

UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE MATEMÁTICA ESTATÍSTICA E COMPUTAÇÃO
CIENTÍFICA
Departamento de Estatística

Cálculo do VaR utilizando Acoplamentos e Teoria de Valores Extremos

Edimilson Costa Lucas

Prof. Dr. Luiz Koodi Hotta
Orientador

Dissertação apresentada junto ao Departamento de Estatística do Instituto de Matemática Estatística e Computação Científica da Universidade Estadual de Campinas, para obtenção do título de Mestre em Estatística.

Campinas
2003

Cálculo do VaR utilizando Acoplamentos e Teoria de Valores Extremos

Este exemplar corresponde à redação final da dissertação devidamente corrigida e defendida por Edimilson Costa Lucas e aprovada pela comissão julgadora.

Campinas, 22 de Setembro de 2003

Luiz Koodi Hotta
Orientador

Banca Examinadora:

1. Prof. Dr. Luiz Koodi Hotta – Orientador — IMECC/UNICAMP
2. Prof. Dr. Pedro Luiz Valls Pereira – IBMEC/SP
3. Prof. Dra. Verónica Andrea González-López – IMECC/UNICAMP

Dissertação apresentada ao Instituto de Matemática Estatística e Computação Científica, UNICAMP, como requisito parcial para obtenção do título de MESTRE em ESTATÍSTICA.

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DO IMECC DA UNICAMP

Lucas, Edimilson Costa
L962c Cálculo do VaR utilizando acoplamentos e teoria de valores
extremos
/Edimilson Costa Lucas. – Campinas, [S.P. :s.n.], 2003.

Orientador: Luiz Koodi Hotta.
Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Computação Científica.

1. Derivativos(Finanças) 2. Análise de séries temporais I. Hotta, Luiz Koodi. II. Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Computação Científica. III. Título.

Resumo

Muitos métodos de estimação têm sido desenvolvidos para estimar o Valor em Risco (VaR) de um portfólio. Uma das maiores dificuldades na estimação do VaR é conseguir modelar a dependência temporal e contemporânea dos ativos. A teoria de acoplamentos, que é uma ferramenta muito importante para modelar a distribuição conjunta multivariada, pode ser útil nessa tarefa. Por outro lado, o VaR está mais ligado a ocorrência de eventos extremos no mercado financeiro e, desta forma, os retornos podem ser analisados sob o ponto de vista da Teoria de Valores Extremos. O objetivo deste trabalho é apresentar um método de aplicação de acoplamentos para estimar o VaR num portfólio. O modelo utilizado foi o $\text{ARMA}(p_1, q_1) + \text{GARCH}(p_2, q_2)$ para modelar as séries de retornos individuais. As inovações dos modelos marginais foram modeladas através da classe de acoplamentos de valores extremos utilizando-se distribuições marginais empíricas e distribuições marginais GPD (Distribuição Generalizada de Pareto).

Essa metodologia foi aplicada a um portfólio com duas séries financeiras. Para essa aplicação foram utilizados dados diários do IBOVESPA, índice da Bolsa de Valores de São Paulo, e dados diários do Merval, índice da Bolsa de Valores da Argentina no período de 04/03/1997 à 07/11/2000 totalizando 900 observações.

Abstract

Many methods of estimation has been developed for estimating the Value at Risk (VaR) of a portfolio. An of great difficult at estimation of VaR is become modelling the temporal and contemporanea dependence of assets. The copula theory, that is a tool very important for modelling the joint multivariate distribution, can be useful in this work. On the other hand, the VaR is more linked in the ocurrence of extremes events at financial market and, so, the returns can be analysed under the view point of the Extreme Value Theory. The aim of this work is to present a method of estimation of copulas functions for estimating the VaR at portfolio. The model used was the $\text{ARMA}(p_1, q_1) + \text{GARCH}(p_2, q_2)$ for modelling the series of individual returns. The innovations of the marginal models were modelled through class of copulas extreme value with empirical and GPD (Generalized Pareto Distribution) marginals distributions.

This methodology were applied to portfolio with two financial series. For this application were used daily data of the IBOVESPA, São Paulo Stock Exchange Index, and daily data of the Merval, Argentina Stock Exchange Index at period 03/04/1997 to 11/07/2000 totaling 900 observations.

Agradecimentos

Agradeço meu orientador, Prof. Dr. Luiz Koodi Hotta, na verdade, mais que um orientador, um profissional brilhante e um grande amigo onde, com certeza, sou uma pessoa de muita sorte por ter tido a oportunidade de aprender e trabalhar ao lado desse grande mestre.

Agradeço aos professores da banca Pedro Valls e Verónica Gonzáles-Lópes.

Agradeço ao Helder pela sua grande ajuda ao longo desse trabalho, principalmente, na parte computacional.

Agradeço ao Prof. Geraldo Botelho por sempre me incentivar desde o período da graduação.

Agradeço meus grandes amigos do mestrado: Anderson (carioca), Mariana (princesinha do Imecc), Luiz Carlos (Mineiro), Roberto Thomé (Pic), Cláudio, Benilton, Tatiana, Fernando, Celeste, Fabiana e Marcelo.

Agradeço aos meus estimados amigos Nelson, Kalango e Alessandro por todo companheirismo, amizade e apoio ao longo desses anos.

Agradeço meus pais, Antonio e Marta por todo apoio incondicional ao longo desse tempo, seja de forma direta ou indireta sempre estiveram ao meu lado.

Agradeço meus irmãos Edilson e Elieser(in memorian) por todo apoio, amizade, carinho e torcida ao longo de toda minha vida.

Agradeço minhas cunhadas Andréa e Penha, meus sobrinhos Lucas, Isabela e Emilyn por toda alegria e carinho proporcionado.

Agradeço ao CNPq pelo apoio financeiro.

Campinas
22 de Setembro de 2003

Edimilson Lucas

Dedico ao meu grande incentivador, irmão, pai e amigo ***Elieser*** (*in memorian*).

Sumário

1	Acoplamentos	4
1.1	Acoplamentos Não-Condicionais	4
1.1.1	Definições e Propriedades	5
1.1.2	Medidas de Dependência	7
1.2	Acoplamentos Condicionais	11
1.3	Acoplamentos de Valores Extremos	14
1.4	Diagnóstico e Acoplamento Empírico	16
1.5	Estimação	17
1.5.1	O Método de Máxima Verossimilhança	18
1.5.2	O método IFM	19
1.5.3	O método CML	20
2	Modelo para o Portfólio	22
2.1	Modelo $\text{ARMA}(p_1, q_1) + \text{GARCH}(p_2, q_2)$	22
2.2	VaR	24
2.3	Distribuição das Inovações	26
2.4	Estimação do Acoplamento de Valor Extremo	27

3	Aplicação	31
3.1	Modelagem das distribuições marginais	31
3.2	Estimação do Acoplamento e Diagnóstico	39
3.3	Cálculo do VaR	43
4	Conclusões	46
A	Teoria de Valores Extremos	47
A.1	Caso Univariado	48
A.1.1	Distribuição Generalizada do Valor Extremo (GEV) . .	49
A.1.2	Domínio de Atração de Fréchet	51
A.1.3	Domínio de Atração de Weibull	52
A.1.4	Domínio de Atração de Gumbel	52
A.1.5	Estimador de Hill para ξ	53
A.1.6	Distribuição Generalizada de Pareto (GPD)	55
A.2	Caso Multivariado	58

Lista de Tabelas

3.1	Estatísticas Descritivas das Séries de Retornos	35
3.2	Estimativas dos Parâmetros para o modelo GARCH(1,1) e respectivas estimativas dos desvios-padrões.	37
3.3	Estimativas dos Parâmetros da GPD e entre parênteses esti- mativas dos desvios-padrões.	38
3.4	Estimativa dos Parâmetros das Acoplamentos e entre parênteses estimativas dos desvios-padrões.	39
3.5	Distâncias entre o acoplamento empírico e os acoplamentos estimados para 20%, 10%, 5% e 1% das maiores perdas	42
3.6	Proporção de observações com perdas do Portfólio maiores que o VaR estimado para $\alpha = 10\%, 5\%$ e 1%. Os números entre parênteses correspondem a quantidade de observações.	44

Lista de Figuras

2.1	Gráfico Ilustrativo para a estimação dos acoplamentos	28
3.1	Séries Financeiras dos índices do BOVESPA e Merval durante o período de 04/03/1997 - 07/11/2000	33
3.2	Retornos IBOVESPA e Merval durante o período de 04/03/1997 - 07/11/2000	34
3.3	Diagrama de dispersão dos retornos Merval versus IBOVESPA	35
3.4	Auto-correlações e auto-correlações parciais dos retornos	36
3.5	Auto-correlações e auto-correlações parciais dos quadrados dos retornos	36
3.6	Auto-correlações dos quadrados das inovações	37
3.7	Resíduos Padronizados - GARCH(1,1)	38
3.8	Curvas de Nível - Acoplamento de Gumbel(GPD) - dados extremos, onde as perdas são dadas a partir do extremo superior direito da figura (1%, 5%, 10%, 15%, e assim por diante)	40

3.9	Curvas de Nível - Acoplamento de Gumbel(Empírica) - dados extremos, onde as perdas são dadas a partir do extremo superior direito da figura (1%, 5%, 10%, 15%, e assim por diante) .	40
3.10	Curvas de Nível - Acoplamento de Gumbel A(GPD) - dados extremos, onde as perdas são dadas a partir do extremo superior direito da figura (1%, 5%, 10%, 15%, e assim por diante) .	41
3.11	Curvas de Nível - Acoplamento de Gumbel A(Empírica) - dados extremos, onde as perdas são dadas a partir do extremo superior direito da figura (1%, 5%, 10%, 15%, e assim por diante)	41
3.12	Gráfico do VaR 5% estimado - Gumbel(GPD)	45

Introdução

Uma das principais preocupações no mercado financeiro é medir a exposição ao risco. Essa preocupação aumentou com a ocorrência de grandes perdas, tais como, Orange County, Proctor and Gamble, NatWest, Barings, Daiwa Bank e Long Term Capital Management. Devido a esses problemas surgiu a necessidade de se encontrar medidas de risco adequadas. A medida de risco mais utilizada atualmente é a metodologia Valor em Risco, mais conhecida como VaR (Value at Risk). Genericamente falando, VaR é a perda máxima esperada de uma carteira de investimentos, em um certo período de tempo, com um nível de significância especificado.

A administração de riscos financeiros envolve modelagem usando distribuições com caudas pesadas, com mudanças de regime, com interdependências complexas, fazendo com que seja necessário o uso de metodologias mais sofisticadas baseadas em suposições mais realistas.

As carteiras são constituídas por uma combinação linear de ativos. Considerando sua composição atual fixa ao longo do tempo poderia ser aplicada a teoria de valores extremos (TVE) univariada à série dos preços da carteira. Em Souza e Silva(1999) essa metodologia é aplicada e comparada a duas séries financeiras: retornos de C-Bond e retornos de Telebrás. Em Pin-

to e Valls Pereira(2002) é verificado a adequação do ajuste às distribuições de probabilidades dos retornos financeiros, seguindo um procedimento semi-paramétrico baseado em duplo-bootstrapping.

Entretanto, em um portfólio cuja composição varia ao longo do tempo, surge a necessidade de se trabalhar com modelos multivariados. Poucos trabalhos foram desenvolvidos para a teoria de valores extremos multivariada, pois não existe uma noção padrão de ordem, e conseqüentemente, não existe uma noção padrão de extremos. Em Embrechts e Huang(1999) é apresentado uma modelagem das caudas no caso bi-variado. Em González-López e Tanaka(2003) é apresentado uma modelagem de dados bi-variados através de acoplamentos arquimedianos generalizados onde as distribuições marginais são escolhidas entre a empírica, gaussiana, GEV(Distribuição Generalizada de Valores Extremos) ou GPD(Distribuição Generalizada de Pareto). Em Bouyé *at al*(2000) é aplicado, no caso-bivariado, acoplamentos de valores extremos no banco de dados LME (London Metal Exchange). Em Patton(2001) o teorema de Sklar não condicional é estendido para o caso condicional e é feito uma aplicação para modelar a distribuição conjunta variando no tempo com dados do dólar americano e o yen.

Nesse trabalho, é proposto um método de aplicação de acoplamentos para estimar o VaR num portfólio com dois ativos. O modelo utilizado foi o $ARMA(p_1, q_1) + GARCH(p_2, q_2)$ para modelar as séries de retornos individuais. As inovações dos modelos marginais foram modeladas através da classe de acoplamentos de valores extremos utilizando-se distribuições marginais

empíricas e distribuições marginais GPD.

O trabalho está dividido da seguinte forma: no capítulo I, é apresentado uma revisão sobre os principais conceitos e propriedades dos acoplamentos; no capítulo II, é apresentado os métodos de estimação dos acoplamentos e distribuições marginais; no capítulo III, é apresentado o modelo e uma aplicação dessa metodologia para dados reais do IBOVESPA e Merval; e o apêndice apresenta uma revisão teórica sobre a teoria de valores extremos.

Capítulo 1

Acoplamentos

A principal dificuldade da teoria de valores extremos multivariada clássica, está na falta de um caminho eficiente para descrever estruturas de dependência multivariada, ou seja, as dependências entre as componentes de um vetor aleatório. O objetivo desse capítulo será procurar uma maneira para descrever a estrutura de dependência existente entre as distribuições marginais das componentes de um vetor aleatório.

Um bom caminho para atingir esse objetivo será dado pelo conceito de acoplamentos, que será introduzido na próxima seção.

1.1 Acoplamentos Não-Condicionais

Nelsen(1998) é um bom livro introdutório sobre acoplamentos, onde são apresentadas suas principais propriedades e algumas aplicações. De acordo com Nelsen, acoplamentos podem ser vistos sob dois pontos de vista: "sob um ponto de vista, acoplamentos são funções que juntam ou "acoplam" funções distribuições conjuntas a suas funções distribuições marginais. Alternativa-

mente, acoplamentos são funções distribuições multivariadas, cujas marginais unidimensionais são distribuições uniformes no intervalo $(0,1)$ ”.

1.1.1 Definições e Propriedades

Nesta seção serão dadas algumas definições e propriedades para uma boa compreensão do estudo de acoplamentos.

Definição 1.1.1. Um acoplamento de dimensão d é uma função de distribuição de um vetor aleatório em $\mathbb{R}^d$ com marginais uniformes $(0, 1)$. Ou seja, uma função acoplamento C é qualquer função $C : [0, 1]^d \rightarrow [0, 1]$ que satisfaça as seguintes propriedades:

- (i) $C(u_1, \dots, u_d)$ é crescente em cada componente;
- (ii) $C(1, \dots, 1, u_i, 1, \dots, 1) = u_i, \forall i \in \{1, \dots, d\}, u_i \in [0, 1]$; e
- (iii) Para todo $(a_1, \dots, a_d), (b_1, \dots, b_d) \in [0, 1]^d$ com $a_i \leq b_i$ tem-se que:

$$\sum_{i_1}^2 \dots \sum_{i_d}^2 (-1)^{i_1 + \dots + i_d} C(u_{1i_1}, \dots, u_{di_d}) \geq 0, \text{ onde } u_{j1} = a_j \text{ e } u_{j2} = b_j$$
para todo $j \in \{1, \dots, d\}$.

A função de acoplamento também é chamada simplesmente de acoplamento. Da definição tem-se que se $u = F(x)$ e $v = G(y)$, onde F e G são funções de distribuições (marginais) das variáveis aleatórias X e Y , respectivamente, então $C(F(x), G(y))$ é uma função de distribuição conjunta das variáveis aleatórias X e Y . Isso é dado formalmente pelo teorema de Sklar que é o teorema fundamental desse capítulo.

Teorema 1.1.1. (*Sklar*)

Seja H uma função de distribuição d -dimensional com marginais $F_1, \dots, F_d$. Então existe um acoplamento C de modo que para todo $\mathbf{x} \in \overline{\mathbb{R}}^d$ onde $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$,

$$H(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d)).$$

Se $F_1, \dots, F_d$ são todas contínuas, então C é único. Reciprocamente, para um acoplamento C e marginais contínuas $F_1, \dots, F_d$ a função H definida em (1.1.1) é uma função de distribuição com marginais $F_1, \dots, F_d$.

Assim, conforme visto no teorema anterior, se todas as marginais forem contínuas o acoplamento C será único.

Como todas as marginais de qualquer distribuição de valor extremo multivariada são contínuas, serão considerados vetores aleatórios com marginais contínuas.

Esse teorema mostra duas formas de se utilizar funções acoplamentos. Nas aplicações em finanças, em geral, ele é utilizado para encontrar a distribuição multivariada a partir das marginais, que devem reproduzir os fatos estilizados e a estrutura de dependência entre as variáveis. Nesse sentido, o problema de modelagem segue dois passos importantes:

- identificação e estimação das marginais; e
- escolha adequada da função acoplamento de forma a representar de maneira conveniente a estrutura de dependência entre as variáveis.

A seguir, são apresentados alguns exemplos de acoplamentos.

Exemplo 1.1.2. *(Acoplamento independente ou acoplamento produto):*

$$C^\perp(u_1, \dots, u_d) = \prod_{i=1}^d u_i.$$

Exemplo 1.1.3. *Considere a distribuição logística bivariada de Gumbel dada por $F(x_1, x_2) = (1 + e^{-x_1} + e^{-x_2})^{-1}$. As marginais correspondentes são dadas por: $F_1(x_1) = (1 + e^{-x_1})^{-1}$ e $F_2(x_2) = (1 + e^{-x_2})^{-1}$. Com essas informações pode-se encontrar a função acoplamento correspondente da seguinte forma:*

$$C(u_1, u_2) = F(F_1^{-1}(u_1), F_2^{-1}(u_2)) = \frac{u_1 u_2}{u_1 + u_2 - u_1 u_2}$$

Exemplo 1.1.4. *(Acoplamento Gaussiano)*

O acoplamento normal ou gaussiano é dado da seguinte forma:

$$C(x, y) = \int_{-\infty}^{\Phi^{-1}(x)} \int_{-\infty}^{\Phi^{-1}(y)} \frac{1}{2\pi(1-\rho^2)^{1/2}} \exp\left\{\frac{-(s^2-2\rho st+t^2)}{2(1-\rho^2)}\right\} ds dt,$$

onde $-1 < \rho < 1$ e Φ é a função de distribuição normal padrão univariada.

1.1.2 Medidas de Dependência

Sendo a distribuição conjunta definida através da distribuição das marginais e da estrutura de dependência, é de fundamental importância o conhecimento da dependência. A medida clássica de correlação linear é adequada apenas para distribuições elípticas que, geralmente, não é o caso dos dados em finanças. Antes de apresentar algumas medidas de dependência é necessário ter uma noção de dependência perfeita entre duas variáveis. Para isso, é necessário compreender o que significa um acoplamento ser menor que outro.

Definição 1.1.2. Diz-se que um acoplamento C_1 é menor que um acoplamento C_2 , e escreve-se $C_1 \prec C_2$, se:

$$\forall(u_1, \dots, u_d) \in \mathbb{R}^d, C_1(u_1, \dots, u_d) \leq C_2(u_1, \dots, u_d).$$

Como acoplamentos são funções de distribuição multivariadas, pode-se aplicar os limites de Fréchet e mostrar que:

$$C^- \prec C \prec C^+$$

onde:

$$C^-(u_1, \dots, u_j, \dots, u_d) = \max(\sum_{j=1}^d u_j + 1 - d, 0)$$

$$C^+(u_1, \dots, u_j, \dots, u_d) = \min(u_1, \dots, u_j, \dots, u_d)$$

Pode-se mostrar que para $d = 2$, C^- e C^+ são acoplamentos, enquanto que, para $d > 2$ somente C^+ é um acoplamento. Mais precisamente, C^- é a função de distribuição do vetor aleatório $(U, 1 - U)$ e C^+ do vetor aleatório (U, U) , onde $U \sim U(0, 1)$.

A distribuição de $(U, 1 - U)$ tem toda a sua massa sobre a diagonal entre $(0, 1)$ e $(1, 0)$, e a distribuição de (U, U) tem toda a sua massa sobre a diagonal entre $(0, 0)$ e $(1, 1)$. Dessa maneira, pode-se dizer que C^- e C^+ descrevem *dependência perfeita positiva* e *dependência perfeita negativa*, respectivamente. Para $d = 2$, C^- é chamado de *acoplamento limitante inferior de Fréchet ou acoplamento "anti-monotônico"*, enquanto que para $d \geq 2$, C^+ é chamado de *acoplamento limitante superior de Fréchet ou acoplamento "co-monotônico"*.

A definição acima de dependência perfeita implica que esse tipo de dependência é uma propriedade do acoplamento, isto é, não é influenciada pelas marginais. Assim, se for considerado um vetor aleatório (X, Y) tendo acoplamento C^- ou C^+ , é suficiente dizer que X e Y são perfeitamente dependentes.

Agora, será dado um resultado que mostra que as funções acoplamentos são invariantes em relação a transformações contínuas crescentes.

Proposição 1.1.5. *Seja $(X_1, \dots, X_d)$ um vetor aleatório com acoplamento C e sejam $T_1, \dots, T_d$ funções estritamente crescentes. Então, $(T_1(X_1), \dots, T_d(X_d))$ também tem acoplamento C .*

A dependência entre variáveis aleatórias é caracterizada inteiramente pelo acoplamento da correspondente distribuição multivariada. Entretanto, para

possibilitar o uso dessa informação na identificação da função de acoplamento adequada, deve-se ter uma forma mais adequada de caracterizar essa dependência, se possível, através de um índice. Dado a invariância do acoplamento a transformações das variáveis aleatórias através de funções monótonas crescentes é natural procurar uma medida que também seja invariante a esse tipo de transformações. Essa propriedade de invariância está intimamente ligada ao conceito de concordância.

Definição 1.1.3. Uma medida numérica k de associação entre duas variáveis aleatórias contínuas X_1 e X_2 cujo acoplamento seja C , denotada por k_{X_1, X_2} ou por k_C é uma **medida de concordância** se satisfaz as seguintes propriedades:

1. k é definida para todo par X_1, X_2 de variáveis aleatórias contínuas;
2. $-1 = k_{X, -X} \leq k_C \leq k_{X, X} = 1$;
3. $k_{X_1, X_2} = k_{X_2, X_1}$;
4. Se X_1 e X_2 são independentes, então $k_{X_1, X_2} = k_{C^\perp} = 0$;
5. $k_{-X_1, X_2} = k_{X_1, -X_2} = -k_{X_1, X_2}$;
6. Se $C_1 \prec C_2$, então $k_{C_1} \leq k_{C_2}$; e
7. Se $\{(X_{1,j}, X_{2,j})\}$ é uma seqüência de variáveis aleatórias contínuas com acoplamentos C_j , e se $\{C_j\} \rightarrow C$ então $k_{C_j} \rightarrow k_C$.

Algumas medidas de concordância que satisfazem a propriedade de invariância são apresentadas no exemplo a seguir.

Exemplo 1.1.6. *Algumas medidas de concordância:*

$$\text{tau de Kendall: } \tau = 4 \int \int C(u_1, u_2) dC(u_1, u_2) - 1$$

$$\text{rho de Spearman: } \rho = 12 \int \int u_1 u_2 dC(u_1, u_2) - 3$$

$$\text{índice Gini: } \gamma = 2 \int \int (|u_1 + u_2 - 1| - |u_1 - u_2|) dC(u_1, u_2)$$

Alguns acoplamentos não cobrem o intervalo $[-1,1]$ de valores possíveis para as medidas de concordância. Por exemplo, os acoplamentos de Gumbel, Gumbel A e Galambos, que serão apresentados mais adiante, não permitem dependência negativa.

Definição 1.1.4. Uma medida numérica δ de associação entre duas variáveis aleatórias contínuas X_1 e X_2 cujo acoplamento seja C , denotada por δ_{X_1, X_2} ou por δ_C é uma **medida de dependência** se satisfaz as seguintes propriedades:

1. δ é definida para todo par X_1, X_2 de variáveis aleatórias contínuas;
2. $0 = \delta_{C^\perp} \leq \delta_C \leq \delta_{C^+} = 1$;
3. $\delta_{X_1, X_2} = \delta_{X_2, X_1}$;
4. $\delta_{X_1, X_2} = \delta_{C^\perp} = 0$ se, e somente se, X_1 e X_2 são independentes;
5. $\delta_{X_1, X_2} = \delta_{C^+} = 1$ se, e somente se, cada uma de X_1 e X_2 é quase-certamente uma função estritamente monótona da outra;
6. Se T_1 e T_2 são quase-certamente funções estritamente monótonas sobre $Im(X_1)$ e $Im(X_2)$, respectivamente, então: $\delta_{T_1(X_1), T_2(X_2)} = \delta_{X_1, X_2}$; e
7. Se $\{(X_{1,j}; X_{2,j})\}$ é uma seqüência de variáveis aleatórias contínuas com acoplamentos C_j , e se $\{C_j\} \rightarrow C$ então $\delta_{C_j} \rightarrow \delta_C$.

Seguem dois exemplos de medidas de dependência:

Exemplo 1.1.7. $\sigma = 12 \int \int |C(u_1, u_2) - C^\perp(u_1, u_2)| du_1 du_2$
 $\Phi^2 = 90 \int \int |C(u_1, u_2) - C^\perp(u_1, u_2)|^2 du_1 du_2$

No estudo de acoplamentos é bom ter atenção às seguintes características encontradas nas séries financeiras:

1. Dependência é normalmente maior nas caudas do que nos valores centrais da distribuição;

2. Dependência é assimétrica para valores positivos e negativos;
3. Dependência pode modificar ao longo do tempo, inclusive com quebras estruturais; e
4. Valores aberrantes podem ocorrer com certa frequência.

Sabe-se que para calcular o VaR há interesse no comportamento nas caudas das distribuições. Assim, é natural que se procure analisar a dependência nesta região. Uma medida de dependência nas caudas entre duas variáveis aleatórias X e Y pode ser dada por:

Definição 1.1.5. Se uma função de acoplamento bivariada C é tal que

$$\lim_{u \rightarrow 1} \frac{\overline{C}(u, u)}{1-u} = \lambda \text{ existe,}$$

então C tem dependência na cauda superior para $\lambda \in (0, 1]$ e não tem dependência na cauda superior quando $\lambda = 0$.

Observação 1.1.1. $\overline{C}$ é a função de sobrevivência conjunta que é da forma:

$$\overline{C}(u_1, u_2) = 1 - u_1 - u_2 + C(u_1, u_2).$$

A medida λ pode ser vista como a probabilidade de uma variável ser extrema dado que a outra também é extrema, ou seja:

$$\lambda(u) = P(U_1 > u | U_2 > u) = \frac{\overline{C}(u, u)}{1-u}$$

1.2 Acoplamentos Condicionais

Nessa seção serão apresentados o conceito de acoplamentos condicionais e o teorema de Sklar para o caso condicional, permitindo utilizar a teoria de acoplamentos para modelar distribuições condicionais.

Definição 1.2.1. Um acoplamento bi-variado, condicionado a um conjunto de informações F , é uma função $C : [0, 1] \times [0, 1] \rightarrow [0, 1]$ com as seguintes propriedades:

- (i) $C(u, 0|F) = C(0, v|F) = 0$, e $C(u, 1|F) = u$ e $C(1, v|F) = v$, $\forall u, v \in [0, 1]$
- (ii) $C(u_2, v_2|F) - C(u_1, v_2|F) - C(u_2, v_1|F) + C(u_1, v_1|F) \geq 0$ para todo $u_1, u_2, v_1, v_2 \in [0, 1]$, de modo que $u_1 \geq u_2$ e $v_1 \geq v_2$.

A primeira condição da definição 1.2.1 mostra o limite inferior da distribuição, e assegura que as distribuições marginais, $C(u, 1|F)$ e $C(1, v|F)$, são uniformes. A segunda condição assegura que a probabilidade de um ponto observado na região $[u_1, u_2] \times [v_1, v_2]$ é não-negativa.

Alternativamente, a partir dessas condições, pode-se definir acoplamento condicional como uma função de distribuição condicional bi-variada de um par de vetores aleatórios (U, V) cujas distribuições marginais são uniformes $(0,1)$. Com essa definição torna-se mais intuitivo quando considera-se a *transformação integral de probabilidade*, ou seja, a variável aleatória $U_t = F(X_t|F)$ é a transformada integral de probabilidade de X_t , e é conhecida ter distribuição uniforme $(0,1)$.

Uma extensão do resultado chave da teoria de acoplamentos é o teorema de Sklar para distribuições condicionais.

Teorema 1.2.1. *Seja H uma função distribuição condicional bi-variada com marginais contínuas F e G , e considere F algum conjunto de informações. Então existe um único acoplamento condicional $C : [0, 1] \times [0, 1] \rightarrow [0, 1]$ de modo que*

$$H(x, y|F) = C(F(x|F), G(y|F)|F), \forall x, y \in \overline{\mathbb{R}}$$

Reciprocamente, se C é um acoplamento condicional e F e G são as funções distribuições condicionais de duas variáveis aleatórias X e Y , então a função H definida acima é uma função de distribuição condicional bi-variada com marginais F e G .

Uma consequência do teorema 1.2.1 garante que é possível construir um acoplamento condicional a partir de qualquer função de distribuição condicional bi-variada, mas primeiro é necessária a definição de *quasi-inversa* de uma função.

Definição 1.2.2. A quasi-inversa, $F^{(-1)}$, de uma função de distribuição F é definida como:

$$F^{(-1)}(u) = \inf\{x : F(x) \geq u\}, u \in [0, 1].$$

Observação 1.2.1. Se F é uma função estritamente crescente então a definição 1.2.2 fornece a função inversa usual de F .

Corolário 1.2.2. *Seja H uma função de distribuição condicional bi-variada com distribuições marginais contínuas, F e G , e sejam $F^{(-1)}$ e $G^{(-1)}$ as funções quasi-inversas das distribuições marginais. Considere F um conjunto de informações. Então existe um único acoplamento condicional $C : [0, 1] \times [0, 1] \rightarrow [0, 1]$ de tal maneira que:*

$$C(u, v|F) = H(F^{(-1)}(u|F), G^{(-1)}(v|F)|F), \forall u, v \in [0, 1]$$

Esse corolário completa a idéia que uma função de distribuição bi-variada pode ser decomposta em três partes: dadas quaisquer duas distribuições marginais e qualquer acoplamento tem-se uma função de distribuição conjunta, e vindo de uma dada distribuição conjunta pode-se extrair as distribuições marginais e um acoplamento.

1.3 Acoplamentos de Valores Extremos

Nesta seção será dada uma versão multivariada do teorema de Fisher-Tippet bem como definição, propriedades e alguns exemplos de acoplamentos de valores extremos. Serão apresentados, também, os acoplamentos de Gumbel e Gumbel A que serão utilizados nesse trabalho.

Sabe-se que se H é uma distribuição de valor extremo, então suas marginais pertencem a família GEV (Distribuição Generalizada de Valores Extremos), ou seja, H_i é do tipo H_{ξ_i} para algum ξ_i , $i = 1, \dots, d$. Ainda mais, como H tem marginais contínuas então terá um único acoplamento C_0 . Para uma revisão sobre definições e outros aspectos teóricos da teoria de valores extremos, ver Apêndice.

Teorema 1.3.1. (*Fisher-Tippet multivariado*)

Suponha $F = (F_1, \dots, F_n) \in MDA(H)$ (Domínio de atração do máximo de H) e $H(\mathbf{x}) = C_0(H_1(x_1), \dots, H_d(x_d))$. Então:

- (i) As marginais H_i , $i = 1, \dots, d$ de H são do tipo H_{ξ_i} , para algum $\xi_i \in \mathbb{R}$ e $F_i \in MDA(H_i)$; e
- (ii) $C_0(\mathbf{u}^t) = C_0^t(\mathbf{u})$, para todo $t > 0$.

Agora, suponha que o acoplamento C satisfaça $C(\mathbf{u}^t) = C^t(\mathbf{u})$ para todo $t > 0$. Tome $F_i \in MDA(H_i)$, $i = 1, \dots, d$ e defina

$$F(\mathbf{x}) = C(F_1(x_1), \dots, F_d(x_d)), \text{ e}$$

$$H(\mathbf{x}) = C(H_1(x_1), \dots, H_d(x_d)).$$

Considere um vetor aleatório $\mathbf{X} \sim F$ e defina o bloco de máximos $\mathbf{M}_n$. Sabe-se que para $k = 1, \dots, d$ existem números $a_{k,n} > 0$ e $b_{k,n}$ tais que

$$F_k^n(a_{k,n}x + b_{k,n}) \rightarrow H_k(x), \text{ quando } n \rightarrow \infty.$$

Sejam $\mathbf{a}_n = (a_{1,n}, \dots, a_{d,n})$ e $\mathbf{b}_n = (b_{1,n}, \dots, b_{d,n})$. Observe que:

$$P[\mathbf{M}_n \leq \mathbf{a}_n x + \mathbf{b}_n] = F^n(\mathbf{a}_n x + \mathbf{b}_n) = C^n(F_1(a_{1,n}x_1 + b_{1,n}), \dots, F_d(a_{d,n}x_d + b_{d,n})) = C(F_1^n(a_{1,n}x_1 + b_{1,n}), \dots, F_d^n(a_{d,n}x_d + b_{d,n}))$$

Logo, $C(H_1(x_1), \dots, H_d(x_d)) = H(\mathbf{x})$, quando $n \rightarrow \infty$, ou seja, $F \in MDA(H)$.

Com essas considerações acima, tem-se o seguinte teorema:

Teorema 1.3.2. *Uma função de distribuição d-variada H com acoplamento associado C é uma distribuição de valor extremo se, e somente se:*

- (i) *As marginais H_i são do mesmo tipo que H_ξ para algum $\xi \in \mathbb{R}$,*
- (ii) *$C(\mathbf{u}^t) = C^t(\mathbf{u})$ para todo $t > 0$.*

Com isto, segue a definição 1.3.1:

Definição 1.3.1. Um acoplamento C que satisfaz $C(\mathbf{u}^t) = C^t(\mathbf{u})$ para todo $t > 0$, é denominado de *acoplamento de valor extremo*.

Exemplo 1.3.3. *(Acoplamento Independente)*

O exemplo mais simples de acoplamento de valor extremo é dado pelo acoplamento independente:

$$C^\perp(\mathbf{u})^t = (\prod_{i=1}^d u_i)^t = C^\perp(\mathbf{u}^t).$$

Exemplo 1.3.4. *(Acoplamento de Galambos)*

Um outro exemplo de acoplamento de valor extremo é o acoplamento de Galambos:

$$C_\delta(u, v) = uv \exp\{[(-\log u)^{-\delta} + (-\log v)^{-\delta}]^{\frac{-1}{\delta}}, \delta \in [0, \infty).$$

Como dito no início dessa seção, os acoplamentos de valores extremos que serão trabalhados são o de Gumbel e Gumbel A.

O **acoplamento de Gumbel** é apresentado da seguinte forma:

$$C(u_1, \dots, u_d) = \exp\{ -[(-\log u_1)^{1/\beta} + \dots + (-\log u_d)^{1/\beta}]^\beta \}$$

onde $0 < \beta \leq 1$ e $0 < u_i < 1$, $i = 1, \dots, d$.

Esse é um acoplamento de valor extremo, pois:

$$\begin{aligned} C(u_1, \dots, u_d)^t &= \exp\{ -t[(-\log u_1)^{1/\beta} + \dots + (-\log u_d)^{1/\beta}]^\beta \} \\ &= \exp\{ -[(-t \log u_1)^{1/\beta} + \dots + (-t \log u_d)^{1/\beta}]^\beta \} \\ &= C(u_1^t, \dots, u_d^t) \end{aligned}$$

O parâmetro $0 < \beta \leq 1$ controla a quantidade de dependência entre as componentes de uma variável aleatória; $\beta = 1$ indica independência e o limite de C para $\beta \rightarrow 0^+$ indica uma dependência perfeita. A dependência de cauda superior é dada por: $\lambda = 2 - 2^\beta$.

O **acoplamento de Gumbel A** tem a seguinte forma:

$$C_\delta(u, v) = uv \exp\left\{ -\delta \frac{\log u \log v}{\log uv} \right\}$$

Também é um acoplamento de valor extremo, pois:

$$C_\delta(u, v)^t = u^t v^t \exp\left\{ -\delta \frac{t^2 \log u \log v}{t \log uv} \right\} = C_\delta(u^t, v^t).$$

Esses dois acoplamentos são simétricos, isto é, $C(u, v) = C(v, u)$.

1.4 Diagnóstico e Acoplamento Empírico

Para selecionar o acoplamento que apresenta um melhor ajuste, não existe uma forma automática de seleção do modelo, mas ferramentas que podem auxiliar nessa escolha. Em Junker *et al*(2003) é apresentado seis métodos para seleção de acoplamentos. Nesse trabalho serão utilizadas duas maneiras

para seleção de acoplamentos: uma é a escolha de uma norma apropriada para medir a distância entre o acoplamento estimado e o acoplamento empírico, e a outra é o método gráfico construindo as curvas de nível do acoplamento estimado e acoplamento empírico, num mesmo gráfico, que também será bastante útil.

Para esse tipo de diagnóstico é preciso apresentar o acoplamento empírico.

Considere os dados amostrais $\chi = \{(x_1^t, \dots, x_N^t)\}_{t=1}^T$.

Assim, o acoplamento empírico é dado por:

$$\hat{C}(\frac{t_1}{T}, \dots, \frac{t_N}{T}) = \frac{1}{T} \sum_{t=1}^T \prod_{i=1}^N I_{[x_i^t \leq x_i^{(t_i)}]}$$

onde $x_n^{(t)}$ são as estatísticas de ordem e $1 \leq t_1, \dots, t_N \leq T$.

A densidade do acoplamento empírico tem a seguinte expressão:

$$\hat{c}(\frac{t_1}{T}, \dots, \frac{t_N}{T}) = \sum_{i_1=1}^2 \dots \sum_{i_n=1}^2 (-1)^{i_1 + \dots + i_n} \hat{C}(\frac{t_1 - i_1 + 1}{T}, \dots, \frac{t_n - i_n + 1}{T}).$$

Seja $\{C_k\}_{1 \leq k \leq K}$ o conjunto de acoplamentos sob consideração. Então a distância utilizada é dada por:

$$\bar{d}_n(\hat{C}, C_k) = [\sum_{t_1=1}^T \dots \sum_{t_n=1}^T (\hat{C}(\frac{t_1}{T}, \dots, \frac{t_n}{T}) - C_k(\frac{t_1}{T}, \dots, \frac{t_n}{T}))^2]^{1/2}$$

1.5 Estimação

Nessa seção serão mostrados os métodos utilizados para a estimação dos parâmetros dos acoplamentos.

1.5.1 O Método de Máxima Verossimilhança

Considere f sendo a densidade da função distribuição conjunta F . Então:

$$f(x_1, \dots, x_n|F) = c(F_1(x_1|F), \dots, F_n(x_n|F)) \prod_{i=1}^n f_i(x_i|F),$$

onde f_i é a densidade univariada da distribuição marginal F_i e c é a densidade do acoplamento dada pela seguinte expressão:

$$c(u_1, \dots, u_n|F) = \frac{\partial C(u_1, \dots, u_n|F)}{\partial u_1 \dots \partial u_n}.$$

Agora, considere um conjunto de T dados empíricos de n retornos $\chi = \{(x_1^t, \dots, x_n^t)\}_{t=1}^T$. Seja $\vartheta = (\vartheta_1, \dots, \vartheta_n, \theta)$ o vetor de parâmetros à serem estimados, onde ϑ_i , $i = 1, \dots, n$ é o vetor de parâmetros da distribuição marginal F_i e θ é o vetor de parâmetros do acoplamento. Assim, a função de log-verossimilhança será a seguinte:

$$l(\vartheta|F, x) = \sum_{t=1}^T \log c(F_1(x_1^t; \vartheta_1|F), \dots, F_n(x_n^t; \vartheta_n|F); \theta) + \sum_{t=1}^T \sum_{i=1}^n \log f_i(x_i^t; \vartheta_i|F). \quad (1.5.1)$$

Dessa forma, o estimador de máxima verossimilhança (EMV) $\hat{\vartheta}$ do vetor de parâmetros ϑ é aquele que maximiza a expressão anterior, ou seja:

$$\hat{\vartheta} = \arg \max l(\vartheta|F, x).$$

O estimador de máxima verossimilhança tem uma distribuição assintoticamente normal, ou seja:

$$\sqrt{T}(\hat{\vartheta}_{ML} - \vartheta) \rightarrow N(0, I^{-1}(\vartheta)),$$

onde $I(\vartheta)$ é a matriz de informação de Fisher.

Esse método apresenta um grande custo computacional para estimar os parâmetros do acoplamento e de suas distribuições marginais conjuntamente. Assim, é necessário a utilização de outros métodos de estimação desses parâmetros conforme serão apresentados nas duas próximas seções.

1.5.2 O método IFM

Para se calcular o VaR, nesse trabalho, um dos modelos que será utilizado, primeiramente, estimará os parâmetros da GPD sobre os resíduos do modelo ARMA+GARCH, para posterior estimação dos parâmetros dos acoplamentos.

De acordo com o método IFM (Inference Functions for Margins), apresentado em Joe and Xu(1996), os parâmetros das distribuições marginais são estimados separadamente dos parâmetros do acoplamento. Em outras palavras, o processo de estimação é dividido em dois passos:

- (i) estima os parâmetros ϑ_i , $i = 1, \dots, n$ das distribuições marginais F_i utilizando o método de máxima verossimilhança:

$$\hat{\vartheta}_i = \arg \max l^i(\vartheta_i | F, x) = \arg \max \sum_{t=1}^T \log f_i(x_i^t | F)$$

onde l^i é a função de log-verossimilhança da distribuição marginal F_i ;

- (ii) estima os parâmetros do acoplamento θ , dado as estimativas encontradas no passo (i):

$$\hat{\theta} = \arg \max l^c(\theta|F) = \arg \max \sum_{t=1}^T \log c(F_1(x_1^t; \hat{\vartheta}_1), \dots, F_n(x_n^t; \hat{\vartheta}_n); \theta|F)$$

onde l^c é a função de log-verossimilhança do acoplamento.

Sob condições usuais de regularidade Patton(2001) mostrou que o estimador de máxima verossimilhança condicional em dois estágios é assintoticamente consistente e normal.

1.5.3 O método CML

Um outro modelo que será utilizado para cálculo do VaR, será a partir de distribuições marginais empíricas estimadas sobre os resíduos do modelo ARMA+GARCH aplicado às séries de retornos. Com essas distribuições marginais empíricas estimadas serão estimados os parâmetros dos acoplamentos.

O método de Máxima Verossimilhança Canônica (CML) é análogo ao método IFM exceto que, na estimação dos acoplamentos, as distribuições marginais são estimadas pelas distribuições marginais empíricas. O processo de estimação está dividido em dois passos:

- (i) transforma o conjunto de dados $(x_1^t, \dots, x_n^t)$, $t = 1, \dots, T$, em variáveis uniformes $(\hat{u}_1^t, \dots, \hat{u}_n^t)$, usando as distribuições empíricas;
- (ii) estima os parâmetros do acoplamento como segue:

$$\hat{\vartheta} = \arg \max \sum_{t=1}^T \ln c(\hat{u}_1^t, \dots, \hat{u}_n^t; \vartheta|F).$$

onde F é um conjunto de informações.

Esse método também é consistente sob condições de regularidade (ver Nelson e Cao(1992)).

Capítulo 2

Modelo para o Portfólio

Esse capítulo tem por objetivo apresentar a forma de como será construído o modelo para um Portfólio. Para simplificar a notação o modelo será apresentado para dois ativos. Primeiramente, as séries de retornos serão modeladas por um modelo ARMA+GARCH. Sobre as inovações do modelo GARCH serão estimados os acomplamentos utilizando distribuições marginais GPD e distribuições marginais empíricas.

2.1 Modelo $\text{ARMA}(p_1, q_1) + \text{GARCH}(p_2, q_2)$

Nessa seção será apresentada apenas uma breve revisão sobre o modelo $\text{ARMA}(p_1, q_1) + \text{GARCH}(p_2, q_2)$. Os modelos ARCH foram introduzidos originalmente por Engle(1982), para modelar a inflação do Reino Unido. Em princípio, o interesse foi tentar modelar as perturbações de modelos mais complexos como os ARMA ou de regressão, com o objetivo de melhorar as estimativas e obter intervalos de confiança mais precisos. Posteriormente, esses modelos passaram a ser fortemente associados à modelagem da variância

condicional das séries financeiras.

Considerando r_{jt} como o retorno do ativo j no instante de tempo t , o modelo do tipo ARMA(p_1, q_1)+GARCH(p_2, q_2) tem a seguinte estrutura geral:

$$\begin{aligned} r_{jt} &= \mu_{jt} + \varepsilon_{jt} \\ \varepsilon_{jt} &= \sigma_{jt}\eta_{jt}, \end{aligned}$$

onde a média μ_{jt} é modelada pelo ARMA(p_1, q_1) e as perturbações ε_{jt} são modeladas pelo GARCH(p_2, q_2), ou seja:

$$\begin{aligned} \mu_{jt} &= \phi_{j0} + \sum_{i=1}^{p_1} \phi_{ji}r_{j(t-i)} - \sum_{i=1}^{q_1} \theta_{ji}\varepsilon_{j(t-i)} \\ \sigma_{jt}^2 &= \alpha_{j0} + \sum_{i=1}^{p_2} \alpha_{ji}\varepsilon_{j(t-i)}^2 + \sum_{k=1}^{q_2} \beta_{jk}\sigma_{j(t-k)}^2, \end{aligned}$$

onde $\{\varepsilon_{jt}\}$ é uma seqüência de variáveis aleatórias iid com média 0 e variância 1, $\alpha_{j0} > 0$, $\alpha_{ji} \geq 0$, $\beta_{jk} \geq 0$, e $\sum_{i=1}^{\max(p_2, q_2)} (\alpha_{ji} + \beta_{ji}) < 1$, $\forall j$.

As restrições anteriores são para garantir que $\sigma_{jt}^2 > 0$ é finito. Nesse trabalho, será considerado que ε_{jt} assume uma distribuição normal $N(0, \sigma_{jt}^2)$.

Por exemplo, o modelo AR(1)+GARCH(1,1) apresenta a seguinte estrutura:

$$\begin{aligned} r_{jt} &= \phi_{j0} + \phi_{j1}r_{j(t-1)} + \varepsilon_{jt} \\ \varepsilon_{jt} &= \sigma_{jt}\eta_{jt}, \quad \eta_{jt} \sim N(0, 1) \\ \sigma_{jt}^2 &= \alpha_j + \beta_j\sigma_{j(t-1)}^2 + \gamma_j\varepsilon_{j(t-1)}^2, \end{aligned}$$

onde $\alpha_j > 0$, $\beta_j > 0$, $\gamma_j > 0$, e $\beta_j + \gamma_j < 1$, $\forall j$.

Esses modelos serão utilizados para explicar a distribuição marginal condicional dos retornos, e a distribuição conjunta das inovações $(\eta_{1t}, \dots, \eta_{nt})$ será modelada através de acoplamentos de valores extremos utilizando distribuições marginais GPD e distribuições marginais empíricas.

2.2 VaR

O Valor em Risco (VaR) é uma estimativa de perda ao longo de um período futuro de tempo, para uma certa probabilidade. Não há regra para o tamanho do período e normalmente as instituições financeiras estimam perda para um dia. Para períodos de d dias, deve ser analisado o comportamento das variações de preços em períodos de d dias, ou seja, de retornos diários para períodos de um dia, de retornos semanais para estimar perdas ao longo dos próximos cinco dias, e assim por diante. Nesse trabalho será estimado perda para um dia, e para calcular o VaR em um portfólio com dois ativos será feito da seguinte forma:

Considere, por exemplo, o modelo AR(1)+GARCH(1,1). Então:

$$r_{jt} = \phi_{j0} + \phi_{j1}r_{j(t-1)} + \sqrt{(\alpha_j + \beta_j\sigma_{j(t-1)}^2 + \gamma_j\varepsilon_{j(t-1)}^2)}\eta_{jt}, j = 1, 2.$$

Considerando que os retornos são pequenos o retorno do portfólio é dado por $r_t \approx \sum_{j=1}^k x_{j,t}r_{j,t}$, onde $x_{j,t}$, $j = 1, \dots, k$ são os pesos dos retornos.

Assim, para o exemplo em questão, atribuindo-se o mesmo peso a cada um dos ativos, a proporção dos retornos que estarão abaixo do VaR pode ser dado da seguinte forma:

$$r_t \approx \frac{r_{1t} + r_{2t}}{2} \leq VaR$$

O valor do VaR pode ser calculado caso se conheça a distribuição conjunta de $(\eta_1, \dots, \eta_n)$. Caso não exista uma solução analítica isso pode ser feito a partir de simulação utilizando o acoplamento.

$$\eta_{1t} \leq \frac{2VaR - \phi_1 r_{1(t-1)} - \phi_2 r_{2(t-1)}}{\sqrt{(\alpha_1 + \beta_1 \sigma_{1(t-1)}^2 + \gamma_1 \varepsilon_{1(t-1)}^2)}} - \frac{\sqrt{(\alpha_2 + \beta_2 \sigma_{2(t-1)}^2 + \gamma_2 \varepsilon_{2(t-1)}^2)}}{\sqrt{(\alpha_1 + \beta_1 \sigma_{1(t-1)}^2 + \gamma_1 \varepsilon_{1(t-1)}^2)}} \eta_{2t}$$

Assim,

$$\eta_{1t} \leq a - b\eta_{2t} \tag{2.2.1}$$

onde:

$$a = \frac{2VaR - \phi_1 r_{1(t-1)} - \phi_2 r_{2(t-1)}}{\sqrt{(\alpha_1 + \beta_1 \sigma_{1(t-1)}^2 + \gamma_1 \varepsilon_{1(t-1)}^2)}},$$

e

$$b = \frac{\sqrt{(\alpha_2 + \beta_2 \sigma_{2(t-1)}^2 + \gamma_2 \varepsilon_{2(t-1)}^2)}}{\sqrt{(\alpha_1 + \beta_1 \sigma_{1(t-1)}^2 + \gamma_1 \varepsilon_{1(t-1)}^2)}}$$

Dessa forma, para calcular o VaR, basta verificar a proporção das inovações que estarão abaixo da equação 2.2.1. Essa equação irá variar de acordo com o tempo.

2.3 Distribuição das Inovações

A Distribuição Generalizada de Pareto (GPD) é uma distribuição preocupada em modelar os excessos além de um certo limiar pré-fixado (para maiores detalhes sobre a distribuição GPD, ver Apêndice). Abaixo desse limiar será utilizada a distribuição empírica, pois o interesse maior está em trabalhar nos extremos das distribuições para o cálculo do VaR, e com esse tipo de distribuição fica mais fácil para simular as inovações.

Inicialmente, será apresentado o estimador de máxima verossimilhança para estimar os parâmetros da GPD, que será utilizada para modelar a distribuição das inovações.

Considere $X_1, X_2, \dots, X_n$ uma amostra iid com função de distribuição F .

Defina $x_0 = \sup\{x : F(x) < 1\}$. Agora, defina a função de distribuição condicional de $X - u$ dado $X > u$ como sendo $F_u(y)$. Assim,

$$F_u(y) = \frac{F(u+y) - F(u)}{1 - F(u)},$$

onde $u < x_0$ e $0 < y < x_0 - u$.

A distribuição GPD é representada por:

$$W_\gamma(y) = 1 - (1 + \gamma y / \psi)^{-1/\gamma}.$$

Pickands(1974) mostrou que a GPD acima é uma boa aproximação de F_u , no sentido que

$$\lim_{u \rightarrow x_0} \sup_{0 < y < x_0 - u} |F_u(y) - W_{\gamma, \psi}(y)| = 0$$

se, e somente se, F pertence ao domínio de atração de uma das três distribuições limites da GEV. (ver Apêndice)

Fixe o limiar u e denote por N o número de excedentes a u . Represente esses excessos por $Y_1, Y_2, \dots, Y_N$, ou seja, $Y = X - u > 0$.

Utilizando o resultado de Pickands, citado anteriormente, com a expressão de F_u tem-se que:

$$1 - (1 + \gamma y / \psi)^{-1/\gamma} = \frac{F(u+y) - F(u)}{1 - F(u)}.$$

Observe que, $\frac{N}{n}$ é um estimador para $1 - F(u)$. Portanto,

$$1 - \hat{F}(u + y) = \frac{N}{n} (1 + \hat{\gamma} y / \hat{\psi})^{-1/\hat{\gamma}},$$

onde $\hat{\gamma}$ e $\hat{\psi}$ são os estimadores de máxima verossimilhança de γ e ψ , respectivamente, baseando-se somente nos excessos $Y_1, Y_2, \dots, Y_N$.

Dessa maneira, os estimadores de máxima verossimilhança minimizam a função de log-verossimilhança:

$$L(\gamma, \psi; y_1, y_2, \dots, y_N) = -N \log(\psi) - \left(\frac{1}{\gamma}\right) \sum_{i=1}^N \log(1 + \frac{\gamma}{\psi} y_i).$$

Maiores detalhes sobre as propriedades desses estimadores podem ser encontrados em Smith(1987).

2.4 Estimação do Acoplamento de Valor Extremo

O procedimento adotado para estimar os parâmetros dos acoplamentos, será apresentado para o caso bivariado.

Considere r_{1t} e r_{2t} as duas séries de retornos financeiros. Primeiramente, será aplicado o modelo $\text{ARMA}(p_1, q_1) + \text{GARCH}(p_2, q_2)$ para explicar a distribuição condicional dos retornos, conforme descrito anteriormente.

Depois de estimados os parâmetros desses modelos, calculam-se os resíduos padronizados $\eta_{jt} = \frac{\varepsilon_{jt}}{\sigma_{jt}}$ para $j = 1, 2$. Sobre esses resíduos η_{jt} , ($j = 1, 2$) são calculados todos os excessos considerando um limiar de 10% sobre as perdas para cada um dos ativos.

A partir dos excessos encontrados para as duas séries, estima-se os parâmetros das GPD's conforme descrito em (2.3).

O restante da distribuição é estimada utilizando-se distribuições marginais empíricas.

A figura 2.1 mostra a situação em estudo para a estimação dos acoplamentos.

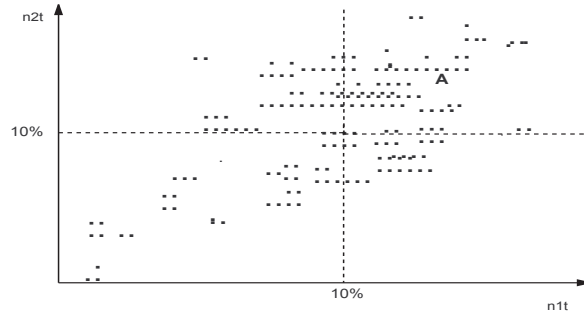


Figura 2.1: Gráfico Ilustrativo para a estimação dos acoplamentos

A função de verossimilhança para o acoplamento será da seguinte forma:

$$l(\theta) = \{P_\theta[(\eta_{1t}, \eta_{2t}) \in A]\}^{n_1} \prod_{A^c} c(F_1(\eta_{1t}), F_2(\eta_{2t}); \theta)$$

onde:

F_1 e F_2 são estimadas via GPD e n_1 representa o número de pontos que pertencem à região A.

Agora, surge o problema de como estimar os parâmetros do acoplamento. Como o interesse é trabalhar somente com valores extremos, não interessa como os pontos se comportam dentro da região A apresentada no gráfico 2.1, pois quando o VaR (5% ou 1%) do portfólio forem calculados, estes não pertencerão à região A, ou seja, para o cálculo do VaR será utilizada apenas a região na qual as distribuições marginais foram estimadas pelas GPD's. Dessa forma, tem-se uma classe maior de acoplamentos que podem ser utilizados para o cálculo do VaR, pois o interesse maior é no extremo da distribuição. Sendo assim, o problema de assimetria de dependência não influencia na estimação dos acoplamentos.

Para se estimar a distribuição referente aos η_{it} 's que pertencem à região A, foi utilizado o Método de simulação de Monte Carlo, no qual para os dados simulados foram contados quantos pertenciam à região A. Fora da região A foi utilizada a distribuição GPD para estimar as marginais e dentro dessa região foi utilizada a distribuição empírica. Dessa forma, foi estimada a probabilidade:

$$P_\theta[(\eta_{1t}, \eta_{2t}) \in A] = \frac{\#A}{total}$$

onde:

$\sharp A$ representa a quantidade de dados simulados que pertencem à A ; e $total$ representa o número total de pontos simulados.

Uma vez estimada a probabilidade anterior e as GPD's para os excessos, agora a verossimilhança (1.5.1) do acoplamento é conhecida. Pelo método IFM, como os parâmetros das marginais já estão estimados, basta estimar os parâmetros do acoplamento via método de máxima verossimilhança.

O outro método de estimação utilizado foi utilizando acoplamentos de valores extremos sobre as inovações, conforme visto no início dessa seção, cujas distribuições marginais são todas empíricas. Pelo método CML, como as distribuições marginais foram estimadas pela empírica, basta estimar os parâmetros do acoplamento via método de máxima verossimilhança.

Capítulo 3

Aplicação

Neste capítulo será realizada uma aplicação de alguns tipos de acoplamentos de valores extremos, comparando-os em relação ao acoplamento empírico, utilizando as séries de retornos diários r_1 e r_2 do IBOVESPA e Merval, respectivamente. Os dados do IBOVESPA foram coletados no site www.ipea.gov.br e os dados do Merval em www.bolsar.com. Na seção 3.1 será apresentado como as distribuições marginais serão modeladas, na seção 3.2 a estimação e diagnóstico dos acoplamentos utilizados serão feitos e na seção 3.3 o modelo construído será utilizado para calcular o VaR.

3.1 Modelagem das distribuições marginais

O IBOVESPA corresponde ao índice de ações do mercado brasileiro enquanto que o Merval corresponde ao índice de ações do mercado argentino.

Os dados foram tomados no período de 04/03/1997 à 07/11/2000 totalizando 900 observações. Para análise dos dados, os dias em que a bolsa não funcionou no Brasil e funcionou na Argentina (e vice-versa) esse tipo

de dado não foi considerado. Foram considerados somente os dias em que as duas bolsas estavam funcionando nos dois países. Dessa forma foram descartados 49 dias onde nenhum desses dias apresentou grandes perdas. No período analisado houve a crise do Leste Asiático (julho de 1997), passando pela moratória Russa (outubro de 1998), pela mudança do regime cambial brasileiro (janeiro de 1999) e pela crise na bolsa Nasdaq (abril de 2000).

A figura 3.1 apresenta os gráficos das séries financeiras, a figura 3.2 apresenta os retornos e a figura 3.3 apresenta o diagrama de dispersão dos retornos do IBOVESPA e MERVAL.

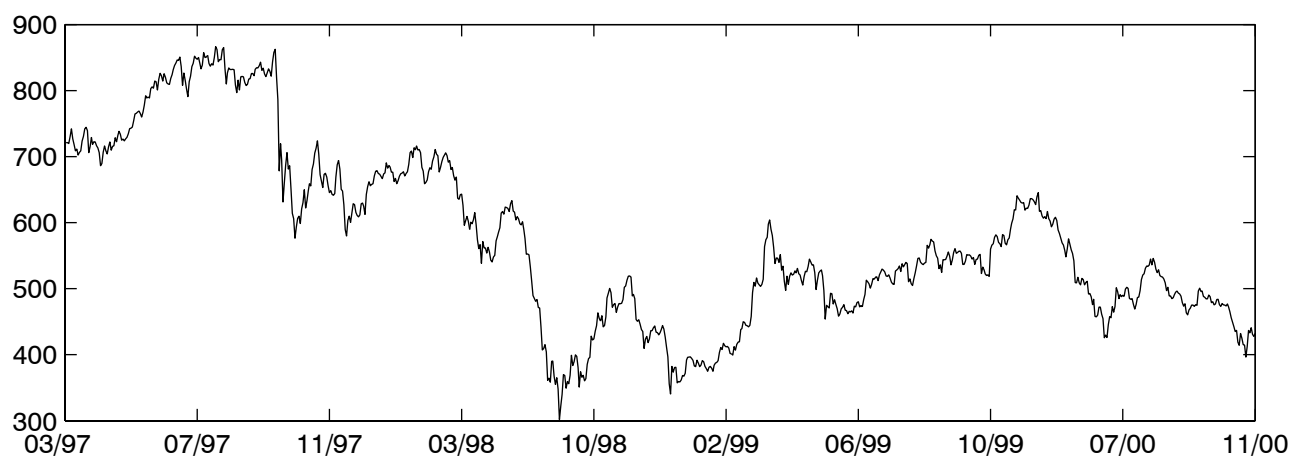
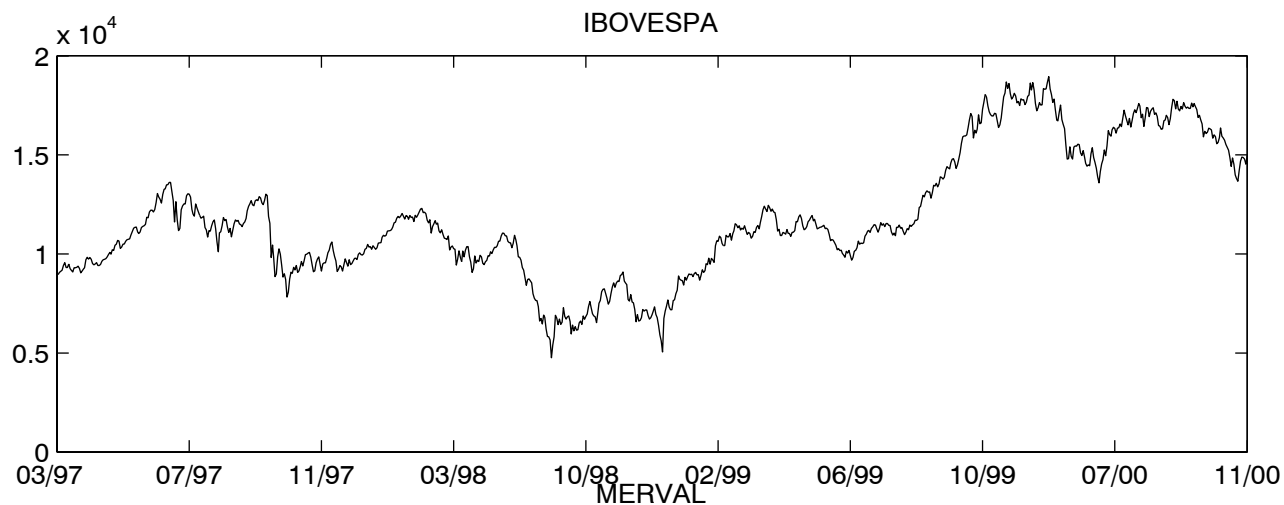


Figura 3.1: Séries Financeiras dos índices do BOVESPA e Merval durante o período de 04/03/1997 - 07/11/2000

A tabela 3.1 apresenta uma análise descritiva univariada das duas séries de retornos.

As estatísticas descritivas indicam, para cada série, que os retornos têm

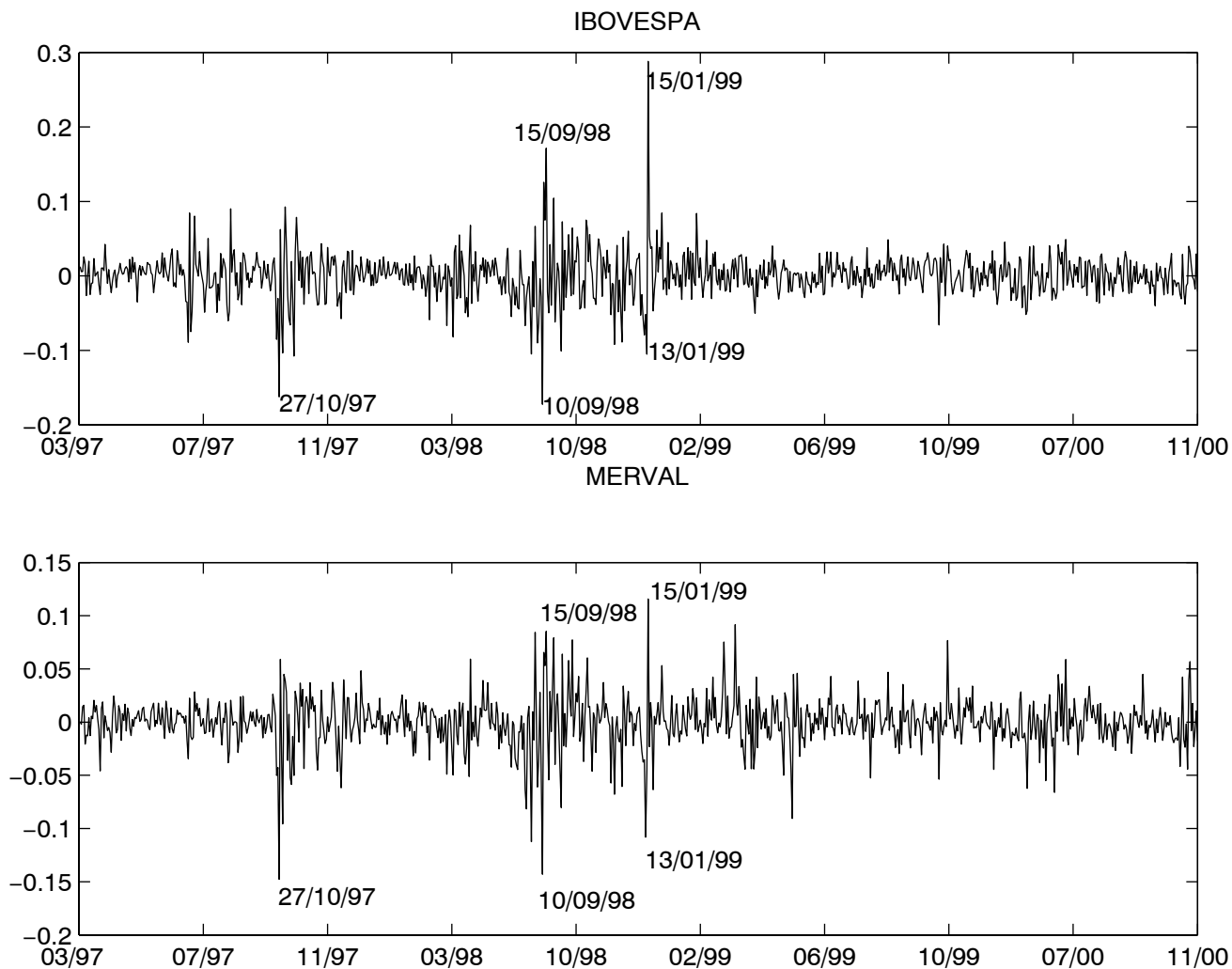


Figura 3.2: Retornos IBOVESPA e Merval durante o período de 04/03/1997 - 07/11/2000

excesso de curtose e distribuições assimétricas.

Pelo que foi apresentado fica confirmado que as duas série apresentam distribuições com caudas pesadas descartando a possibilidade de distribuição normal.

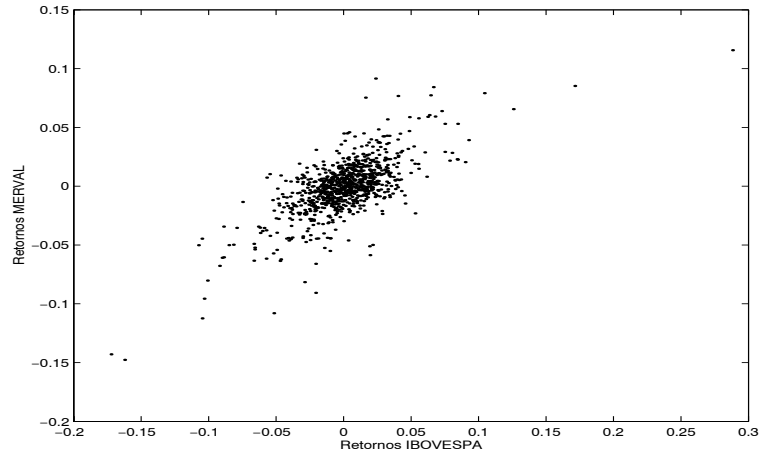


Figura 3.3: Diagrama de dispersão dos retornos Merval versus IBOVESPA

Estatísticas	IBOVESPA	Merval
média	0,00055	-0,00059
mediana	0,00146	0,00025
erro padrão	0,00101	0,00079
desvio-padrão	0,03042	0,02364
excesso de curtose	12,58184	5,92814
assimetria	0,59857	-0,62118
mínimo	-0,17229	-0,14765
máximo	0,28818	0,11565

Tabela 3.1: Estatísticas Descritivas das Séries de Retornos

Observação 3.1.1. A partir daqui, para a estimação dos parâmetros, os retornos foram multiplicados por -1.

Os gráficos das auto-correlações dos retornos apresentado na figura 3.4 para verificar a existência de correlação no nível. O teste de Lyung-Box foi não-significativo, com 10% de significância, indicando não ser necessário o emprego dos modelos ARMA. Os valores-p associados aos retornos do IBOVESPA e Merval foram 0,13 e 0,17, respectivamente. Esse teste, com

10% de significância, não rejeitou a hipótese de não correlação, nos primeiros *lags*. Os gráficos de auto-correlações dos quadrados dos retornos dados na figura 3.5 sugerem o uso do modelo GARCH(1,1).

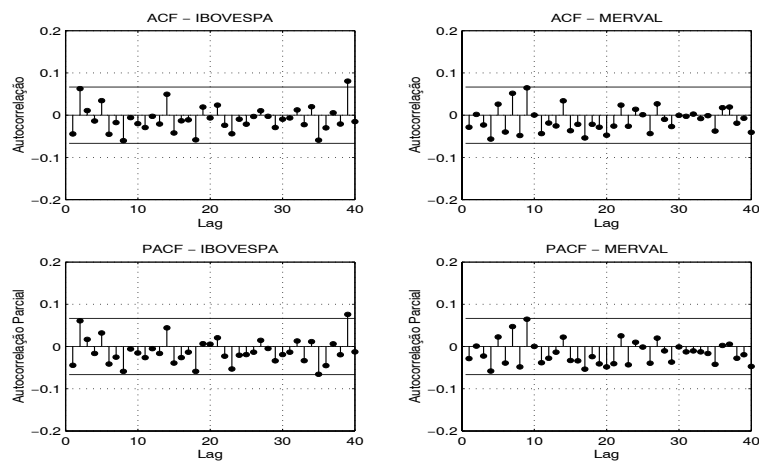


Figura 3.4: Auto-correlações e auto-correlações parciais dos retornos

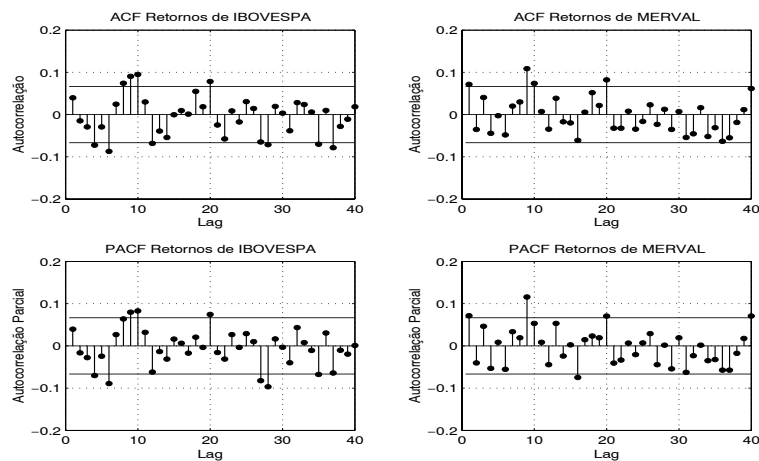


Figura 3.5: Auto-correlações e auto-correlações parciais dos quadrados dos retornos

Os parâmetros estimados do modelo GARCH(1,1) para as duas séries são apresentados na tabela 3.2.

Parâmetros	IBOVESPA	Desvio-padrão	MERVAL	Desvio-padrão
α	$3,7513 \cdot 10^{-5}$	$8,5937 \cdot 10^{-6}$	$2,8423 \cdot 10^{-5}$	$5,6956 \cdot 10^{-6}$
β	0,2117	0,0226	0,1621	0,0147
γ	0,7539	0,0269	0,7952	0,0198

Tabela 3.2: Estimativas dos Parâmetros para o modelo GARCH(1,1) e respectivas estimativas dos desvios-padrões.

A figura 3.6 apresenta os gráficos das auto-correlações e auto-correlações parciais do quadrado das inovações, para verificar se, de fato, o modelo GARCH(1,1) explicou toda a dependência dos quadrados dos retornos.

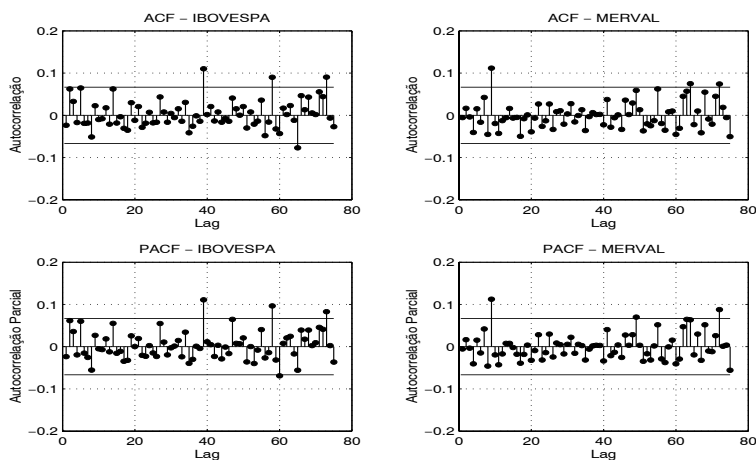


Figura 3.6: Auto-correlações dos quadrados das inovações

Pela análise das auto-correlações e auto-correlações parciais do quadrado

das inovações, verifica-se que o modelo GARCH(1,1) explicou toda a dependência dos quadrados dos retornos.

A figura 3.7 apresenta o gráfico dos resíduos padronizados das duas séries em questão:

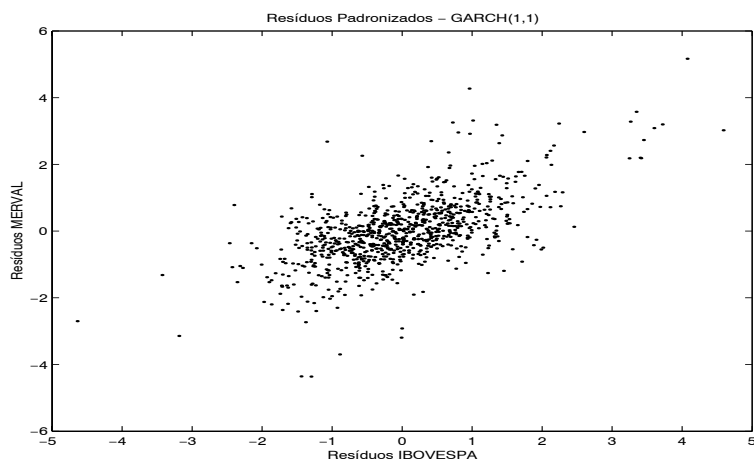


Figura 3.7: Resíduos Padronizados - GARCH(1,1)

A partir dos resíduos estimados, considerando um limiar de 10% para determinação dos excessos, os parâmetros estimados da GPD para as séries são dados na tabela 3.3.

Parâmetros GPD	IBOVESPA	MERVAL
limiar	1,2171	1,1548
γ	0,1499 (0,0143)	0,0150 (0,0086)
ψ	0,5417 (0,0098)	0,8118 (0,0145)

Tabela 3.3: Estimativas dos Parâmetros da GPD e entre parênteses estimativas dos desvios-padrões.

3.2 Estimação do Acoplamento e Diagnóstico

Na seção anterior as distribuições marginais foram estimadas. Agora, utilizando a metodologia descrita no capítulo 2 foi possível obter as estimativas dos parâmetros dos acoplamentos que são apresentadas na tabela 3.4.

Acoplamentos (Marginais)	Parâmetros
Gumbel (GPD)	0,5304 (0,0478)
Gumbel (Empírica)	0,6008 (0,0467)
Gumbel A (GPD)	0,8235 (0,0526)
Gumbel A (Empírica)	0,9579 (0,0183)

Tabela 3.4: Estimativa dos Parâmetros das Acoplamentos e entre parênteses estimativas dos desvios-padrões.

Com as estimativas desses parâmetros, foram construídos os gráficos das curvas de nível comparando os acoplamentos estimados com o acoplamento empírico considerando somente o extremo superior, onde se encontram as maiores observações referentes a 20% dos dados.

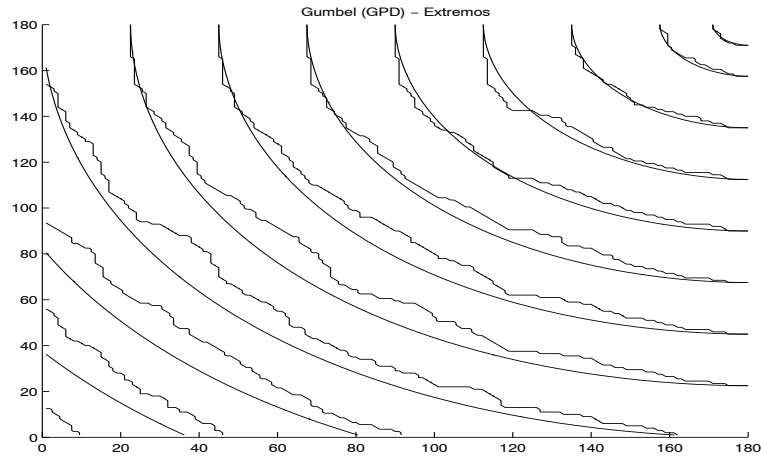


Figura 3.8: Curvas de Nível - Acoplamento de Gumbel(GPD) - dados extremos, onde as perdas são dadas a partir do extremo superior direito da figura (1%, 5%, 10%, 15%, e assim por diante)

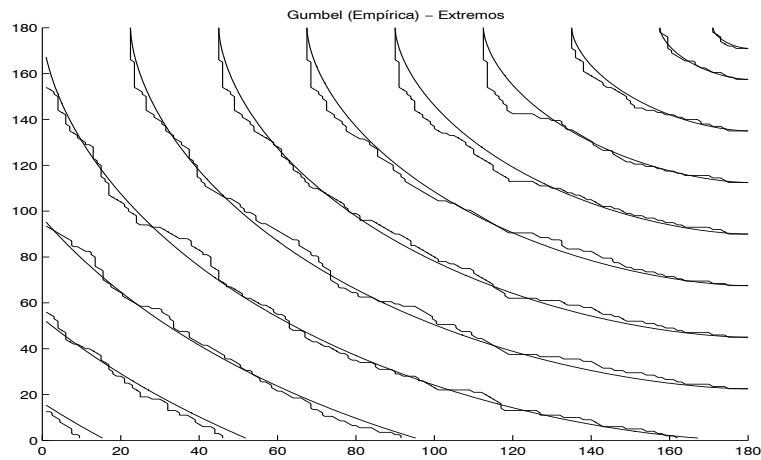


Figura 3.9: Curvas de Nível - Acoplamento de Gumbel(Empírica) - dados extremos, onde as perdas são dadas a partir do extremo superior direito da figura (1%, 5%, 10%, 15%, e assim por diante)

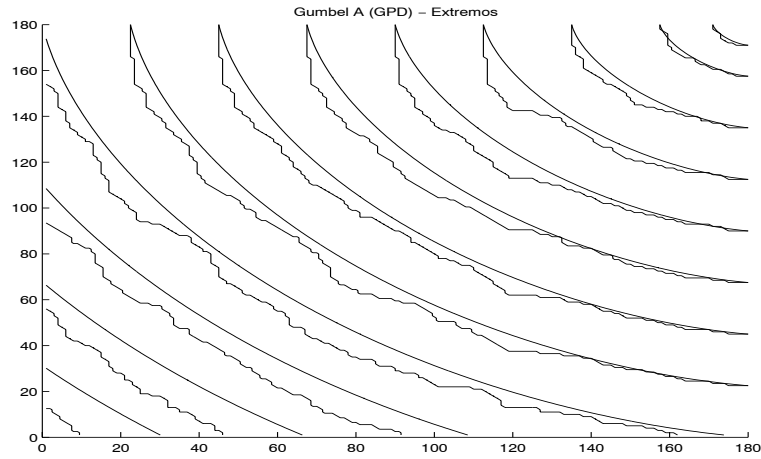


Figura 3.10: Curvas de Nível - Acoplamento de Gumbel A(GPD) - dados extremos, onde as perdas são dadas a partir do extremo superior direito da figura (1%, 5%, 10%, 15%, e assim por diante)

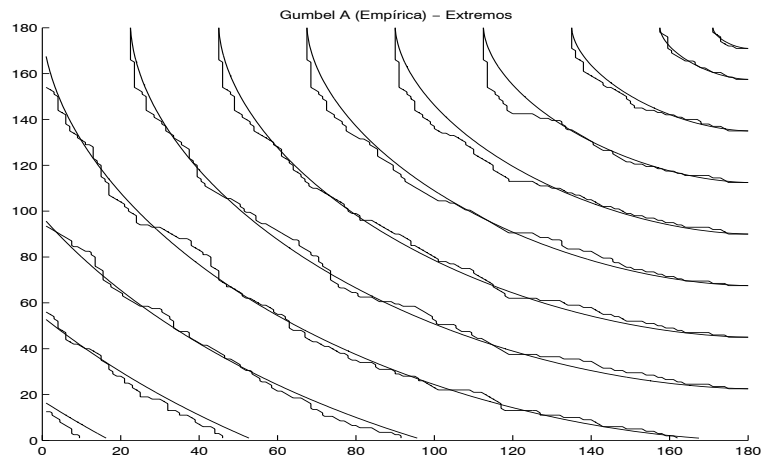


Figura 3.11: Curvas de Nível - Acoplamento de Gumbel A(Empírica) - dados extremos, onde as perdas são dadas a partir do extremo superior direito da figura (1%, 5%, 10%, 15%, e assim por diante)

Como pode-se observar pelos gráficos anteriores, quando é utilizado a GPD para estimar as distribuições marginais, verifica-se que os acoplamentos ficam bem ajustados nos extremos, enquanto que, no centro da distribuição o ajuste fica ruim. Confirmando, que a distribuição GPD está interessada apenas nos extremos da distribuição.

Como o interesse principal é trabalhar com os extremos, foram calculadas as distâncias dos acoplamentos em relação ao acoplamento empírico utilizando somente esses extremos, lembrando que os dados foram multiplicados por -1. Lembrando que, a medida de distância utilizada foi:

$$\bar{d}_n(\hat{C}, C_k) = [\sum_{t_1=1}^T \cdots \sum_{t_n=1}^T (\hat{C}(\frac{t_1}{T}, \dots, \frac{t_n}{T}) - C_k(\frac{t_1}{T}, \dots, \frac{t_n}{T}))^2]^{1/2}$$

Onde: C_k corresponde ao acoplamento em consideração e $\hat{C}$ corresponde ao acoplamento empírico.

As distâncias encontradas estão na tabela 3.5.

Acoplamentos (Marginais)	20%	10%	5%	1%
Gumbel (GPD)	1,0500	0,2452	0,0899	0,0041
Gumbel (Empírica)	0,3806	0,2218	0,0713	0,0056
Gumbel A (GPD)	1,1083	0,4365	0,1559	0,0079
Gumbel A (Empírica)	0,3856	0,2283	0,0738	0,0057

Tabela 3.5: Distâncias entre o acoplamento empírico e os acoplamentos estimados para 20%, 10%, 5% e 1% das maiores perdas

Utilizando-se acoplamentos com distribuições marginais empíricas estes apresentam menores distâncias em relação ao acoplamento empírico, quando trabalha-se com 20%, 10% e 5% das maiores perdas. Acoplamentos com

distribuições marginais GPD se ajustam melhor que os acoplamentos com marginais empíricas para 1% das maiores perdas quando utiliza-se o acoplamento de Gumbel.

3.3 Cálculo do VaR

O VaR pode ser estimado através de simulação. Para verificar se os valores estimados condizem com os dados pode-se utilizar a seguinte metodologia: um modelo é ajustado às primeiras n_0 observações. Através de simulação das inovações (número de simulação igual a 2000) é estimado o VaR de um dia a vários níveis de confiança onde verifica se a perda excede o valor do VaR estimado. Observado o próximo valor dos retornos aplica-se novamente os passos anteriores, tomando-se sempre uma janela de n_0 observações para estimar o próximo VaR. Essa operação é repetida até o final da série. Feito isso, pode-se comparar a proporção de valores fora dos limites estimados. Entretanto, como esse método é muito trabalhoso e considerando que não se espera muita mudança ao mudar apenas poucas observações, o modelo foi estimado apenas a cada 50 observações, tomando-se $n_0 = 450$. Dessa forma, o modelo estima apenas 9 vezes. Verifique que dessa forma, o mesmo conjunto de valores simulados das inovações podem ser utilizados 50 vezes já que o modelo é o mesmo. Entretanto, o VaR estimado é modificado a cada passo pois a média e variância condicional são alteradas.

A figura 3.12 apresenta um gráfico do VaR (5%) estimado para uma visualização dos pontos que estão abaixo dessa estimativa. Na abscissa do

gráfico, encontram-se as projeções (denotadas por x) dos retornos que estão abaixo do VaR estimado.

A proporção de valores observados fora dos limites são apresentados na tabela 3.6.

	Marginais					
	Empirica			GPD		
	10%	5%	1%	10%	5%	1%
Acoplamento						
Gumbel	0,1210 (54)	0,0798 (36)	0,0254 (11)	0,0811 (36)	0,0427 (19)	0,0075 (4)
Gumbel A	0,1982 (89)	0,1225 (55)	0,0270 (12)	0,1314 (59)	0,0724 (32)	0,025 (11)

Tabela 3.6: Proporção de observações com perdas do Portfólio maiores que o VaR estimado para $\alpha = 10\%, 5\%$ e 1% . Os números entre parênteses correspondem a quantidade de observações.

Conforme verifica-se na tabela 3.6, quando utiliza-se marginais empíricas, o risco fica subestimado utilizando-se os dois acoplamentos em questão. Estimando as marginais via GPD, obtem-se melhores resultados para as estimativas do VaR. Observa-se, ainda, que o acoplamento de Gumbel apresenta, para esse caso, uma melhor performance que o acoplamento Gumbel A.

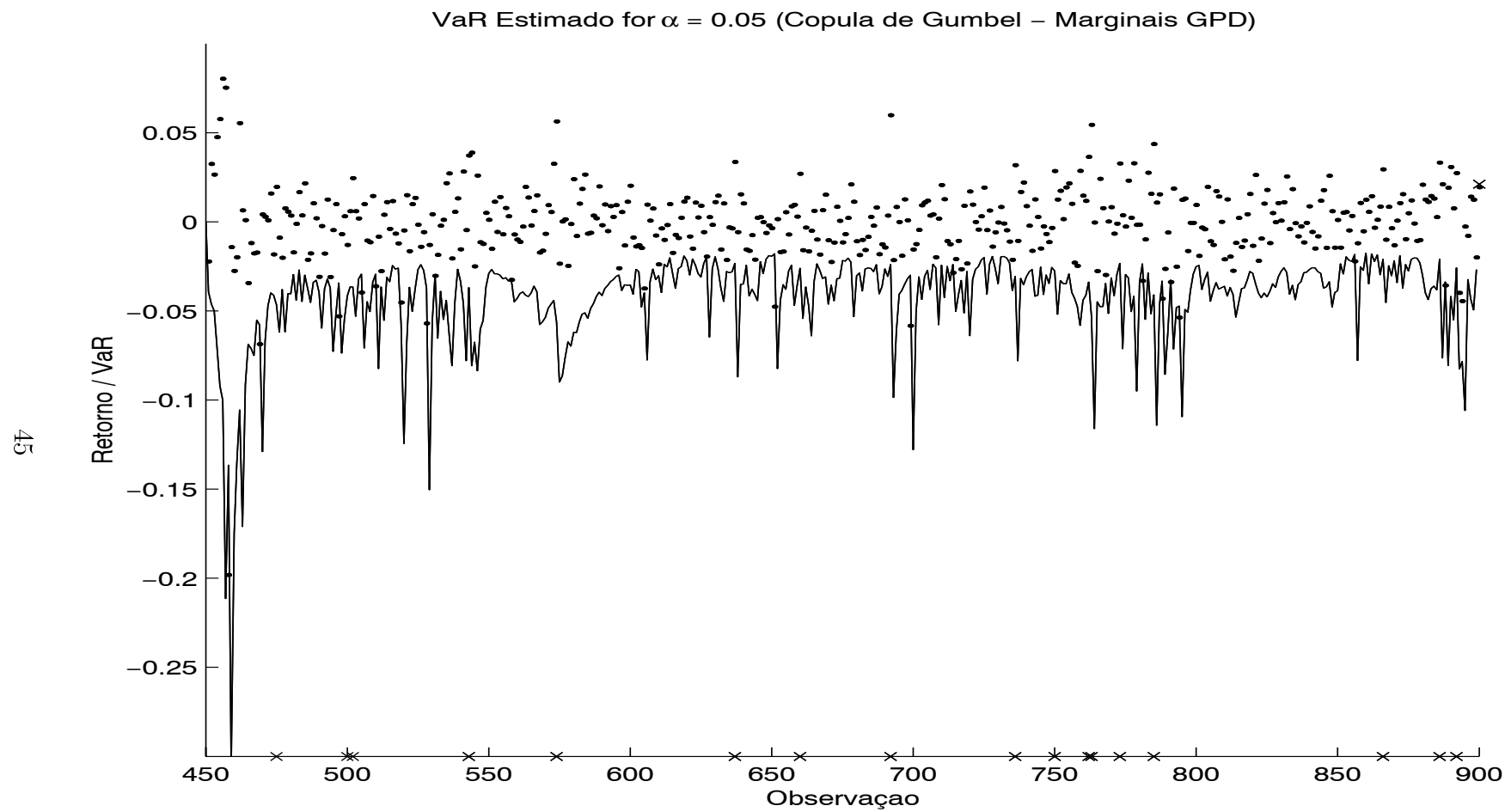


Figura 3.12: Gráfico do VaR 5% estimado - Gumbel(GPD)

Capítulo 4

Conclusões

A principal proposta desse trabalho, foi aplicar as teorias de acoplamentos e de valores extremos para calcular o VaR num portfólio. O método de estimação dos acoplamentos proposto proporcionou bons resultados quando foi utilizada a distribuição GPD para estimativa das marginais quando comparada com a utilização das marginais empíricas. O método mostrou, também, que a estimação claramente enfatiza o ajuste na cauda da distribuição. Esse procedimento aumenta a quantidade de funções de acoplamento que podem ser usadas não sendo, necessariamente, acoplamentos com dependência assimétrica. Foi utilizado o software Matlab 6.0 para execução dos programas e construção dos gráficos aqui apresentados. O grande problema que surgiu nesse trabalho, foi o tempo computacional gasto para execução das rotinas que foi muito alto. Para pesquisas futuras, a extensão desses conceitos para um portfólio com três ou mais ativos e a utilização de outros tipos de acoplamentos serão de grande utilidade prática para o cálculo do VaR.

Apêndice A

Teoria de Valores Extremos

As caracterizações dos domínios de atração das distribuições de valores extremos são úteis, pois para se trabalhar com as distribuições GPD a função de distribuição F de uma variável aleatória X , tem que pertencer ao domínio de atração de uma das três distribuições de valores extremos.

A Teoria de Valores Extremos está preocupada com a análise estatística de eventos raros ou extremos, ou seja, com a análise de valores máximos ou mínimos de uma variável aleatória, estatísticas de ordem e valores que excedem um certo limite. Com a sua utilização pode-se estimar probabilidades com razoável precisão nos limites encontrados dentro da amostra, e mesmo além deles. Assim, modelar valores extremos significa ajustar apenas a cauda da distribuição dos dados originais, ressaltando o caráter pouco informativo que se pode ter no interior da distribuição para as ocorrências de baixíssima frequência.

A.1 Caso Univariado

Esta seção tem por objetivo mostrar os principais resultados da Teoria de Valores Extremos univariada, resultados esses que serão, posteriormente, estendidos para o caso multivariado.

Considere uma seqüência de variáveis aleatórias iid $\{X_1, X_2, \dots, X_n\}$ com uma função de distribuição comum F . Defina M_n e L_n como o máximo e mínimo da amostra, respectivamente. Ou seja:

$$\begin{aligned} M_n &= \max\{X_1, X_2, \dots, X_n\}, \\ L_n &= \min\{X_1, X_2, \dots, X_n\}. \end{aligned}$$

O objetivo é estudar o comportamento assintótico de M_n e L_n . Todos os resultados serão referentes somente ao máximo, pois:

$$\max\{X_1, X_2, \dots, X_n\} = -\min\{-X_1, -X_2, \dots, -X_n\}.$$

Assim, a distribuição do máximo é determinada da seguinte forma:

$$P[M_n \leq x] = P[X_1 \leq x, X_2 \leq x, \dots, X_n \leq x] = F^n(x).$$

Defina $x_0 = \sup\{x : F(x) < 1\}$. Agora, observe que $F^n(x) \rightarrow 0, \forall x < x_0$ e $F^n \rightarrow 1, \forall x \geq x_0$, conseqüentemente, não existe uma distribuição limite não-degenerada para o máximo, a menos que o mesmo seja normalizado.

Definição A.1.1. Se existem constantes $a_n > 0$, $b_n \in \mathbb{R}$, de modo que

$$\lim_{n \rightarrow \infty} F^n(a_n x + b_n) = H(x)$$

para alguma função de distribuição não-degenerada H , então H é conhecida como uma *distribuição do valor extremo do máximo* e F é dita pertencer ao *domínio de atração do máximo* de H , ou seja, $F \in MDA(H)$.

Definição A.1.2. Duas funções de distribuição $H(x)$ e $G(x)$ são do mesmo tipo, se existem números reais a e b , com $b > 0$, tais que $G(x) = H(a + bx)$, para todo $x \in \mathbb{R}$.

O resultado central da teoria de valores extremos é dado a seguir:

Teorema A.1.1. *Seja $\{X_n\}$ uma seqüência de variáveis aleatórias iid. Se existem constantes $a_n > 0$, $b_n \in \mathbb{R}$ e alguma função distribuição não-degenerada H de modo que*

$$\frac{M_n - b_n}{a_n} \rightarrow^d H$$

então H pertence a um dos três tipos de distribuição:

Fréchet: ($\alpha > 0$)

$$\Phi_\alpha(x) = \begin{cases} 0 & , x \leq 0 \\ \exp(-x^{-\alpha}) & , x > 0 \end{cases}$$

Weibull: ($\alpha > 0$)

$$\Psi_\alpha(x) = \begin{cases} \exp(-(-x)^\alpha) & , x \leq 0 \\ 1 & , x > 0 \end{cases}$$

Gumbel:

$$\Lambda(x) = \exp(-e^{-x}) \quad , x \in \mathbb{R}$$

Definição A.1.3. As funções distribuições Φ_α , Ψ_α e Λ são chamadas de *distribuições de valor extremo*.

A.1.1 Distribuição Generalizada do Valor Extremo (GEV)

Uma dúvida que pode surgir até aqui, é como identificar a distribuição correta do extremo de uma determinada amostra, entre as três possíveis. Após uma mudança paramétrica locação/escala pode-se representar os três casos padrão das distribuições de valor extremo numa única família paramétrica

que engloba as três distribuições. Esta família é conhecida como Distribuição Generalizada do Valor Extremo (GEV) e é definida como:

$$H_{\xi}(x) = \begin{cases} \exp\{-(1 + \xi x)^{-\frac{1}{\xi}}\} & \text{se } \xi \neq 0 \\ \exp\{-\exp(-x)\} & \text{se } \xi = 0 \end{cases}$$

onde: $1 + \xi x > 0$.

O parâmetro ξ pode ser interpretado como um parâmetro de cauda e a ligação entre as três distribuições padrão com a GEV é dada da seguinte forma:

$\xi = \alpha^{-1} > 0$ corresponde a distribuição de Fréchet Φ_{α} ,

$\xi = 0$ corresponde a distribuição de Gumbel Λ ,

$\xi = -\alpha^{-1} < 0$ corresponde a distribuição de Weibull Ψ_{α} .

É conhecida na literatura a relação entre Gumbel, Fréchet e Weibull que é dada por:

$$X \sim Frechet(\alpha) \Leftrightarrow \ln(X^{\alpha}) \sim Gumbel \Leftrightarrow -X^{-1} \sim Weibull(\alpha)$$

A seguir, tem-se algumas definições e proposição que serão úteis para caracterização de domínios de atração.

Definição A.1.4. Seja X uma variável aleatória com função de distribuição F . Então $\overline{F}(x) = P[X > x] = 1 - F(x)$ é chamada de função de sobrevivência de F .

Proposição A.1.2. A função de distribuição F pertence ao domínio de atração do máximo da distribuição de valor extremo H se, e somente se, existem constantes $a_n > 0$ e $b_n \in \mathbb{R}$ tais que:

$$\lim_{n \rightarrow \infty} n\overline{F}(a_n x + b_n) = -\log H(x), x \in \mathbb{R}.$$

Quando $H(x) = 0$ o limite é interpretado como ∞ .

Definição A.1.5. Uma função medida $U : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ é dita de *variação regular no infinito* com índice $\rho \geq 0$, se para $x > 0$,

$$\lim_{t \rightarrow \infty} \frac{U(tx)}{U(t)} = x^\rho.$$

ρ é definido como coeficiente de variação e se $\rho = 0$, U é definida como uma *função de baixa variação* e será representada por $L(x)$.

Exemplo A.1.3. *Funções de baixa variação: $\log(1+x)$, $\log \log(e+x)$*

Observação A.1.1. Para uma possível escolha das constantes normalizadoras ver, por exemplo, Embrechts, Klüppelberg e Mikosch(1997).

Nas três próximas subseções serão apresentadas algumas caracterizações dos domínios de atração de Fréchet, Weibull e Gumbel. As demonstrações dos resultados podem ser encontradas em Embrechts, Klüppelberg e Mikosch (1997).

A.1.2 Domínio de Atração de Fréchet

Teorema A.1.4. *A função de distribuição F pertence ao domínio de atração de Φ_α , $\alpha > 0$ se, e somente se:*

$$\overline{F}(x) = x^\alpha L(x)$$

para alguma função de baixa variação $L(x)$.

Teorema A.1.5. *(Gnedenko) Uma função de distribuição F pertence ao domínio de atração do máximo de Φ_α se, e somente se,*

$$\lim_{t \rightarrow \infty} \frac{\overline{F}(tx)}{\overline{F}(t)} = x^{-\alpha}, x > 0.$$

Exemplo A.1.6. *(Distribuição de Pareto): Para a distribuição de Pareto tem-se que:*

$$\overline{F}(x) = \left(\frac{k}{k+x}\right)^\alpha.$$

Observe que:

$$\lim_{t \rightarrow \infty} \frac{\overline{F}(tx)}{\overline{F}(t)} = x^{-\alpha} \Rightarrow F \in MDA(\Phi_\alpha).$$

A.1.3 Domínio de Atração de Weibull

Dar condições sobre F para pertencer ao domínio de atração de Weibull é similar ao caso anterior. Uma primeira diferença que se encontra entre funções pertencentes ao domínio de atração de Fréchet e pertencentes ao domínio de atração de Weibull é que este último tem ponto final a direita finito. ($x_0 < \infty$)

Teorema A.1.7. *A função de distribuição F pertence ao domínio de atração do máximo de Ψ_α , $\alpha > 0$ se, e somente se, $x_0 < \infty$ e $F(x_0 - x^{-1}) = x^{-\alpha}L(x)$ para alguma função de baixa variação L .*

Teorema A.1.8. *(Gnedenko): A função de distribuição F pertence ao domínio de atração do máximo de Ψ_α se, e somente se,*

$$\lim_{t \rightarrow 0^+} \frac{\overline{F}(x_0 + tx)}{\overline{F}(x_0 - t)} = (-x)^\alpha, \quad x < 0.$$

Exemplo A.1.9. *(Distribuição Uniforme $(0,1)$)*

Neste caso, tem-se que $x_0 = 1$ e para $x \in (0, 1)$, $F(x) = x$ e $\overline{F}(1 - x^{-1}) = x^{-1}$. Portanto, $F \in MDA(\Psi_1)$.

A.1.4 Domínio de Atração de Gumbel

Foi visto que o domínio de atração de Fréchet contém somente distribuições com caudas pesadas e o de Weibull contém somente funções de distribuições com ponto final direito finito. Agora, o domínio de atração de Gumbel consiste de uma grande variedade de distribuições, permitindo que o ponto direito final seja finito ou infinito.

Neste caso, uma ferramenta útil é representada pelas funções de von Mises definidas abaixo:

Definição A.1.6. Seja F uma função de distribuição com ponto final direito $x_0 \leq \infty$. Suponha que exista algum $z < x_0$ de modo que F tenha a seguinte representação:

$$\overline{F}(x) = c \cdot \exp\left(-\int_z^x \frac{1}{a(t)} dt\right), z < x < x_0,$$

onde c é uma constante positiva, $a(\cdot)$ é uma função positiva e absolutamente contínua com densidade a' e $\lim_{x \rightarrow x_0^-} a'(x) = 0$. Então, F é denominada de *função de von Mises* e a função $a(\cdot)$ de função auxiliar de F .

Proposição A.1.10. *Toda função de von Mises pertence ao domínio de atração de Gumbel.*

Exemplo A.1.11. *(Distribuição Exponencial)*

Para a distribuição exponencial tem-se que $\overline{F}(x) = \exp(-\lambda x)$ para $x \geq 0$ e $\lambda > 0$. Portanto, F é uma função de von Mises com função auxiliar $a(x) = \lambda^{-1}$, conseqüentemente, F pertence ao domínio de atração de Gumbel.

A.1.5 Estimador de Hill para ξ

A partir de agora, ao invés de se assumir que os máximos provém de uma GEV, pode-se relaxar essa hipótese para que os máximos observados sejam iid e provenientes de uma distribuição F que pertença ao domínio de atração do máximo de H_ξ .

Assim, surge a seguinte questão: "Como estimar $\xi = \frac{1}{\alpha} > 0$?"

Para responder tal questão, suponha uma amostra $X_1, \dots, X_n$ de variáveis aleatórias iid cuja distribuição F pertença ao domínio de atração de Fréchet. Nesse caso, como visto no teorema (A.1.4), segue que F é da forma $F(x) = x^{-\alpha} L(x)$, onde $x > 0$ e $L(x)$ uma função de baixa variação.

Supondo a função $L(x) = c$, onde c é uma constante, o estimador de Hill pode ser obtido como o estimador de máxima verossimilhança, considerando

a função de distribuição $F(x)$ válida para um domínio onde $x > u > 0$, com $u = X_{k,n}$, onde $X_{k,n}$ é uma estatística de ordem k . Os estimadores de máxima verossimilhança para α e c , condicionados a um valor de k , e para uma amostra de tamanho n , serão dados por:

$$\hat{\alpha}_{k,n} = \left(\frac{1}{k} \sum_{j=1}^k \ln X_{j,n} - \ln X_{k,n} \right)^{-1} \text{ e } \hat{c} = \frac{k}{n} (X_{k,n})^{\hat{\alpha}_{k,n}}.$$

É possível obter consistência fraca quando tem-se uma seqüência de variáveis aleatórias iid. Para $k \rightarrow \infty$ e $\frac{k}{n} \rightarrow 0$, $n \rightarrow \infty$ tem-se que $\hat{\alpha}_{k,n} \rightarrow \alpha$ em probabilidade. Sob restrições adicionais, é possível obter consistência forte (convergência quase-certa), e ainda normalidade assintótica.

Conforme visto, o estimador de Hill depende do valor de k , a última estatística de ordem utilizada. É a partir do valor desse k onde inicia-se a cauda da distribuição, a partir de onde a teoria de valor extremo torna-se válida para a descrição da cauda da distribuição.

Na literatura, não existe nenhuma forma definida para a escolha adequada de k . A estimativa do índice de cauda é altamente sensível à escolha de k , principalmente em amostras consideradas pequenas, uma vez que a variância do estimador cai com o aumento do tamanho da amostra, e o conseqüente aumento de k . Da mesma forma, é possível mostrar a existência de um vício crescente com a escolha de k . Assim, ao aumentar-se o valor de k , simultaneamente, aumenta-se o vício e reduz-se a variância do estimador de Hill, indicando que deve existir um ponto ótimo k^* entre vício e variância.

Uma maneira para encontrar o valor desse k^* ótimo seria pelo método

gráfico, o qual consiste em se traçar o gráfico conhecido como Hill-plot e procurar a região mais próxima das caudas onde o valor do índice de caudas estimado se estabiliza. O gráfico é constituído dos pares $\{(k, \hat{\alpha}_{k,n}) : k = 1, \dots, n\}$. A idéia é que à medida que um k maior é escolhido, avança-se em direção ao interior da distribuição e menos peso é dado às observações da cauda, o real interesse da modelagem. Entretanto, um valor pequeno de k incorpora muito poucas observações extremas e produz um estimador com alta variância.

Em Danielsson e de Vries(1997b) é generalizada a metodologia de bootstrap proposta por Hall(1990) para a estimação do erro quadrático médio (EQM) do estimador de Hill.

Na metodologia de Hall o bootstrapping é feito sobre o EQM do estimador de Hill, enquanto que na metodologia de Danielsson et al é feito sobre uma estatística auxiliar que converge a mesma taxa que o estimador de Hill. Ao contrário da metodologia de Hall onde não se conhece o valor ao qual converge o estimador de Hill, na metodologia geral sabe-se que a estatística auxiliar converge a zero.

A.1.6 Distribuição Generalizada de Pareto (GPD)

Na seção anterior, foi visto uma metodologia para modelar a distribuição dos máximos coletados em blocos de um mesmo tamanho. Agora, será apresentada a Distribuição Generalizada de Pareto (GPD) que está preocupada em modelar os excessos de uma variável aleatória X além de um limiar u , ou

seja, modelar $X - u$. É apresentado uma definição formal de excessos.

Definição A.1.7. Considere $X_1, X_2, \dots, X_N$ variáveis aleatórias iid com função de distribuição F_X onde $F_X \in MDA(H_\varepsilon)$ para algum $\varepsilon \in \mathbb{R}$. Serão denominados de **excessos** aqueles valores de X_i tais que $X_i \geq u$, onde u é um limiar pré-definido.

Denota-se por N_u o número de excedentes do limiar u , ou seja, $N_u = \sum_{i=1}^N I(X_i \geq u)$, onde $I(X_i \geq u) = \begin{cases} 1 & \text{se } X_i \geq u \\ 0 & \text{c.c.} \end{cases}$

Os valores $X_i - u \geq 0$, que são os excessos além do limiar u , serão representados por $Y_1, Y_2, \dots, Y_{N_u}$.

Assim como as distribuições de valores extremos são as distribuições limite para o máximo, a distribuição generalizada de Pareto é a forma paramétrica para distribuições limite de excessos.

As distribuições generalizadas de Pareto são a Exponencial, a Pareto e a Beta, cuja representações são da seguinte forma:

Exponencial:

$$W_0(y) = \begin{cases} 0 & , y < 0 \\ 1 - e^{-y} & , y \geq 0 \end{cases}$$

Pareto, $\alpha > 0$:

$$W_{1,\alpha}(y) = \begin{cases} 0 & , y < 1 \\ 1 - y^{-\alpha} & , y \geq 1 \end{cases}$$

Beta, $\alpha < 0$:

$$W_{2,\alpha}(y) = \begin{cases} 0 & , y < -1 \\ 1 - (-y)^{-\alpha} & , -1 \leq y \leq 0 \\ 1 & , y > 0 \end{cases}$$

A distribuição generalizada de Pareto (GPD) representa as três distribuições em uma única forma, sob a γ -parametrização: $W_\gamma(y) = 1 - (1 + \gamma y)^{-1/\gamma}$

A família de locação e escala $W_{\gamma,\delta,\psi}$ pode ser obtida substituindo-se o argumento y por $\frac{y-\delta}{\psi}$, onde $\delta \in \mathbb{R}$ e $\psi > 0$.

Dentro dessa família $W_{\gamma,\delta,\psi}$, a distribuição de maior interesse nas aplicações práticas é a $W_{\gamma,0,\psi}$ para $\gamma \geq 0$. Desta forma, a partir desse ponto, quando for referida a distribuição GPD, será considerada a $W_{\gamma,0,\psi}$ com $\gamma \geq 0$ e $y \geq 0$.

A relação existente entre a GEV e a GPD é dada por:

$$W(x) = 1 + \ln H(x), \text{ para } \ln H(x) > -1.$$

Para estimar os parâmetros γ e ψ da GPD, vários procedimentos têm sido propostos na literatura. Por exemplo, Smith(1987) propôs um estimador baseado nos estimadores de máxima verossimilhança que seria altamente eficiente se o limiar escolhido fosse ótimo. Neste trabalho será utilizado esse estimador de máxima verossimilhança (EMV), conforme será discutido no capítulo 3.

Observação A.1.2. Os EMV $(\hat{\gamma}, \hat{\psi})$ possuem suas boas propriedades quando $\gamma > -\frac{1}{2}$. (ver Smith(1987))

Para se encontrar o limiar ótimo u para a definição dos excessos tem-se o mesmo problema quanto à escolha do número ótimo k de estatísticas de ordem a serem usadas pelo estimador de Hill. Neste trabalho, será utilizado um limiar u tal que 10% das maiores observações fiquem acima dele, conforme foi utilizado em Gonzalez-Lopez and Tanaka(2003).

A.2 Caso Multivariado

Após a introdução dos principais resultados da Teoria de Valores Extremos univariada, não parece natural uma extensão desses resultados para o caso multivariado. Não existe uma maneira única para extensão desses resultados. Será considerado que a distribuição de cada componente da distribuição multivariada seja uma distribuição de valor extremo.

Definição A.2.1. Considere $\mathbf{X}_i$ $i = 1, 2, \dots, d$ vetores aleatórios iid com função de distribuição multivariada F . Defina o máximo como sendo $\mathbf{M}_n = (M_{1,n}, \dots, M_{d,n})$ onde $M_{j,n} = \max\{X_{1,j}, \dots, X_{n,j}\}$, $j = 1, \dots, d$.

Observação A.2.1. Como no caso univariado, estudar o comportamento de $\mathbf{L}_n = (L_{1,n}, \dots, L_{d,n})$, é similar ao estudo de $\mathbf{M}_n$.

Observação A.2.2. Sobre algumas notações que serão utilizadas, será utilizado que: para $k \in \mathbb{R}$ e $\mathbf{a}, \mathbf{b}, \mathbf{x}, \mathbf{y} \in \mathbb{R}^d$ escreve-se $\mathbf{x} < \mathbf{y}$ e $\mathbf{a}\mathbf{x} + \mathbf{b}$ para denotar as relações $x_i < y_i$ para todo $i = 1, \dots, d$ e o vetor $(a_1x_1 + b_1, \dots, a_dx_d + b_d)$, respectivamente. Mais ainda, $\mathbf{x}/\mathbf{y} = (x_1/y_1, \dots, x_d/y_d)$, e se $k, t \in \mathbb{R}$ então $\mathbf{x} + k$ é o vetor $(x_1 + k, \dots, x_d + k)$ e $\mathbf{x}^t = (x_1^t, \dots, x_d^t)$.

O interesse está em estudar o comportamento assintótico de $\mathbf{M}_n$. Como no caso univariado, o objetivo da teoria de valores extremos multivariada está em procurar condições sobre a função de distribuição F de $\mathbf{X}_i$ sob as quais existam seqüências $\mathbf{a}_n$ e $\mathbf{b}_n$ de modo que $P[\mathbf{M}_n \leq \mathbf{a}_n\mathbf{x} + \mathbf{b}_n] = F^n(\mathbf{a}_n\mathbf{x} + \mathbf{b}_n)$, convirja para uma função de distribuição não-degenerada d -variada $\mathbf{H}$ quando $n \rightarrow \infty$.

Definição A.2.2. Suponha que existam seqüências $\mathbf{a}_n > 0$, $\mathbf{b}_n \in \mathbb{R}^d$, tais que $\lim_{n \rightarrow \infty} F^n(\mathbf{a}_n\mathbf{x} + \mathbf{b}_n) = \mathbf{H}(\mathbf{x})$ para alguma função de distribuição não-degenerada $\mathbf{H}$, então $\mathbf{H}$ é denominada como distribuição de valor extremo multivariada e F pertence ao domínio de atração de $\mathbf{H}$.

Conhecido que a convergência em distribuição de $\mathbf{X}_n$ para $\mathbf{Y}$ implica na convergência em distribuição de $X_{i,n}$ para Y_i para cada $i = 1, \dots, d$ aplicando-se o teorema de Fisher-Tippet, pode-se concluir que se a distribuição limite acima existe então cada uma das marginais pertence a família GEV. Conseqüentemente, as seqüências normalizadoras $\mathbf{a}_n$ e $\mathbf{b}_n$ podem ser determinadas componente a componente usando os teoremas univariados.

Caracterizações de alguns domínios de atração podem ser feitas de maneira similar ao caso univariado utilizando algumas adaptações. Para tais caracterizações veja, por exemplo, Marshall e Olkin(1987).

No entanto, problemas podem surgir quando se trabalha com a distribuição multivariada num caminho mais geral, isto é, quando o interesse é compreender as relações existentes entre várias componentes de um vetor aleatório. Nesse sentido, ainda faltam muitos resultados. Esta dificuldade está, principalmente, na descrição da dependência entre as variáveis.

Referências

- [1] Bouyé, E., Durrleman, V., A.Nikeghbali, Riboulet, G., e Roncalli, T. (2001). Copulas: an open field for risk management. *Groupe de Recherche Opérationnelle, Crédit Lyonnais, Working Paper*.
- [2] Bouyé, E., Durrleman, V., Nikeghbali, A., Riboulet, G., e Roncalli, T. (2000). Copulas for finance - a reading guide and some applications. *Groupe de Recherche Opérationnelle, Crédit Lyonnais, Working Paper*.
- [3] Cherubini, U. e Luciano, E. (2000). Value at risk trade-off and capital allocation with copulas. Technical report, University of Bologna.
- [4] Danielsson, J. e de Vries, C. G. (1997b). Beyond the sample: extreme quantile and probability with applications to financial data. Technical report, Tinbergen Institute, Rotterdam.
- [5] Davidson, R. e MacKinnon, J. (1993). *Estimation and inference in econometrics*. Oxford University Press.

- [6] Embrechts, P., de Haan, L., e Huang, X. (1999). Modelling multivariate extremes. Technical report, Econometric Institute, Erasmus University, Rotterdam.
- [7] Embrechts, P., Klüppelberg, C., e Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*. Springer, New York.
- [8] Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 50:987–1007.
- [9] González-López, V. A. e Tanaka, N. I. (2003). Bi-variate data modeling through generalized archimedean copula. Technical Report RT-MAE, IME - USP.
- [10] Hall, P. (1990). Using the bootstrap to estimate mean squared error and select smoothing parameter in nonparametric problems. *Journal of Multivariate Analysis*, 32:177–203.
- [11] Joe, H. e Xu, J. J. (1996). The estimation method of inference functions for margins for multivariate models. Technical Report 166, Department of Statistics, University of British Columbia.
- [12] Junker, M., Szimayer, A., e Wagner, N. (2003). Nonlinear term struture dependence: Copula functions, empirics and risk implications. Working paper, Munich University of Technology.

- [13] Marshall, A. e Olkin, A. (1987). Domains of attraction of multivariate extreme value distributions. *Annals of Probability*, 11:168–175.
- [14] Nelsen, R. B. (1998). *An Introduction to Copulas*, volume 139 of *Lectures Notes in Statistics*. Springer Verlag, New York.
- [15] Nelson, D. B. e Cao, C. O. (1992). Inequality constraints in the univariate garch model. *Journal of Business and Economic Statistics*, 10:229–235.
- [16] Palaro, H. P. e Hotta, L. K. (2003). Using the conditional copula to estimate value at risk. Documento de trabalho, IMECC-Unicamp.
- [17] Patton, A. (2002). Estimation of copula models for time series of possibly different lengths. Working paper, University of California, San Diego.
- [18] Patton, A. J. (2001). Modelling time-varying exchange rate dependence using the conditional copula. Discussion Paper 09, Department of Economics, University of California, San Diego.
- [19] Pickands, J. (1975). Statistical inference using extreme order statistics. *The Annals of Statistics*, 3(1):119–131.
- [20] Pinto, F. C. e Valls Pereira, P. L. (2002). Teoria de valores extremos: aplicações em valor em risco. Documento de trabalho, IBMEC.
- [21] Smith, R. L. (1987). Estimating tails of probability distributions. *The Annals of Statistics*, 15(3):1174–1207.

- [22] Souza, L. A. R. e Silva, M. E. (1999). Teoria de valores extremos para cálculo de var. Finance lab, Faculdade de Economia e Administração, Universidade de São Paulo, São Paulo.