem todas medidas na mesma escala (percentual das classes de ativos na exposição da carteira). Tampouco houve necessidade de se realizar o procedimento *stepwise*, visto que não se está explorando quais seriam as variáveis mais adequadas para discriminação entre as categorias. As variáveis foram escolhidas com base na teoria da Análise de Estilo, logo, puderam ser inseridas em conjunto e simultaneamente no modelo.

Os parâmetros estimados para as funções classificatórias encontram-se disponíveis na Tabela 18. Ressalta-se que estas funções são ajustadas às probabilidades prévias, porém

são estimadas assumindo custos iguais para erros de classificação entre os grupos. Assumir custos iguais é na presente dissertação o mais adequado realmente, pois não há teoria nem indícios que indiquem o contrário.

O capítulo de metodologia mostrou que a operacionalização da previsão é extremamente simples: basta saber quais são as sensibilidades do retorno do fundo às classes de ativos para alimentar as funções. Cada função fornece então um valor discriminante (*discriminant score*). A categoria que apresentar o maior valor é aquela na qual o objeto deve ser alocado (Sharma, 1996).

Variável	Categoria				
	1	2	3	4	5
Constante	-329,89	-325,07	-37,45	-316,14	-327,99
$P_{i1}$	655,51	667,29	211,48	661,70	692,85
$P_{i2}$	590,81	653,26	196,11	604,91	643,60
$P_{i3}$	567,72	616,07	210,26	596,78	642,68
$P_{i4}$	656,17	584,80	182,16	569,21	595,88
$P_{i5}$	547,91	566,72	168,98	616,32	570,92

**Tabela 18 - Parâmetros estimados para as funções classificatórias**

No cômputo geral da amostra de estimação, 98,3% dos fundos foram classificados corretamente. Sabe-se, porém, que tal medida é sujeita a viés e não deve ser utilizada como critério de avaliação da capacidade preditiva do modelo (Sharma, 1996 e Hair et al., 2005). Quando se aplica as funções estimadas à amostra de validação, elimina-se este viés e chega-se finalmente a uma medida imparcial e confiável do desempenho esperado em futuras classificações. Esta medida atingiu um índice de previsão correta igual a 99,6%. Isto se traduz por apenas 1 fundo equivocadamente classificado dos 230 compondo a amostra de validação, um resultado a primeira vista impressionante.

De modo a certificar a qualidade da previsão, foram realizados testes estatísticos sobre a significância do número de acertos geral e por categoria na amostra de validação. É imprescindível que assim se proceda porque os acertos poderiam não ser diferentes do que se atingiria numa separação aleatória dos fundos desta amostra em cinco grupos com tamanhos equivalentes aos de suas probabilidades prévias.

O número de acertos obtido com a aplicação da Discriminante pode ser encontrado a partir da Tabela 19, que apresenta a matriz de previsão por categoria para as amostras de

estimação e de validação, assim como para a amostra total. A Tabela 20, por sua vez, apresenta os respectivos percentuais de acerto e erro de previsão referentes à tabela imediatamente anterior. Convém frisar que os elementos nas diagonais correspondentes às submatrizes de cada amostra representam os acertos.

Amostra	Categoria original	Categoria prevista				
		1	2	3	4	5
Estimação	1	65	0	0	0	0
	2	0	205	1	0	0
	3	0	0	6	0	0
	4	0	0	0	13	0
	5	0	5	0	0	49
Validação	1	43	0	0	0	0
	2	0	138	0	0	0
	3	0	0	4	0	0
	4	0	0	0	8	1
	5	0	0	0	0	36
Total	1	108	0	0	0	0
	2	0	343	1	0	0
	3	0	0	10	0	0
	4	0	0	0	21	1
	5	0	5	0	0	85

**Tabela 19 - Matriz de previsão por tipo de amostra**

Amostra	Categoria original	Categoria prevista				
		1	2	3	4	5
Estimação	1	100,0	0,0	0,0	0,0	0,0
	2	0,0	99,5	0,5	0,0	0,0
	3	0,0	0,0	100,0	0,0	0,0
	4	0,0	0,0	0,0	100,0	0,0
	5	0,0	9,3	0,0	0,0	90,7
Validação	1	100,0	0,0	0,0	0,0	0,0
	2	0,0	100,0	0,0	0,0	0,0
	3	0,0	0,0	100,0	0,0	0,0
	4	0,0	0,0	0,0	88,9	11,1
	5	0,0	0,0	0,0	0,0	100,0
Total	1	100,0	0,0	0,0	0,0	0,0
	2	0,0	99,7	0,3	0,0	0,0
	3	0,0	0,0	100,0	0,0	0,0
	4	0,0	0,0	0,0	95,5	4,5
	5	0,0	5,6	0,0	0,0	94,4

**Tabela 20 - Percentuais de erro e acerto de previsão por tipo de amostra**

O teste de significância desempenhado foi o sugerido em Huberty *apud* Sharma (1996) e baseia-se na construção de uma estatística  $Z$  que segue aproximadamente distribuição Normal sob a hipótese nula de igualdade entre os acertos aleatórios e aqueles gerados pela discriminante. A estatística é dada por:

$$Z = \frac{(o - e)\sqrt{n}}{\sqrt{e(n - e)}},$$

onde:

$o$  é o número de classificações corretas,

$e$  é o número esperado de classificações corretas de acordo com as probabilidades prévias,

$n$  é o número de observações.

Os resultados dos testes encontram-se disponíveis na Tabela 21. Absolutamente todos rejeitaram a hipótese nula com vasta folga, o que oferece suporte definitivo para a utilização da Análise Discriminante mesmo neste ambiente de constatada violação de premissas. Esta excelente capacidade preditiva reflete em última instância o bom trabalho feito pelo algoritmo de aglomeração quando da definição do número de grupos e de suas fronteiras na Análise de *Clusters*. Logo, há plena segurança de que as fronteiras são consistentes e efetivamente segregam grupos de fundos com estilos significativamente diferentes (por consequência, de relação risco-retorno distintas entre si). Nem mesmo a dispersão identificada nos *boxplots* (Gráfico 7), com a ocorrência inclusive de observações extremas em algumas categorias, foi capaz de prejudicar a precisão da Análise Discriminante. Mais uma vez, cabe enfatizar que a avaliação de desempenho foi efetuada para dados fora da amostra, estando, portanto, livre de viés.

Categoria	Estatística-Z	Significância
1	13,67	0,000
2	9,59	0,000
3	15,03	0,000
4	13,15	0,000
5	13,93	0,000
Geral	17,64	0,000

**Tabela 21 - Resultados dos Testes de Huberty**

A Análise Discriminante se mostrou, portanto, uma prática e poderosa ferramenta que permite evitar os custos computacionais e o tempo necessário à aplicação da Análise de *Clusters* a todos os fundos na população. Basta aplicar rapidamente as funções classificatórias estimadas nesta dissertação para saber a que categoria um determinado fundo pertence. O único requisito para isto é conhecer os pesos  $P_{ij}$ .

Por outro lado, deve-se ter em mente que o universo de fundos não é estático. Além de existirem fundos que alteram políticas de investimento ao longo de suas vidas, novos fundos são criados diariamente. A aplicação intertemporal do modelo, pelo fato de se tratar de extrapolação estatística, gerará problemas se alguns destes fundos surgirem com uma proposta de atuação significativamente diferente das já existentes. Estes fundos formariam provavelmente uma sexta categoria não contemplada na estimação da discriminante. A fim de corrigir este problema, a Análise de *Clusters* deveria ser rodada novamente e em seguida deveriam ser construídas novas funções discriminantes.

Conclui-se daí que o modelo discriminante tem prazo de validade. O revés é que não se pode determinar *a priori* qual é este prazo. Profissionais que acompanhem de perto o setor talvez estejam aptos a indicar com certa propriedade o momento adequado para se atualizar a metodologia, visto que dispõem de acesso constante a informações novas e especializadas. São estes profissionais, inclusive, que em geral trazem as propostas inovadoras a fim de aproveitarem algum nicho de mercado potencialmente lucrativo. De qualquer forma, retornar-se-ia ao problema de subjetividade. Uma alternativa que pode substituir ou complementar a opinião de especialistas e reduzir a subjetividade é a utilização de gráficos que ilustrem a distância euclidiana entre os fundos suspeitos e os centróides das categorias. Quando esta distância for muito discrepante, será recomendável rodar novamente a Análise de *Clusters*, pois somente ela pode confirmar empiricamente a hipótese de uma nova categoria e eliminar qualquer traço de julgamento pessoal.

#### **5.4. Avaliação do desempenho prático do modelo**

Em seu processo de planejamento financeiro, o investidor anseia por instrumentos que o auxiliem a explicar melhor os rendimentos futuros de sua carteira. Esta seção visa então

comparar o modelo proposto com o modelo da ANBID em termos de qualidade preditiva de retornos.

Os testes foram realizados para fora da amostra. Isso significa que a avaliação ocorreu sobre um período diferente daquele utilizado para estimação do modelo. Relembrando, a Análise de Estilo tomou por base o intervalo com início em 02/04/2007 e fim em 31/07/2007. Os resultados encontrados foram aplicados para estimação dos retornos dos 13 dias úteis subsequentes, ou seja, até 17/08/2007, data em terminava a série temporal disponível ao autor.

Dos 574 fundos inicialmente considerados para estimação, três deixaram de existir durante o período de validação e foram eliminados. Sendo assim, sobraram 571. Cabe enfatizar que a perda destes fundos não trouxe prejuízo relevante à análise visto que faziam parte de categorias bem representadas em ambos os modelos de classificação. No sistema ANBID, dois destes fundos estavam na categoria “Multimercado sem Renda Variável” – cuja quantidade passou de 30 para 28 fundos – ao passo que o terceiro estava na categoria “Multimercado com Renda Variável com Alavancagem”, que passou de 198 fundos para 197. Do ponto de vista do novo modelo de classificação proposto, todos estavam na categoria “Referência em Selic”. Seu número de integrantes passou de 344 para 341.

Duas metodologias de avaliação distintas foram empregadas. A primeira é semelhante à realizada por Brown e Goetzmann (1997), enquanto que a segunda é semelhante à de Pattarin et al. (2004).

Na primeira, o que se fez foi estimar, para cada sistema de classificação, uma regressão em corte transversal para cada um dos dias subsequentes ao “período-base do estudo”. Os retornos dos fundos num determinado dia foram tratados como sendo a variável dependente. No tocante às variáveis explicativas, se incluiu, além do intercepto, variáveis binárias que refletiam as categorias dos fundos. Cumpre destacar que nas regressões efetuadas para o novo modelo proposto, a categoria de absolutamente todos os fundos na amostra foi determinada por previsão, isto é, através das funções classificatórias, e não pelos resultados da Análise de *Clusters*.

Foi assumido que os fundos apresentaram estabilidade em seus estilos durante o intervalo selecionado para teste e, logo, não suscitaram nenhum remanejamento entre categorias. O objetivo do teste era saber o quanto, em média, do retorno futuro dos fundos podia ser explicado possuindo unicamente a informação das categorias às quais pertenciam. A medida  $R^2$ -Ajustado auxiliou nesta tarefa e serviu como critério de comparação entre os modelos. Através de um teste-t em par para médias bicaudal, avaliou-se a capacidade explicativa de cada um dos modelos ao longo dos 13 dias. Foi constatado que, estatisticamente, são equivalentes.

Na segunda metodologia, estimou-se qual seria o retorno dos fundos em cada um destes dias, usando para isso os retornos das classes de ativos nas datas correspondentes e as respectivas exposições médias (centróides) de cada categoria a estas classes<sup>33</sup>. O estilo de um fundo qualquer foi tomado como sendo igual ao estilo médio da categoria na qual se inseria. Sendo assim, o retorno estimado para dois fundos dentro da mesma categoria foi obviamente o mesmo. Em seguida, avaliou-se, para cada dia, o  $R^2$  (não-ajustado). Os resultados mostraram superioridade estatística do modelo proposto.

A Tabela 22 sumariza o que foi encontrado tanto para a primeira como para a segunda metodologia de avaliação. Embora as diferenças observadas na média sejam simétricas quando se compara uma metodologia com a outra, a significância presente no teste-t da segunda metodologia é devida majoritariamente a um altíssimo coeficiente de correlação entre as séries de  $R^2$  dos dois modelos, o qual atingiu 0,97, enquanto que para a primeira metodologia foi de 0,86. Este efeito ofuscou inclusive o efeito dos desvios-padrão, que são mais expressivos nas séries da segunda metodologia.

A diferença estatisticamente significativa pode parecer não possuir relevância prática, porém conclusão nesse sentido seria enganosa. Há de se lembrar que os dados em mãos são diários. Qualquer pequena diferença nesta escala, quando anualizada, pode ganhar peso considerável.

---

<sup>33</sup> Assumiu-se que os centróides encontrados pela Análise de *Clusters* são iguais aos da população e que se manteriam constantes ao longo do tempo. Por este motivo, empregou-se eles mesmos para todo o período de teste, ao invés de calcular centróides para os agrupamentos previstos pela Discriminante.

Descrição	Metodologia	
	Primeira*	Segunda**
Modelo proposto***	0,86 (0,07)	0,84 (0,08)
Modelo ANBID***	0,87 (0,05)	0,83 (0,08)
Diferença****	-0,01	0,01
Estatística-t	-1,51	2,67
P-valor	0,16	0,02

\* Utiliza R<sup>2</sup>-Ajustado.

\*\* Utiliza R<sup>2</sup>.

\*\*\* R<sup>2</sup> médio (desvio-padrão).

\*\*\*\* Assume-se que a diferença apresenta distribuição Normal.

**Tabela 22 - Testes-t sobre a diferença nos R<sup>2</sup> médios entre o modelo proposto e o da ANBID**

Esta evidência reforça a superioridade do modelo proposto (ao menos quando comparado à classificação ANBID), indicando que sua implementação traria benefícios aos investidores. Além de apresentar estrutura muito mais simples de se entender, o que reduz os custos do processo analítico e decisório, ainda permite, através de um conteúdo informacional mais eficiente, trazer melhores resultados em termos de explicação de rendimentos futuros. Cabe salientar também que o desempenho do modelo, de acordo com o R<sup>2</sup>, é bastante alto, evidência que, inclusive, suporta a premissa de estabilidade dos estilos. Obviamente, tal conclusão está limitada pelo curto intervalo de tempo utilizado para teste.

## 6. CONCLUSÃO

Desde a década de 90, a indústria de fundos de investimento vem aumentando seu peso na economia brasileira, atraindo como cotistas milhões de indivíduos e empresas. Dada sua elevada importância, propostas que confirmam maior eficiência ao seu funcionamento são particularmente interessantes.

O presente trabalho mostrou que oportunidades nesse sentido existem com relação aos modelos vigentes de classificação de fundos. Modelos de classificação são bem-vindos porque ajudam a trazer ordem a partir do caos, organizando e sintetizando a informação sobre a imensa oferta de fundos disponível aos investidores. As deficiências dos dois sistemas existentes no Brasil, o da CVM e o da ANBID, não são poucas. Todas, porém, são derivadas da falta de imparcialidade na forma como foram elaborados.

O modelo aqui proposto busca remediar esse problema. Sua qualidade principal é a objetividade promovida por sua metodologia, a qual é composta por três etapas: Análise de Estilo, Análise de *Clusters* e Análise Discriminante. A primeira etapa do modelo lhe confere aderência à teoria de investimentos e, por conseguinte, garante adequado tratamento da relação risco-retorno dos fundos.

Outro ponto forte da aplicação da Análise de Estilo é que assim se captura o real comportamento de investimento do gestor. Ao focar exclusivamente em uma informação de difícil manipulação, o rendimento do fundo, permite evitar ou até mesmo eliminar o risco de desvio moral (*moral hazard*) ao qual está sujeita a classificação auto-declarada.

O desvio moral está associado ao incentivo que gestores de fundos têm em divulgar uma carteira marginalmente mais conservadora do que a que realmente possuem, a fim de serem classificados em uma categoria de investimentos menos arriscada. Assumindo maiores riscos do que divulgam, poderiam também alçar maiores retornos do que os fundos classificados na mesma categoria. Sendo assim, aumentariam as chances de atingir posições vencedoras no *ranking* de rentabilidades daquela categoria.

Em países com mercado pouco maduro e estrutura de controle precária, isso é especialmente preocupante. Uma medida de controle que provavelmente reduziria bastante o risco moral seria a separação das pessoas do gestor e do administrador, respectivamente: o responsável pela decisão de investimento e o responsável pelo registro, processamento e divulgação das transações realizadas. Muito embora a prestação dos dois serviços simultaneamente não seja vedada no atual ambiente brasileiro, o país está menos suscetível a tal problema, pois está equipado com outros sistemas de supervisão e monitoramento razoavelmente sólidos.

A Instrução CVM nº 409/4 obriga os administradores, através de seu Artigo 71, a enviarem, via rede mundial de computadores, a composição da carteira em informes diários, além dos balancetes mensais do fundo. Adicionalmente, no Artigo 84, a CVM obriga também que os fundos de investimento sejam auditados em bases anuais por auditores independentes. A ANBID, por sua vez, também estabelece, através de seu Código, a obrigatoriedade de envio de dados eletrônicos em ambiente de rede. Sobre as

carteiras informadas pelos administradores, emprega filtros para todos os fundos abertos a fim de assegurar que estão classificados corretamente.

Ainda assim, os modelos vigentes não conseguem extinguir o risco de desvio moral por completo, pois sempre existirá a possibilidade de gestores mal-intencionados divulgarem uma carteira não correspondente às posições efetivamente assumidas. A verdade é que de nada adiantam os filtros se as informações originais transmitidas estão corrompidas. De fato, a CVM faz questão de ressaltar que o sistema de consulta pública sobre composição de carteiras disponibilizado em sua página na *internet* é alimentado por informações prestadas pelos administradores e que não há garantia de que as mesmas sejam verdadeiras.

Embora o modelo de classificação proposto torne muita remota a chance de desvio moral – o qual seria possível de ser mantido apenas por períodos curtos de tempo – não se sugere aqui o descarte da estrutura de controle existente no Brasil. Decerto, ela poderia desempenhar papel complementar ao modelo. Serviria, por exemplo, para criticar a estimação efetuada via Análise de Estilo, objetivando confrontar os parâmetros estimados para as carteiras com as posições divulgadas pelo gestor/administrador. Aparecendo divergências muito salientes entre estimação e auto-declaração, se justificariam análises aprofundadas até que fossem esclarecidas, pois, conforme explicado na seção 5.1, existe a possibilidade de que os baixos  $R^2$  observados para alguns fundos sejam devidos a fragilidades na modelagem. Não obstante, pelos cuidados tomados ao longo da dissertação para mitigar tais fragilidades, suspeita-se que a maioria destes fundos sejam fundos de gestão ativa, ou seja: com predominância de uma estratégia de seleção específica de ativos.

Ao contrário do que possa parecer, a presença de baixos  $R^2$  para alguns fundos não significa necessariamente uma deficiência do modelo. A capacidade que possui de diferenciar entre fundos de gestão passiva e ativa é na verdade uma característica benéfica, pois esta é uma informação valiosa que pode ser prestada ao investidor, inclusive por viabilizar comparações efetivas entre as taxas de administração cobradas pelos fundos. O revés é que não é possível determinar exclusivamente através do modelo qual é a causa do baixo  $R^2$ . Por este motivo, reitera-se a sugestão de aliar a

aplicação do modelo à estrutura de controle já existente no Brasil. Desta forma, seria possível sanar completamente o problema.

A maneira como foi estruturada a segunda etapa do estudo traz uma alternativa ao que muito provavelmente é a maior falha dos sistemas de classificação vigentes: o processo de definição do número de categorias e a abrangência de cada uma. Em ambas as instituições (CVM e ANBID), as categorias surgem de discussões entre os integrantes dos grupos de trabalho responsáveis por defini-las. Tal processo está invariavelmente contaminado, em maior ou menor grau, por subjetividade. Isso culmina no estabelecimento de critérios arbitrários para repartição do universo de fundos que não refletem fronteiras reais entre categorias.

Um prejuízo grave que daí decorre é a falsa impressão de diversificação que as categorias de seus modelos de classificação geram ao investidor. Uma vez que não são representativas dos diferentes grupos existentes na população de fundos, é possível que fundos alocados em categorias distintas sejam na verdade muito similares em termos de risco e retorno. Sendo assim, nada garante que uma carteira composta por aplicações distribuídas em mais de uma categoria esteja efetivamente diversificada.

No modelo proposto, aplica-se um algoritmo de aglomeração hierárquico para determinar o número natural de categorias. Em seguida, refina-se a Análise de *Clusters* pelo método de *k*-médias. Foram encontradas, na amostra extraída para execução dos trabalhos empíricos, apenas cinco categorias. Cabe ressaltar que, em virtude da amostra ter sido colhida aleatoriamente, permite extrapolação para o restante da população, desde que empregada metodologia apropriada.

É neste contexto, inclusive, é que se faz útil e prática a Análise Discriminante. Ela habilita a identificação estatística das características amostrais com base nas variáveis independentes e, a partir daí, a classificação dos fundos não contemplados durante a Análise de *Clusters*. Sua utilização, ao mesmo tempo em que viabilizou economizar tempo e recursos computacionais, terminou por servir também como um teste sobre a consistência das fronteiras encontradas para as categorias, as quais se mostraram bastante sólidas e confiáveis no sentido de promoverem adequada distinção de estilo.

Vale enfatizar que o número de categorias encontrado foi inferior ao apresentado pela classificação da CVM e muitíssimo inferior ao da ANBID. Em última instância, isso representa outra grande qualidade do modelo: simplicidade. Com uma lista reduzida de categorias, a análise e tomada de decisão do investidor fica facilitada.

A última qualidade apresentada pelo modelo corresponde a sua habilidade em explicar retornos futuros dos fundos com base exclusivamente na informação das categorias as quais pertencem. Fora o fato de que apresenta um alto poder explicativo, ainda é superior em desempenho ao modelo ANBID, contra o qual foi testado.

Agregando todas as melhorias proporcionados pelo modelo, tem-se: (a) redução de custos de agência ligados a desvio moral; (b) diversificação efetiva da carteira; (c) redução dos custos envolvidos no processo de seleção de fundos; (d) maior capacidade de explicar retornos futuros. Logo, é provável que sua adoção fomente ainda mais o crescimento da indústria, pois atrairá indivíduos que antes não percebiam valor na aplicação em fundos. Investidores que possuíam custos marginais de entrada levemente acima dos benefícios marginais esperados podem experimentar uma reversão na ordem de magnitude destas variáveis e, conseqüentemente, virem a se tornar participantes, caso a diferença resultante entre as duas variáveis seja maior do que o custo de oportunidade.

Com relação aos cotistas pré-existentes, se vislumbra mudanças na inclinação da restrição orçamentária que podem ensejar maior propensão a poupar do que consumir e/ou rearranjo da carteira de investimentos, passando a atribuir maior peso aos fundos na sua composição do que a demais ativos.

Por fim, vale lembrar que o modelo proposto é perfeitamente flexível e ajustável a qualquer alteração na legislação tributária, pois se baseia, para a classificação dos fundos, em seus rendimentos brutos, logo, rendimentos anteriores à apuração do imposto de renda.

### **6.1. Limitações do estudo**

As limitações do estudo foram todas mencionadas ao longo do documento. Na presente seção, retoma-se esta discussão.

Em primeiro lugar, como todo modelo de classificação, a proposta do trabalho está sujeita a cada dia que passa a ficar desatualizada. Estratégias significativamente inovadoras de gestores visando atender nichos inexplorados de mercado podem surgir a qualquer momento. Naturalmente, o modelo teria que ser repetido desde o início a fim de gerar funções classificatórias adaptadas a nova realidade. Desta forma, a aplicação intertemporal das funções aqui estimadas deve ser feita com cuidado.

Ainda associado à dinâmica no estilo de investimento dos fundos, é capaz que alguns gestores alterem seu comportamento não para estratégias totalmente novas, mas sim para se movimentarem entre as categorias já identificadas. Muito embora se suspeite que gestores mantenham razoável estabilidade em seus estilos visando reduzir custos de transação e também atender à legislação e aos anseios dos investidores por fundos com perfil bem definido, esta é uma preocupação que surge. Se a movimentação entre categorias fosse algo comum e corriqueiro, a informação prestada pelas categorias seria inútil, pois um fundo ora teria uma classificação, ora teria outra. O teste de previsão de retornos futuros mostrou alto  $R^2$  médio e com baixa amplitude de variação, o que é evidência contra a instabilidade dos estilos. Entretanto, há de se admitir que os testes preditivos foram feitos sobre um período bastante curto.

A segunda limitação do modelo está ligada aos índices ausentes na Análise de Estilo, tanto os ausentes por força maior quanto aqueles descartados como consequência da delimitação do escopo. Índices de crédito privado (debêntures e notas promissórias emitidas publicamente), índices ligados a ativos imobiliários (CRIs) e à volatilidade de opções não foram tratados ou por deficiência do mercado brasileiro ou por característica natural impeditiva do ativo. Adicionalmente, não foram incluídos índices de ativos internacionais. Caso haja algum fundo com aplicações significativas em sua carteira em alguns desses índices, o termo residual da estimação do estilo estará indevidamente superavaliado. Um potencial mitigador do problema é a eventual existência de alta correlação entre esses índices com os índices contemplados no modelo.

Com relação aos parâmetros internacionais, convém salientar que vários índices são apurados ao redor do mundo, ao menos os mais representativos, e estão disponíveis para coleta. A justificativa para excluí-los do modelo apresentado foi que se identificou uma maneira de torná-lo mais parcimonioso, sem que isto representasse riscos materiais à

generalização das conclusões. Porém, na hipótese de os gestores brasileiros começarem a se interessar por mercados externos, é necessária imediata inserção desses índices durante o processo de análise.

Da mesma forma, se alguma nova classe de ativos for criada no mercado financeiro brasileiro, o modelo deveria ser reestimado a fim de considerá-la.

## **6.2. Direcionamentos futuros**

Uma primeira oportunidade de pesquisa futura que vem à mente seria a realização de testes intertemporais do modelo visando avaliar se sua capacidade preditiva se mantém.

Outra possível alternativa seria a estimação dinâmica do estilo dos fundos. Uma vez estimadas as séries temporais dos parâmetros de estilo, poderiam ser elaborados testes individuais de estacionariedade sobre cada uma destas séries. Fundos com ao menos um parâmetro avaliado como não-estacionário seriam então classificados como “fundos mutantes”. Esta seria uma informação valiosa ao investidor, que provavelmente evitaria tais fundos, devido ao fato de não oferecerem previsibilidade nenhuma da relação risco-retorno e também por prejudicarem completamente o processo de diversificação da carteira. A fim de implementar a estimação dinâmica, sugere-se a utilização do modelo em Espaço-Estado mencionado por Pizzinga (2004).

Uma terceira hipótese de pesquisa seria incluir classes de ativos internacionais ao modelo desde já (sem aguardar que gestores demonstrem maior interesse por mercados externos) e verificar se a distribuição dos  $R^2$  calculados para avaliar o grau de ajuste das estimações de estilo dos fundos destoa muito do encontrado na presente dissertação (Gráfico 6). Caso se conclua que não, significaria que as suspeitas levantadas aqui se fazem valer, ou seja: os cuidados tomados durante a fase de definição de escopo mitigaram/eliminaram os riscos de distorção nos resultados.

Por fim, outros métodos de Análise de *Clusters* mais sofisticados podem ser aplicados – com vistas à avaliação da robustez do modelo – na fase de identificação do número de categorias e definição de fronteiras, tais como: *fuzzy clustering* ou *simulated annealing*.

## 7. REFERÊNCIAS BIBLIOGRÁFICAS

ANBID; **Deliberação ANBID nº 32/07**; disponível no endereço eletrônico [http://www.anbid.com.br/documentos\\_download/auto\\_regulacao/fundos/deliberacoes\\_e\\_pareceres/Deliberacao32-anexo1.pdf](http://www.anbid.com.br/documentos_download/auto_regulacao/fundos/deliberacoes_e_pareceres/Deliberacao32-anexo1.pdf), acessado em 25 de março de 2008.

ANDIMA; **Metodologia do IMA**; disponível no endereço eletrônico [http://www.andima.com.br/ima/arqs/ima\\_cartilha.pdf](http://www.andima.com.br/ima/arqs/ima_cartilha.pdf), acessado em 14 de março de 2008.

BOECHAT, D.; 2008; **Emissões de valores mobiliários consolidam tendência de alta**; Retrospectiva ANDIMA 2007. Disponível no endereço eletrônico [http://www.andima.com.br/publicacoes/arqs/2007\\_valores.pdf](http://www.andima.com.br/publicacoes/arqs/2007_valores.pdf), acessado em 04 de abril de 2008.

BROWN, S.; GOETZMANN, W.; 1997; **Mutual Fund Styles**; Journal of Financial Economics, v. 43, p. 373-399.

CHEVALIER, J.; ELLISON, G.; 1999; **Carreer Concerns of Mutual Fund Managers**; Quarterly Journal of Economics, v. 114, n. 2, p. 389-432.

CVM; **Sistema de Consulta a Carteiras de Fundos**; disponível no endereço eletrônico <http://cvmweb.cvm.gov.br/SWB/Sistemas/SCW/CPublica/FormBuscaPartic.aspx?TpConsulta=5>, acessado em 20 de abril de 2008.

DERMINE, J.; ROLLER, L.; 1992; **Economies of Scale and Scope in French Mutual Funds**; Journal of Financial Intermediation, v. 2, i. 1, p. 83-93.

ELTON, E.; GRUBER, M.; BLAKE, C.; 2003; **Incentive Fees and Mutual Funds**; The Journal of Finance, v. 58, n. 2, p. 779-804.

FRALEY, C.; RAFTERY, A.; 1998; **How Many Clusters? Which Clustering Method? Answers Via Model-Based Cluster Analysis**; University of Washington Technical Report n. 329.

GABAIX, X.; LAIBSON, D.; 2003; **Some Industrial Organization with Boundedly Rational Consumers**; Working Paper; *apud* MAHONEY, P.; 2004; **Manager-Investor Conflicts in Mutual Funds**; Journal of Economic Perspectives, v. 18, n.2, p. 161-182.

GALLO, J.; LOCKWOOD, L.; 1997; **Benefits of Proper Style Classification of Equity Portfolio Managers**; Journal of Portfolio Management, v. 23, p. 47-55.

HAIR Jr., J.; ANDERSON, R.; TATHAM, R.; BLACK, W.; BABIN, B.; 2005; **Multivariate Data Analysis**; Prentice Hall, New Jersey.

HUBERTY, C.; 1984; **Issue in the Use and Interpretation of Discriminant Analysis**; Psychological Bulletin, v. 95, n. 1, p. 156-171; *apud* SHARMA, S.; 1996; Applied Multivariate Techniques; John Wiley & Sons, New York.

KIM, M.; SHUKLA, R.; TOMAS, M.; 2000; **Mutual Fund Objective Misclassification**; Journal of Economics and Business, v. 52, p. 309-323.

LE SOURD, V.; 2006; **Return-Based Style Analysis: an Answer to the Difficulties of Implementing Holding-Based Style Analysis**; Funds Europe, November; disponível em [http://www.edhec-risk.com/latest\\_news/featured\\_analysis/RISKArticle.2007-05-21.0503?newsletter=yes](http://www.edhec-risk.com/latest_news/featured_analysis/RISKArticle.2007-05-21.0503?newsletter=yes), acessado em 08 de março de 2007.

MALHOTRA, D.; MARTIN, R.; RUSSEL, P.; **Determinants of Cost Efficiencies in the Mutual Fund Industry**; Review of Financial Economics, v. 16, p. 323-334.

MAMAYSKY, H.; SPIEGEL, M.; 2001; **A Theory of Mutual Funds: Optimal Fund Objectives and Industry Organization**; Yale ICF Working Paper n. 00-50. Disponível em [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=281000](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=281000), acessado em 03 de março de 2008.

MAHONEY, P.; 2004; **Manager-Investor Conflicts in Mutual Funds**; Journal of Economic Perspectives, v. 18, n.2, p. 161-182.

MAYES, T.; JAY, N.; THURSTON, R.; 2000; **A Returns-Based Style Analysis Examination of Asset Classes**; Journal of Financial Planning, v. 13, n. 8, p. 94-104.

MINISTÉRIO DA FAZENDA, Secretaria do Tesouro Nacional; **Relatório Anual da Dívida Pública Federal de 2007**; Brasília, 2008. Disponível em [www.tesouro.fazenda.gov.br/divida\\_publica/downloads/Relatorio\\_Divida\\_2007.pdf](http://www.tesouro.fazenda.gov.br/divida_publica/downloads/Relatorio_Divida_2007.pdf), acessado em 20 de março de 2008.

MORENO, D.; MARCO, P.; OLMEDA, I.; 2006; **Self-Organizing Maps Could Improve the Classification of Spanish Mutual Funds**; European Journal of Operational Research, v. 174, p. 1039-1054.

NEUBERT, A.; **Knowing Your Indexes: Background, Definitions and Rules**; In: FABOZZI, F.; 1997; **Professional Perspectives on Indexing**; *apud* VARGA, G.; 1999; **Índices de Renda Fixa para o Brasil**; Resenha BM&F, n. 131, p. 55-60.

NORUSIS, M.; 1988; **SPSS/PC+ Advanced Statistics™ V2.0**; SPSS Inc., Chicago.

NORUSIS, M.; 2008; **SPSS 16.0 Statistical Procedures Companion**; Prentice Hall, New Jersey.

PARWADA, J.; 2003; **Trends and Determinants of Australian Managed Funds Transaction Costs**; Accounting and Finance, v. 43, p. 345-363.

PATTARIN, F.; PATERLINI, S.; MINERVA, T.; 2004; **Clustering Financial Time Series: an Application to Mutual Fund Style Analysis**; Computational Statistics & Data Analysis, v. 47, p. 353-372.

PIZZINGA, A.; 2004; **Modelos em Espaço de Estado com Restrições nas Componentes de Interesse: Aplicações em Análise Dinâmica de Estilo para Fundos de Investimento Brasileiros**; Dissertação de Mestrado; PUC-Rio.

SHARMA, S.; 1996; **Applied Multivariate Techniques**; John Wiley & Sons, New York.

SHARPE, W.; 1966; **Mutual Fund Performance**; Journal of Business, v. 39, n. 1 part 2, p. 119-138.

SHARPE, W.; 1992; **Asset Allocation: Management Style and Performance Measurement**; Journal of Portfolio Management, v. 18, n. 2, p. 7-19.

SIRRI, E.; TUFANO, P.; **Costly Search and Mutual Fund Flows**; The Journal of Finance, v. 53, n. 5; 1998.

VARGA, G.; VALLI, M.; 1998; **Análise de Estilo Baseada no Retorno**; Revista da ANBID, Dezembro; p. 12-16.

VARGA, G.; 1999; **Índices de Renda Fixa para o Brasil**; Resenha BM&F, n. 131, p. 55-60.

XUE-JUN, J.; XIAO-LAN, Y.; 2004; **Empirical Study on Mutual Fund Objective Classification**; Journal of Zhejiang University SCIENCE, v. 5, n. 5, p. 533-538.

## **APÊNDICE A – Fundos pesquisados no sistema de consulta a carteiras da CVM**

- Africa Fundo de Investimento Multimercado Agressivo
- BB Multimercado Alocação Agressivo Private FIC FI
- BB Top RF Agressivo Pré Fundo de Investimento Renda Fixa Longo Prazo
- HSBC FIC FI Multimercado Longo Prazo Estratégia Agressivo
- HSBC Fundo de Investimento Multimercado Previdenciário Agressivo
- HSBC Fundo de Investimento Multimercado Previdenciário Agressivo - VGBL
- Itaú Multimercado Agressivo - Fundo de Investimento
- Itaú Multimercado Agressivo – FIC FI
- Itaú Personnalité Multimanager Agressivo 30 Multimercado – FIC FI
- Itaú Personnalité Multimanager Agressivo Multimercado – FIC FI
- Itaú Personnalité Multimercado Agressivo – FIC FI
- Itaú Private Alocação Agressivo Multimercado – FIC FI
- Real FIC FI Multimercado Nederland Agressivo
- Realprev Agressivo Superior FIC FI Multimercado
- Safra Multicarteira Agressivo - Fundo de Investimento Multimercado
- SI Agressivo FIC FI Multimercado
- UBS Pactual Juros Agressivo Ativo FIC FI - Multimercado
- UBS Pactual Multicarteira Agressivo FIC FI Multimercado
- UBS Pactual Multigestores Agressivo FIC FI - Multimercado
- ZZ Agressivo FIC FI Multimercado
- BB Top Arbitragem Alavancado Fundo De Investimento Multimercado Longo Prazo
- Bradesco FIC FI em Ações Ibovespa Alavancado
- Bram Fundo de Investimento em Ações Ibovespa Alavancado
- BB Multimercado Alocação Arrojado Private FIC FI
- BB Multimercado Arrojado LP 10 Mil FIC FI

## **APÊNDICE A – Continuação**

- BB Multimercado Arrojado LP Estilo FIC FI
- BB Multimercado Arrojado LP FIC FI
- BB Regime Próprio Arrojado Fundo de Investimento Referenciado DI
- BB Top Multi Arrojado LP Fundo de Investimento Multimercado
- BB Top RF Arrojado Fundo de Investimento Renda Fixa Longo Prazo
- HSBC Fundo de Investimento Multimercado Longo Prazo Xian Arrojado
- Itaú Flexprev Private Arrojado V25 Multimercado- FIC FI
- Itaú Multimercado Arrojado - Fundo de Investimento
- Itaú Multimercado Arrojado – FIC FI
- Itaú Personnalité Multimercado Arrojado – FIC FI
- Real Investimento Personalizado Van Gogh Arrojado FIC FI Multimercado

## APÊNDICE B – Código do programa em Matlab para Análise de Estilo

```
% =====  
% Análise de Estilo (obs: código aplicável somente a 6 classes de ativos)  
% Usa o M-File: somquad.m  
% =====  
% início do programa  
clear all  
close all  
  
ativos=xlsread('ativos.xls');  
fundos=xlsread('fundos.xls');  
  
[ta J] = size(ativos);  
[tf N] = size(fundos);  
  
% Define as restrições da fmincon  
% =====  
% Cria a matriz A e o vetor b de restrições sobre a alavancagem conjunta em  
% mais de um ativo (SubA3 escrita manualmente e aplicável apenas a 6 ativos)  
% -----  
for z=1:J-1  
    SubAz=zeros(J-1,J);  
    for i=1:J-z  
        SubAz(i,z)=-1;  
        SubAz(i,i+z)=-1;  
    end  
    SSubAz(:,z)=SubAz;  
end  
  
SubA2=[SSubAz(1:J-1,1);SSubAz(1:J-2,2);SSubAz(1:J-3,3);SSubAz(1:J-4,4);...  
    SSubAz(1:J-5,5)];  
  
SubA3=[-1 -1 -1 0 0 0; -1 -1 0 -1 0 0; -1 -1 0 0 -1 0; -1 -1 0 0 0 -1; -1 0 -1 -1 0 0;...  
    -1 0 -1 0 -1 0; -1 0 -1 0 0 -1; -1 0 0 -1 -1 0; -1 0 0 -1 0 -1; -1 0 0 0 -1 -1;...  
    0 -1 -1 -1 0 0; 0 -1 -1 0 -1 0; 0 -1 -1 0 0 -1; 0 -1 0 -1 -1 0; 0 -1 0 -1 0 -1;...  
    0 -1 0 0 -1 -1; 0 0 -1 -1 -1 0; 0 0 -1 -1 0 -1; 0 0 -1 0 -1 -1; 0 0 0 -1 -1 -1];  
  
A=[SubA2;SubA3];  
  
% Calcula o fatorial de J para dar entrada na combinação (J tomados i a i)  
% que encontra o número de linhas da matriz A e do vetor b  
fator=ones(1,J);  
  
for i=2:J  
    fator(i)=i*fator(i-1);  
end  
  
CombinJ2a2=fator(J)/(fator(2)*fator(J-2));  
CombinJ3a3=fator(J)/(fator(3)*fator(J-3));
```

## APÊNDICE B – Continuação

```

b2 = ones(CombinJ2a2,1)*0.07;
b3 = ones(CombinJ3a3,1)*0.08;

b=[b2;b3];
% -----

lb = ones(J,1)*-0.05;
ub = ones(J,1)*1.05;
Aeq = ones(1,J);
beq = 1;
x0 = zeros(J,1);
%=====

% Define opções para o algoritmo da fmincon
options=optimset('TolCon',1e-70,'TolFun',1e-70,...
    'MaxFunEvals',100000,'MaxSQPIter',20000000,...
    'DiffMinChange',1e-70,...
    'MaxPCGIter',20000000,'TolPCG',0.00000005,'MaxIter',100000);

% Inicia o loop para estimação dos pesos para cada um dos N fundos
%=====
for i = 1:N
    fundos_i = fundos(:,i);
    % Roda a função com restrições, utilizando a função somquad (M-File construído
    % para calcular o somatório dos quadrados)
    parametros_i = fmincon(@(parametros_i)somquad(parametros_i,...
        ativos,fundos_i),x0,A,b,Aeq,beq,lb,ub,[],options);
    parametros(i,:)=parametros_i';

    % Calcula o R2 da estimação para cada fundo i
    RQuad(i)=var(ativos*parametros_i)/var(fundos_i);
end
%=====

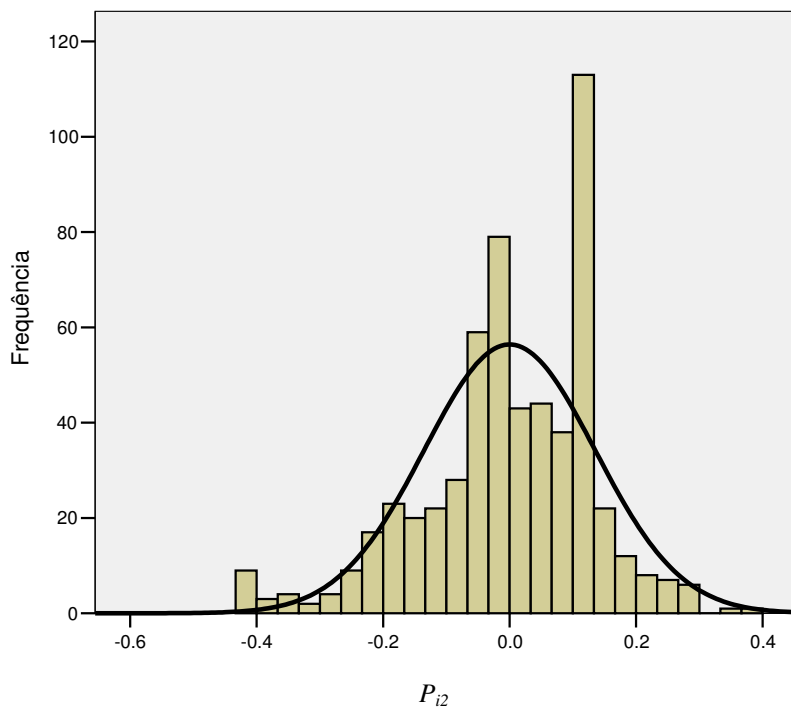
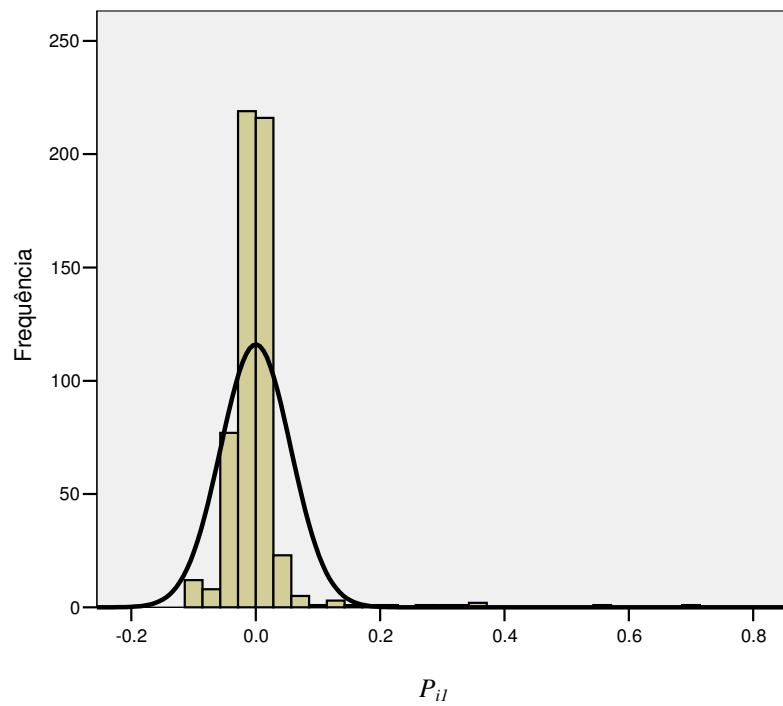
% registra os outputs
xlswrite('parametros.xls',parametros)
xlswrite('RQuad.xls',RQuad)
% fim do programa

% =====
% M-File: somquad.m >>> Calcula o somatório dos quadrados

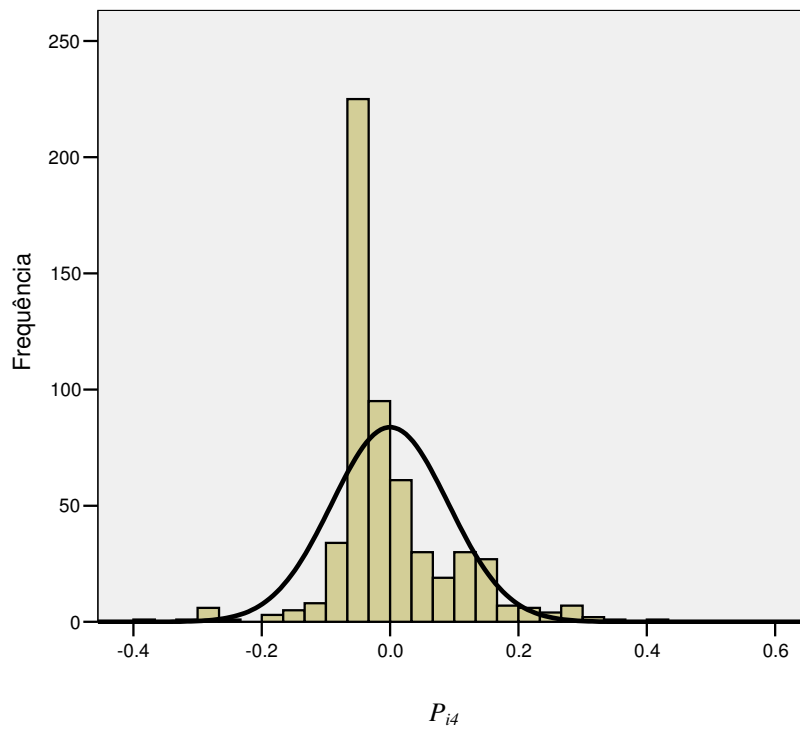
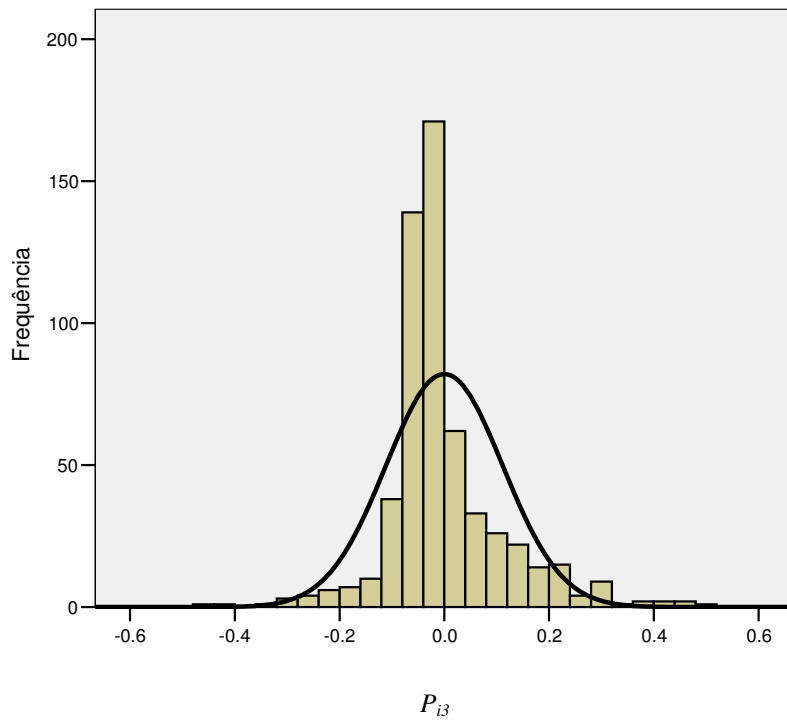
%calcula a soma do quadrado dos erros entre os retornos reais e os
%estimados
function f=somquad(parametros_i,ativos,fundos)
f = (fundos - ativos*parametros_i)*(fundos - ativos*parametros_i);
% =====

```

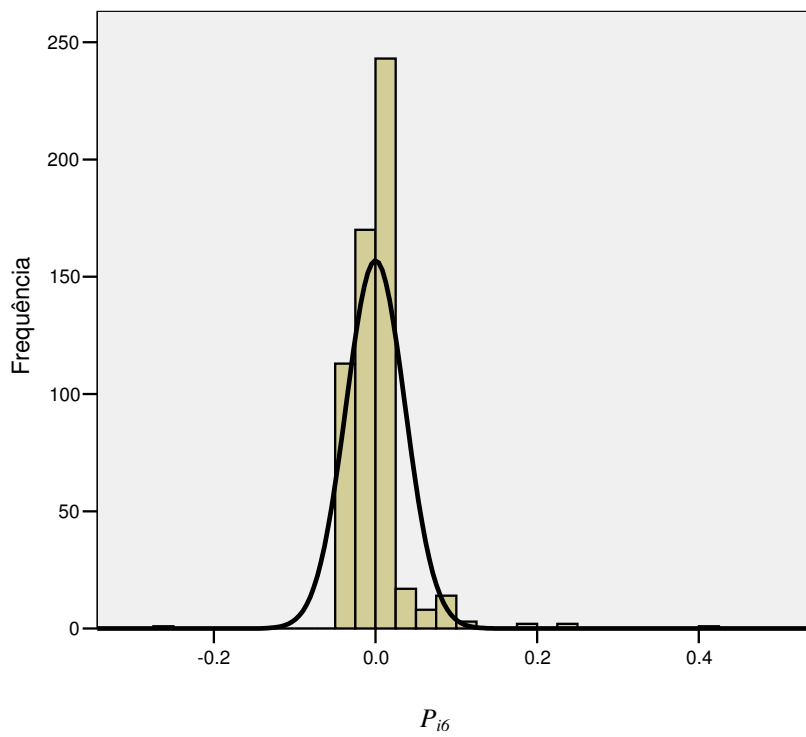
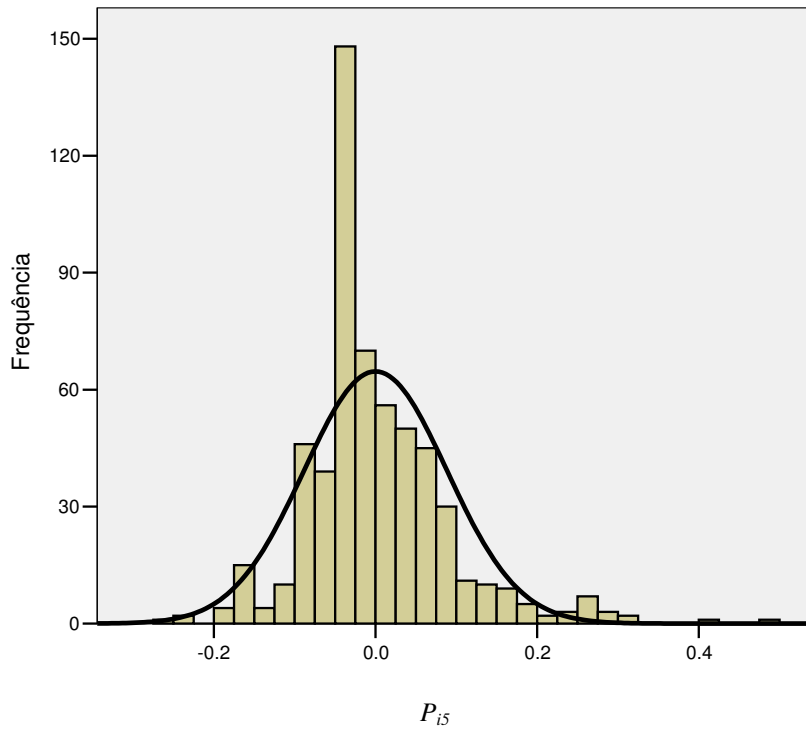
## APÊNDICE C – Histogramas das variáveis independentes da análise discriminante



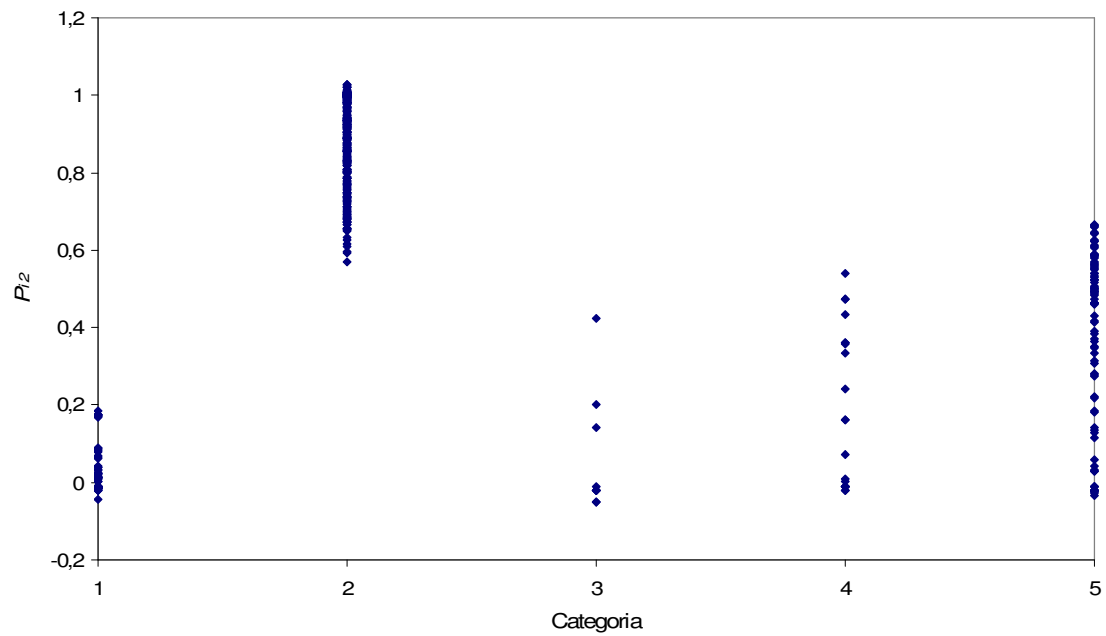
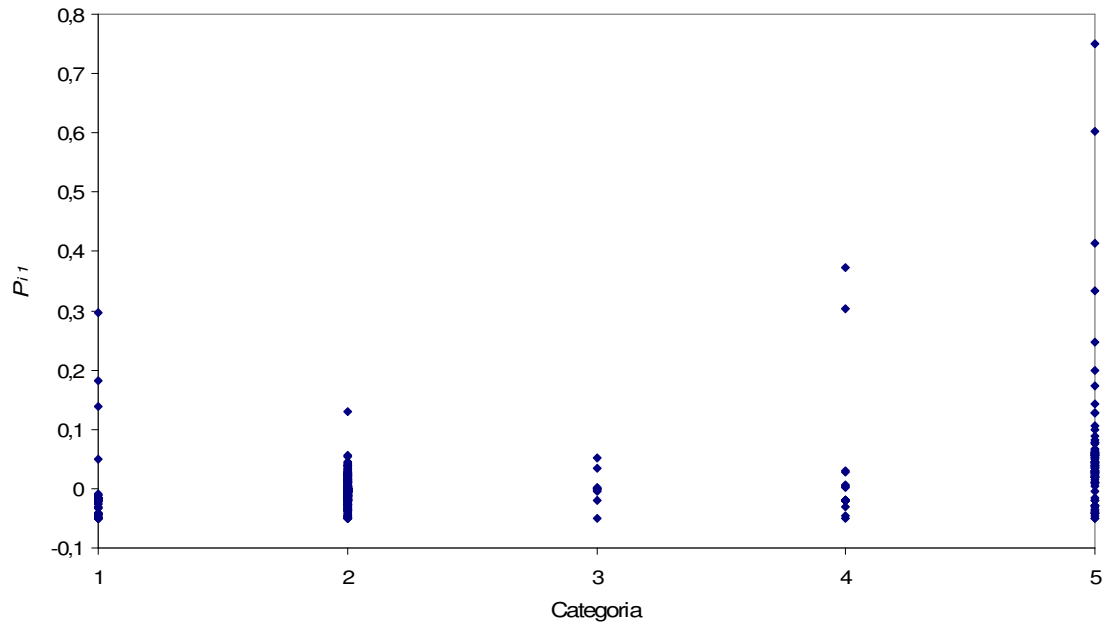
## APÊNDICE C – Continuação



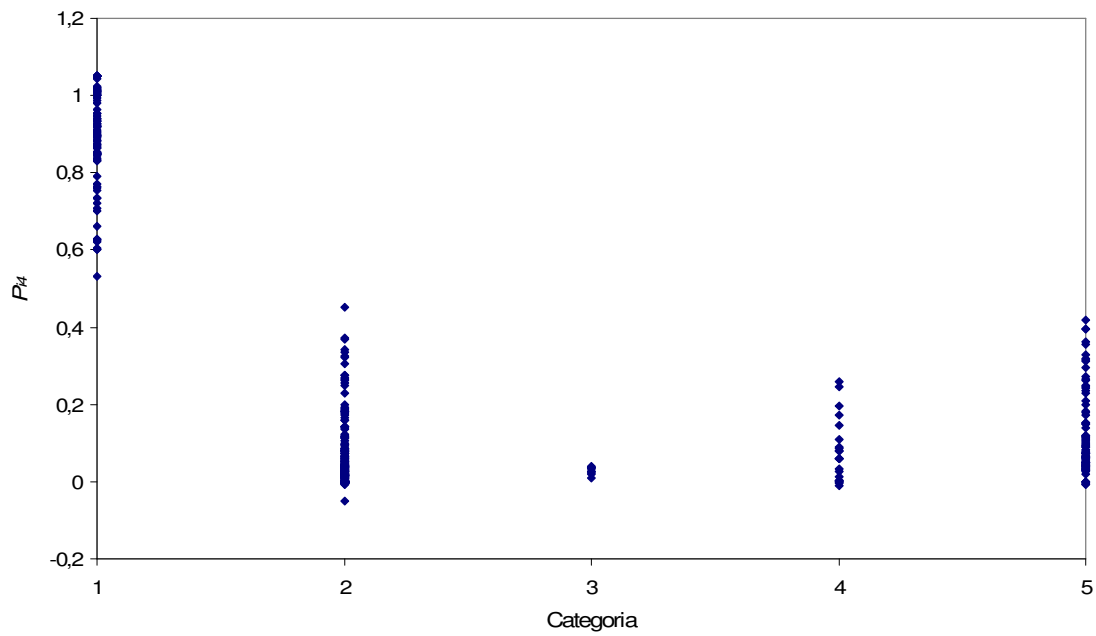
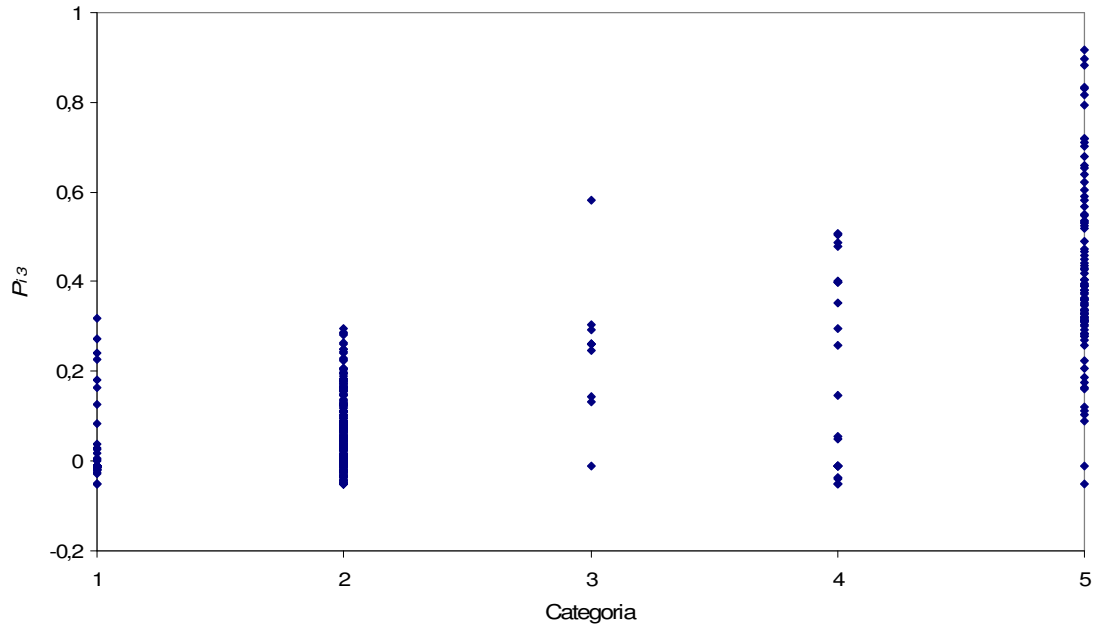
## APÊNDICE C – Continuação



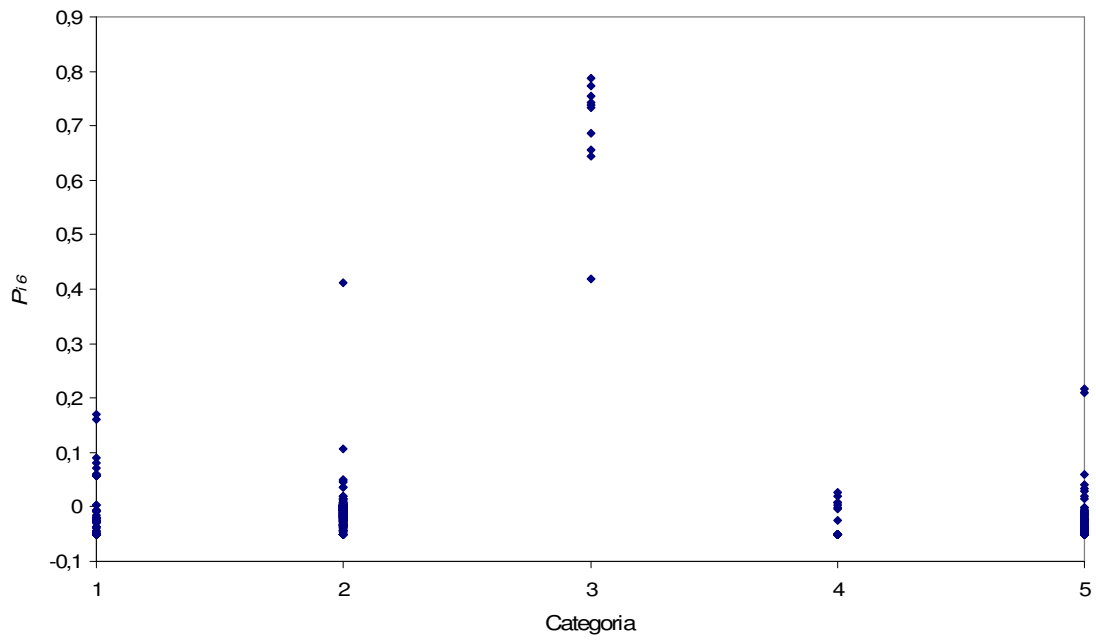
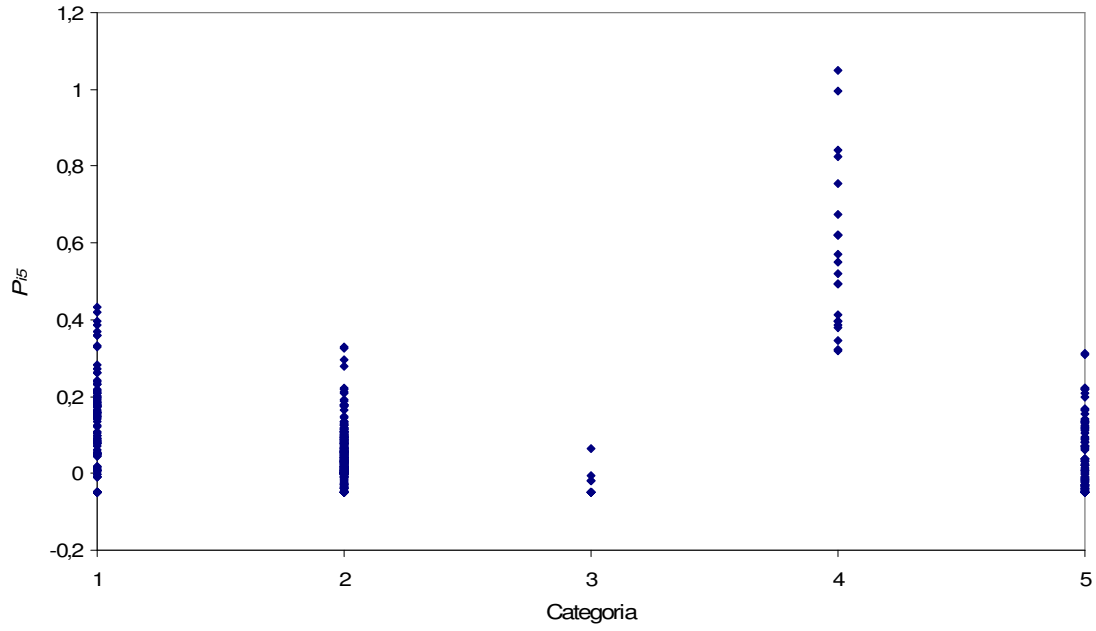
## APÊNDICE D – Gráficos de dispersão das variáveis discriminantes por categoria



## APÊNDICE D – Continuação



## APÊNDICE D – Continuação



# Livros Grátis

( <http://www.livrosgratis.com.br> )

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)  
[Baixar livros de Literatura de Cordel](#)  
[Baixar livros de Literatura Infantil](#)  
[Baixar livros de Matemática](#)  
[Baixar livros de Medicina](#)  
[Baixar livros de Medicina Veterinária](#)  
[Baixar livros de Meio Ambiente](#)  
[Baixar livros de Meteorologia](#)  
[Baixar Monografias e TCC](#)  
[Baixar livros Multidisciplinar](#)  
[Baixar livros de Música](#)  
[Baixar livros de Psicologia](#)  
[Baixar livros de Química](#)  
[Baixar livros de Saúde Coletiva](#)  
[Baixar livros de Serviço Social](#)  
[Baixar livros de Sociologia](#)  
[Baixar livros de Teologia](#)  
[Baixar livros de Trabalho](#)  
[Baixar livros de Turismo](#)