

FUNDAÇÃO GETÚLIO VARGAS
ESCOLA DE ECONOMIA DE SÃO PAULO

GABRIEL CAMPOS PÉRGOLA

Seguro contra risco de downside de uma carteira: uma proposta híbrida
frequentista-Bayesiana com uso de derivativos

SÃO PAULO

2012

GABRIEL CAMPOS PÉRGOLA

Seguro contra risco de downside de uma carteira: uma proposta híbrida
frequentista-Bayesiana com uso de derivativos

Dissertação apresentada à Escola de Economia de São
Paulo da Fundação Getúlio Vargas, como requisito para
obtenção do título de Mestre em Economia

Campo de conhecimento:

Finanças Quantitativas

Orientador: Prof. Dr. Juan Carlos Ruilova Terán

SÃO PAULO

2012

Pérgola, Gabriel Campos.

Seguro contra risco de downside de uma carteira: uma proposta híbrida frequentista-Bayesiana com uso de derivativos / Gabriel Campos Pérgola. - 2013.

68 f.

Orientador: Juan Carlos Ruilova Terán.

Dissertação (MPFE) - Escola de Economia de São Paulo.

1. Derivativos (Finanças). 2. Programa de Seguro de Carteira (Finanças). 3. Teoria bayesiana de decisão estatística. 4. Algoritmos. I. Terán, Juan Carlos Ruilova. II. Dissertação (MPFE) - Escola de Economia de São Paulo. III. Título.

CDU 336.763

GABRIEL CAMPOS PÉRGOLA

Seguro contra risco de downside de uma carteira: uma proposta híbrida
frequentista-Bayesiana com uso de derivativos

Dissertação apresentada à Escola de Economia de São
Paulo da Fundação Getúlio Vargas, como requisito para
obtenção do título de Mestre em Economia

Campo de conhecimento:

Finanças Quantitativas

Data de aprovação:

____/____/____

Banca examinadora:

Prof. Dr. Juan Carlos Ruilova Terán (Orientador)
FGV-EESP

Prof. Dr. Alessandro Martim Marques
FGV-EESP

Prof. Dr. Oswaldo Luiz do Valle Costa
USP-EPUSP

RESUMO

Seguros de carteiras proporcionam aos gestores limitar o risco de *downside* sem renunciar a movimentos de *upside*. Nesta dissertação, propomos um arcabouço de otimização de seguro de carteira a partir de um modelo híbrido frequentista-Bayesiano com uso de derivativos. Obtemos a distribuição conjunta de retornos regulares através de uma abordagem estatística frequentista, uma vez removidos os *outliers* da amostra. A distribuição conjunta dos retornos extremos, por sua vez, é modelada através de Redes Bayesianas, cuja topologia contempla os eventos que o gestor considera crítico ao desempenho da carteira. Unindo as distribuições de retornos regulares e extremos, simulamos cenários futuros para a carteira. O seguro é, então, otimizado através do algoritmo Evolução Diferencial. Mostramos uma aplicação passo a passo para uma carteira comprada em ações do Ibovespa, utilizando dados de mercado entre 2008 e 2012.

Palavras-chave: Seguro de carteira; Modelos híbridos; Modelos multivariados de retornos; Identificação de *outliers*; *Minimum Covariance Determinant*; Distribuição Hiperbólica Generalizada; Redes Bayesianas; Simulação; Otimização; Algoritmo Evolução Diferencial.

ABSTRACT

Portfolio insurance allows a manager to limit downside risk while allowing participation in upside markets. The purpose of this dissertation is to introduce a framework to portfolio insurance optimization from a hybrid frequentist-Bayesian approach. We obtain the joint distribution of regular returns from a frequentist statistical method, once the outliers have been identified and removed from the data sample. The joint distribution of extreme returns, in its turn, is modelled by a Bayesian network, whose topology reflects the events that can significantly impact the portfolio performance. Once we link the regular and extreme distributions of returns, we simulate future scenarios for the portfolio value. The insurance subportfolio is then optimized by the Differential Evolution algorithm. We show the framework in a step by step example for a long portfolio including stocks participating in the Bovespa Index (Ibovespa), using market data from 2008 to 2012.

Keywords: Portfolio insurance; Hybrid methods; Multivariate return models; Outlier detection; Minimum Covariance Determinant; Generalized Hyperbolic Distribution; Bayesian Nets; Simulation; Optimization; Differential Evolution Algorithm.

Lista de Figuras

2.1	Evolução do determinante da matriz de covariância em função do número de pontos removidos. Fonte: elaboração própria.	20
2.2	Um exemplo de aplicação para 3 ativos (IBOV, BVMF3 e GOLL4). Em azul, os retornos regulares (diários), representando o núcleo da amostra; em vermelho os retornos extremos detectados pelo MCD. Fonte: elaboração própria.	21
2.3	Ajuste da DHG e da Normal à densidade dos log-retornos diários de VALE5. Fonte: elaboração própria.	25
2.4	QQ-Plot dos log-retornos diários de VALE5: uma análise quantílica para a DHG e a distribuição Normal. Por essa metodologia, os retornos empíricos melhor se ajustam à distribuição cuja linha apresenta menor desvio em relação aos pontos da amostra. Fonte: elaboração própria.	26
3.1	Rede Bayesiana adotada para a aplicação. Com exceção dos nós de mesmo mercado ou setor, todos os eventos <i>down</i> são independentes dos eventos <i>up</i> , e vice-versa. Fonte: elaboração própria.	33
3.2	Distribuição das magnitudes das probabilidades de ocorrência dos eventos conjuntos de cada parte da RB (<i>up</i> e <i>down</i>), cada qual contendo 128 eventos. .	36
4.1	<i>Scatterplots</i> dos log-retornos de 3 ativos pela abordagem frequentista (esquerda) e híbrida frequentista-Bayesiana (direita). Fonte: elaboração própria.	40

5.1	Velocidade de convergência do algoritmo Evolução Diferencial para otimizações de seguro ótimo com uma média de 300 cenários de <i>downside</i> e 15 opções disponíveis.	
	Fonte: elaboração própria.	44
6.1	Histogramas de retornos simulados da carteira sem seguro (primeiro gráfico) e com seguro (segundo gráfico). O terceiro histograma corresponde ao seguro da carteira isolado.	
	Fonte: elaboração própria.	50
6.2	Desempenho da carteira da aplicação-exemplo com e sem seguro a partir de dados de mercado entre Agosto 2009 e Fevereiro 2012 (Valor inicial das carteiras = 100).	
	Fonte: elaboração própria.	53
C.1	Evolução da diversidade de opções de ações negociadas na Bovespa para o intervalo da amostra de dados.	
	Fonte: elaboração própria.	65

Lista de Tabelas

3.1	Classificação setorial para os ativos da amostra	29
3.2	Probabilidades marginais para os eventos <i>Alta</i> e <i>Baixa</i> do Mercado de ações	35
3.3	Probabilidades condicionais para as 12 arestas da RB	35
3.4	Retornos diários extremos, <i>up</i> e <i>down</i> , por ativo	38
6.1	Desempenho realizado das estratégias para o período e dados utilizados . . .	47
6.2	Carteira de ativos proposta para exemplificar a aplicação	49
6.3	Número de <i>outliers</i> e quantidade de <i>downsides</i> por rebalanceamento da aplicação-exemplo	51
6.4	Resumo do seguro para cada etapa do rebalanceamento da aplicação-exemplo	52
A.1	Descrição dos <i>tickers</i> dos ativos utilizados	63
B.1	Estatística descritiva dos log-retornos diários dos preços (ajustados para proventos) dos ativos da amostra de dados (%)	64
C.1	Representatividade de cada ativo subjacente da base de dados a partir do número de negócios com opções para o período entre Agosto de 2008 e Abril de 2012	66

Lista de Siglas

BOVESPA - Bolsa de Valores de São Paulo
CDF - Função distribuição acumulada
DAG - Directed Acyclic Graph
DHG - Distribuição hiperbólica generalizada
ED - Evolução Diferencial
IBOV - Índice Ibovespa
IGG - Inversa Gaussiana Generalizada
MCD - Minimum Covariance Determinant
MID - Preço médio
NBN - Naive Bayesian Network
PDF - Função densidade de probabilidade
RB - Rede Bayesiana

Conteúdo

1	Introdução	12
1.1	Resumo do problema	12
1.2	Dados utilizados	14
2	Modelagem frequentista	16
2.1	Identificação dos retornos regulares	16
2.2	Modelagem dos retornos regulares	22
2.2.1	Distribuição hiperbólica generalizada multivariada	24
3	Modelagem bayesiana	27
3.1	Definição e modelagem dos eventos extremos	27
3.2	Redes Bayesianas	31
3.2.1	Distribuição multivariada dos retornos extremos	36
4	Modelagem híbrida frequentista-bayesiana	39
4.1	Unindo retornos regulares e extremos	39
5	Otimização do seguro	41
5.1	Definição do problema de seguro ótimo	41
5.2	O algoritmo Evolução Diferencial	43
6	Resultados	45
6.1	Um exemplo passo a passo da aplicação	45
7	Conclusão	54
	Apêndice A Descrição dos tickers utilizados	63
	Apêndice B Estatísticas descritivas	64
	Apêndice C Evidências de liquidez para opções de ações no Brasil	65

1. Introdução

1.1 Resumo do problema

Ao longo das três últimas décadas, o mercado financeiro tem observado *crashes* de impacto e frequência cada vez maiores (Sornette (2004), Pasquariello (2012)), o que fortalece nossa intuição de esperar que um indivíduo que toma decisões de investimentos, *e.g.* um gestor de carteiras, deseje minimizar a probabilidade de perda de valor dos ativos de risco aos quais expõe suas carteiras. Intuitivamente, um gestor de uma carteira de ações, por exemplo, poderá adquirir opções de venda dos ativos da carteira para assegurar um determinado nível máximo de perda de valor.

Ao adquirir um seguro para o automóvel ou para a casa, por exemplo, o indivíduo expressa o desejo de receber uma determinada compensação material caso ocorra algum cenário que o desfavoreça; em contrapartida, desembolsa algum valor financeiro, tão maior quanto maiores sejam as recompensas que receberá em caso de sinistro e tanto maior quanto maior for a probabilidade de sinistro. De maneira análoga, o gestor de uma carteira de investimentos incrementará seus custos ao adquirir proteção contra cenários desfavoráveis, o que contribui negativamente para a performance da carteira.

Dessa maneira, com o constante aperfeiçoamento e complexidade do mercado financeiro, especialmente o de derivativos, o interesse em desenvolver estratégias eficazes e otimizadas de seguro de carteiras se torna patente. Como limitar o risco de *downside*¹ e, *pari passu*, usufruir de potenciais *upsides*? Como proteger a carteira contra cenários pré-determinados e com o menor custo financeiro possível? Como estruturar um seguro que contemple as crenças do gestor sobre potenciais eventos que podem impactar significativamente o valor da carteira?

As perguntas acima são algumas das motivações para a criação do arcabouço desta dissertação. O trabalho se distingue da literatura clássica de *portfolio insurance* (Leland e Rubinstein (1976), Rubinstein e Leland (1981), Perold (1986), Black e Jones (1987)) pois se

¹Risco de *downside* é aquele associado às perdas financeiras, *e.g.* no contexto de uma carteira é a diferença, quando negativa, entre o retorno realizado e o esperado.

aplica a um conjunto de ativos definidos, sem alterar o peso dos componentes da carteira. Do melhor de nosso conhecimento, este é o primeiro trabalho de otimização de seguro de carteira através de uma abordagem híbrida frequentista-Bayesiana.

Esquemáticamente, dados: a) uma carteira de ativos; b) um determinado nível de *downside* que se queira evitar; c) o *mindset* do gestor quanto a eventos que possam prejudicar o desempenho da carteira; d) um conjunto de derivativos disponíveis, pergunta-se, qual é o melhor subconjunto de derivativos e suas respectivas quantidades – em termos de efetividade de proteção e custo financeiro – que deve-se adquirir para que o valor da carteira de ativos não seja menor que o nível máximo de *downside* determinado pelo gestor ao fim de um horizonte temporal?

Esta dissertação enfrenta este problema decompondo-o em duas etapas. Primeiramente, é essencial quantificar o risco dos ativos de maneira acurada, o que suscita o uso de ferramentas estatísticas que vão além do paradigma gaussiano que sustenta o arcabouço clássico da teoria de finanças, como Bachelier (1900), Markowitz (1959) e Black e Scholes (1973). Em segundo lugar, a existência de riscos extremos - cuja modelagem entendemos ir além da estatística clássica ou frequentista - deve ser levado em conta no intento de quantificar os riscos dos ativos (Sornette (2006)).

Modelos quantitativos, em geral, são baseados em análises estatísticas das distribuições dos retornos dos ativos, que, por sua vez, são oriundos de séries históricas. Muitos artigos e discussões existem a respeito da descrição adequada dos retornos dos ativos e qual distribuição estatística melhor se ajusta aos dados. A premissa por trás desses modelos é que se espera que as respostas sejam encontradas através dos dados. Porém, essa discussão se torna opaca ao lidarmos com eventos de cauda, ou '*black swans*', *i.e.* eventos que podem criar consequências de dimensão sistêmica (Taleb (2010)).

Dessa maneira, seguindo a estrutura proposta por Rebonato e Denev (2011), propomos um método de proteção de carteira que integra probabilidades subjetiva e frequentista, na qual se assume a premissa de que em condições normais de mercado, *i.e.* removidos os *outliers* do conjunto de dados, os retornos dos ativos são bem comportados, de modo que se possa usufruir de técnicas estatísticas frequentistas para descrever a distribuição conjunta dos retornos dos ativos.

Em contrapartida, admite-se que retornos extremos são provenientes de eventos específicos (*e.g.* crises, do ponto de vista sistemático e quebras de empresas, *ceteris paribus*, do ponto de vista específico) e, assim, regularidades nesses eventos são muito difíceis de se descrever através de um modelo quantitativo baseado em dados históricos, fazendo que, em condições extremas de mercado, um *approach* frequentista não seja sustentável. Dessa forma, propomos que a modelagem dos retornos extremos seja resultado do entendimento do gestor da carteira de “como o mundo funciona”.

Assim, a modelagem dos retornos dos ativos será decomposta em duas etapas: a dos retornos extremos terá um método mais subjetivo, valendo-se de uma abordagem bayesiana; a dos retornos regulares, será orientada pela estatística clássica ou frequentista, através de uma distribuição de probabilidade multivariada.

A modelagem dos retornos regulares (Seção 2) e a dos retornos extremos (Seção 3) são, então, unidas para obtermos uma distribuição híbrida frequentista-Bayesiana (Seção 4). Com um processo gerador de retornos holístico, definimos o problema de seguro ótimo de uma carteira (Seção 5) e apresentamos a relevância prática da aplicação através dos resultados de um exemplo passo a passo (Seção 6) utilizando dados de mercado entre 2008 e 2012.

1.2 Dados utilizados

A base de dados utilizada consiste em uma amostra de preços diários de ações e opções de empresas cotadas na Bolsa de Valores de São Paulo (BOVESPA) e do preço diário do Índice Ibovespa (IBOV), entre o período de Agosto de 2008 e Abril de 2012, englobando um conjunto de 14 log-retornos de ativos ao longo de 905 dias. Os preços utilizados para as ações e o Índice são os de fechamento.

A amostra é composta pelos seguintes ativos (por *tickers*): AMBV4, BBAS3, BBDC4, BVMF3, CMIG4, GOAU4, GOLL4, IBOV, ITUB4, NATU3, PETR4, TRPL4, USIM5, VALE5. As cotações livres de ajustes e as ajustadas para proventos (dividendos, bonificações, direitos de subscrição, etc.) foram obtidas pelo sistema de informação ECONOMÁTICA.

A descrição dos *tickers* dos ativos estão no Apêndice A. As estatísticas descritivas dos log-retornos diários dos ativos para o período da amostra estão no Apêndice B.

Os dados relativos às opções (preço, *strike*², vencimento e ativo subjacente) foram obtidos a partir das Cotações Históricas do site da BOVESPA³. Os preços utilizados foram os preços médios (*MID*)⁴. A amostra de dados das opções é composta por “opções de venda” (*puts*)⁵ de estilo Europeu⁶ que tiveram ao menos 1 negócio no dia e que possuam ativo subjacente pertencente ao conjunto de ações supracitado, contemplando assim uma amostra de 31.586 observações (uma média de 35 opções disponíveis por dia).

O período utilizado inicia em 2008 devido ao fato que anteriormente o mercado de opções de ações brasileiro era considerado bem menos líquido, concentrando-se basicamente em dois nomes: PETR4 e VALE5 (Mello (2009), Carvalho (2012)). A escolha dos nomes das ações que compõem a amostra de dados se deve ao fato de as opções destas empresas apresentarem, para o período da amostra, continuidade de negociação e estarem entre as mais negociadas para o período da base de dados.

No Apêndice C encontra-se evidências de liquidez para o mercado de opções de ações da Bovespa: a Figura C.1 mostra a evolução da diversidade de opções de ações negociadas na Bovespa para o período entre Agosto de 2008 e Abril de 2012 através da média diária do número de opções distintas em que houve ao menos um negócio; a Tabela C.1 mostra o número de negócios com opções por ativo subjacente da amostra de dados, que representa 93,6% do total do número de negócios com opções de ações da Bovespa para o período da base de dados.

²Preço de exercício.

³<http://www.bmfbovespa.com.br/>. Acesso em 21 abril 2012.

⁴Escolhemos os preços médios para as opções com o intuito de mitigar a assimetria entre os horários de fechamento do mercado de opções e de ações, bem como reduzir eventuais distorções que ocorrem no fechamento do pregão (Almeida, Yoshino e Schirmer (2003), Mello (2009)).

⁵Não utilizamos “opções de compra” (*calls*) pois a aplicação, por simplicidade, contempla apenas carteiras com posições compradas (veja Seções 5 e 6).

⁶De exercício apenas na data de vencimento.

2. Modelagem frequentista

A modelagem dos retornos¹ dos ativos será decomposta em duas etapas: a dos retornos extremos terá um método mais subjetivo, valendo-se de uma abordagem bayesiana; a dos retornos regulares, desenvolvida neste capítulo, será orientada pela estatística clássica ou frequentista, através de uma distribuição de probabilidade multivariada. Assim, o primeiro passo é isolar as observações regulares das extremas na amostra multivariada de retornos.

2.1 Identificação dos retornos regulares

Uma proposta simples para separar os retornos regulares dos extremos seria definir, arbitrariamente, um número de desvios-padrão sobre o qual se entenda que os retornos dos ativos, sob condições normais de mercado, não ultrapassem. Então, eliminar-se-iam todos aqueles retornos da amostra que estão além desse nível, restando assim apenas aqueles que seriam considerados regulares ou não-extremos. Existe, entretanto, métodos mais objetivos e sofisticados que nos permite remover os *outliers* da série de retornos.

A remoção de *outliers* de amostras de dados é um tema presente há mais de 40 anos na literatura acadêmica, como se nota pelos artigos de Ferguson (1961) e Grubbs (1969). Ambos os trabalhos apresentam procedimentos estatísticos formais para detectar observações que são extremas ao ponto de não poderem ser explicadas por alguma distribuição de probabilidade sugerida. Dessa maneira, o êxito desses procedimentos está condicionado à correta hipótese do processo gerador dos dados e, além disso, se concentram no caso unidimensional, o que não se adequa ao objetivo desta dissertação, que exigirá um método que possua tratamento multidimensional.

Para tanto, seguimos o método *Minimum Covariance Determinant* (MCD), introduzido por Leroy e Rousseeuw (1987), para encontrar os retornos extremos das séries. O MCD é

¹Como considera Cappa e Valls (2010), a maioria dos estudos financeiros modelam retornos e não os preços dos ativos pelas propriedades estatísticas interessantes que os retornos apresentam, *i.e.* estes são, em geral, estacionários, ergódicos e compreendem uma medida livre de escala.

utilizado num amplo leque de ciências, tais como finanças, química, engenharia e medicina. Para esta última, por exemplo, uma aplicação recente é o trabalho de Prastawa et al. (2004), que utiliza o MCD num processo de detecção de tumores cerebrais. Para a aplicação do MCD, utilizamos o algoritmo de Rousseeuw e Van Driessen (1999).

O ponto de partida do MCD é a possibilidade de se definir, a partir da média e covariância de uma amostra multivariada, o elipsoide que melhor se ajusta às observações², *i.e.* o menor elipsoide dentre os que englobam todos pontos da amostra. No entanto, para amostras com *outliers*, o elipsoide tenderá a possuir estimadores de média e covariância de acurácia mais pobre devido ao esforço de levar em conta as observações distantes do núcleo da amostra. Há, ademais, outro resultado fundamental para a existência do MCD: o volume do elipsoide é proporcional à raiz quadrada do determinante da matriz de covariância da amostra³.

Dessa maneira, através do algoritmo que apresentamos a seguir, pode-se encontrar a observação que, se removida da amostra, produz a maior redução do determinante e, logo, do volume do elipsoide. Removido esse maior *outlier*⁴ da amostra, repete-se o algoritmo e se retira o segundo maior *outlier*. O procedimento deve ser repetido sob monitoramento da evolução da magnitude do determinante, com o intuito de aferir quantas observações são necessárias remover da amostra para que o determinante se estabilize. Isso feito, restará na amostra apenas o núcleo das observações, que, para a aplicação em tela, são os retornos regulares.

Entretanto, esse procedimento requer análise e interação humana, pois o ponto de parada do algoritmo está sujeito à avaliação (subjetiva) da estabilização do determinante da matriz de covariância. Dessa maneira, para fundamentarmos o ponto de parada do algoritmo com embasamento estatístico, aplicamos o método de Filzmoser (2004), que computa o ponto de parada do MCD a partir da amostra multivariada de dados. O princípio do método está em medir a dispersão da distribuição nas caudas em relação a uma Normal multivariada.

O método de Filzmoser se baseia em que, para uma amostra normal p -variada, os valores dos quadrados das distâncias de Mahalanobis (definido no Apêndice D) são aproximadamente distribuídos segundo uma distribuição chi-quadrado com p graus de liberdade (χ_p^2), possibili-

²Veja Meucci (2007, p. 54).

³Ver Meucci (*ibid.*) para a demonstração.

⁴Utilizamos o termo “maior *outlier*” para nos referir ao ponto mais distante do centro do elipsoide.

tando, assim, para fins de detecção de *outliers*, a comparação entre as caudas da distribuição da amostra e as caudas da distribuição χ_p^2 .

Embora o método tenha a premissa de uma distribuição multivariada normal nas caudas, Filzmoser (2005) faz um estudo de desempenho para diversas distribuições e constata que o método é “bastante estável” (p. 136; tradução nossa). Ademais, o método se mostra útil para diversos casos empíricos: amostras com caudas grossas, assimetria e de diferentes dimensões.

Assim, seguindo a aplicação de Rebonato e Denev (2011) e o desenvolvimento teórico de Meucci (2007), apresentamos abaixo em passos esquemáticos como encontrar os *outliers* da série multivariada de retornos.

1. Para cada ativo i pertencente à carteira, seu retorno (logarítmico)⁵, r_i^k , referente ao dia t_k , é definido como

$$r_i^k = \ln \left(\frac{p_i^k}{p_i^{k-1}} \right)$$

na qual p_i^k e p_i^{k-1} representam o preço do ativo i nos dias t_k e t_{k-1} , respectivamente.

2. Para uma amostra multivariada de retornos, de tamanho N $\{\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N\}$, sendo cada vetor de retornos $\mathbf{r}_{n(1 \leq n \leq N)}$ de dimensão m , definimos a matriz $N \times m$:

$$U = \begin{pmatrix} \mathbf{r}_1 - \bar{\mathbf{r}} \\ \vdots \\ \mathbf{r}_N - \bar{\mathbf{r}} \end{pmatrix}$$

na qual $\bar{\mathbf{r}}$, uma matriz $1 \times m$, é a média da amostra m -dimensional dos retornos \mathbf{r} .

3. A matriz de covariância pode, então, ser calculada da seguinte maneira:

$$\widehat{Cov} = \frac{1}{N} U'U$$

na qual U' denota a matriz U transposta.

⁵Os termos “retorno” e “log-retorno” serão utilizados intercambiadamente ao longo da dissertação.

4. Devemos, agora, encontrar o retorno $\mathbf{r}_{n(1 \leq n \leq N)}$ tal que se removido do conjunto $\{\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N\}$ o determinante da matriz de covariância, $|\widehat{Cov}|$, se reduza o máximo⁶. Para isso, seguimos o seguinte resultado de Poston et al. (1997):

$$|U'_{(-n)}U_{(-n)}| = (1 - \lambda_n)|U'U|$$

na qual $U_{(-n)}$ é a matriz U após a remoção de sua n -ésima linha e λ_n representa o n -ésimo elemento da diagonal da matriz $N \times N$ ⁷:

$$\lambda_n = (U(U'U)^{-1}U')_{(nn)}.$$

Logo, removendo a observação correspondente ao maior valor de λ_n , retiramos o vetor de retornos que corresponde ao maior *outlier* da série multivariada $\{\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N\}$. Obviamente, isso corresponde em remover a n -ésima linha da matriz U .

5. Como ponto de parada do algoritmo, as sugestões de Meucci (2007) e Denev (2011) carecem de objetividade, apoiando-se basicamente na inspeção visual ou qualitativa da evolução do determinante. Dessa maneira utilizamos alternativamente o método de Filzmoser (2004), que estima o ponto de separação entre o núcleo dos dados e os *outliers* a partir da distribuição das distâncias de Mahalanobis. Assim, itera-se o algoritmo (das etapas 2 a 4) o número de vezes correspondente à quantidade de *outliers* que o método indicar⁸.

Um exemplo do resultado desse procedimento é mostrado nas Figuras 2.1 e 2.2. A Figura 2.1 mostra a evolução da magnitude do determinante da matriz de covariância em função do número de pontos removidos. Percebemos que para as diversas amostras de 253 observações da aplicação-exemplo, a remoção de cerca de 30 observações é suficiente para estabilizar a matriz de covariância (veja Tabela 6.3).

⁶Utilizamos a notação $|M|$ para o determinante da matriz M .

⁷Utilizamos a notação $M_{(ij)}$ para designar o elemento da matriz M localizado na i -ésima linha e j -ésima coluna.

⁸O método de Filzmoser (2004) é executado através do pacote *mvoutlier: Multivariate outlier detection based on robust methods*, executado no ambiente do *software* R (R Development Core Team (2005)). Veja Filzmoser e Gschwandtner (2012) para a documentação; <http://cran.r-project.org/web/packages/mvoutlier/> para baixar o pacote (Acesso em 20 nov. 2012).

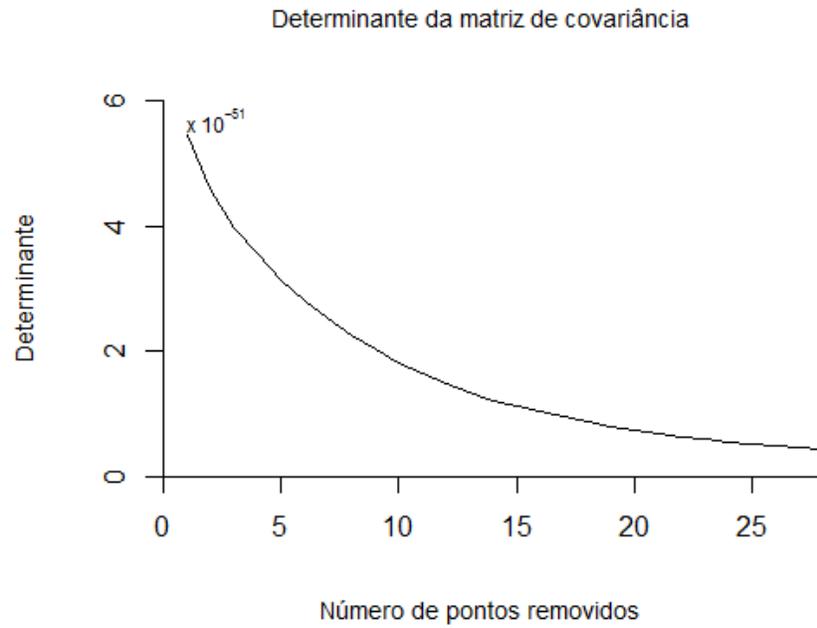


Figura 2.1: Evolução do determinante da matriz de covariância em função do número de pontos removidos.

Fonte: elaboração própria.

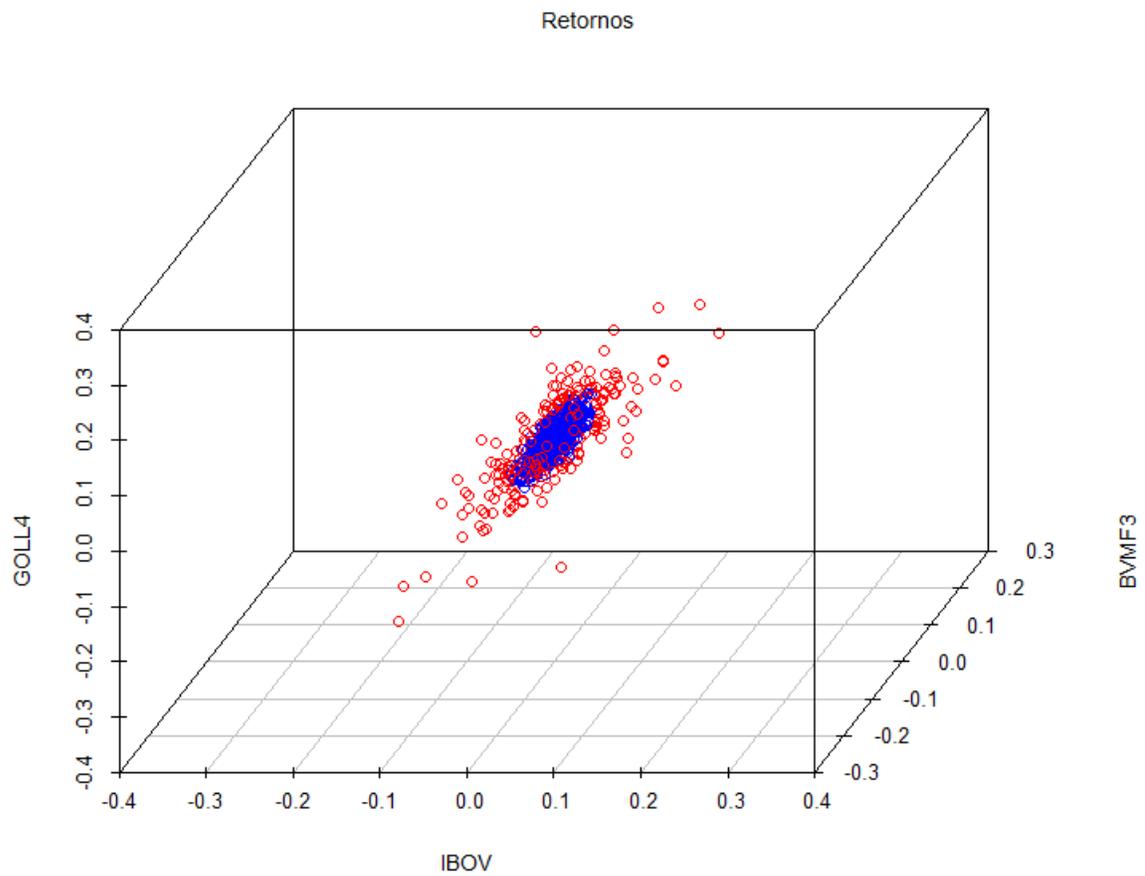


Figura 2.2: Um exemplo de aplicação para 3 ativos (IBOV, BVMF3 e GOLL4). Em azul, os retornos regulares (diários), representando o núcleo da amostra; em vermelho os retornos extremos detectados pelo MCD.

Fonte: elaboração própria.

2.2 Modelagem dos retornos regulares

Uma vez identificados os retornos regulares, sua estimação será realizada através de um modelo paramétrico⁹ que se ajuste à distribuição empírica de seus retornos. Por se tratar de um modelo multivariado, sua construção pode ser realizada por dois métodos:

1. Ajustar as distribuições paramétricas univariadas individualmente para cada ativo, selecionar uma cópula¹⁰ para representar a estrutura de dependência entre as distribuições marginais e, então, obter a distribuição conjunta; ou
2. Estimar diretamente uma distribuição paramétrica multivariada, contemplando numa só estrutura a modelagem de todos os ativos e a estrutura de dependência entre eles.

No contexto de uma modelagem paramétrica, o primeiro método oferece a possibilidade de estimar as distribuições marginais com diferentes distribuições de probabilidade. Além disso, a seleção da distribuição da cópula oferece a possibilidade de modelar diferentes padrões de dependência entre os ativos. Entretanto, como destaca Fischer et al. (2009), é incipiente em finanças a comprovação empírica de eficácia da aplicação de cópulas de dimensões grandes (maior que 2) com distribuição diferente da Normal ou t-student. Além disso, dado as diversas formas funcionais possíveis, não há regras gerais para uma seleção ótima no caso paramétrico (Panchenko (2005)).

O segundo método, por sua vez, é baseado explicitamente em distribuições multivariadas, o que requer uma distribuição suficientemente flexível para ajustar as distribuições marginais de cada ativo e, ao mesmo tempo, capturar a estrutura de dependência entre eles.

Joe (1997) destaca algumas propriedades desejáveis em um modelo multivariado: a) ter um formato (*shape*) apropriado de densidade para as distribuições marginais; b) ter uma

⁹Outras abordagens estatísticas na modelagem de séries de retornos financeiros, como a não-paramétrica e a semi-paramétrica, também são possíveis. Adotamos a abordagem paramétrica por sua parcimônia operacional e tratabilidade analítica. Para uma breve discussão sobre abordagens estatísticas na modelagem de séries de retornos, veja Cont (2001).

¹⁰Cópulas são funções que permitem representar uma distribuição conjunta como uma função de distribuições marginais, conforme estabelece o teorema de Sklar (1959). Para uma referência abrangente sobre cópulas, veja Nelsen (1999). Para uma revisão da literatura em aplicações de cópulas na modelagem de séries financeiras, veja Patton (2009).

estrutura de dependência adequada; c) ser o mais parcimonioso quanto possível; d) ter parâmetros que devem ser facilmente interpretáveis; e) possuir representação analítica (ou computacionalmente calculável) das distribuições acumulada (CDF) e de densidade (PDF). Joe expõe, ainda, que geralmente não é possível satisfazer todas essas propriedades, o que nos leva a ponderar a importância de cada propriedade para o problema em questão e abandonar uma ou mais delas.

Outro ponto a ser levado em conta para a escolha do método e da distribuição é que os retornos regulares também podem apresentar caudas grossas (*fat tails*)¹¹. Isso é consequência de fenômenos que ocorrem regularmente nos mercados financeiros, *e.g.* volatilidade estocástica, leptocurtose e *jumps*. Dessa maneira, a distribuição de retornos diários de ativos financeiros exibirão, geralmente, caudas mais grossas e maior concentração em torno da média em relação a uma distribuição Normal, na qual estão baseados os modelos clássicos de finanças, tais como CAPM, APT e Black e Scholes¹². Assim, a modelagem dos retornos diários de ativos financeiros seguindo uma distribuição Normal nem sempre é uma hipótese razoável¹³, como mostram Ané e Kharoubi (2003) e Embrechts e Straumann (2002).

Considerando o exposto, uma distribuição candidata à modelagem dos retornos regulares é a hiperbólica generalizada (DHG; discutida na próxima Seção), que, além de demonstrar evidências de contemplar esses fatos estilizados presentes nos retornos de ativos financeiros, possui muitas das propriedades desejáveis para um modelo multivariado, o que a tornou uma ferramenta bastante popular na modelagem de retornos de ativos (Borak, Misiolek e Weron (2011)).

Dessa maneira, para a aplicação em tela, optamos modelar os retornos regulares através do segundo método exposto acima com o uso da DHG. A escolha de estimar diretamente uma distribuição paramétrica multivariada se deve basicamente à maior parcimônia e simplicidade

¹¹É sutil porém importante ressaltar que a remoção dos *outliers* foi realizada em um espaço dimensional específico. Assim, alguns pontos considerados *outliers* pelo MCD em uma dimensão específica podem não sê-lo em outras dimensões, e vice-versa. Um exemplo gráfico pode ser visto em Denev (2011, p. 22)

¹²Para uma revisão dos modelos clássicos de finanças, veja Dimson e Mussavian (1999).

¹³Uma característica empírica frequentemente discutida é que as distribuições de retornos de ativos financeiros mais se aproximam da distribuição gaussiana a medida que se incrementa o horizonte temporal observado (*e.g.* semanal, mensal).

operacional que a abordagem via cópulas. Não obstante, o processo gerador de retornos regulares pode ser modelado de maneira distinta sem prejuízo ao arcabouço como um todo.

2.2.1 Distribuição hiperbólica generalizada multivariada

A classe das Distribuições Hiperbólicas Generalizadas (DHG)¹⁴ aparece pela primeira na literatura em Barndorff-Nielsen (1977), ao aplicar a sub-classe Hiperbólica no estudo da distribuição do tamanho de partículas de areia sujeitas à ação do vento. Um ano depois, em Barndorff-Nielsen (1978), os conceitos foram generalizados para a classe DHG.

A primeira aplicação da DHG em finanças é a de Eberlein e Keller (1995), que apresenta a modelagem de retornos de ações alemãs com bastante acurácia em relação a outras distribuições¹⁵. Eberlein e Prause (2000) aplicam a DHG à derivativos e medidas de risco, concluindo que “a flexibilidade da DHG permite descrever as distribuições de retornos de maneira quase perfeita” (p. 24; tradução nossa); Bauer (2000), por sua vez, aplicando a DHG à mensuração de riscos de mercado, também comprova o melhor ajuste desta distribuição para as séries de retornos em relação à distribuição normal.

Na última década foram desenvolvidas algumas aplicações para testar a aderência da DHG aos retornos dos ativos do mercado brasileiro. Barbachan, Schuschny e Silva (2001) apresentam evidências empíricas da aplicabilidade da distribuição hiperbólica (uma sub-classe da DHG) para modelar os retornos do índice Ibovespa, bem como para os retornos de outras ações brasileiras. Fajardo e Farias (2004) apresentam resultados semelhantes ao mostrar que a DHG “se adequa muito bem” (p. 269; tradução nossa) aos log-retornos dos ativos de maior liquidez do índice Ibovespa. Barbachan e Ornelas (2003) propõem um modelo para precificar opções de IDI e constata que a DHG se adequa melhor aos retornos dos títulos pré-fixados do que a modelagem com a distribuição Normal, que é a mais utilizada pelo mercado.

Por ser obtida como uma mistura na média e na variância da Normal, uma vantagem da

¹⁴O nome da distribuição é oriundo do fato de sua log-densidade ser uma hipérbole. A log-densidade da distribuição normal, por exemplo, é uma parábola. Um exemplo gráfico pode ser visto em Fonseca (2004) e Borak, Misiorek e Weron (2011).

¹⁵Como curiosidade, segundo Eberlein e Keller (1995), o próprio Barndorff-Nielsen foi quem sugeriu pessoalmente a ele a aplicação da DHG em séries de retornos financeiros.

família DHG é a de ser facilmente estendida a muitos casos particulares e distribuições limites. Alguns exemplos são a Hiperbólica, Gama, Normal, Normal Inversa Gaussiana, t-Student, Laplace Assimétrica e *Variance Gamma*. Outra vantagem é a capacidade de descrever assimetrias, caudas pesadas e fornecer, numa estrutura multivariada, dependência não-linear. Desenvolvemos a especificação analítica da DHG no Apêndice D, indicando referências para outras demonstrações e mais detalhes sobre sua parametrização.

Em 2007 surgiu a primeira versão do pacote *ghyp*: *A package on the generalized hyperbolic distribution and its special cases*¹⁶, que provê diversas funcionalidades para as DHG uni e multivariadas, contemplando grande parte de suas sub-classes de distribuições. As estimações são calculadas numericamente por máxima verossimilhança, sendo McNeil, Frey e Embrechts (2005) a principal referência teórica.

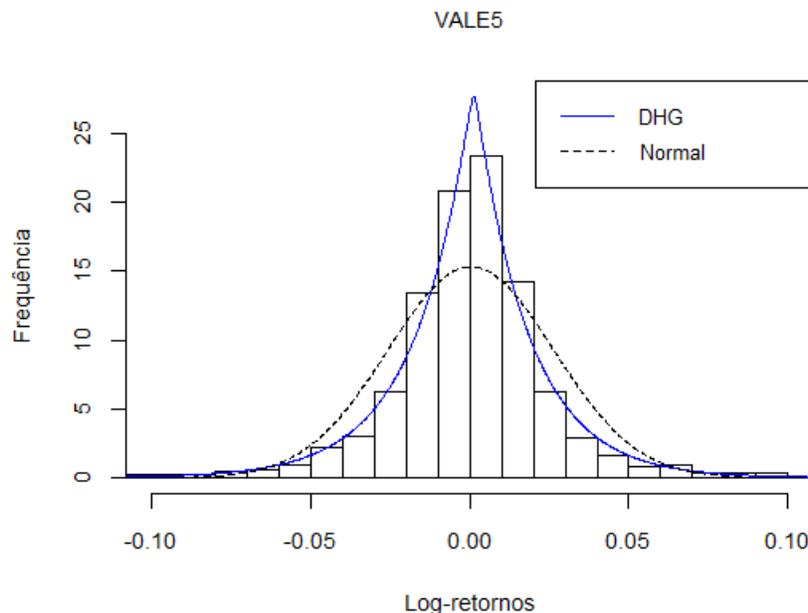


Figura 2.3: Ajuste da DHG e da Normal à densidade dos log-retornos diários de VALE5.

Fonte: elaboração própria.

Assim, desenvolvemos a modelagem das séries dos retornos regulares desta dissertação através do pacote *ghyp*, executado no ambiente do *software* R (R Development Core Team

¹⁶Veja Breymann e Luthi (2011) para a documentação; <http://cran.r-project.org/web/packages/ghyp/> para baixar o pacote (Acesso em 9 set. 2012.).

(2005)). As Figuras 2.3 e 2.4 mostram o ajuste executado a partir deste aplicativo para a série de log-retornos de uma ação da base de dados.

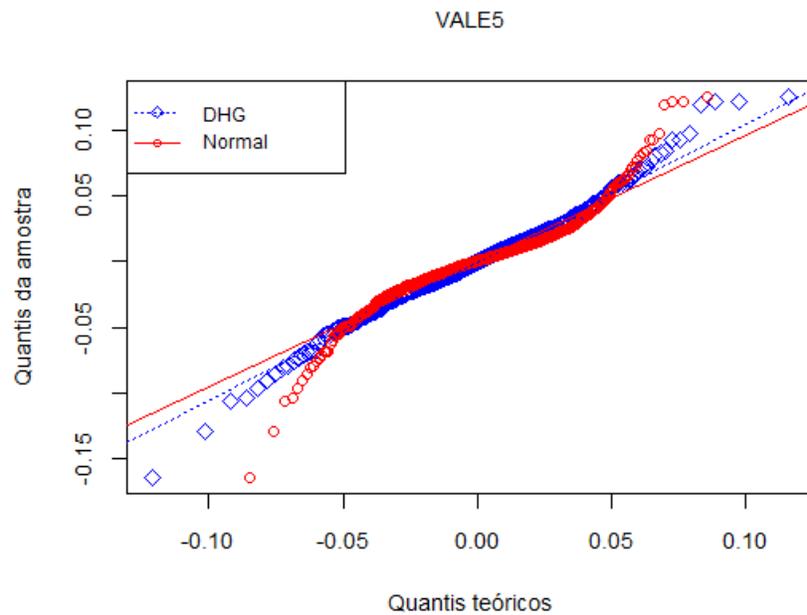


Figura 2.4: QQ-Plot dos log-retornos diários de VALE5: uma análise quantílica para a DHG e a distribuição Normal. Por essa metodologia, os retornos empíricos melhor se ajustam à distribuição cuja linha apresenta menor desvio em relação aos pontos da amostra.

Fonte: elaboração própria.

3. Modelagem bayesiana

Na Seção anterior identificamos os *outliers* da amostra e escolhemos uma distribuição de probabilidade flexível para modelar a parte regular dos retornos, que, por construção, descreverá o núcleo da distribuição de retornos frequentista-Bayesiana. Todo o desenvolvimento, até então, foi realizado através de uma abordagem estatística frequentista, *i.e.* um método baseado em dados amostrais.

Para completar o arcabouço proposto, efetuamos neste capítulo a modelagem dos retornos não-regulares ou extremos. Tais retornos representarão respostas do mercado a eventos incomuns (mas prováveis), cujas probabilidades de ocorrência serão definidas de maneira subjetiva. Para tanto, a distribuição conjunta dos retornos extremos, que, como se pode intuir, constituirá as caudas da distribuição holística, será construída a partir de uma abordagem Bayesiana, seguindo o desenvolvimento de Rebonato (2010) e Denev (2011).

3.1 Definição e modelagem dos eventos extremos

Os eventos extremos devem ser uma proposta do gestor e representarão o esforço de seu sistema cognitivo em qualificar as principais ameaças ao desempenho de sua carteira. A definição dos eventos deverá não só contemplar a experiência do gestor como também o ambiente macroeconômico e as condições atuais do mercado, destacando-se prováveis acontecimentos que possam causar fortes impactos na carteira, em termos de ganhos ou perdas. Nesta abordagem, modelaremos os eventos como variáveis aleatórias booleanas, *i.e.* poderão assumir os valores lógicos ‘verdadeiro’ (**V**) ou ‘falso’ (**F**), indicando sua ocorrência ou não.

Rebonato (2010) sugere dois métodos (que podem ser combinados) para a seleção de eventos a partir de uma determinada carteira:

1. **Top-down**: seleciona-se eventos que se baseiam em uma análise macro, *e.g.* que tenha relevância à economia, ao cenário geopolítico ou, ainda, algum evento de caráter mais amplo, como uma catástrofe natural, por exemplo. Naturalmente, o gestor deve selecionar eventos que acredite que tragam fortes consequências ao desempenho da

carteira. Uma forte injeção de liquidez do governo para uma posição em Índice de ações (*e.g.* Futuros de Ibovespa), forte alta do preço do petróleo para ações de companhias aéreas e casos de doença como a da ‘vaca louca’ para ações de empresas frigoríficas podem ser eventos escolhidos através do método *top-down*.

2. ***Bottom-up***: a partir do conhecimento das idiossincrasias da composição da carteira, ataca-se diretamente suas vulnerabilidades de acordo com o grau de severidade desejado. Por exemplo: concentrações setoriais, posições de valor relativo (*e.g.* *long-short*) e posições alavancadas.

A seleção dos eventos é o primeiro passo para a construção da distribuição conjunta dos retornos extremos. Para uma discussão mais extensa e profunda sobre a análise e seleção de eventos de *stress* para uma carteira de ativos, veja Rebonato (2010).

Para a aplicação em tela e seu conjunto de ativos da amostra de retornos (conforme a Seção 1.2), utilizamos, por simplicidade, o menor número de eventos que englobassem, tanto de maneira sistêmica quanto específica, movimentos expressivos de alta (retornos positivos) e de baixa (retornos negativos) dos preços dos ativos.

Dessa maneira, agrupamos os ativos de acordo com suas classificações setoriais, conforme a Tabela 3.1, e definimos:

- Dois eventos genéricos que impactam fortemente todos os ativos, rotulados *Forte alta do mercado de ações* e *Forte baixa do mercado de ações*; e
- Dois eventos genéricos para cada setor, denominados *Forte alta das ações do setor* e *Forte baixa das ações do setor*.

Definidos os eventos marginais (ou elementares), a atenção deve voltar-se aos eventos conjuntos. Uma vez que obtivermos os valores das probabilidades dos eventos conjuntos (que comporão a distribuição multivariada de eventos extremos), poderemos, obviamente, responder a qualquer consulta probabilística sobre a ocorrência dos eventos.

De maneira geral, para v variáveis aleatórias booleanas, representando os eventos marginais E_1, E_2, \dots, E_v , um evento conjunto C_i será qualquer combinação das realizações \mathbf{V} ou \mathbf{F} para os v eventos individuais:

$$C_i = \{E_1 = e_1, E_2 = e_2, \dots, E_v = e_v\}$$

na qual e_1, e_2, \dots, e_v assumem os valores **V** ou **F**, conforme composição do evento conjunto C_i . Dessa maneira, com v eventos marginais definidos, haverá 2^v eventos conjuntos, estando cada um deles associado a uma probabilidade de ocorrência $p(i)$.

Tabela 3.1: Classificação setorial para os ativos da amostra

Ativo	Setor	Ativo	Setor
IBOV	-	BVMF3	Financeiro
GOLL4	Transporte	VALE5	Materiais básicos
NATU3	Consumo não-cíclico	USIM5	Materiais básicos
AMBV4	Consumo não-cíclico	GOAU4	Materiais básicos
BBAS3	Financeiro	PETR4	Petróleo
BBDC4	Financeiro	CMIG4	Utilidade pública
ITUB4	Financeiro	TRPL4	Utilidade pública

Fonte: Classificações setoriais obtidas do *site* da BM&FBOVESPA e trabalhadas pelo autor. Disponível em <http://www.bmfbovespa.com.br> (Acesso em 12 setembro 2012).

É importante enfatizar que dois ou mais eventos conjuntos não podem ocorrer ao mesmo tempo. Formalmente, os eventos conjuntos C_i devem ser mutuamente exclusivos ou disjuntos. Para isso, duas propriedades básicas da teoria das probabilidades devem ser respeitadas:

- i) $p(i) \geq 0, i = 1, 2, \dots, 2^v$
- ii) $\sum_{i=1}^{2^v} p(i) = 1$

Visualmente, com $v = 3$, pode-se adotar a seguinte notação para a tabela de probabilidades conjuntas:

	E_1	E_2	E_3	Probabilidade conjunta
C_1	0	0	1	$p(1)$
C_2	0	1	0	$p(2)$
C_3	1	0	1	$p(3)$
C_4	0	1	1	$p(4)$
C_5	1	0	1	$p(5)$
C_6	1	1	0	$p(6)$
C_7	1	1	1	$p(7)$
C_8	0	0	0	$p(8)$

A primeira linha da tabela indica que para o evento conjunto C_1 , cuja probabilidade de ocorrência é $p(1)$, apenas o evento elementar E_3 ocorre. O mesmo pode ser interpretado para as demais linhas. Note que a penúltima linha indica que todos os 3 eventos elementares ocorrem, enquanto a última linha, por sua vez, está associada a não-ocorrência dos 3 eventos elementares. Como, para nossa aplicação, os eventos estão relacionados a situações extremas, $p(8)$ deverá estar próximo de 1 (de ocorrência bastante provável) e $p(7)$, por sua vez, deverá ser muito próximo de zero.

Mesmo modelando as variáveis aleatórias num espaço discreto dicotômico, a tarefa de determinar a distribuição conjunta dos eventos extremos, *i.e.* o conjunto das probabilidades $[p(i)]$, será tão complexa quanto maior o número de dimensões e o conseqüente incremento do número de possíveis relacionamentos entre os eventos elementares. A complexidade computacional e conceitual fica evidente ao considerar o tamanho da tabela de probabilidades conjuntas que, para determiná-la, precisaríamos de $2^v - 1$ equações (considerando a propriedade de que todas elas devem somar 1).

Assim, para chegarmos às probabilidades conjuntas (em geral, um exercício nada trivial para ser estimado por um ser humano, principalmente para eventos de rara ocorrência), teremos de injetar no problema mais informações de como entendemos que o mundo funciona, que, no contexto das distribuições de probabilidade, pode ser intuído como definir a independência e condicionalidade entre as variáveis aleatórias. Entretanto, como poderíamos definir, esquematicamente, e sem perdermos o conforto intuitivo da aplicação, a estrutura causal entre os eventos elementares?

Neste sentido, necessitamos de uma técnica para sistematizar nosso conhecimento causal e traduzí-lo para a linguagem probabilística, o que naturalmente nos direciona às Redes Bayesianas como ferramenta para explorarmos o problema.

3.2 Redes Bayesianas

Rede Bayesiana (RB) é um modelo gráfico para representar relacionamentos probabilísticos entre variáveis e realizar inferências entre elas. RBs possuem a vantagem de codificar o conhecimento de um *expert* de uma maneira fácil de construir e de interpretar, explicitando os relacionamentos relevantes entre as variáveis e descrevendo, univocamente, uma distribuição conjunta de probabilidade.

De maneira mais precisa, para um domínio de variáveis $V = \{X_1, \dots, X_v\}$, a distribuição de probabilidade conjunta V é a distribuição de probabilidade sobre todos os estados de V . Uma RB para V representará a distribuição de probabilidade conjunta de V e consistirá de um conjunto local de distribuições de probabilidade condicionais combinado a um conjunto de definições de independência condicional, que, por sua vez, permitirão a descrição da distribuição conjunta global a partir das distribuições locais.

Uma RB pode ser definida como um grafo acíclico direcionado¹ (*Directed Acyclic Graph*; DAG) no qual, para nossa aplicação, cada vértice estará associado a uma variável aleatória booleana (um evento extremo) e uma tabela de probabilidades condicionais. Formalmente, uma Rede Bayesiana B sobre um conjunto finito de variáveis V possui dois componentes:

1. Um DAG $G = (V, A)$, no qual V e A representam o conjunto de vértices e as arestas direcionadas do grafo, respectivamente. Ao discutir relações entre variáveis, a seguinte semântica é utilizada: se existe uma aresta de V_i para V_j , dizemos que V_i é pai de V_j (e V_j é filho de V_i); se existe um caminho direcionado de V_i para V_j , dizemos que V_i é ancestral

¹Um grafo G é um par ordenado $G = (V, A)$, no qual V é um conjunto não-vazio de nós (ou vértices) e A é um conjunto de pares dos vértices de V , denominados arestas (que conectam um vértice ao outro). Se as arestas são constituídas de pares ordenados de vértices, o grafo é direcionado. Se para qualquer vértice $V_i \in V$ não existe um caminho direcionado que começa e termina em V_i , então o grafo é acíclico direcionado. O grafo da Figura 3.1 é um exemplo de grafo acíclico direcionado.

de V_j (e V_j é descendente de V_i). Desse modo, o grafo descreverá, qualitativamente, a topologia da RB.

2. Uma tabela de probabilidades S . Para cada variável $V_i \in V$ uma tabela de probabilidades é especificada contendo a probabilidade condicional da variável dado seus pais, *i.e.* $p(V_i|Pa_{V_i})$ (para os vértices sem pais², haverá apenas probabilidades marginais). Pode-se entender, assim, que a tabela S descreverá, quantitativamente, a força das relações entre os vértices.

O grafo G e a tabela de probabilidades S se vinculam através da condição de Markov, na qual todas variáveis da RB são condicionalmente independentes de todos os seus não-descendentes dado o conjunto de todos os seus pais:

$$P(E_1, E_2, \dots, E_v) = \prod_{i=1}^v P(E_i|Pa_{E_i})$$

A condição de Markov pode ser deduzida da composição do conceito de independência condicional³ e a seguinte regra do cálculo de probabilidades:

$$P(E_v, \dots, E_1) = P(E_v|E_{v-1}, \dots, E_1) * P(E_{v-1}|E_{v-2}, \dots, E_1) * \dots * P(E_1)$$

que, por sua vez, é derivada diretamente da definição de probabilidade condicional,

$$P(X|Y) = \frac{P(X, Y)}{P(Y)}.$$

Rebonato (2010) e Denev (2011) sugerem que as Redes Bayesianas com vértices que possuam no máximo dois pais simplificam enormemente o problema sob a ótica conceitual e computacional, além de manter o problema mais próximo de nossas habilidades cognitivas. Desse modo, para a aplicação de nosso interesse, confeccionamos a RB conforme a Figura 3.1, na qual sintetizamos os termos *Forte alta* e *Forte baixa* através das palavras *up* e *down*,

²Os vértices “Mercado de ações *up*” e “Mercado de ações *down*” são os únicos vértices sem pais do grafo presente na Figura 3.1.

³Dado um evento C , os eventos A e B são condicionalmente independentes se $P(A|B, C) = P(A|C)$ e $P(B|A, C) = P(B|C)$. Em palavras, dado o conhecimento da ocorrência ou não de C , conhecer a ocorrência ou não do evento B não adiciona informação quanto à ocorrência de A e vice-versa. Veja Dawid (1979) e Pearl (2009).

respectivamente. Por simplicidade, optamos, ainda, por uma composição de *Naive Bayesian Networks* (NBN), que consiste de uma RB com apenas um pai e diversos filhos (Hand e Yu (2001)).

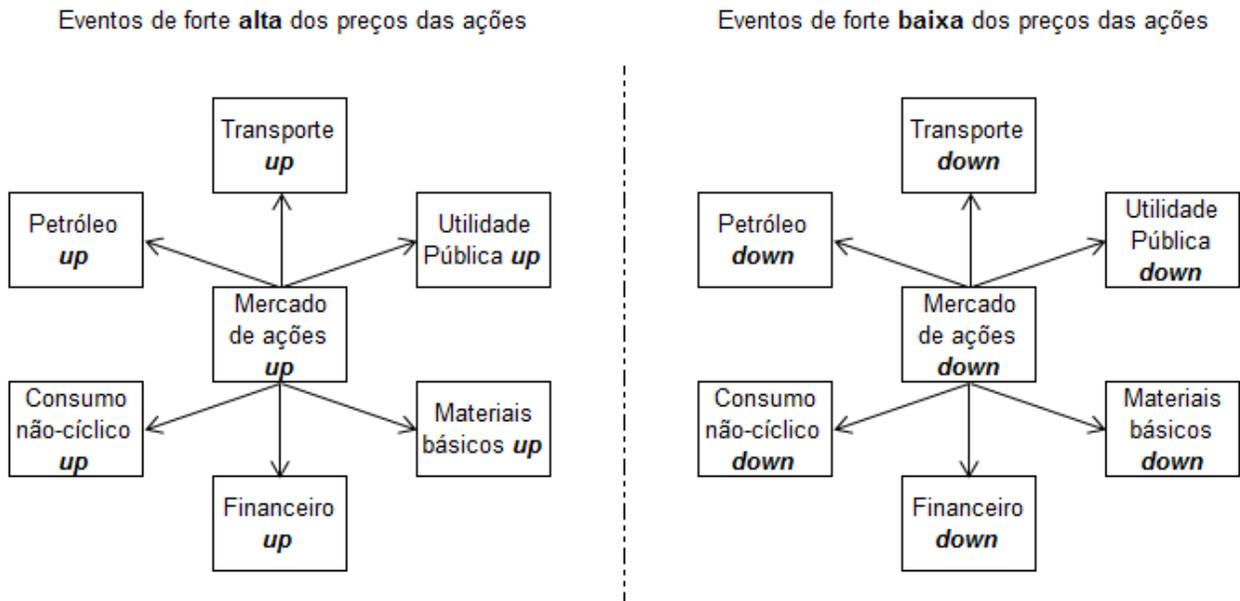


Figura 3.1: Rede Bayesiana adotada para a aplicação. Com exceção dos nós de mesmo mercado ou setor, todos os eventos **down** são independentes dos eventos **up**, e vice-versa.

Fonte: elaboração própria.

O poder de uma RB está no fato de podermos interpretar a existência de uma aresta entre dois vértices como um vínculo causal entre as variáveis aleatórias relacionadas, deixando de contar apenas com o contexto associativo das probabilidades. Com esse intuito, adicionamos as informações de causalidade entre os eventos extremos e estamos mais próximos de obter a distribuição conjunta. O próximo passo, então, é definirmos a intensidade dessas relações causais, que serão atribuídas através das probabilidades condicionais dos eventos.

Para a estimação de probabilidades marginais podemos e devemos aproveitar as informações implícitas de mercado (quando disponíveis), capturando-as através, por exemplo, de preços de opções⁴ e *spreads* de crédito. Para as probabilidades condicionais, entretanto,

⁴Para uma breve revisão de modelos para extração de informações incorporadas aos preços de opções, veja Jackwerth (1999).

difícilmente informações implícitas de mercado serão úteis, o que nos remete a estimações subjetivas (Denev (2011) e Rebonato (2010)).

Estimar subjetivamente probabilidades condicionais pode ser mais simples do que fazê-lo para as probabilidades marginais dos mesmos eventos (Rebonato (ibid.)), *e.g.* estimar a probabilidade de ocorrência de uma forte baixa do setor financeiro ou de uma forte baixa do mercado de ações (eventos “Financeiro *down*” e “Mercado de ações *down*” da RB da Figura 3.1, respectivamente) pode ser bastante difícil, ao passo que a probabilidade condicional da ocorrência de uma forte alta nas ações do setor financeiro *dado uma forte alta no mercado de ações* é mais confortável para se estimar de maneira subjetiva.

Denev (2011) cita, ainda, uma estratégia para se estimar as probabilidades condicionais simples, *i.e.* $P(A|B)$: começar imaginando se a ocorrência do evento B diminui ou aumenta a chance de ocorrência do evento A um pouco, muito, não faz diferença ou o faz quase certo de ocorrer. Dessa maneira, pode-se alocar as probabilidades condicionais estimadas em blocos de confiança, cada qual com uma determinada probabilidade de ocorrência.

Todavia, devido à possível dificuldade cognitiva para se atribuir as probabilidades condicionais, é possível que elas sejam incoerentes, o que significa não existir uma distribuição conjunta a partir das quais as probabilidades condicionais possam ter sido extraídas. Para assegurarmos que nossas estimativas são coerentes, é salutar verificá-las junto a algumas regras derivadas dos axiomas da teoria da probabilidade, como por exemplo⁵:

i) $0 \leq P(A|B) \leq 1$

ii) $P(A|B) \leq \frac{P(A)}{P(B)}$

Se as probabilidades condicionais estimadas são coerentes, elas obedecerão as regras acima; caso contrário, o gestor deverá revisá-las⁶. Para averiguar as probabilidades marginais, por sua vez, pode-se compará-las com informações históricas, uma vez que a análise deve concentrar-se sobre a ordem de grandeza da frequência de ocorrência do evento, *e.g.* uma ocorrência por ano, uma a cada 5 anos, uma a cada século.

⁵Veja Moskowitz e Sarin (1983) para mais regras de consistência para probabilidades condicionais estimadas.

⁶Veja Denev (2011) e Rebonato (2010, cap. 9).

Para nossa aplicação, dada a simples estrutura da RB proposta, devemos estimar 12 probabilidades condicionais e 2 probabilidades marginais, sendo 6 condicionais e 1 marginal para a parte de Alta (*up*) da RB e as demais para a parte de Baixa dos preços da RB (*down*)⁷. As estimativas propostas estão nas Tabelas 3.2 e 3.3.

Tabela 3.2: Probabilidades marginais para os eventos *Alta* e *Baixa* do Mercado de ações

Mercado	Probabilidade <i>up</i>	Probabilidade <i>down</i>
Mercado de ações	0,75%	1%

Fonte: Autor.

Para RBs mais sofisticadas (*e.g.*, que possuam vértices com mais de dois pais), há técnicas avançadas para verificar a coerência das probabilidades marginais e condicionais estimadas, bem como para estimar intervalos para probabilidades condicionais duplas ou de ordens superiores, via, por exemplo, Programação Linear. Veja Kwiatkowski e Rebonato (2011), Moskowitz e Sarin (1983), Rebonato (2010) e Denev (2011).

Tabela 3.3: Probabilidades condicionais para as 12 arestas da RB

Setor	P(Setor_{<i>up</i>} Mercado_{<i>up</i>})	P(Setor_{<i>down</i>} Mercado_{<i>down</i>})
Consumo não-cíclico	45%	35%
Financeiro	55%	60%
Materiais básicos	50%	55%
Petróleo	55%	60%
Transporte	55%	60%
Utilidade pública	45%	35%

Fonte: Autor.

Para uma introdução sobre RBs, veja Charniak (1991). Para uma abordagem teórica mais extensa e uma discussão aprofundada sobre RBs e causalidade, sugerimos Pearl (1988, 2009) e Williamson (2005).

⁷Dada a topologia da RB, a quantidade de estimações para se obter a distribuição conjunta pode ser calculada utilizando-se a condição de Markov.

3.2.1 Distribuição multivariada dos retornos extremos

Uma vez que definimos a topologia da RB e estimamos as probabilidades marginais e condicionais necessárias, a distribuição conjunta pode ser facilmente obtida através da condição de Markov e de outras regras oriundas dos axiomas básicos da probabilidade, tais como⁸

i) $P(A|B) + P(\tilde{A}|B) = 1$

ii) $P(\tilde{A}|\tilde{B}) + P(A|\tilde{B}) = 1$

Assim, primeiramente tratamos de calcular cada metade da RB individualmente, *i.e.* a parte **up** e a parte **down**, uma vez que podemos considerá-las independentes. Por meio de histogramas, dispomos a distribuição das magnitudes das probabilidades de ocorrência dos eventos conjuntos (um total de 128, 2^7 , para cada parte da RB) na Figura 3.2.

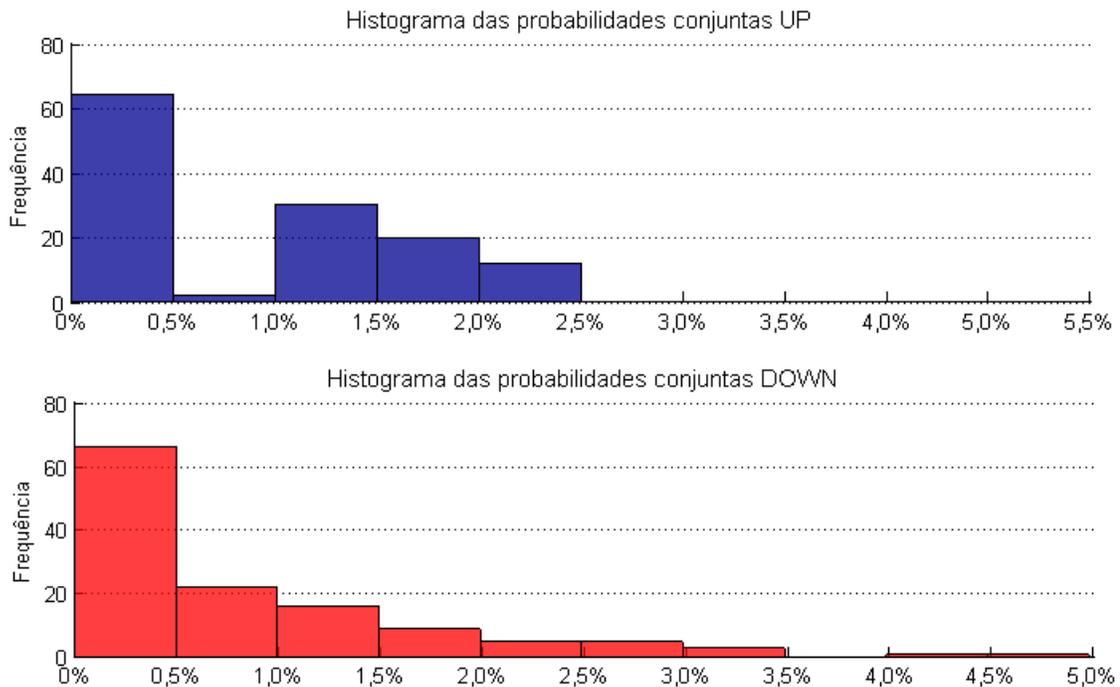


Figura 3.2: Distribuição das magnitudes das probabilidades de ocorrência dos eventos conjuntos de cada parte da RB (*up* e *down*), cada qual contendo 128 eventos.

⁸Utilizamos a notação \tilde{A} para denotar o complemento de A .

Para obtermos, finalmente, a distribuição completa de eventos extremos, teremos de considerar toda a topologia da RB proposta. Para os 14 vértices que definimos, temos uma tabela de probabilidades conjuntas de $2^{14} = 16.384$ eventos conjuntos possíveis. Entretanto, como cada evento elementar possui duas configurações possíveis na RB (*up* e *down*), há diversos eventos conjuntos impossíveis de ocorrer, *e.g.* todas as combinações que possuam *Transporte up* e *Transporte down* ocorrendo simultaneamente. Logo, se eliminamos essas impossibilidades, obtemos uma tabela de probabilidades conjuntas com 2.187 eventos (conjuntos) possíveis⁹.

Ao unirmos as duas distribuições conjuntas, *up* e *down*, obtemos a distribuição completa ao multiplicar, para cada evento possível, as probabilidades conjuntas de cada parte da RB, uma vez que não existem arestas entre seus vértices e, assim, não são influenciadas umas pelas outras.

Dessa maneira, considerando todas as probabilidades estimadas até aqui e as etapas descritas nesta Seção, temos que a distribuição multivariada dos eventos extremos corresponderá à $(1 - p_0) = 15,5\%$ da densidade da distribuição holística frequentista-Bayesiana, aproximadamente. Seu complemento, p_0 , que corresponde ao evento conjunto no qual nenhum evento elementar ocorre, possuirá retornos (regulares) modelados por uma distribuição de probabilidade ajustada através de dados históricos, como exposto na Seção 2.

Caso o gestor possua convicção de que as probabilidades condicionais e marginais estimadas sejam razoáveis mas a probabilidade conjunta $P(\tilde{E}_1, \tilde{E}_2, \dots, \tilde{E}_v) = p_0$ não corresponda ao esperado, pode-se rebalancear o peso da distribuição extrema (de maneira uniforme) no contexto da distribuição holística, frequentista-Bayesiana, através de um fator k , como descrito na próxima Seção.

⁹É fácil ver que para esse tipo específico de topologia (NBN), teremos um número de eventos conjuntos possíveis proporcionais ao cubo do número de setores/mercados: para 1 setor, teremos 3 eventos possíveis do total de 4 combinações (eliminamos o evento conjunto no qual os dois eventos elementares ocorrem); para 2 setores, teremos 9 eventos possíveis para as 16 combinações; para n setores, teremos 3^n combinações possíveis de eventos conjuntos.

Definimos a realização de cada evento conjunto por um conjunto determinístico (fixo) de retornos extremos, conforme Tabela 3.4¹⁰. Por simplicidade, esses log-retornos correspondem às maiores magnitudes de alta e baixa dos preços de cada ativo da base de dados, entre os anos 2005 e 2011¹¹. Não obstante, os retornos marginais extremos dos ativos podem ser estocásticos, *i.e.* modelados através de uma distribuição de probabilidade, sem qualquer prejuízo à abordagem ou à aplicação.

Tabela 3.4: Retornos diários extremos, *up* e *down*, por ativo

Ativo	Retorno <i>up</i>	Retorno <i>down</i>
AMBV4	20%	-12%
BBAS3	13%	-15%
BBDC4	22%	-24%
BVMF3	21%	-13%
CMIG4	16%	-14%
GOAU4	13%	-16%
GOLL4	11%	-13%
IBOV	16%	-23%
ITUB4	13%	-15%
NATU3	18%	-16%
PETR4	17%	-16%
TRPL4	14%	-12%
USIM5	11%	-12%
VALE5	19%	-17%

Fonte: Autor.

¹⁰Para os eventos conjuntos nos quais os eventos elementares de um determinado setor/mercado assumam o valor ‘falso’ (**F**), definimos, por simplicidade, os retornos dos ativos pertencentes a esse setor/mercado igual a zero.

¹¹Estendemos, excepcionalmente, o período da base de dados apresentada na Seção 1.2 com o intuito de contemplar uma amostra de retornos mais significativa.

4. Modelagem híbrida frequentista-bayesiana

As duas seções anteriores descrevem a modelagem de duas distribuições conjuntas: a) a dos retornos regulares, modelada através de uma distribuição multivariada da classe das hiperbólicas. O ajuste da distribuição se dá pelos retornos históricos dos ativos após a remoção dos *outliers*; b) a dos retornos extremos, modelada subjetivamente através de uma Rede Bayesiana, cuja topologia contempla os eventos que o gestor considera crítico ao desempenho de sua carteira.

Para completar a modelagem híbrida frequentista-Bayesiana, o próximo passo é unir as duas distribuições conjuntas a um único processo gerador de retornos.

4.1 Unindo retornos regulares e extremos

A função densidade de probabilidade n -variada dos retornos regulares, $f_{regulares}(r_1, r_2, \dots, r_n)$, é representada pela formulação matemática da DHG apresentada no Apêndice D; a dos retornos extremos, por sua vez, dado que definimos cada evento extremo conjunto como uma realização de um conjunto de retornos pré-definidos, conforme a Tabela 3.4, podemos associá-los a uma distribuição multivariada Delta de Dirac¹:

$$f_{extremos}(r_1, r_2, \dots, r_n) = \frac{1}{1 - p_0} \sum_{i=1}^{2^n - 1} p(i) \delta(r_1 - L_1^i, r_2 - L_2^i, \dots, r_n - L_n^i)$$

Como a probabilidade $P(\tilde{E}_1, \tilde{E}_2, \dots, \tilde{E}_n) = p_0$ da distribuição discreta de eventos extremos corresponde à realização de nenhum evento extremo, mas sim de um retorno regular (regidos pela abordagem frequentista), o termo $\frac{1}{1 - p_0}$ atua como um normalizador, estabelecendo que a soma de todos os eventos extremos conjuntos possíveis seja igual a um.

O somatório do lado direito da igualdade percorre cada um dos eventos extremos conjuntos, cada qual ponderado por uma probabilidade $p(i)$, que será multiplicada pela função

¹Em estatística, a função Delta de Dirac pode ser utilizada para representar uma função densidade de probabilidade discreta, *e.g.* a PDF de $f(x)$ com os n elementos $x = \{x_1, \dots, x_n\}$ de probabilidades p_1, \dots, p_n pode ser descrita por $f(x) = \sum_{i=1}^n p_i \delta(x - x_i)$. Para mais aplicações em estatística, veja Chakraborty (2008).

Delta de Dirac, *i.e.* por 1 quando as perdas associadas ao i -ésimo evento extremo e ao j -ésimo ativo, L_j^i , coincidirem com os retornos r_j , com $j = 1, 2, \dots, n$.

Logo, podemos combinar as duas distribuições através de um fator k para obter a distribuição holística, frequentista-Bayesiana:

$$f(r_1, r_2, \dots, r_n) = (1 - k)f_{regulares}(r_1, r_2, \dots, r_n) + \frac{k}{1 - p_0} \sum_{i=1}^{2^n - 1} p(i)\delta(r_1 - L_1^i, r_2 - L_2^i, \dots, r_n - L_n^i)$$

Caso o gestor esteja suficientemente confiante de que a probabilidade $(1 - p_0)$ reflète o peso da frequência dos retornos regulares no modelo completo (retornos regulares e extremos), o fator $k = (1 - p_0)$ deve ser utilizado. Dessa maneira, todas as probabilidades dos eventos extremos conjuntos serão idênticas àquelas implícitas na topologia da RB utilizada.

Por outro lado, o gestor poderá tratar o valor de k como um parâmetro livre, fazendo com que os retornos regulares possuam uma massa de probabilidade $(1 - k)$ e os extremos, renormalizados uniformemente, com massa de probabilidade total k .

A Figura 4.1 proporciona uma visão da abordagem clássica (utilizando apenas os retornos regulares) *versus* a abordagem híbrida frequentista-Bayesiana (retornos regulares e extremos). Os gráficos contemplam 5.000 simulações de log-retornos diários.

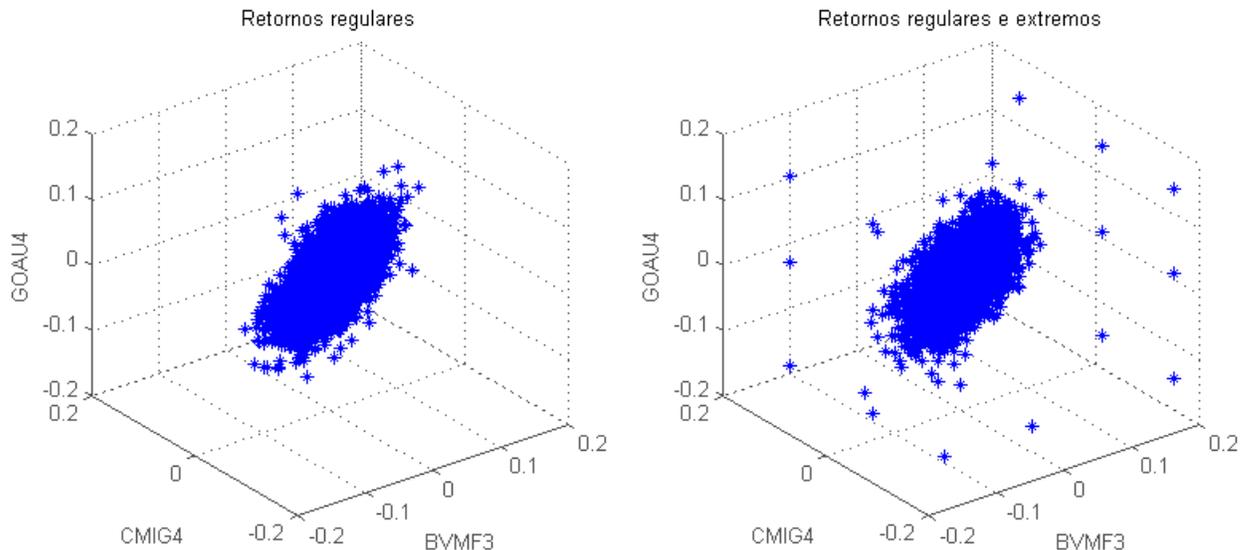


Figura 4.1: *Scatterplots* dos log-retornos de 3 ativos pela abordagem frequentista (esquerda) e híbrida frequentista-Bayesiana (direita).

Fonte: elaboração própria.

5. Otimização do seguro

As seções anteriores descrevem a modelagem de retornos de ativos a partir de uma abordagem híbrida frequentista-Bayesiana. O modelo, dessa forma, absorve dados históricos de mercado e percepções de mundo do gestor da carteira. Com esse ferramental, é possível estimar retornos para a carteira de ativos para qualquer horizonte de tempo, permitindo, assim, encontrar quais derivativos (e em quais quantidades) otimizam a proteção da carteira contra cenários de *downside* com o menor custo possível.

5.1 Definição do problema de seguro ótimo

Definimos um seguro ótimo para uma carteira de ativos, de valor inicial V^0 , como um instrumento que tem o objetivo de mitigar as possibilidades do valor da carteira ao fim de um horizonte temporal determinado ser menor que o nível máximo de *downside* definido pelo gestor. Nesse contexto, o problema é, então, encontrar quais derivativos e suas quantidades melhor suportam esse objetivo minimizando o custo relativo à aquisição desses instrumentos.

A subcarteira ótima de derivativos será aquela com o menor custo possível que, no contexto do arcabouço modelado, terá resultados que contrabalançarão as perdas superiores ao nível máximo de *downside* da carteira de ativos, incrementando as possibilidades de que o desempenho da carteira de ativos seja superior ao nível máximo de perda estabelecido pelo gestor.

Por questão de simplicidade, serão tratadas apenas carteiras com posições compradas em ativos e derivativos do tipo “opção de venda” (*puts*). Logo, podemos assumir que as quantidades a otimizar deverão possuir a restrição de serem maiores ou iguais a zero.

O principal insumo para a otimização será o conjunto de valores da carteira, V_s^t , ao fim do horizonte de tempo desejado, T , obtido por meio de S simulações, *i.e.* V_s^T , para $s = 1, 2, \dots, S$. O gestor, ao adotar o seguro, deseja proteger o valor da carteira contra um *downside* superior a uma constante α (definido em termos percentuais), ou seja, por meio do seguro deseja-se mitigar as possibilidades do valor da carteira ao final do período (sem considerar o custo com

as opções) ser menor que $V_{min}^T = (1 - \alpha) * V^0$.

Denotaremos por X_s^t o valor de mercado da subcarteira de derivativos no instante t para a simulação s . O vetor com as quantidades de cada *put* será representado por Q_j , com $j = 1, 2, \dots, J$, denotando as *puts* disponíveis para negociação e selecionadas pelo gestor¹; os preços de mercado das *puts* serão representadas pela matriz P_j^t .

Como o foco do problema é sobre os cenários em que $V_s^T < V_{min}^T$, denotamos o subconjunto de cenários em que essa desigualdade ocorre como V_d^T , sendo $d = 1, 2, \dots, D$, com D ($D \leq S$) representando o número de cenários em que ocorre um *downside* superior à constante α determinada pelo gestor.

Com o intuito de fornecer mais uma possibilidade à aplicação, sujeitaremos a otimização à restrição orçamentária do custo do seguro, que denominaremos W (em unidades monetárias). Dessa maneira, uma vez que ao adquirir opções de venda ocorre um dispêndio de caixa, o gestor poderá limitar o valor monetário investido no seguro.

Assim, no instante de tempo $t = 0$, o gestor possui uma determinada carteira comprada em ativos, de valor V^0 , e quer evitar que seu valor após um determinado período T , V^T , seja inferior à V_{min} . Para isso, ele dispõe no mercado de J *puts*, cada qual com um preço P_j^0 e com *strikes*, vencimentos e ativos subjacentes distintos. Executa-se, então, S simulações, cada qual com T retornos diários para cada ativo da carteira e para os demais ativos subjacentes das J opções disponíveis. Com isso, podemos calcular o seguro ótimo, Q_j , através do conjunto V_d^T e X_d^T .

O problema a resolver tem, assim, a seguinte formulação:

$$\text{Maximizar } \sum_{d=1}^D \min(X_d^T + V_d^T - V_{min}^T; 0) - \sum_{j=1}^J (Q_j * P_j^0)$$

Sujeito a: $Q_j \geq 0$, para $j = 1, 2, \dots, J$

$$\sum_{j=1}^J (Q_j * P_j^0) \leq W$$

No primeiro somatório, o mínimo entre zero e o valor abaixo do nível de *downside* determinado limita o incremento da quantidade de *puts* a suportar apenas cenários em que o

¹O número de *puts*, J , pode ser maior que o número de ativos, n , devido a existência de distintos *strikes* e vencimentos.

valor da carteira de ativos mais o valor da subcarteira de derivativos seja inferior ao valor de *downside* máximo, lembrando que $X_d^T = \sum_{j=1}^J (Q_j * P_j^T)$. O segundo somatório, por sua vez, representa o custo do seguro, adquirido no momento $t = 0$.

Assim como um grande número de problemas em finanças, a solução ótima para o problema acima é difícil de obter-se deterministicamente, sendo necessária a aplicação de outros tipos de algoritmos, como não determinísticos e probabilísticos. Isso ocorre devido à natureza não-linear e descontínua de sua função objetivo. Dessa maneira, uma alternativa é a aplicação de métodos heurísticos, como Algoritmos Genéticos e Evolutivos (Gilli e Schumann (2009), Melo e Delbem (2009)).

5.2 O algoritmo Evolução Diferencial

O algoritmo Evolução Diferencial (ED) é um dos Algoritmos Evolutivos (Back, Fogel e Michalewicz (2000)) existentes mais eficientes para resolver problemas de otimização contínua global (Melo e Delbem (2009)). Introduzido por Storn e Price (1997), o algoritmo é eficaz para encontrar ótimos globais de funções objetivo descontínuas e não-diferenciáveis. Desde sua criação, o ED tem apresentado bons resultados em diversos campos da ciência, como finanças, física computacional e pesquisa operacional (Mullen et al. (2011)).

Pertencendo à classe dos Algoritmos Genéticos, o ED possui inspiração biológica, aplicando operações de seleção, mutação e recombinação (*crossover*) sobre uma população, com o intuito de minimizar² uma função objetivo ao longo de sucessivas gerações. Para mais detalhes sobre algoritmos genéticos e suas operações, veja Goldberg (1989).

Basicamente, o algoritmo inicia definindo a primeira geração aleatoriamente, respeitando-se um intervalo definido ou utilizando valores fornecidos. Cada nova geração é criada, então, a partir de mutações de membros da população atual. Isso é realizado através do parâmetro de mutação v , criado a partir de três membros da população que são escolhidos aleatoriamente, x_1 , x_2 e x_3 , sendo F o fator de diferenciação, tipicamente um valor entre 0 e 1:

²Grande parte da literatura de otimização tem como convenção o estudo de problemas de minimização. Na prática, para a otimização do seguro ótimo, formulado como uma maximização, basta invertermos o sinal do retorno da função objetivo.

$$v = x_1 + F * (x_2 - x_3)$$

Cada membro da população sofre, então, mutação com probabilidade de recombinação $CR \in [0, 1]$. A probabilidade CR controla a fração dos parâmetros que são copiados do mutante para cada possível novo membro da população. Por fim, cada novo filho criado pelo processo de mutação é avaliado pela função objetivo e, caso possua valor menor ou igual ao membro anterior, este é sobreposto pelo membro resultante da mutação.

Para mais detalhes sobre o ED, veja Price, Storn e Lampinen (2006) e Storn e Price (1997).

Em 2005 surgiu a primeira versão do pacote *DEoptim: Global optimization by Differential Evolution*³, que provê uma função para otimização global através do algoritmo ED. Através dele, desenvolvemos todas as simulações e construções de seguro ótimo dessa dissertação, executado no ambiente do *software* R (R Development Core Team (2005)). A Figura 5.1 mostra cinco exemplos da velocidade de convergência do algoritmo para otimizações de seguro, dado uma média de 300 cenários de *downside* e 15 opções disponíveis.

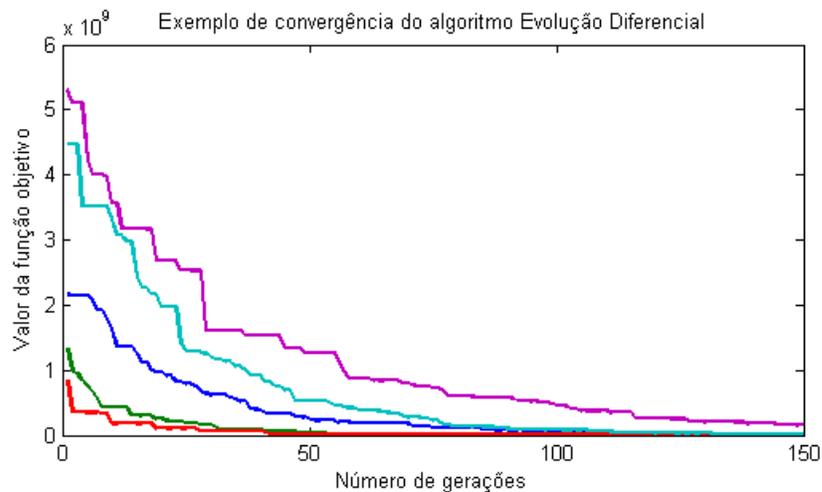


Figura 5.1: Velocidade de convergência do algoritmo Evolução Diferencial para otimizações de seguro ótimo com uma média de 300 cenários de *downside* e 15 opções disponíveis.

Fonte: elaboração própria.

³Veja Ardia et al. (2012) para a documentação técnica; <http://cran.r-project.org/web/packages/DEoptim/> para baixar o pacote (Acesso em 18 nov. 2012).

6. Resultados

6.1 Um exemplo passo a passo da aplicação

Para ilustrar os pontos fortes e fracos da aplicação, apresentamos nesta Seção um exemplo passo a passo, complementando partes do arcabouço com detalhes que aproximam o modelo da prática e sugerindo premissas que podem ser úteis em sua implementação.

A carteira de ativos utilizada (um dado exógeno ao modelo) é o conjunto equiponderado de todos os ativos presentes na amostra de dados descrita na Seção 1.2, adquiridos pelos preços de fechamento (ajustados para proventos) do dia 26/08/2009. Para aproximar o exemplo da prática de mercado de renda variável brasileiro, as ações são compradas em quantidades múltiplas de 100. O Índice Ibovespa (IBOV) é comprado pelo valor financeiro equivalente em pontos (R\$ 1,00 por ponto de Índice) em uma quantidade¹. Dessa maneira, obtemos a carteira conforme a Tabela 6.2, sem considerar qualquer custo de transação.

O exemplo será desenvolvido utilizando-se todo o intervalo da amostra de dados disponível (conforme Seção 1.2) e se baseará em uma estratégia *buy-and-hold*, sendo o seguro da carteira (uma subcarteira de *puts*) rebalanceado a cada 30 dias (“período”). Dessa maneira, no instante inicial (26/08/2009) e a cada rebalanceamento do seguro realizamos as etapas elencadas abaixo. A notação utilizada é a mesma da Seção 5.1, “Definição do problema de seguro ótimo”.

1. Através do MCD, identificamos os *outliers* da amostra de retornos dos 253 dias imediatamente anteriores. O procedimento adotado é o descrito na Seção 2.1.
2. Identificados os *outliers* da janela de retornos de 253 dias, os retornos regulares são, então, utilizados para a calibração da DHG multivariada, conforme Seção 2.2. Os retornos extremos e a distribuição conjunta de eventos extremos são constantes por todo o intervalo, conforme Seção 3.2.1. Definimos, subjetivamente, o fator k da distribuição

¹Por simplicidade, assumimos que o IBOV é negociado como um *Exchange Traded Fund* (ETF), que é uma unidade negociável em Bolsa que pode representar um Índice de Ações, entre outros.

holística frequentista-Bayesiana (Seção 4.1) igual a 1%.

3. Definimos $\alpha = 10\%$. Desse modo, a configuração otimizada do seguro mitigará as possibilidades do valor da carteira ao final do período de 30 dias (quando se chega ao próximo rebalanceamento ou ao final do intervalo total) ser menor que 90% do valor da carteira no instante do rebalanceamento ou ao início da estratégia.
4. Definimos $W = 10\%$, *i.e.* o valor máximo do seguro é de 10% do valor da carteira no instante do rebalanceamento ou no início da estratégia.
5. As opções selecionadas e que estarão disponíveis para a calibração do seguro são todas as opções de venda da base de dados (*puts*) de estilo Europeu, conforme Seção 1.2, cuja data de vencimento é igual ou superior a 30 dias a partir do instante em que se otimiza o seguro. Não são considerados os custos de transação (corretagem, emolumentos, etc) com as opções.
6. Definimos $S = 5.000$ simulações, sendo que em cada uma delas são gerados os 30 retornos aleatórios² correspondentes ao período, através do modelo de retornos multivariado híbrido frequentista-Bayesiano. Com os retornos simulados dos ativos, obtemos o conjunto de valores da carteira ao fim do período, V_s^T e, logo, o conjunto V_d^T .
7. Otimiza-se, então, o seguro ótimo através do conjunto Q_j como descrito na Seção 5.1. Por simplicidade, o valor de mercado das *puts* é calculado como seu valor intrínseco (diferença, se positiva, entre o *strike* e o preço do ativo subjacente) mais uma aproximação linear de seu valor extrínseco (diferença entre o preço de mercado da opção e seu valor intrínseco), estimado como o produto entre seu valor extrínseco inicial (no momento da compra) e a fração entre o número de dias até o vencimento e o número de dias entre o dia da compra e o vencimento.
8. A cada rebalanceamento, todas *puts* pertencentes ao seguro calibrado anteriormente são vendidas. Da mesma maneira que as ações, as opções são sempre adquiridas em quantidades múltiplas de 100.

²A geração de números aleatórios foi executada através do pacote *randtoolbox* (Chalabi et al. (2009)) no ambiente do *software* R (R Development Core Team (2005)).

Para os 21 rebalanceamentos realizados em todo o intervalo, uma média de 24 observações (máximo de 44 e mínimo de 13) *outliers* foram identificados, o que representa uma média de 9,5% do tamanho das amostras de 253 dias de retornos. Com o parâmetro $\alpha = 10\%$, realizou-se uma média de 250 cenários de *downside* (máximo de 735 e mínimo de 89), V_d^T , correspondendo a 5% dos 5.000 cenários simulados em cada período de rebalanceamento. A Tabela 6.3 descreve esses resultados para cada etapa de rebalanceamento do seguro.

A Tabela 6.4 sumariza o resultado da otimização a cada rebalanceamento, mostrando a quantidade de opções distintas incorporadas ao seguro a partir do número de opções disponíveis na data, bem como o custo total do seguro e a respectiva fração do orçamento máximo, W , utilizado. Ademais, pode-se ver o resultado de perdas ou ganhos (P&L) acumulado dos seguros ao longo da trajetória da estratégia em todo o intervalo.

Para o período da aplicação e com todas as demais premissas e parâmetros adotados, o resultado final realizado (com dados de mercado, conforme Seção 1.2) da carteira *long-only* para a estratégia *buy-and-hold* com e sem seguro se dá, supondo valores iniciais normalizados iguais a 100, como mostra a Tabela 6.1.

Tabela 6.1: Desempenho realizado das estratégias para o período e dados utilizados

Carteira	Valor da carteira				Retorno
	Inicial	Mínimo	Máximo	Final	
Sem seguro	100,00	92,30	136,95	136,75	36,75%
Com seguro	100,00	95,31	135,79	130,78	30,78%
IBOV	100,00	84,25	126,36	118,16	18,16%

Fonte: elaboração própria.

Percebe-se que o retorno final da carteira com seguro é inferior à carteira sem seguro, porém, de maneira relevante, maior que um potencial *benchmark*, o Índice Ibovespa (IBOV). Mais importante, a carteira com seguro fornece o maior valor mínimo entre as 3 carteiras analisadas. O desempenho completo das carteiras com e sem seguro durante todo o intervalo do exemplo é apresentado na Figura 6.2.

Para reforço visual do *modus operandi* da otimização, a Figura 6.1 mostra, em seu primeiro gráfico, o histograma dos retornos da carteira de ativos resultantes das 5.000 simulações do re-

balanceamento de 30/01/2012. A partir dessa distribuição (simulada) de retornos obtivemos o seguro, cuja distribuição de retornos está no terceiro histograma. Por fim, se incorporamos o seguro à carteira de ativos, obtemos a distribuição de retornos como mostrada no segundo histograma. A distribuição da carteira sem seguro, de retorno médio de $-0,78\%$, passa a ter um retorno médio de $5,11\%$ com o seguro e, assim, não exibe nenhum retorno inferior à 10% , nível máximo de *downside* de nosso exemplo.

Tabela 6.2: Carteira de ativos proposta para exemplificar a aplicação

Ativo	Quantidade	Preço de compra (R\$)	Valor presente (R\$)	Proporção da carteira
BVMF3	5400	10,60	57.246,70	7,24%
CMIG4	2800	20,56	57.559,53	7,28%
GOAU4	2200	26,01	57.224,18	7,23%
BBDC4	2400	23,41	56.189,29	7,10%
NATU3	2100	26,48	55.616,20	7,03%
GOLL4	3200	17,76	56.837,54	7,19%
PETR4	1900	30,10	57.198,54	7,23%
ITUB4	1900	30,03	57.049,85	7,21%
BBAS3	2500	22,24	55.594,76	7,03%
AMBV4	2200	25,16	55.344,59	7,00%
VALE5	1900	29,46	55.971,11	7,08%
USIM5	2600	21,89	56.903,20	7,19%
TRPL4	1500	36,36	54.545,31	6,90%
IBOV	1	57.766,00	57.766,00	7,30%
Totais	-	-	791.046,79	100,00%

Fonte: elaboração própria a partir dos dados da Seção 1.2.

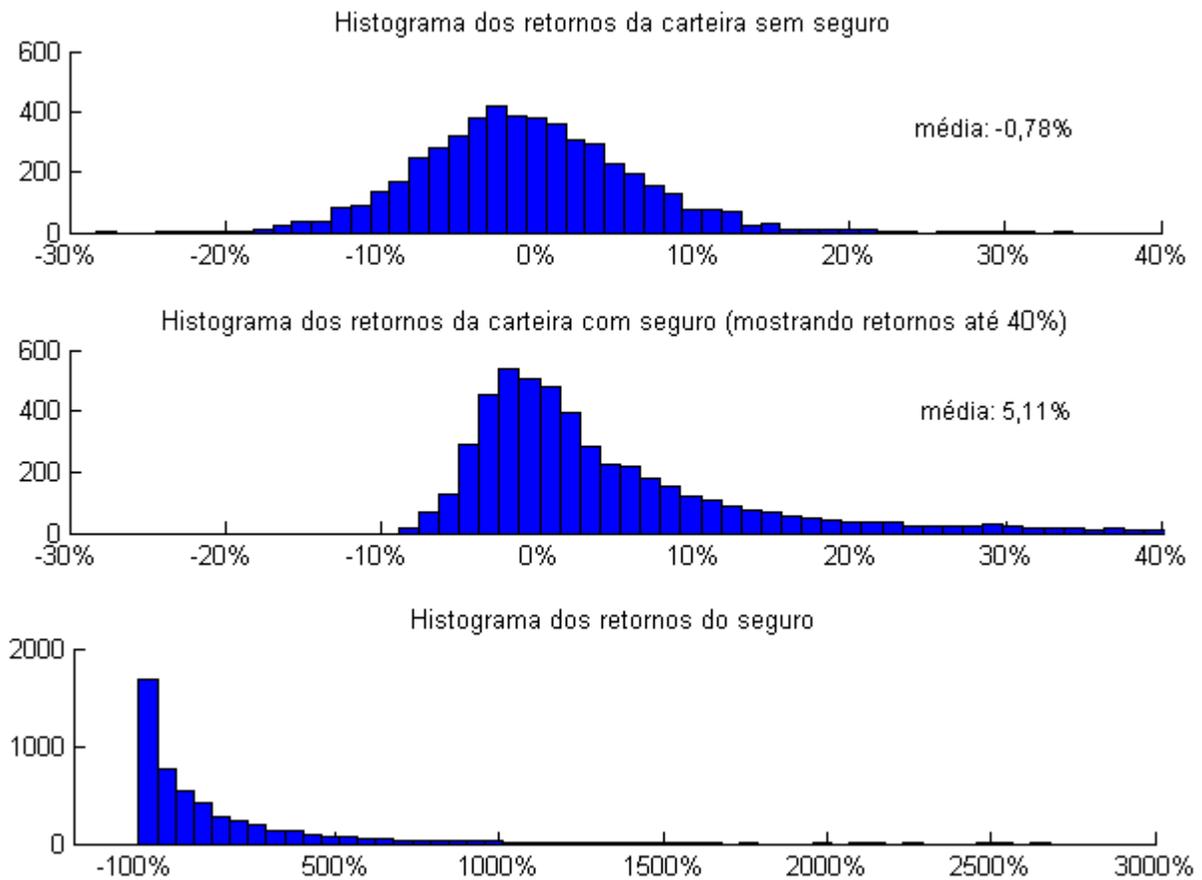


Figura 6.1: Histogramas de retornos simulados da carteira sem seguro (primeiro gráfico) e com seguro (segundo gráfico). O terceiro histograma corresponde ao seguro da carteira isolado.

Fonte: elaboração própria.

Tabela 6.3: Número de *outliers* e quantidade de *downsides* por rebalanceamento da aplicação-exemplo

Data	Número de <i>outliers</i>	Quantidade de <i>downsides</i>
26/ago/09	41	735
08/out/09	44	353
24/nov/09	38	251
11/jan/10	26	232
25/fev/10	22	138
09/abr/10	23	89
24/mai/10	23	113
06/jul/10	25	201
18/ago/10	26	174
30/set/10	25	154
16/nov/10	24	136
29/dez/10	20	142
11/fev/11	20	200
29/mar/11	18	220
12/mai/11	16	229
24/jun/11	14	196
05/ago/11	13	270
19/set/11	22	307
01/nov/11	22	364
15/dez/11	23	369
30/jan/12	25	375

Fonte: elaboração própria.

Tabela 6.4: Resumo do seguro para cada etapa do rebalanceamento da aplicação-exemplo

Data	Opções		Custo do Seguro (R\$)	Orçamento utilizado	P&L acumulado dos seguros (R\$)
	Disponíveis	Adquiridas			
26/ago/09	12	2	63.485	80%	0
08/out/09	8	1	80.136	92%	-56.578
24/nov/09	12	9	72.298	78%	-10.564
11/jan/10	11	4	30.785	32%	-13.042
25/fev/10	19	7	34.049	38%	-1.470
09/abr/10	11	4	26.385	27%	-15.313
24/mai/10	15	6	30.240	36%	33.362
06/jul/10	16	8	43.748	50%	27.956
18/ago/10	16	10	46.103	48%	12.441
30/set/10	16	6	49.415	50%	97.634
16/nov/10	17	12	49.287	49%	67.832
29/dez/10	15	10	45.095	45%	-12.203
11/fev/11	10	3	92.901	97%	-40.983
29/mar/11	13	5	28.917	30%	-78.645
12/mai/11	12	6	43.580	46%	-74.898
24/jun/11	16	9	44.832	50%	-57.305
05/ago/11	44	32	35.713	46%	68.862
19/set/11	26	18	40.261	47%	5.214
01/nov/11	25	19	42.667	49%	49.990
15/dez/11	31	18	43.868	49%	18.624
30/jan/12	32	31	47.657	49%	-64.302
14/mar/12					-47.229

Fonte: elaboração própria.

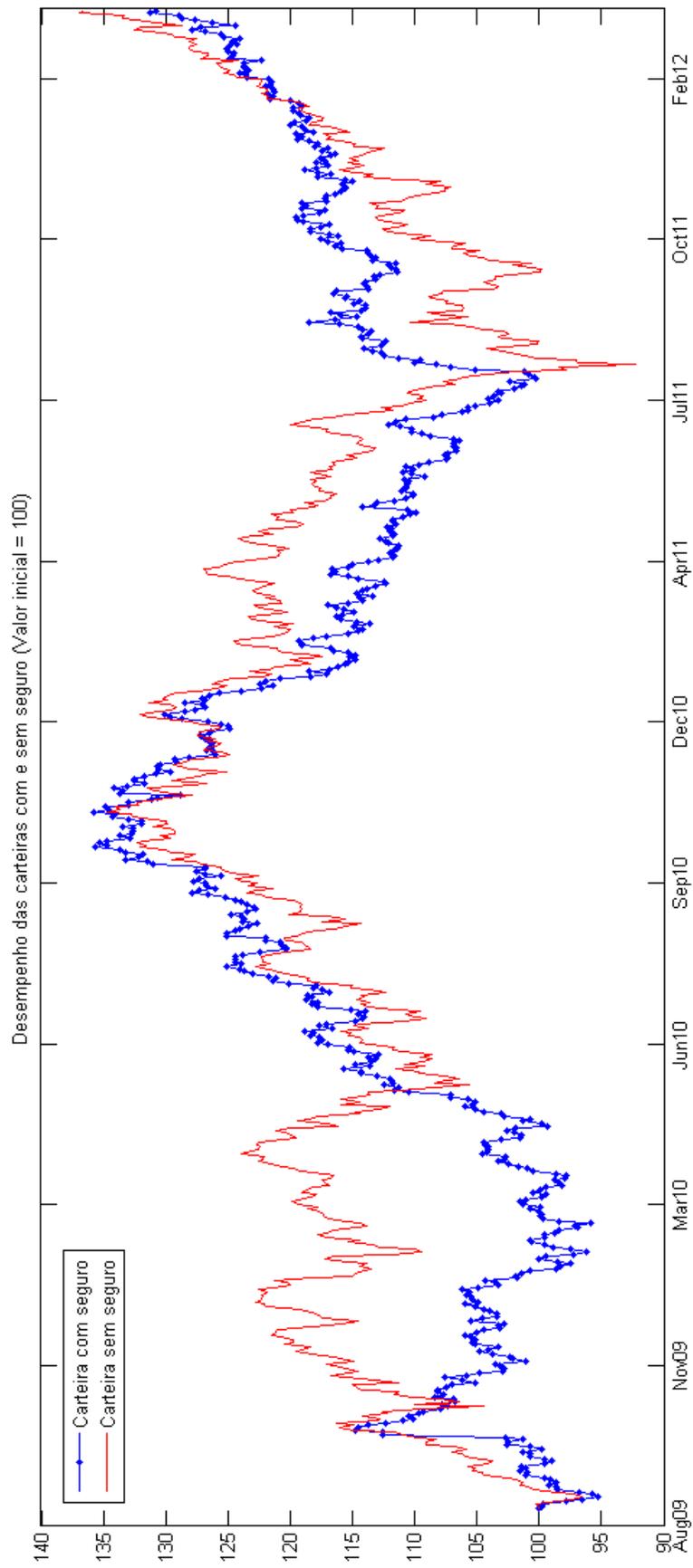


Figura 6.2: Desempenho da carteira da aplicação-exemplo com e sem seguro a partir de dados de mercado entre Agosto 2009 e Fevereiro 2012 (Valor inicial das carteiras = 100).

Fonte: elaboração própria.

7. Conclusão

O obstáculo primário para o investimento em um *downside hedge* de uma carteira com diversos ativos é o custo. Nesta dissertação, formulamos o problema de seguro ótimo de maneira a encontrar o subconjunto de opções de venda (*puts*) de menor custo que ofereça ganhos que contrabalanceie as perdas da carteira de ativos em cenários que seu valor seja inferior ao nível máximo de *downside* estabelecido pelo gestor. A otimização contempla, ainda, a possibilidade de restrição orçamentária ao custo total do seguro.

A aplicação possui flexibilidade em grande parte de seus parâmetros, contribuindo com a intuição do gestor sem trazer dificuldades operacionais: a) carteira de ativos é insumo da aplicação, sem qualquer restrição de nomes ou pesos dos ativos; b) nível de *downside* máximo é definido pelo gestor; c) horizonte de tempo para a calibração do seguro é parâmetro livre do modelo; d) conjunto de derivativos disponíveis ao seguro é parâmetro exógeno, de mercado, possibilitando, por exemplo, considerar apenas os mais líquidos no momento do cálculo do seguro; e) não há restrição para o modelo de apreçamento das opções nos cenários simulados.

Os cenários prospectivos para os retornos da carteira são gerados a partir de um arcabouço híbrido frequentista-Bayesiano, *i.e.* segregamos a modelagem de retornos de ativos em regulares e extremos: para o primeiro, utilizamos a família das distribuições hiperbólicas generalizadas, que tem promissora e descripto fatos estilizados de retornos financeiros em diversos mercados e instrumentos financeiros ao longo das duas últimas décadas; para o último, obtemos sua distribuição conjunta de maneira parcimoniosa através de uma Rede Bayesiana, cuja topologia e probabilidades marginais e condicionais dos nós (eventos extremos) são atribuídas subjetivamente pelo gestor.

A abordagem híbrida para o modelo de retornos apresenta diversas vantagens: a) independência no *modus operandi* dos retornos regulares e extremos (*e.g.* poder-se-ia modelar a estrutura de dependência dos retornos regulares com cópulas, sem prejuízos ou alterações à estrutura dos retornos extremos); b) tamanho da janela de retornos utilizada para a detecção de *outliers* e ajuste da distribuição dos retornos regulares é definido pelo gestor; c) flexibilidade para definir os vínculos causais entre eventos extremos (arestas da rede baye-

siana) e as probabilidades subjetivas através de uma perspectiva *bottom-up* ou *top-down*;

d) possibilidade de atribuir distribuições de probabilidades diversas aos eventos extremos;

e) flexibilidade para definir a fração da distribuição dos retornos extremos na distribuição holística.

A modelagem Bayesiana, em especial, acrescenta significativa robustez à aplicação a medida que possibilita a construção de uma estrutura complexa, *i.e.* a distribuição multivariada de retornos extremos, através de relações simples e estimações intuitivas. Para o exemplo apresentado nesta dissertação, a partir da topologia da Rede Bayesiana proposta e apenas 14 probabilidades estimadas (12 condicionais e 2 marginais), obtivemos uma distribuição multivariada com um total de 2.187 eventos conjuntos.

O problema de seguro ótimo é formulado de modo a contemplar apenas carteiras com posições compradas em ativos e derivativos do tipo “opções de venda” (*puts*). Contudo, a aplicação pode ser facilmente expandida para considerar carteiras com posições vendidas em ativos (posições *short*) ou até mesmo carteiras *long-short*. Para tanto, basta simplesmente incluir opções de compra (*calls*) ao universo de derivativos disponíveis ao seguro. Além disso, trabalhos futuros podem contemplar custos de transação, principalmente em aplicações que possuam diversos rebalanceamentos do seguro ao longo do tempo.

Devido à natureza não-linear e descontínua da função objetivo da otimização, sua solução é difícil de obter-se deterministicamente, sendo necessária a aplicação de outros tipos de algoritmos, como não determinísticos e probabilísticos. Dessa maneira, aplicamos o algoritmo Evolução Diferencial através do pacote *DEoptim* (Ardia et al. (2012)), que possui um histórico de bons resultados em diversos campos da ciência, como finanças, física computacional e pesquisa operacional.

O estudo de seguros de carteiras, sobretudo para dados do mercado brasileiro, é um campo com diversas possibilidades abertas de exploração. Trabalhos futuros podem considerar, entre outros:

- comparar a eficácia da aplicação do seguro entre carteiras de ativos ponderadas por valor, carteiras equiponderadas e carteiras dinamicamente rebalanceadas;
- monitoramento das “gregas” (derivadas dos preços das opções em relação ao preço do

ativo subjacente, volatilidade implícita, tempo, taxa de juros, etc) a fim de administrar dinamicamente os impactos de resultado da carteira de derivativos ao longo da vida do seguro;

- comparar o desempenho do seguro com intervalos menores de calibração da subcarteira de derivativos, bem como para a identificação dos *outliers* e ajuste da distribuição de retornos regulares;
- incorporar custos de transação aos rebalanceamentos do seguro;
- modelar os parâmetros α , W e k (nível máximo de *downside*, valor máximo do seguro e fração dos retornos extremos na distribuição holística, respectivamente) em função de um modelo de detecção de regime de volatilidade, com o objetivo de aumentar a eficiência do seguro, principalmente em momentos calmos de mercado; e
- testes mais robustos de efetividade do seguro, como critérios de dominância estocástica da carteira com seguro sobre a carteira sem seguro, como proposto em Annaert, Osselaer e Verstraete (2009).

Bibliografia

- Abramowitz, M. e I. A. Stegun (1968). *Handbook of Mathematical Functions*. Dover Publ., New York.
- Almeida, L.A., J. Yoshino e PPS Schirmer (2003). «Derivativos de renda fixa no Brasil: modelo Hull-White». Em: *Pesquisa e planejamento econômico* 33.2.
- Ané, T. e C. Kharoubi (2003). «Dependence structure and risk measure». Em: *Journal of Business* 76, pp. 411–438.
- Annaert, J., S.V. Osselaer e B. Verstraete (2009). «Performance evaluation of portfolio insurance strategies using stochastic dominance criteria». Em: *Journal of Banking & Finance* 33.2, pp. 272–280.
- Ardia, D. et al. (2012). *Package 'DEoptim'*.
- Bachelier, L. (1900). *Théorie de la spéculation*. Gauthier-Villars.
- Back, T., D. Fogel e Z. Michalewicz (2000). *Evolutionary Computation 1 Basic Algorithms and Operators*. Institute of Physics Publishing, Bristol e Philadelphia.
- Barbachan, J.F., A.R. Schuschny e A.C. Silva (2001). «Lévy processes and the Brazilian Market». Em: *Brazilian Review of Econometrics* 21.2, pp. 263–289.
- Barbachan, J.S.F. e J.R.H. Ornelas (2003). «Apreçamento de opções de IDI usando distribuições hiperbólicas generalizadas». Em: *ECONOMIA APLICADA* 7.4.
- Barndorff-Nielsen, O.E. (1977). «Exponentially decreasing distributions for the logarithm of particle size». Em: *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences* 353.1674, pp. 401–419.
- (1978). «Hiperbolic distributions and distributions on hyperbolae». Em: *Scandinavian Journal of Statistics* 5, pp. 151–157.
- Bauer, C. (2000). «Value at Risk using hyperbolic distributions». Em: *Journal of Economics and Business* 52, pp. 455–467.
- Black, F. e R. Jones (1987). «Simplifying Portfolio Insurance». Em: *The Journal of Portfolio Management* 14, pp. 48–51.

- Black, F. e M. Scholes (1973). «The Pricing of Options and Corporate Liabilities». Em: *Journal of Political Economy* 81, pp. 637–654.
- Borak, S., A. Misiorek e R. Weron (2011). «Models for heavy-tailed asset returns». Em: *Statistical Tools for Finance and Insurance*, p. 21.
- Breymann, W. e D. Luthi (2011). *ghyp: A package on generalized hyperbolic distributions*.
- Cappa, L. e P.L. Valls (2010). «Modelando a volatilidade dos retornos de Petrobrás usando dados de alta frequência». Em: *Textos para Discussão da Escola de Economia de São Paulo da Fundação Getúlio Vargas* 258.
- Carvalho, M. (2012). «Performance das melhores ideias dos gestores de fundos de ações brasileiros». Em: *Dissertação (mestrado) - Fundação Getulio Vargas, Escola de Pós-Graduação em Economia*. P. 41.
- Chakraborty, S. (2008). «Some Applications of Dirac’s Delta Function in Statistics for More than One Random Variable». Em: *Applications and Applied Mathematics* 3.1, 42:54.
- Chalabi, Y. et al. (2009). *Package ‘randtoolbox’*.
- Charniak, E. (1991). «Bayesian networks without tears». Em: *AI Magazine* 12.4, pp. 50–63.
- Cont, R. (2001). «Empirical properties of asset returns: stylized facts and statistical issues». Em: *Quantitative Finance* 1, pp. 223–236.
- Dawid, A. P. (1979). «Conditional independence in Statistical Theory». Em: *Journal of the Royal Statistical Society. Series B (Methodological)* 41.1, pp. 1–31.
- Denev, A. (2011). «Coherent Asset Allocation and Diversification in the Presence of Stress Events». Em: *A thesis submitted in partial fulfillment of the MSc in Mathematical Finance, University of Oxford*.
- Dimson, E. e M. Mussavian (1999). «Three centuries of asset pricing». Em: *Journal of Banking and Finance* 23, pp. 1745–1769.
- Eberlein, E. e U. Keller (1995). «Hyperbolic distributions in finance». Em: *Bernoulli: Official Journal of the Bernoulli Society of Mathematical Statistics and Probability* 1.3, pp. 281–299.
- Eberlein, E. e K. Prause (2000). «The generalized hyperbolic model: financial derivatives and risk measures». Em: *H. Geman, D. Madan, S. Pliska, T. Vorst (Eds.), Mathematical Finance-Bachelier Congress*. Citeseer, pp. 245–267.

- Embrechts, P. A. McNeil e D. Straumann (2002). «Correlation and dependence in risk management: Properties and pitfalls». Em: *M. A. H. Dempster (ed.): Risk Management: Value at Risk and Beyond*. 176–223.
- Fajardo, J. e A. Farias (2004). «Generalized Hyperbolic Distribution and Brazilian Data». Em: *Brazilian Review of Econometrics* 24.2, pp. 249–271.
- Ferguson, T. S. (1961). «On the rejection of outliers». Em: *Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability*, pp. 253–287.
- Filzmoser, P. (2004). «A multivariate outlier detection method». Em: *Proceedings of the Seventh International Conference on Computer Data Analysis and Modeling* 1.1989.
- (2005). «Identification of Multivariate Outliers: A Performance Study». Em: *Austrian Journal of Statistics* 34.2, pp. 127–138.
- Filzmoser, P. e M. Gschwandtner (2012). *Package mvoutlier: Multivariate outlier detection based on robust methods*. <http://cran.r-project.org/web/packages/mvoutlier/>.
- Fischer, M. et al. (2009). «An empirical analysis of multivariate copula models». Em: *Quantitative Finance* 9.7, pp. 839–854.
- Fonseca, T. (2004). «Análise bayesiana de referência para a classe de distribuições hiperbólicas generalizadas». Tese de doutoramento. Dissertação de Mestrado, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil.
- Gilli, M. e E. Schumann (2009). «Heuristic Optimisation in Financial Modelling». Em: *COMISEF Working Papers Series* 7.
- Goldberg, D.E. (1989). «Genetic algorithms in search, optimization, and machine learning». Em:
- Grubbs, F. E. (1969). «Procedures for detecting outlying observations in samples». Em: *Technometrics* 11.1, pp. 1–21.
- Hand, D. J. e K. Yu (2001). «Idiot’s Bayes - Not So Stupid After All?» Em: *International Statistical Review* 69.3, pp. 385–398.
- Jackwerth, J. C. (1999). «Option Implied Risk-Neutral Distributions and Implied Binomial Trees: A Literature Review». Em: *Journal of Derivatives* 7.2, pp. 66–82.
- Joe, H. (1997). *Multivariate Models and Dependence Concepts*. Chapman e Hall London.

- Jorgensen, B. (1982). *Statistical properties of the generalized inverse Gaussian distribution*. Volume 9 de Lectures Notes in Statistics. Heidelberg: Springer.
- Kwiatkowski, J. e R. Rebonato (2011). «A coherent aggregation framework for stress testing and scenario analysis». Em: *Applied Mathematical Finance* 18.2, pp. 139–154.
- Leland, H. E. e M. Rubinstein (1976). «The evolution of portfolio insurance». Em: *Portfolio Insurance: A Guide to Dynamic Hedging*.
- Leroy, A.M. e P.J. Rousseeuw (1987). «Robust regression and outlier detection». Em: *Wiley Series in Probability and Mathematical Statistics, New York: Wiley, 1987* 1.
- Markowitz, H.M. (1959). «Portfolio Selection: Efficient Diversification of Investments». Em: *Cowles Foundation Monograph* 16.
- McNeil, A.J., R. Frey e P. Embrechts (2005). *Quantitative Risk Management: concepts, techniques, and tools*. Princeton University Press.
- Mello, A. (2009). «Volatilidade Implícita das Opções de Ações: Uma Análise Sobre a Capacidade de Previsão do Mercado sobre a Volatilidade Futura». Em: *Dissertação (mestrado) - Escola Economia de São Paulo*, p. 44.
- Melo, V. e A. Delbem (2009). *Arquitetura de Amostragem Inteligente para Determinação de Regiões Promissoras em Problemas de Otimização Numérica Global*.
- Meucci, A. (2007). *Risk and Asset Allocation*. Springer.
- Moskowitz, H. e R.K. Sarin (1983). «Improving the consistency of conditional probability assessment for forecasting and decision making». Em: *Management Science* 29.6, pp. 735–749.
- Mullen, K. et al. (2011). *DEoptim: An R Package for Global Optimization by Differential Evolution*. <http://CRAN.R-project.org/package=DEoptim>.
- Nelsen, R.B. (1999). *An introduction to Copulas*. Springer, New York.
- Panchenko, V. (2005). «Goodness-of-fit test for copulas». Em: *Physica A: Statistical Mechanics and its Applications* 355.1, pp. 176–182.
- Paoletta, M. (2007). *Intermediate probability: a computational approach*. Wiley, Chichester.
- Pasquariello, P. (2012). «Financial Market Dislocations». Em: *Available at SSRN 1769771*.
- Patton, A.J. (2009). «Copula-based models for financial time series». Em: *Handbook of financial time series*, pp. 767–785.

- Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann Publishers Inc.
- (2009). *Causality: Models, Reasoning, and Inference*. Cambridge University Press, Cambridge, UK.
- Perold, A.F. (1986). «Constant proportion portfolio insurance». Em: *Harvard Business School*.
- Poston, W. et al. (1997). «A deterministic method for robust estimation of multivariate location and shape». Em: *Journal of Computational and Graphical Statistics* 6.3, pp. 300–313.
- Prastawa, M. et al. (2004). «A brain tumor segmentation framework based on outlier detection». Em: *Medical Image Analysis* 8.3, pp. 275–283.
- Price, K.V., R.M. Storn e J.A. Lampinen (2006). *Differential Evolution: A Practical Approach to Global Optimization*. Springer-Verlag, Berlin, 2nd edition.
- R Development Core Team (2005). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, <http://www.r-project.org>.
- Rebonato, R. (2010). *Coherent Stress Testing: A Bayesian Approach to the Analysis of Financial Stress*. John Wiley Sons Ltd, United Kingdom.
- Rebonato, R. e A. Denev (2011). «Coherent Asset Allocation and Diversification in the Presence of Stress Events». Em: *Available at SSRN 1824207*.
- Rousseeuw, P. e K. Van Driessen (1999). «A fast algorithm for the minimum covariance determinant estimator». Em: *Technometrics* 41.3.
- Rubinstein, M. e H.E. Leland (1981). «Replicating options with positions in stock and cash». Em: *Financial Analysts Journal* 37.4, pp. 63–72.
- Sklar, A. (1959). «Fonctions de répartition à n dimensions et leurs marges». Em: *Publ. Inst. Statist. Univ. Paris* 8.1, p. 11.
- Sornette, D. (2004). *Why stock markets crash: critical events in complex financial systems*. Princeton University Press.
- Sornette, Y.M.D. (2006). *Extreme financial risks*.
- Storn, R. e K. Price (1997). «Differential Evolution - A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces». Em: *Journal of Global Optimization* 11, 341:359.

Taleb, N.N. (2010). *The black swan: The impact of the highly improbable*. Random House Trade Paperbacks.

Williamson, J. (2005). *Bayesian nets and causality*. Oxford University Press, 2005.

A. Descrição dos tickers utilizados

Tabela A.1: Descrição dos *tickers* dos ativos utilizados

Ativo	Descrição	Ativo	Descrição
IBOV	Índice Ibovespa	BVMF3	BMFBOVESPA ON NM
GOLL4	GOL PN	VALE5	VALE R DOCE PNA
NATU3	NATURA ON NM	USIM5	USIMINAS PNA
AMBV4	AMBEV PN	GOAU4	GERDAU MET PN
BBAS3	BRASIL ON	PETR4	PETROBRAS PN
BBDC4	BRADESCO PN	CMIG4	CEMIG PN
ITUB4	ITAUUNIBANCO PN	TRPL4	TRAN PAULIST PN

Fonte: Descrições obtidas do *site* da BM&FBOVESPA. Disponível em <http://www.bmfbovespa.com.br> (Acesso em 12 setembro 2012).

B. Estatísticas descritivas

Tabela B.1: Estatística descritiva dos log-retornos diários dos preços (ajustados para proventos) dos ativos da amostra de dados (%)

Ativo	Média	Mediana	Máximo	Mínimo	D.Padrão	Assimetria	Curtose
AMBV4	0,0080	0,0000	15,7686	-22,5895	3,4623	-0,1707	5,9809
BBAS3	0,0871	0,1066	10,9434	-12,9896	1,9529	-0,0385	4,9437
BBDC4	-0,0541	-0,0863	17,6704	-15,9603	3,0896	0,0774	3,5847
BVMF3	0,0370	0,0914	19,9889	-12,2123	2,4667	0,6797	8,2063
CMIG4	0,1069	0,0543	12,6041	-6,6226	2,1343	0,2929	1,8197
GOAU4	-0,0323	-0,0697	21,7723	-24,3600	3,6767	-0,2284	6,7800
GOLL4	-0,0343	0,0369	12,9750	-14,8035	2,5803	-0,2341	5,1678
IBOV	0,0237	0,0000	21,0039	-12,9425	2,7472	0,7156	7,9879
ITUB4	0,0334	0,0000	18,8256	-16,6832	2,7647	0,0764	7,1695
NATU3	0,1675	0,1030	10,9063	-11,9629	1,9131	0,0408	5,6002
PETR4	0,0283	0,0744	12,5657	-16,4432	2,6035	-0,0757	5,8127
TRPL4	-0,0869	-0,0991	16,6281	-15,9676	3,0946	0,1141	4,4216
USIM5	0,0719	0,0794	15,7629	-13,9651	1,8700	-0,2495	12,0057
VALE5	0,0124	0,0797	13,6794	-12,0961	2,1407	0,0567	7,1732

Fonte: elaboração própria a partir dos dados apresentados na seção 1.2.

C. Evidências de liquidez

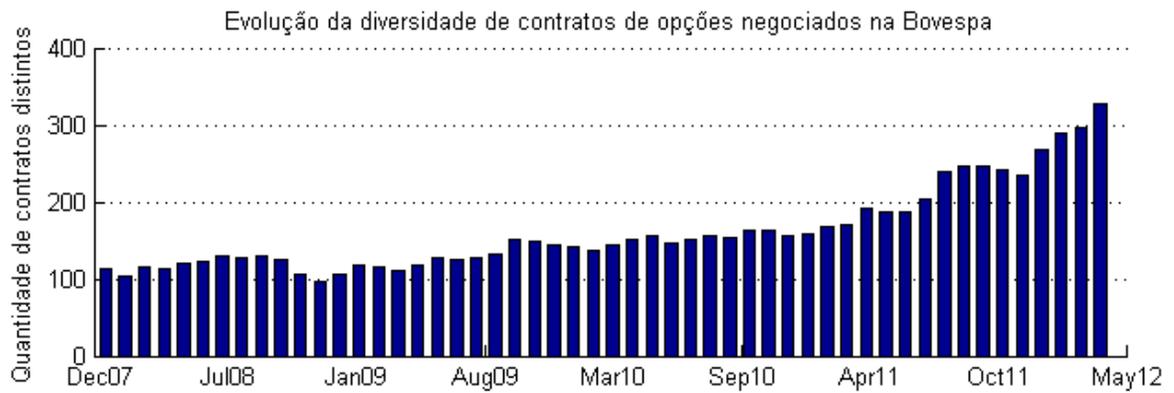


Figura C.1: Evolução da diversidade de opções de ações negociadas na Bovespa para o intervalo da amostra de dados.

Fonte: elaboração própria.

Tabela C.1: Representatividade de cada ativo subjacente da base de dados a partir do número de negócios com opções para o período entre Agosto de 2008 e Abril de 2012

Ativo Subjacente	Quantidade de negócios	% do total de negócios
VALE5	39.108.920	50,345%
PETR4	31.803.548	40,941%
BVMF3	917.550	1,181%
ITUB4	568.415	0,732%
USIM5	123.569	0,159%
BBDC4	101.994	0,131%
BBAS3	78.206	0,101%
AMBV4	4.698	0,006%
CMIG4	1.764	0,002%
GOAU4	1.751	0,002%
GOLL4	866	0,001%
NATU3	808	0,001%
Outras	4.969.697	6,398%
Total	77.681.786	100,000%

Fonte: elaboração própria a partir dos dados apresentados na seção 1.2.

D. Distribuição Hiperbólica Generalizada

A família das distribuições hiperbólicas generalizadas (DHG) é obtida através de uma mistura na média e na variância da Normal. O vetor aleatório \mathbf{X} terá uma DHG multivariada se

$$\mathbf{X} = \boldsymbol{\mu} + W\boldsymbol{\gamma} + \sqrt{W}\mathbf{A}\mathbf{Z}$$

na qual

- i) $\mathbf{Z} \sim N_k(\mathbf{0}, \mathbf{I}_k)$, um vetor de variáveis aleatórias (v.a.) iid com distribuição normal padrão (média zero e variância um);
- ii) $\mathbf{A} \in \mathbb{R}^{d \times k}$, uma matriz de valores constantes;
- iii) $\boldsymbol{\mu}, \boldsymbol{\gamma} \in \mathbb{R}^d$, vetores de parâmetros;
- iv) $W \geq 0$, uma variável aleatória independente de \mathbf{Z} e com distribuição Inversa Gaussiana Generalizada (IGG), i.e. $W \sim N^-(\lambda, \chi, \psi)$. A densidade de W é dada por

$$f(x) = \frac{\chi^{-\lambda}(\sqrt{\chi\psi})^\lambda}{2K_\lambda(\sqrt{\chi\psi})} x^{\lambda-1} \exp(-\frac{1}{2}(\chi x^{-1} + \psi x)), \quad x \geq 0$$

com K_λ representando a função de Bessel modificada de terceiro tipo; ver Abramowitz e Stegun (1968) para mais detalhes sobre a função de Bessel. Para mais detalhes sobre a IGG, ver Jorgensen (1982).

Os parâmetros da DHG definidos como acima permitem as seguintes interpretações:

- λ, χ e ψ determinam a forma da distribuição ao alocar mais ou menos peso ao centro ou à cauda. Em geral, quanto maiores esses parâmetros, mais próxima a distribuição será da Normal.
- $\boldsymbol{\mu}$ é o parâmetro de locação.
- $\Sigma = \mathbf{A}\mathbf{A}'$ é a matriz de variância-covariância (*dispersion-matrix*).
- $\boldsymbol{\gamma}$ é o parâmetro de assimetria. Se igual a zero, então a distribuição é simétrica em torno de $\boldsymbol{\mu}$.

Uma vez que $\mathbf{X}|W = w \sim N_d(\boldsymbol{\mu} + w\boldsymbol{\gamma}, w\Sigma)$, a densidade da DHG pode ser encontrada pela mistura de $\mathbf{X}|W$ em relação a W ,

$$f_{\mathbf{X}}(\mathbf{x}) = \int_0^\infty f_{\mathbf{X}|W}(\mathbf{x}|w) f_W(w) dw$$

$$f_{\mathbf{X}}(\mathbf{x}) = \int_0^\infty \frac{e^{(\mathbf{x}-\boldsymbol{\mu})'\Sigma^{-1}\boldsymbol{\gamma}}}{(2\pi)^{\frac{d}{2}}|\Sigma|^{\frac{1}{2}}w^{\frac{d}{2}}} \exp\left\{-\frac{Q(\mathbf{x})}{2w} - \frac{\boldsymbol{\gamma}'\Sigma^{-1}\boldsymbol{\gamma}}{2/w}\right\} f_W(w) dw$$

$$f_{\mathbf{X}}(\mathbf{x}) = c \frac{K_{\lambda-\frac{d}{2}}(\sqrt{(\chi + Q(\mathbf{x}))(\psi + \boldsymbol{\gamma}'\Sigma^{-1}\boldsymbol{\gamma})}) e^{(\mathbf{x}-\boldsymbol{\mu})'\Sigma^{-1}\boldsymbol{\gamma}}}{(\sqrt{(\chi + Q(\mathbf{x}))(\psi + \boldsymbol{\gamma}'\Sigma^{-1}\boldsymbol{\gamma})})^{\frac{d}{2}-\lambda}}$$

Sendo:

- $c = \frac{(\sqrt{\psi/\chi})^\lambda (\psi + \boldsymbol{\gamma}'\Sigma^{-1}\boldsymbol{\gamma})^{\frac{d}{2}-\lambda}}{(2\pi)^{\frac{d}{2}}|\Sigma|^{\frac{1}{2}}K_\lambda(\sqrt{\chi\psi})}$, uma constante normalizadora; e
- $Q(\mathbf{x}) = (\mathbf{x} - \boldsymbol{\mu})'\Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu})$ denotando a distância de Mahalanobis.

Assim, considerando a notação exposta, a DHG multivariada é parametrizada como $\mathbf{X} \sim DHG_d(\lambda, \chi, \psi, \boldsymbol{\mu}, \Sigma, \boldsymbol{\gamma})$. Há outras parametrizações possíveis, veja Eberlein e Prause (2000) e Paoletta (2007).

Para um referencial teórico abrangente, veja McNeil, Frey e Embrechts (2005). Para uma revisão teórica em português, veja Fonseca (2004).