



Impa Instituto Nacional de Matemática Pura e Aplicada

Obtenção da Distribuição de Probabilidade do Preço de Ativos Financeiros Via Máxima Entropia

Autor: Elder Rodrigo Barbosa Campos

Orientador: Vinicius V. L. Albani

Rio de Janeiro
Outubro de 2014

Para as duas pessoas que sempre estiveram ao meu lado e que, com muita paciência e amor, me ajudam a sonhar e realizar: Maria Helena, minha mãe,
Wanda Maria, minha avó.

Agradecimentos

Agradeço a todos aqueles que me ajudaram e contribuíram para o desenvolvimento deste trabalho. Menciono aqui a equipe da Divisão de Ensino, representada pelo funcionário Josenildo Pedro, e todos os professores, que sempre se preocuparam em transmitir o conhecimento da melhor forma. Além disso, faço uma menção especial aos professores Jorge Zubelli e Vinícius Albani por todos os ensinamentos fornecidos e, principalmente, pela paciência, compreensão e confiança demonstradas. Obrigado a todos!

Abstract

We obtain an algorithm for finding the maximum entropy distribution of a financial asset, given a finite set of European call option prices, for a particular maturity. The main references are the articles of Cassio Neri and Lorenz Schneider [11] and [12]. The aim of calibrating an asset distribution is to evaluate, for example, the price of other European derivatives, with more complex payoffs, as well as to interpolate and to extrapolate the market implied volatility. These procedures are fundamental in hedging strategies and this algorithm seems to be useful and easy to implement. We also give some numerical examples with market option prices traded in the BM&F Bovespa, available at [2].

Key words: Entropy . Asset Distribution . Option Pricing

Sumário

Índice	vii
Introdução	1
1 Preço de Opções e Máxima Entropia	3
2 O Formato da Distribuição de Máxima Entropia	7
2.1 O Teorema de Csiszár	7
2.2 O Método Heurístico e o Princípio de não Arbitragem	10
3 Questões Práticas Sobre o Modelo	15
3.1 Representação da Densidade de Máxima Entropia e Condição de Não Arbitragem	15
3.2 Entropia como Função das Opções Digitais	17
4 O Algoritmo	23
5 Exemplos Numéricos	25
5.1 Testes	26
5.1.1 Valores Sintéticos	26
5.1.2 Valores Reais	27
5.2 Análise de Incerteza dos Parâmetros	27
5.2.1 Estimção da Incerteza	28
5.2.2 Resultados Práticos da Análise de Incerteza	28
5.3 Gráficos	29
5.3.1 Teste de Arbitragem	29
5.3.2 Gráfico da Densidade	32
5.3.3 Preços Extrapolados das Opções de Compra para Valores Sintéticos	34
5.3.4 Preços Extrapolados das Opções de Compra para Valores Reais	36
5.3.5 Volatilidade Implícita para Valores Sintéticos	38
5.3.6 Volatilidade Implícita para Valores Reais	40
6 Conclusão	43
A Definições Importantes	45
A.1 Martingais	45
A.2 Multiplicadores de Lagrange	46
B O Conceito de Entropia de Shannon	47
B.1 Sistemas de Comunicação	47
B.2 Geradores Discretos e o Conceito de Entropia	48

C Códigos de Matlab	51
Bibliografia	59

Introdução

O presente trabalho possui as opções de compra, que são um tipo de derivativos, como objetos principais. Por este motivo, a partir de definições presentes em [9] e [8], podemos dizer que:

Opção de Compra Uma opção de compra é um derivativo que permite ao detentor da opção o direito (mas não a obrigação) de comprar uma determinada quantidade do ativo em um dado instante já definido T (denominado maturidade) por certo valor pré acordado K (chamado strike).

Além disso, podemos tratar as chamadas Opções Européias ou Americanas. Na primeira, a opção só pode ser exercida na maturidade enquanto que, na segunda, a opção pode ser exercida em qualquer instante até a maturidade. Um importante resultado deduzido em [9] mostra que exercer uma opção de compra americana que não paga dividendos antes da maturidade não vale a pena. Ou seja, opções de compra americanas e européias possuem o mesmo preço. Por este motivo, construiremos todo o trabalho com a modelagem de opções européias, que são mais fáceis de se manipular, mas no fim podemos dizer que a precificação obtida também se aplica às opções americanas.

Assim sendo, este trabalho se baseia nos artigos de Cassio Neri e Lorenz Schneider [11] e [12] e também na dissertação de Ramilijaona [13], em que a partir de preços de opções de compra (não importa se americana ou européia) praticados no mercado, para uma dada maturidade, é possível obter uma função distribuição de probabilidade do ativo associado. Desta forma, a partir da distribuição de probabilidade, podemos extrapolar o preço da opção para outros strikes com a mesma maturidade. Na verdade, com esta calibragem extremamente valiosa, podemos fazer a precificação de derivativos mais complexos ou estratégias de hedging [8].

Na verdade, conforme será visto ao longo de todo o trabalho, devido ao formato relativamente simples da distribuição nós conseguimos explorar todo tipo de parâmetro do ativo. Outros bons exemplos além do preço da opção para qualquer strike são a volatilidade implícita e a derivada parcial do preço da opção em função do strike. Além disso, usando o mesmo processo para maturidades distintas podemos extrapolar uma superfície de volatilidade do ativo, algo extremamente útil e importante no mercado.

Motivação

Em geral, assume-se que o preço de uma dado ativo como função do tempo, descreve um processo estocástico semi-martingale. Essas seriam as chamadas circunstâncias normais obedecidas pelo preço do ativo em um mercado "bem comportado". No Apêndice, temos definições dos conceitos de Procesos Estocásticos, Martingais, Super-Martingais e Sub-Martingais. Porém, com relação aos martingais, uma ideia intuitiva sobre eles é que dada a posição atual do processo, na média, a posição futura é a atual.

Modelos financeiros, como Black e Scholes, podem sofrer em circunstâncias anormais. Ou seja, um determinado modelo que prevê bem o comportamento de um mercado em ambientes

normais pode estar completamente errado em situações atípicas, como as que ocorrem em momentos de crise econômica ou política.

Uma forma de retificar isto é usar métodos de estimação Bayesiana, os quais utilizam informação a priori acerca dos parâmetros desconhecidos a serem calibrados. Uma outra opção, que será a abordada neste projeto, cria, através do uso de métodos de estimação, funções de densidade de probabilidade via máxima entropia. A entropia, conceito presente em [6] e que será definida mais formalmente adiante, pode ser vista como uma medida de quão não enviesada é uma distribuição de probabilidade dada.

Assim sendo, tais distribuições tendem a ser menos dependentes de modelos pré definidos que funcionam somente em circunstâncias de equilíbrio. Ou seja, ao encontrar uma distribuição que maximiza a entropia a partir de dados vindos do mercado, nós podemos dizer que obtivemos uma distribuição de probabilidade que está de acordo com os dados do mercado (em outras palavras, não depende de modelos pré definidos), mas que ao mesmo tempo é a menos enviesada possível. Tais fatos não são observados em estimações Bayesianas.

Estrutura do Trabalho

O método de determinação da distribuição de probabilidade via máxima entropia se baseia na coleta de preços de opção de compra do ativo para um número finito de strikes K_i e com uma determinada maturidade T . Feito isso, a distribuição do preço do ativo será determinada para cada intervalo $[K_i, K_{i+1}]$. Assim sendo, observaremos que esta distribuição está de acordo com os dados reais do mercado e ainda consegue ser não enviesada.

Cabe ressaltar que os preços do mercado devem satisfazer uma condição de não arbitragem que será deduzida ao longo deste trabalho. Ou seja, antes de começar o método utilizando os preços vindos do mercado, um teste deve ser feito, onde cada preço e strike que não satisfaçam a condição de não arbitragem são excluídos.

Após o exposto acima, o presente trabalho tem por objetivo implementar o algoritmo proposto por Cassio Neri e Lorenz Schneider [11] e [12] com dados reais do mercado. Além das vantagens já expostas, o método proposto possui forte embasamento teórico e pode ser facilmente resolvido numericamente.

Todavia, antes desse ponto de resolução numérica com dados reais faremos uma revisão bibliográfica dos teoremas que embasam formalmente o método, além de implementações prévias com preços teóricos vindos de Black e Scholes.

Por fim, comentários serão feitos sobre os resultados obtidos no que se referem à sua consistência e praticidade. Cabe ressaltar também que todos os códigos e gráficos gerados na implementação numérica estão apresentados no apêndice.

Capítulo 1

Preço de Opções e Máxima Entropia

Neste capítulo veremos como se calcula o preço de uma opção de compra a partir da densidade de probabilidade do preço do ativo. Além disso, veremos a definição de dois conceitos importantes: *Opções Digitais* e *Entropia*. A partir disto, o objetivo central aqui será mostrar que o problema global de encontrar a função densidade de probabilidade em $[0, \infty[$ que atende, para uma dada maturidade T , aos preços de opções de compra com strikes estritamente crescente $K_0 = 0, K_1, \dots, K_n, K_{n+1} = \infty$, é equivalente a encontrar densidades nos intervalos $[K_i, K_{i+1}[$.

Definição 1.1 (Opção Digital). *Uma opção de compra digital ou binária, segundo [8], pode apresentar duas formas:*

cash-or-nothing call *é um aquela que paga um valor pré definido se o valor do ativo estiver acima do strike e zero se estiver abaixo.*

asset-or-nothing call *é um aquela que paga o próprio valor do ativo se ele estiver acima do strike e zero se estiver abaixo.*

Definição 1.2 (Entropia de Shannon). *A entropia associada a uma densidade de probabilidade g é dada por $E(g) := - \int_0^\infty g(x) \ln g(x) dx$*

No apêndice, temos a apresentação mais detalhada da origem do conceito de entropia de Shannon. Baseado no artigo do próprio Shannon [16] faremos inicialmente uma contextualização da origem (que começou na teoria da informação), depois uma breve demonstração de um teorema que define a entropia para um caso particular e, finalmente, faremos um paralelo com nosso trabalho mostrando que ao encontrar a densidade g que maximiza a entropia, estamos determinando uma função que concorda com os dados do mercado e é a menos enviesada possível.

Agora, antes de prosseguir, vamos mostrar como calcular o preço das opções envolvidas. A principal condição necessária para a construção do algoritmo é o conceito de ausência de *oportunidade de arbitragem*. Em [9], página 85, a definição de oportunidade de arbitragem é apresentada formalmente, mas uma ideia intuitiva disto é a seguinte: uma oportunidade de arbitragem ocorre quando um investidor, com capital inicial nulo, possui probabilidade não negativa de terminar com capital não nulo em um instante $T > 0$.

Ainda em [9], pode-se ver que a ausência de oportunidade de arbitragem no mercado implica na existência de uma medida de probabilidade neutra ao risco e isto, devido à representação de Feynman-Kac, permite que o preço de uma opção de compra, com strike K e maturidade T , seja dado pela seguinte equação:

$$C(K, T) = \tilde{\mathbb{E}}^g [e^{-rT} f(S(T))] = \int_0^\infty e^{-rT} f(s) g(s) ds \quad (1.0.1)$$

onde $f(S(T))$ é o payoff da opção e g é função densidade de probabilidade associada a medida neutra ao risco.

Cabe ressaltar que o payoff de uma opção de compra é dado por $f(S(T)) = (S(T) - K)^+$. Além disso, trabalharemos com opções digitais do tipo cash-or-nothing com payoff dado por $f(S(T)) = I_{S(T) > K}$. Logo, de (1.0.1), teremos:

$$C(K, T) = \tilde{\mathbb{E}}^g [e^{-rT}(S(T) - K)^+] = \int_K^\infty e^{-rT}(s - K)g(s)ds \quad (1.0.2)$$

$$D(K, T) = \tilde{\mathbb{E}}^g [e^{-rT}I_{S(T) > K}] = \int_K^\infty e^{-rT}g(s)ds \quad (1.0.3)$$

Assim sendo, seja fixada uma maturidade T e considere uma sequência de strikes estritamente crescente $K_0 = 0, K_1, \dots, K_n, K_{n+1} = \infty$. Além disso, sejam também os preços não descontados de opções de compra e digitais

$$\tilde{C}_i = C(K_i, T)/DF(0, T), \quad \tilde{D}_i = D(K_i, T)/DF(0, T)$$

nos mesmos strikes e com a mesma maturidade, onde o termo $DF(0, T) = e^{-rT}$ denota o fator de desconto. Por convenção adotaremos que

$$\tilde{C}_{n+1} = \tilde{D}_{n+1} = K_{n+1}\tilde{D}_{n+1} = 0$$

Assumindo estarmos sob as condições de existência de uma medida neutra ao risco, queremos determinar a densidade g correspondente ao preço do ativo $S(T)$ que maximiza a entropia $E(g) = -\int_0^\infty g(x) \ln g(x) dx$ sob as condições (devido à equação (1.0.2) e ao payoff da opção de compra)

$$\mathbb{E}^g[(S(T) - K_i)^+] = \tilde{C}_i \quad i.e. \quad \int_{K_i}^\infty (s - K_i)g(s)ds = \tilde{C}_i \quad (1.0.4)$$

além de (devido à equação (1.0.3) e ao payoff da opção digital usada)

$$\mathbb{E}^g[I_{S(T) > K_i}] = \tilde{D}_i \quad i.e. \quad \int_{K_i}^\infty g(s)ds = \tilde{D}_i \quad (1.0.5)$$

para todo $i = 0, 1, \dots, n$. Agora, observe que a condição (1.0.5) gera

$$\int_{K_i}^{K_{i+1}} g(s)ds = \tilde{D}_i - \tilde{D}_{i+1} \forall i = 0, 1, \dots, n \quad (1.0.6)$$

Além disso, considere um derivativo que paga $S(T)$ se $K_i < S(T) < K_{i+1}$ e zero, caso contrário. Ou seja, seu payoff é dado por $f(S(T)) = S(T)I_{K_i < S(T) < K_{i+1}}$. Logo, pela equação (1.0.1), tem-se que o preço descontado desse derivativo será dado por $\tilde{\mathbb{E}}^g [S(T)I_{K_i < S(T) < K_{i+1}}] = \int_{K_i}^{K_{i+1}} sg(s)ds$. Para reescrever esta última integral podemos abrir (1.0.4) e usar (1.0.5) para obter:

$$\int_{K_i}^{K_{i+1}} sg(s)ds = (\tilde{C}_i + K_i\tilde{D}_i) - (\tilde{C}_{i+1} + K_{i+1}\tilde{D}_{i+1}) \forall i = 0, 1, \dots, n \quad (1.0.7)$$

Após todo o exposto acima, podemos finalmente cumprir o objetivo deste capítulo, que é mostrar que encontrar a função densidade de probabilidade g , que maximiza a entropia, em $[0, \infty[$ e que atende, para uma dada maturidade T , aos preços de opções de compra com strikes estritamente crescente $K_0 = 0, K_1, \dots, K_n, K_{n+1} = \infty$, é equivalente a encontrar densidades nos intervalos $[K_i, K_{i+1}[$.

Ou seja, devemos calcular g sob as condições globais (1.0.4) e (1.0.5). Porém, provaremos que as condições locais (1.0.6) e (1.0.7) são equivalentes às condições globais e que, além disso,

$$-\int_0^\infty g(x) \ln g(x) dx = \sum_{i=0}^n \left(-\int_{K_i}^{K_{i+1}} g(x) \ln g(x) dx \right)$$

e, portanto, basta maximizar $-\int_{K_i}^{K_{i+1}} g(x) \ln g(x) dx$ sujeito a (1.0.6) e (1.0.7) em cada intervalo.

O próximo teorema prova o que queremos no momento, mas antes de apresentá-lo algumas definições serão feitas e uma proposição será enunciada e demonstrada.

Seja M^+ o conjunto das funções positivas na medida de Borel definidas em $[0, \infty[$. Defina

$$H := \left\{ g \in M^+ \mid \int_{K_i}^\infty g(x) dx = \tilde{D}_i, \int_{K_i}^\infty (x - K_i) g(x) dx = \tilde{C}_i, \forall i = 0, 1, \dots, n \right\}$$

e, $\forall i = 0, 1, \dots, n$,

$$H_i := \left\{ g \in M^+ \mid \int_{K_i}^{K_{i+1}} g(x) dx = \tilde{D}_i - \tilde{D}_{i+1}, \int_{K_i}^{K_{i+1}} x g(x) dx = (\tilde{C}_i + K_i \tilde{D}_i) - (\tilde{C}_{i+1} + K_{i+1} \tilde{D}_{i+1}) \right\}$$

Proposição 1.1. $H = \bigcap_{i=0}^n H_i$

Demonstração. é consequência imediata de (1.0.4), (1.0.5), (1.0.6) e (1.0.7). \square

Assim, $\forall i = 0, 1, \dots, n$ definimos

$$E_i(g) := -\int_{K_i}^{K_{i+1}} g(x) \ln g(x) dx \quad \forall g \in H_i.$$

O próximo teorema é central neste capítulo e vai provar a equivalência entre os problemas de encontrar g em $[0, \infty[$ e g_i em cada intervalo $[K_i, K_{i+1}[$.

Teorema 1.1. *Se g maximiza E em H , então g é o maximizador de E_i em H_i . Além disso, se g é o maximizador de E_i em $H_i, \forall i = 0, 1, \dots, n$, então g maximiza E em H .*

Demonstração. Seja g um maximizador de E em H , e seja $h \in H_i$. Defina

$$\tilde{g} = g \cdot (1 - \mathbb{I}_{[K_i, K_{i+1}[}) + h \cdot \mathbb{I}_{[K_i, K_{i+1}[}$$

Como $\tilde{g} = h$ em $[K_i, K_{i+1}[$, temos que $\tilde{g} \in H_i$. Além disso, para $j \neq i$, nós temos $\tilde{g} = g$ em $[K_j, K_{j+1}[$, e portanto $\tilde{g} \in H_j$. Logo, da Proposição 1.1 segue que $\tilde{g} \in H$. Com isso, da maximização de $\mathbb{E}(g)$, nós temos que $E(g) - E(\tilde{g}) \geq 0$.

Observe que, $E(\tilde{g}) = E(g) - E_i(g) + E_i(h)$, e portanto

$$E_i(g) - E_i(h) = E(g) - E(\tilde{g}) \geq 0.$$

Daí segue que g maximiza E_i em H_i .

Alternativamente, suponha que g maximiza E_i em $H_i, \forall i = 0, 1, \dots, n$. Seja $h \in H$. Segue que

$$E(g) = \sum_{i=0}^n E_i(g) \geq \sum_{i=0}^n E_i(h) = E(h),$$

o que significa que g maximiza E em H . \square

Acabamos de cumprir o objetivo principal deste capítulo que foi mostrar que o problema principal de encontrar a distribuição g de máxima entropia em $[0, \infty[$ é equivalente à solução de problemas mais simples de encontrar densidades g_i nos intervalos $[K_0, K_1[, [K_1, K_2[, \dots, [K_n, K_{n+1}[$. Este é o resultado do Teorema 1.1.

Os próximos passos são mostrar a forma de g em cada um desses intervalos e propor um método numérico para encontrar os parâmetros necessários. Assim sendo, primeiramente vamos construir a parte formal que determina como g deve ser e depois vamos usar multiplicadores de lagrange para obter as equações que geram o método numérico de determinação de parâmetros usando opções de compra e opções digitais.

Capítulo 2

O Formato da Distribuição de Máxima Entropia

2.1 O Teorema de Csiszár

Um dos grandes problemas da formalização dos resultados, conforme visto em [4] é que os métodos usuais normalmente se baseiam na aplicação de multiplicadores de lagrange e isso requer que a função objetivo seja continuamente diferenciável. Porém, a entropia de Boltzmann - Shannon é finita apenas para densidades em

$$\varrho := \{g \in L^1(0, \infty) \mid g \ln g \in L^1(0, \infty)\}$$

que possui interior vazio em $L^1(0, \infty)$. Logo, o maximizador não é um ponto interior do conjunto ϱ . Além disso, a entropia não é continuamente diferenciável já que não é contínua em lugar algum.

Segundo Neri-Schneider, apresentaremos uma abordagem deste problema baseada em um resultado de Csiszár [7].

Assumindo que nenhuma densidade a priori é dada, estamos interessados na entropia de g em I

$$E_I(g) := - \int_I g(x) \ln g(x) dx$$

em que $I \subset [0, \infty[$ é um intervalo. Contudo, o resultado de Csiszár usa uma entropia relativa de uma função densidade de probabilidade g com respeito a uma medida de probabilidade R em I

$$E_I(g \mid R) := - \int_I g(x) \ln g(x) dR(x).$$

De certa forma, podemos dizer que queremos uma entropia relativa com respeito a medida de Lebesgue que, em $[0, \infty[$, não é uma medida de probabilidade.

Para $i = 0, 1, \dots, n - 1, I = [K_i, K_{i+1}[$ é limitado. Neste caso, podemos usar $dR(x) = (K_{i+1} - K_i)^{-1} dx$. Porém, isso não é possível para o caso global $I = [0, \infty[$ ou para o último intervalo $I = [K_n, \infty[$, dado que para estes casos não existe constante normalizadora que transforme a medida de Lebesgue em uma medida de probabilidade.

Todavia, aplicaremos o resultado de Csiszár e não será necessário fazer qualquer distinção acerca de intervalos finitos ou infinitos.

Proposição 2.1. *Seja $I \subseteq [0, \infty[$ um intervalo. Defina $m(x) = \theta e^{-x}, \forall x \in I$, onde $\theta > 0$ é uma constante de normalização, tal que $dR(x) = m(x)dx$ seja uma medida de probabilidade em I . Sejam também a_0 e a_1 constantes reais e M^+ o conjunto das funções positivas na medida de Borel definidas em $[0, \infty[$. Então, a relação $g \mapsto g/m$ é uma bijeção entre os conjuntos*

$$\Omega := \left\{ g \in M^+ \mid \int_I g(x)dx = a_0, \int_I xg(x)dx = a_1 \right\}$$

e

$$\tilde{\Omega} := \left\{ \tilde{g} \in M^+ \mid \int_I \tilde{g}(x)dR(x) = a_0, \int_I x\tilde{g}(x)dx = a_1 \right\}$$

Além disso, g é o maximizador de E_I em Ω se, e somente se, g/m é maximizador de $E_I(\cdot \mid R)$ em $\tilde{\Omega}$

Demonstração. Defina $\Psi : M^+ \rightarrow M^+$, tal que $\Psi(g) := g/m$. Como m é estritamente positiva, podemos dizer que temos uma bijeção bem definida.

Vamos mostrar agora que Ψ preserva algumas funcionalidades lineares. Seja $g \in M^+$ e $f : I \rightarrow \mathbb{R}$. Então, temos que

$$\int_I f(x) (\Psi(g)(x)) dR(x) = \int_I f(x) \frac{g(x)}{m(x)} dR(x) = \int_I f(x)g(x)dx.$$

Em particular, aplicando este resultado a $f(x) \equiv 1$ e a $f(x) \equiv x$, segue imediatamente que Ψ mapeia Ω em $\tilde{\Omega}$.

Para completar a demonstração é suficiente mostrar que, se $g, h \in \Omega$, então

$$E_I(g) - E_I(h) \geq 0 \Leftrightarrow E_I(\Psi(g) \mid R) - E_I(\Psi(h) \mid R) \geq 0$$

Na verdade, o resultado é mais forte e as desigualdades acima são igualdades. Isto é equivalente a mostrar que $E_I(\Psi(g) \mid R) - E_I(g)$ não depende de $g \in \Omega$. Temos que,

$$\begin{aligned} E_I(\Psi(g) \mid R) &= - \int_I (\Psi(g)(x)) \ln (\Psi(g)(x)) dR(x) \\ &= - \int_I \left(\frac{g(x)}{m(x)} \right) \ln \left(\frac{g(x)}{m(x)} \right) dR(x) \\ &= - \int_I (g(x)) \ln \left(\frac{g(x)}{m(x)} \right) dx \\ &= - \int_I (g(x)) \ln (g(x)) dx + \int_I (g(x)) \ln (m(x)) dx \\ &= E_I(g) + \ln \theta \int_I g(x)dx - \int_I xg(x)dx \\ &= E_I(g) + a_0 \ln \theta - a_1. \end{aligned}$$

□

Agora podemos enunciar o resultado de Csiszár [7] que será fundamental para obtermos o formato de g.

Teorema 2.1 (Csiszár). *Seja R uma medida de probabilidade em um espaço (X, \mathcal{H}) . Seja $\{f_\gamma\}_{\gamma \in \Gamma}$ um conjunto arbitrário de funções reais \mathcal{H} - mensuráveis em X e sejam $\{a_\gamma\}_{\gamma \in \Gamma}$ constantes reais. Seja também \mathcal{E} o conjunto de todas as probabilidades P em (X, \mathcal{H}) para as quais as integrais $\int f_\gamma dP$ existem e são iguais a a_γ ($\gamma \in \Gamma$). Então existe $Q \in \mathcal{E}$ tal que $Q \ll R$ e sua derivada de Radon-Nikodym possui a forma*

$$\frac{\partial Q}{\partial R}(x) = ce^{q(x)} \quad \forall x \in I \quad (2.1.1)$$

onde $c > 0$ e q pertence ao espaço gerado pelas funções f_γ . Logo,

$$\int \left(\frac{\partial Q}{\partial R}(x) \right) \ln \left(\frac{\partial Q}{\partial R}(x) \right) dR(x) \leq \int \left(\frac{\partial P}{\partial R}(x) \right) \ln \left(\frac{\partial P}{\partial R}(x) \right) dR(x)$$

para todo $P \in \mathcal{E}$ tal que $P \ll Q$.

Antes de enunciar o teorema central deste trabalho, que pode ser encontrado em [11], vamos apresentar algumas equações que surgem do uso de multiplicadores de lagrange e que serão deduzidas mais tarde neste trabalho. Para $i = 0, 1, \dots, n-1$, definimos:

$$\alpha_i = \beta_i \cdot \left(\frac{\tilde{D}_i - \tilde{D}_{i+1}}{e^{\beta_i K_{i+1}} - e^{\beta_i K_i}} \right) \quad (2.1.2)$$

$$\frac{K_{i+1}e^{\beta_i K_{i+1}} - K_i e^{\beta_i K_i}}{e^{\beta_i K_{i+1}} - e^{\beta_i K_i}} - \frac{1}{\beta_i} = \frac{(\tilde{C}_i + K_i \tilde{D}_i) - (\tilde{C}_{i+1} + K_{i+1} \tilde{D}_{i+1})}{\tilde{D}_i - \tilde{D}_{i+1}} \quad (2.1.3)$$

Assim, podemos enunciar o teorema central deste trabalho que estabelece o formato da distribuição em cada intervalo $[K_i, K_{i+1}[$:

Teorema 2.2. *Seja $i \in \{0, 1, \dots, n\}$, $I = [K_i, K_{i+1}[$. Sejam α_i e β_i como nas equações (2.1.2) e (2.1.3). Então, $g : I \rightarrow \mathbb{R}$ dada por*

$$g(x) = \alpha_i e^{\beta_i x}, \quad \forall x \in I,$$

maximiza E_i em \mathcal{H} .

Demonstração. Sejam $a_0 = \tilde{D}_{i+1} - \tilde{D}_i$ e $a_1 = (\tilde{C}_i + K_i \tilde{D}_i) - (\tilde{C}_{i+1} + K_{i+1} \tilde{D}_{i+1})$. Sejam ainda m, R, Ω e $\tilde{\Omega}$ conforme a Proposição 2.1. Observe que para estas escolhas de I, a_0 e a_1 , temos que $\Omega = \mathcal{H}_i$ e $E_I = E_i$.

Sejam $X = I$, \mathcal{H} é a σ - álgebra de Borel de subconjuntos de I , $\Gamma = \{0, 1\}$, $f_\gamma \equiv a_0 x^\gamma$ ($\gamma \in \Gamma$) e \mathcal{E} é definido conforme o teorema de Csiszár, isto é, \mathcal{E} o conjunto de todas as probabilidades P em (X, \mathcal{H}) para as quais as integrais $\int f_\gamma dP$ existem e são iguais a a_γ ($\gamma \in \Gamma$)

Dado $\tilde{h} \in \tilde{\Omega}$, defina a medida $P_{\tilde{\Omega}}$ por $dP_{\tilde{\Omega}}(x) = a_0^{-1} \tilde{h}(x) dR(x)$. Assim, da definição de $\tilde{\Omega}$, segue que $P_{\tilde{\Omega}} \in \mathcal{E}$ e $P \ll Q$, então $a_0 \partial P / \partial R \in \tilde{\Omega}$. Logo,

$$\begin{aligned} - \int_I \frac{\partial P_{\tilde{h}}}{\partial R}(x) \ln \frac{\partial P_{\tilde{h}}}{\partial R}(x) dR(x) &= -a_0^{-1} \int_I \tilde{h}(x) \ln \tilde{h}(x) dR(x) + \ln(a_0) \\ &= a_0^{-1} E_I(\tilde{h} \mid R) + \ln(a_0) \end{aligned} \quad (2.1.4)$$

Pela definição de α_i e β_i , temos que $g \in \mathcal{H}_i$ e $\mathcal{H}_i = \Omega$. Logo, pela Proposição 2.1 temos que $\tilde{g} = g/m \in \tilde{\Omega}$. Além disso,

$$\tilde{g}(x) = \alpha_i \theta^{-1} e^{(\beta_i+1)x}, \quad \forall x \in I$$

Seja $Q = P_{\tilde{g}}$. Segue que $Q \in \mathcal{E}$, e sua derivada de Radon-Nikodym com respeito a R (que é $a_0^{-1} \tilde{g}$) tem a forma dada em (2.1.1) com $c = a_0^{-1} \alpha_i \theta^{-1}$ e $q = (\beta_i + 1) a_0^{-1} f_1$. Portanto, o teorema de Csizár nos dá

$$\int_I \frac{\partial Q}{\partial R}(x) \ln \frac{\partial Q}{\partial R}(x) dR(x) \leq \int_I \frac{\partial P}{\partial R}(x) \ln \frac{\partial P}{\partial R}(x) dR(x),$$

para todo $P \in \mathcal{E}$, tal que $P \ll Q$. Em particular, para todo $\tilde{h} \in \tilde{\Omega}$. Logo de (2.1.4) obtemos que

$$E_I(\tilde{g} | R) \geq E_I(\tilde{h} | R).$$

Concluimos que \tilde{g} maximiza $E_I(\cdot | R)$ em $\tilde{\Omega}$ e, novamente, pela Proposição 2.1, temos que g é o maximizador de E_i em H_i . \square

2.2 O Método Heurístico e o Princípio de não Arbitragem

Vamos agora obter as equações (2.1.2) e (2.1.3), que são as equações que fornecem os valores de α_i e β_i , usando multiplicadores de Lagrange. Porém, antes de prosseguir, lembremos:

Problema Central Queremos determinar, para cada intervalo $[K_0, K_1[, [K_1, K_2[, \dots, [K_n, K_{n+1}[$, densidades g_i que maximizam a entropia

$$E_i(g) = - \int_{K_i}^{K_{i+1}} g(x) \ln g(x) dx,$$

sujeito às condições

$$\int_{K_i}^{\infty} g(s) ds = \tilde{D}_i$$

e

$$\int_{K_i}^{\infty} (s - K_i) g(s) ds = \tilde{C}_i.$$

Cabe ressaltar que estamos utilizando o resultado do Teorema 1.1 para dizer que encontrar g é equivalente a encontrar cada g_i .

Assim sendo, devido ao objetivo de maximizar E_i sujeito às condições mencionadas acima, definimos os funcionais:

$$\mathcal{F}(g) := - \int_{K_i}^{K_{i+1}} g(x) \ln g(x) dx,$$

$$\mathcal{G}(g) := \int_{K_i}^{K_{i+1}} g(x) dx \quad e$$

$$\mathcal{H}(g) := \int_{K_i}^{K_{i+1}} x g(x) dx.$$

Portanto, por multiplicadores de Lagrange, devemos resolver a equação

$$\delta\mathcal{F}(g) + \lambda_1\delta\mathcal{G}(g) + \lambda_2\delta\mathcal{H}(g) = 0.$$

Por derivada de Fréchet, cuja definição se encontra no Apêndice, segue que

$$\int_{K_i}^{K_{i+1}} (\ln g(x) + 1)\delta g(x)dx = \int_{K_i}^{K_{i+1}} (\lambda_1 + \lambda_2 x)\delta g(x)dx$$

Logo, no intervalo $[K_i, K_{i+1}[$, temos que

$$\ln g(x) + 1 = \lambda_1 + \lambda_2 x$$

Assim, fazendo $\alpha_i := e^{\lambda_1 - 1}$ e $\beta_i = \lambda_2$, temos que:

$$g(x) = \alpha_i e^{\beta_i x} \quad x \in [K_i, K_{i+1}[. \quad (2.2.1)$$

Usando esta forma explícita de g em (1.0.6) e (1.0.7) obtemos:

$$\alpha_i \int_{K_i}^{K_{i+1}} e^{\beta_i x} dx = \tilde{D}_i - \tilde{D}_{i+1} \quad (2.2.2)$$

$$\alpha_i \int_{K_i}^{K_{i+1}} x e^{\beta_i x} dx = (\tilde{C}_i + K_i \tilde{D}_i) - (\tilde{C}_{i+1} + K_{i+1} \tilde{D}_{i+1}) \quad (2.2.3)$$

Para todo $i = 0, 1, \dots, n$.

Assim, para $i < n$, resolvendo as equações (2.2.2) e (2.2.3) usando integração por partes, obtemos as equações (2.1.2) e (2.1.3), ou seja

$$\alpha_i = \beta_i \cdot \left(\frac{\tilde{D}_i - \tilde{D}_{i+1}}{e^{\beta_i K_{i+1}} - e^{\beta_i K_i}} \right) e$$

$$\frac{K_{i+1} e^{\beta_i K_{i+1}} - K_i e^{\beta_i K_i}}{e^{\beta_i K_{i+1}} - e^{\beta_i K_i}} - \frac{1}{\beta_i} = \frac{(\tilde{C}_i + K_i \tilde{D}_i) - (\tilde{C}_{i+1} + K_{i+1} \tilde{D}_{i+1})}{\tilde{D}_i - \tilde{D}_{i+1}}$$

Falta agora mostrar que essas soluções são únicas.

Antes disso, definamos

$$\Theta(\beta; K_i, K_{i+1}) := \frac{K_{i+1} e^{\beta K_{i+1}} - K_i e^{\beta K_i}}{e^{\beta K_{i+1}} - e^{\beta K_i}} - \frac{1}{\beta} \quad (2.2.4)$$

$$\Theta'(\beta; K_i, K_{i+1}) := \frac{1}{\beta^2} - (K_{i+1} - K_i)^2 \frac{e^{\beta(K_{i+1} + K_i)}}{(e^{\beta K_{i+1}} - e^{\beta K_i})^2} \quad (2.2.5)$$

Proposição 2.2. *Seja $i \in \{0, 1, \dots, n\}$. Se não existir oportunidade de arbitragem gerada por $\tilde{D}_i, \tilde{D}_{i+1}, \tilde{C}_i$ e \tilde{C}_{i+1} , então existe uma única solução (α_i, β_i) para as equações (2.2.2) e (2.2.3).*

Demonstração. Defina

$$\bar{K} := \frac{(\tilde{C}_i + K_i \tilde{D}_i) - (\tilde{C}_{i+1} + K_{i+1} \tilde{D}_{i+1})}{\tilde{D}_i - \tilde{D}_{i+1}}$$

Primeiramente, vamos provar que $K_i < \bar{K} < K_{i+1}$. Isto pode ser visto comparando o preço de três derivativos. Eles pagam, respectivamente, $K_i, S(T)$ e K_{i+1} se $K_i < S(T) < K_{i+1}$ e zero caso contrário. Os payoffs desses derivativos são dados por:

$$f_1(S(T)) = K_i \mathbb{I}_{K_i < S(T) < K_{i+1}}$$

$$f_2(S(T)) = S(T) \mathbb{I}_{K_i < S(T) < K_{i+1}}$$

$$f_3(S(T)) = K_{i+1} \mathbb{I}_{K_i < S(T) < K_{i+1}}$$

Sob as condições de não arbitragem, segue imediatamente que o segundo derivativo é mais caro que o primeiro e mais barato que o terceiro. Para comprovar isso, podemos supor que, por exemplo, o preço do derivativo 2 é menor ou igual ao do derivativo 1 e construir uma estratégia de arbitragem a partir disso. Mais precisamente, um investidor poderia comprar o segundo derivativo e vender o primeiro (observe que o capital inicial no investidor é não positivo, pois o preço do derivativo comprado é menor ou igual que o vendido). Assim, na maturidade T , caso tenhamos $K_i < S(T) < K_{i+1}$ ele terá de gastar K_i , mas receberá $S(T) > K_i$. Ou seja, o investidor partirá de um capital não positivo e mesmo assim terá probabilidade não nula de terminar com capital não nulo. Logo, segue que o segundo derivativo é mais caro que o primeiro e, de forma análoga, que o terceiro é mais caro que o segundo.

Além disso, com os payoffs $f_1(S(T))$, $f_2(S(T))$ e $f_3(S(T))$ e a partir das equações (1.0.6) e (1.0.7), segue que:

$$K_i(\tilde{D}_i - \tilde{D}_{i+1}) < (\tilde{C}_i + K_i \tilde{D}_i) - (\tilde{C}_{i+1} + K_{i+1} \tilde{D}_{i+1}) < K_{i+1}(\tilde{D}_i - \tilde{D}_{i+1}).$$

Assim, da definição de \bar{K} o resultado segue da inequação acima.

Agora, vamos provar que se $K_i < \bar{K} < K_{i+1}$ então a solução (α_i, β_i) é única. Primeiro, considere $i < n$. Sabemos que (2.2.2) e (2.2.3) equivalem a (2.1.2) e (2.1.3). Sem perda de generalidade, vamos assumir que $K_i = 0$ e $K_{i+1} = 1$. Assim, essa mudança de variáveis gera

$$(\beta, \bar{K}) \leftrightarrow (x, \lambda), \quad x := \beta(K_{i+1} - K_i), \quad \lambda := \frac{\bar{K} - K_i}{K_{i+1} - K_i}$$

Com isso, a partir da equação (2.2.4), podemos dizer que $\Theta(\beta; K_i, K_{i+1}) = \bar{K}$ em $\Theta(x; 0, 1) = \lambda$, com $\lambda \in (0, 1)$. Usando a regra de L'Hospital obtemos a função $F: \mathbb{R} \rightarrow \mathbb{R}$ dada por

$$F(x) := \begin{cases} \frac{e^x}{e^x - 1} - \frac{1}{x} & x \neq 0 \\ \frac{1}{2} & x = 0 \end{cases}$$

é uma extensão contínua de $\Theta(\cdot; 0, 1)$. É fácil observar que $\lim_{x \rightarrow -\infty} F(x) = 0$ e $\lim_{x \rightarrow +\infty} F(x) = 1$.

Logo, a equação $F(x) = \lambda$ possui uma solução. Para provar que ela é única, devemos mostrar que F é estritamente crescente. Novamente, pela regra de L'Hospital, nós obtemos que F é diferenciável em $x = 0$ e $F'(0) = 1/12$. Para $x \neq 0$, temos que

$$F'(x) = \frac{1}{x^2} - \frac{e^x}{(e^x - 1)^2}$$

Lembre que $x^{-1} \sinh(x) > 1, \forall x \in \mathbb{R} \setminus \{0\}$. Assim,

$$\frac{e^{x/2} - e^{-x/2}}{x} > 1 \Rightarrow \frac{e^x - 1}{x} > e^{x/2} \Rightarrow \left(\frac{e^x - 1}{x} \right)^2 > e^x \Rightarrow F'(x) > 0.$$

Portanto, F é estritamente crescente.

Por fim, considere o caso $i = n$. As equações (2.2.2) e (2.2.3) se tornam:

$$\alpha_n \int_{K_n}^{\infty} e^{\beta_n x} dx = \tilde{D}_n$$

$$\alpha_n \int_{K_n}^{\infty} x e^{\beta_n x} dx = (\tilde{C}_n + K_n \tilde{D}_n)$$

Pela primeira equação segue que $\beta_n < 0$. Resolvendo as duas equações acima, chegamos a:

$$\alpha_n = -\frac{\beta_n \tilde{D}_n}{e^{\beta_n K_n}}, \beta_n = -\frac{\tilde{D}_n}{\tilde{C}_n}$$

□

Capítulo 3

Questões Práticas Sobre o Modelo

O objetivo principal deste capítulo, é construir as condições formais que permitem a obtenção de um algoritmo capaz de encontrar a Função Densidade de Probabilidade de Máxima Entropia, a partir dos preços das opções de compra de uma sequência de strikes estritamente crescente $K_0 = 0, K_1, \dots, K_n, K_{n+1} = \infty$.

Como um resumo do que foi feito até agora, podemos dizer que, a distribuição de máxima entropia, que a partir daqui será chamada de *DME*, pode ser obtida da seguinte forma: para um dado conjunto de strikes $K_0 = 0, K_1, \dots, K_n, K_{n+1} = \infty$ com seus respectivos preços de opção de compra (supondo ausência de arbitragem) C_i , temos os seguintes passos:

Passo 1: adicionar o preço de opções digitais com os mesmos strikes

Passo 2: obter as constantes α_i e β_i para cada intervalo $[K_i, K_{i+1}[$ a partir das equações (2.1.2) e (2.1.3).

As equações que fornecem α_i e β_i são relativamente fáceis de se resolver numericamente, mas o método ainda possui um ponto importante a ser abordado: como obter os preços das opções digitais com os mesmos strikes das opções de compra?

Na verdade, em princípio, essa questão gera um grande problema na implementação do método, pois trata-se da inserção de uma condição extra. Outros métodos como o que é obtido em [4] são mais complicados e instáveis de se resolver numericamente, mas podem ser resolvidos sem o uso das opções digitais.

Logo, podemos pensar que estamos simplesmente trocando as limitações, mas isso não é verdade. O fato fundamental que será desenvolvido neste capítulo é que os preços das opções digitais não precisam ser extraídos do mercado.

3.1 Representação da Densidade de Máxima Entropia e Condição de Não Arbitragem

Sob a medida neutra ao risco temos que o preço não descontado de uma opção digital de strike K_i é dado pela equação (1.0.5). É fato conhecido que o payoff de uma opção de compra com strike K_i dado pela equação (1.0.4) é uma função estritamente convexa em $K_i \in [0, \infty[$. Além disso, a relação entre esses payoffs é dada por:

$$\mathbb{E}^g[I_{S(T) > K_i}] = -\frac{d}{dK} \mathbb{E}^g[(S(T) - K_i)^+]$$

Portanto, da relação acima e da convexidade do payoff da opção de compra, segue a importante relação:

$$-\frac{\tilde{C}_i - \tilde{C}_{i-1}}{K_i - K_{i-1}} > \mathbb{E}^g[I_{S(T) > K_i}] > -\frac{\tilde{C}_{i+1} - \tilde{C}_i}{K_{i+1} - K_i} \quad \forall i = 1, 2, \dots, n \quad (3.1.1)$$

Em particular, podemos dizer que o preço das opções digitais, usadas na obtenção da DME do Teorema 2.2, também satisfaz (3.1.1).

Em [11] mostra-se que, sob a condição (3.1.1), existe uma única densidade g em $[0, \infty[$ que maximiza a entropia e satisfaz as equações $\mathbb{E}^g[(S(T) - K_i)^+] = \tilde{C}_i$ e $\mathbb{E}^g[I_{S(T) > K_i}] = \tilde{D}_i$. Vamos apresentar (sem demonstrar) uma proposição presente em [14], que ajuda [11] e [12] a demonstrar a unicidade da densidade g de máxima entropia. Porém, antes disso, é importante definir algumas expressões que serão importantes nos próximos passos:

$$c_i(\beta) := \begin{cases} \ln\left(\frac{e^{\beta K_{i+1}} - e^{\beta K_i}}{\beta}\right), & i < n \text{ e } \beta \neq 0 \\ \ln(K_{i+1} - K_i), & i < n \text{ e } \beta = 0 \\ \ln\left(-\frac{e^{\beta K_i}}{\beta}\right), & i = n \text{ e } \beta < 0 \end{cases} \quad (3.1.2)$$

Assim, suas duas primeiras derivadas são:

$$c'_i(\beta) := \begin{cases} \frac{K_{i+1}e^{\beta K_{i+1}} - K_i e^{\beta K_i}}{e^{\beta K_{i+1}} - e^{\beta K_i}} - \frac{1}{\beta}, & i < n \text{ e } \beta \neq 0 \\ \frac{K_{i+1} + K_i}{2}, & i < n \text{ e } \beta = 0 \\ K_i - \frac{1}{\beta}, & i = n \text{ e } \beta < 0 \end{cases}$$

e

$$c''_i(\beta) := \begin{cases} -(K_i - K_{i+1})^2 \frac{e^{\beta(K_{i+1} + K_i)}}{(e^{\beta K_{i+1}} - e^{\beta K_i})^2} + \frac{1}{\beta^2}, & i < n \text{ e } \beta \neq 0 \\ \frac{(K_{i+1} - K_i)^2}{12}, & i < n \text{ e } \beta = 0 \\ \frac{1}{\beta^2}, & i = n \text{ e } \beta < 0. \end{cases}$$

Relembremos a definição de Transformada de Legendre de uma função convexa e apresentemos a proposição citada anteriormente, que está presente em [14].

$$f^*(x) := \sup_{y \in \mathbb{R}} \{xy - f(y)\}, \quad y \in \mathbb{R}$$

Proposição 3.1. *Seja f uma função diferenciável e estritamente convexa em \mathbb{R} . São válidos os seguintes resultados.*

- (i) $xy \leq f(x) + f^*(y)$ com a igualdade ocorrendo se, e somente se, $y = f'(x)$
- (ii) $[f^*]' = [f']^{-1}$
- (iii) f^* é convexa e $[f^*]^* = f$

As funções c_1, \dots, c_n definidas acima são diferenciáveis e estritamente convexas. Em [11] é demonstrado que c'_i é continuamente diferenciável e estritamente crescente, e que $c'_i(\beta)$ tende a K_i e K_{i+1} quando β tende a $+\infty$ e $-\infty$, respectivamente. Para $i < n$ temos que $c'_i(0) = (K_i + K_{i+1})/2$ e da Proposição 3.1 (ii)-(iii) segue que $(K_i + K_{i+1})/2$ é raiz de $[c_i]^*$ e mínimo de c_i^* . Além disso, da Proposição 3.1 (i)-(ii) seguem as equivalências

$$[c_i^*]'(K) = \beta \Leftrightarrow K = c'_i(\beta) \Leftrightarrow c_i(\beta) + c_i^*(K) = \beta K \quad (3.1.3)$$

Finalmente, é demonstrado em [11] que sob, a condição de não arbitragem (3.1.1), para todo $\tilde{D} \in \Omega$, existe uma única densidade $g_{\tilde{D}}$ que maximiza a entropia e satisfaz os preços das opções digitais e de compra. Além disso, para todo $i = 0, 1, \dots, n$, temos em [12]:

(i) A probabilidade de $S(T)$ estar em $[K_i, K_{i+1}[$ é

$$p_i := \mathbb{P}_{\tilde{D}}^g[K_i \leq S(T) < K_{i+1}] = \int_{K_i}^{K_{i+1}} g_{\tilde{D}}(x) dx = \tilde{D}_i - \tilde{D}_{i+1} \quad (3.1.4)$$

(ii) A esperança condicional de $S(T)$ dado que ela está em algum intervalo $[K_i, K_{i+1}[$ é

$$\begin{aligned} \bar{K}_i &:= \mathbb{E}^{g_{\tilde{D}}}[S(T) \mid K_i \leq S(T) < K_{i+1}] \\ &= \frac{1}{p_i} \int_{K_i}^{K_{i+1}} x g_{\tilde{D}}(x) dx \\ &= \frac{(\tilde{C}_i + K_i \tilde{D}_i) - (\tilde{C}_{i+1} + K_{i+1} \tilde{D}_{i+1})}{p_i} \end{aligned} \quad (3.1.5)$$

Além disso, o Teorema 2.2 reescrito com a notação das equações (3.1.2), diz que, em cada intervalo $[K_i, K_{i+1}[$, a densidade $g_{\tilde{D}}$ tem a forma $g_{\tilde{D}}(x) = \alpha_i e^{\beta_i x}$, onde

$$\alpha_i = p_i e^{-c_i(\beta_i)} \quad (3.1.6)$$

$$c'_i(\beta_i) = \bar{K}_i \quad (3.1.7)$$

Observemos que no cálculo de α_i e β_i , a equação (3.1.7) não pode ser resolvida explicitamente. Contudo, na prática ela pode ser resolvida de forma fácil e rápida utilizando o método de Newton-Rapson unidimensional, pois a expressão analítica de c'_i é conhecida. Normalmente, com no máximo três iterações já chegamos ao resultado com um bom nível de precisão.

3.2 Entropia como Função das Opções Digitais

O objetivo principal deste capítulo é obter a DME sem a necessidade de retirar do mercado o preço das opções digitais. Isto será construído nos próximos passos.

Considerando os preços das opções de compra fixos, vamos variar $\tilde{D}_i = \mathbb{E}^g[I_{S(T) > K_i}]$ nos intervalos dados por (3.1.1). Seja então o retângulo aberto n -dimensional $\Omega \subset \mathbb{R}^n$ que contém todos os vetores $\tilde{D} = (\tilde{D}_1, \dots, \tilde{D}_n)$ satisfazendo (3.1.1). Para cada $\tilde{D} \in \Omega$, denote por $g_{\tilde{D}}$ a DME obtida através do Teorema 2.2.

Seja ainda $\mathcal{G} = \{g_{\tilde{D}} \mid \tilde{D} \in \Omega\}$ a família de todas as DMEs geradas deste modo. O objetivo é obter a densidade g^* que possui a maior entropia dentre todas as densidades em \mathcal{G} . Porém, para cumprir este objetivo, usaremos o próximo teorema, que mostra como escrever a entropia em função de $\tilde{D} \in \Omega$.

Teorema 3.1. Para todo $\tilde{D} \in \Omega$ a entropia de $g_{\tilde{D}}$ pode ser expressa como

$$H(g_{\tilde{D}}) = - \sum_{i=0}^n p_i \ln p_i - \sum_{i=0}^n p_i c_i^*(\bar{K}_i)$$

em que p_i e \bar{K}_i são dados exclusivamente pelos preços das opções segundo as equações (3.1.4) e (3.1.6), para todo $i = 0, 1, \dots, n$.

Demonstração. Seja $\tilde{D} \in \Omega$. Para cada $i = 0, 1, \dots, n$, sejam p_i , \bar{K}_i , α_i e β_i conforme as equações (3.1.4), (3.1.6), (3.1.6) e (3.1.7), respectivamente. Fixando $i \in 1, \dots, n$ e denotando por $H_i(g_{\tilde{D}})$ a entropia no intervalo $[K_i, K_{i+1}[$, into é,

$$H_i(g_{\tilde{D}}) := - \int_{K_i}^{K_{i+1}} g_{\tilde{D}}(x) \ln g_{\tilde{D}}(x) dx = - \ln \alpha_i \int_{K_i}^{K_{i+1}} g_{\tilde{D}}(x) dx - \beta_i \int_{K_i}^{K_{i+1}} x g_{\tilde{D}}(x) dx$$

Assim, de (3.1.4) e (3.1.6) segue que

$$H_i(g_{\tilde{D}}) = -p_i(\ln \alpha_i + \beta_i \bar{K}_i)$$

Usando (3.1.6), temos que, para cada $i = 0, 1, \dots, n$

$$H_i(g_{\tilde{D}}) = -p_i(\ln p_i - c_i(\beta_i) + \beta_i \bar{K}_i)$$

Finalmente, de (3.1.7) e (3.1.3) temos que, para cada $i = 0, 1, \dots, n$:

$$H_i(g_{\tilde{D}}) = -p_i \ln p_i - p_i c_i^*(\bar{K}_i)$$

Com isso, a entropia em $[0, \infty[$, que pode ser vista como a soma das entropias em cada intervalo $[K_i, K_{i+1}[$, é encontrada somando-se os termos da última equação e o resultado segue. \square

Assim, a expressão de $H(g_{\tilde{D}})$ dada pelo Teorema 3.1 pode ser dividida em duas partes, uma discreta e outra contínua:

Discreta

$$H^d(g_{\tilde{D}}) := \sum_{i=0}^n H_i^d(g_{\tilde{D}}), \text{ onde } H_i^d(g_{\tilde{D}}) := -p_i \ln p_i, \quad \forall i = 0, \dots, n \quad (3.2.1)$$

Contínua

$$H^c(g_{\tilde{D}}) := \sum_{i=0}^n H_i^c(g_{\tilde{D}}), \text{ onde } H_i^c(g_{\tilde{D}}) := -p_i c_i^*(\bar{K}_i), \quad \forall i = 0, \dots, n \quad (3.2.2)$$

O termo $H^d(g_{\tilde{D}})$ pode ser visto como uma entropia discreta de Shannon, cujo valor máximo ocorre quando todos os p_i são iguais, ou seja, quando todos os intervalos possuem igual probabilidade de ocorrer. Já o termo $H^c(g_{\tilde{D}})$, para $i < n$, atinge seu máximo quando cada \bar{K}_i está no ponto médio do seu intervalo correspondente, ou seja, quando o valor esperado de $S(T)$ em cada intervalo é o seu ponto médio.

Com isso, conseguimos mostrar que, fixados os preços das opções de compra, $H(g_{\tilde{D}})$, $H^d(g_{\tilde{D}})$ e $H^c(g_{\tilde{D}})$ só dependem dos preços das opções digitais. Este é um passo fundamental para chegarmos ao algoritmo central do trabalho. Os próximos teoremas mostrarão que a função $H(g_{\tilde{D}})$ do Teorema 3.1 possui ponto crítico, que ele é único e de máximo.

Teorema 3.2. *Como função das opções digitais, $H : \Omega \rightarrow \mathbb{R}$ é diferenciável e, para todo $\tilde{D} \in \Omega$, temos que*

$$\frac{\partial H}{\partial \tilde{D}_i}(\tilde{D}) = \ln g_{\tilde{D}}(K_i^-) - \ln g_{\tilde{D}}(K_i^+), \quad \forall i = 1, \dots, n,$$

onde

$$g_{\tilde{D}}(K_i^-) := \lim_{x \rightarrow K_i^-} g_{\tilde{D}}(x) = \alpha_{i-1} e^{\beta_{i-1} K_i} \quad e \quad g_{\tilde{D}}(K_i^+) := \lim_{x \rightarrow K_i^+} g_{\tilde{D}}(x) = \alpha_i e^{\beta_i K_i}$$

com α_i e β_i dados pelas equações (3.1.6) e (3.1.7), respectivamente, para todo $i = 0, \dots, n$

Demonstração. Seja $\tilde{D} \in \Omega$. Para cada $i = 0, 1, \dots, n$, sejam p_i , \bar{K}_i , α_i e β_i conforme as equações (3.1.4), (3.1.6), (3.1.6) e (3.1.7), respectivamente. Fixando $i \in 1, \dots, n$ e usando (3.1.4) e (3.1.6), temos que:

$$\begin{aligned} \frac{\partial p_i}{\partial \tilde{D}_i} &= 1, \\ \frac{\partial p_{i-1}}{\partial \tilde{D}_i} &= -1, \\ \frac{\partial \bar{K}_{i-1}}{\partial p_{i-1}} &= \frac{K_{i-1} - \bar{K}_{i-1}}{p_{i-1}} \quad e \\ \frac{\partial \bar{K}_i}{\partial p_i} &= \frac{K_i - \bar{K}_i}{p_i}. \end{aligned}$$

Começando pela parte discreta, de (3.2.1), temos que

$$\begin{aligned} \frac{\partial H^d}{\partial \tilde{D}_i}(\tilde{D}) &= \frac{\partial}{\partial \tilde{D}_i}[-p_{i-1} \ln p_{i-1} - p_i \ln p_i] \\ &= \frac{\partial p_{i-1}}{\partial \tilde{D}_i}(-\ln p_{i-1} - 1) + \frac{\partial p_i}{\partial \tilde{D}_i}(-\ln p_i - 1) \\ &= \ln p_{i-1} - \ln p_i \end{aligned} \tag{3.2.3}$$

Já para a parte contínua, de (3.2.2), temos que

$$\begin{aligned} \frac{\partial H^c}{\partial \tilde{D}_i}(\tilde{D}) &= \frac{\partial}{\partial \tilde{D}_i}[-p_{i-1}c_{i-1}^*(\bar{K}_{i-1}) - p_i c_i^*(\bar{K}_i)] \\ &= \frac{\partial p_{i-1}}{\partial \tilde{D}_i}(-c_{i-1}^*(\bar{K}_{i-1}) - p_{i-1}[c_{i-1}^*]'(\bar{K}_{i-1})\frac{\partial \bar{K}_{i-1}}{\partial p_{i-1}}) \\ &\quad + \frac{\partial p_i}{\partial \tilde{D}_i}(-c_i^*(\bar{K}_i) - p_i[c_i^*]'(\bar{K}_i)\frac{\partial \bar{K}_i}{\partial p_i}) \\ &= c_{i-1}^*(\bar{K}_{i-1}) + p_{i-1}[c_{i-1}^*]'(\bar{K}_{i-1})\left(\frac{\bar{K}_{i-1} - K_{i-1}}{p_{i-1}}\right) \\ &\quad - c_i^*(\bar{K}_i) - p_i[c_i^*]'(\bar{K}_i)\left(\frac{\bar{K}_i - K_i}{p_i}\right) \end{aligned} \tag{3.2.4}$$

Assim, usando (3.1.3) e (3.1.7) chegamos a $[c_{i-1}^*]'(\bar{K}_{i-1}) = \beta_{i-1}$ e $[c_i^*]'(\bar{K}_i) = \beta_i$. Portanto,

$$\frac{\partial H^c}{\partial \tilde{D}_i}(\tilde{D}) = c_{i-1}^*(\bar{K}_{i-1}) - \beta_{i-1}(\bar{K}_{i-1} - K_{i-1}) - c_i^*(\bar{K}_i) - \beta_i(K_i - \bar{K}_i) \tag{3.2.5}$$

Novamente, usando (3.1.3) temos que

$$\frac{\partial H^c}{\partial \tilde{D}_i}(\tilde{D}) = -c_{i-1}(\beta_{i-1}) + \beta_{i-1}K_i + c_i(\beta_i) - \beta_iK_i \tag{3.2.6}$$

Juntando as equações (3.2.3) e (3.2.4) e usando a equação (3.1.6) a demonstração se conclui. \square

Corolário 1. Para qualquer $\tilde{D} \in \Omega$ a densidade $g_{\tilde{D}} \in \mathcal{G}$ é contínua se, e somente se, \tilde{D} é ponto crítico de $H : \Omega \rightarrow \mathbb{R}$, isto é

$$\frac{\partial H}{\partial \tilde{D}_i}(\tilde{D}) = 0, \quad \forall i = 1, \dots, n$$

Demonstração. Isto é consequência direta do Teorema 3.2 e do fato de que em cada intervalo $g_{\tilde{D}}$ é exponencial. \square

Neste ponto da teoria se faz necessário relembrar o que tem sido feito até o momento e quais são os próximos passos. Para encontrar a DME, conseguimos um método simples numericamente, mas com uma limitação séria que era a necessidade dos preços de opções digitais com os mesmos strikes das opções de compra. A partir desta dificuldade, começamos a estudar o papel dessas opções digitais no valor da entropia até conseguirmos explicitar, para um dado conjunto de strikes com seus respectivos valores de opção de compra, H em termos de \tilde{D} .

Na verdade, estamos construindo um algoritmo que será capaz de obter a DME usando somente os preços das opções de compra, sem as opções digitais. Para tal, falta mostrar que, de fato, o ponto crítico de $H(\tilde{D})$ é de máximo e encontrar o formato de sua matriz Hessiana.

Vamos começar pelo formato da matriz Hessiana.

Teorema 3.3. *A função $H : \Omega \rightarrow \mathbb{R}$ é duas vezes diferenciável e sua matriz Hessiana, em qualquer $\tilde{D} \in \Omega$, é simétrica e tridiagonal com entradas dada por*

$$\begin{aligned} \frac{\partial^2 H}{\partial \tilde{D}_i^2}(\tilde{D}) &= -\frac{1}{p_{i-1}} - \frac{1}{p_i} - \frac{(K_i - \bar{K}_{i-1})^2}{p_{i-1}c''_{i-1}(\beta_{i-1})} - \frac{(K_i - \bar{K}_i)^2}{p_i c''_i(\beta_i)}, \quad \forall i = 1, \dots, n \\ \frac{\partial^2 H}{\partial \tilde{D}_i \partial \tilde{D}_{i+1}}(\tilde{D}) &= \frac{1}{p_i} - \frac{(\bar{K}_i - K_i)(K_{i+1} - \bar{K}_i)}{p_i c''_i(\beta_i)} \quad \forall i = 1, \dots, n-1 \end{aligned}$$

onde p_i , \bar{K}_i e β_i são dados por (3.1.4), (3.1.6) e (3.1.7), respectivamente, para todo $i = 0, 1, \dots, n$.

Demonstração. Seja $\tilde{D} \in \Omega$. Para cada $i = 0, 1, \dots, n$, seja p_i , \bar{K}_i e β_i dadas pelas equações (3.1.4), (3.1.6) e (3.1.7), respectivamente. O fato de $H''(\tilde{D})$ ser tridiagonal e simétrica é facilmente observado a partir do Teorema 3.2. Fixando um $i \in \{1, 2, \dots, n\}$ vamos calcular a segunda derivada de H^d . Para a diagonal, temos que

$$\frac{\partial^2 H^d}{\partial \tilde{D}_i^2}(\tilde{D}) = \frac{\partial}{\partial \tilde{D}_i} [\ln p_{i-1} - \ln p_i] = \frac{1}{p_{i-1}} \frac{\partial p_{i-1}}{\partial \tilde{D}_i} - \frac{1}{p_i} \frac{\partial p_i}{\partial \tilde{D}_i} = -\frac{1}{p_{i-1}} - \frac{1}{p_i}.$$

Se $i < n$, as demais entradas serão

$$\frac{\partial^2 H^d}{\partial \tilde{D}_i \partial \tilde{D}_{i+1}}(\tilde{D}) = -\frac{\partial}{\partial \tilde{D}_{i+1}} [\ln p_i] = -\frac{1}{p_i} \frac{\partial p_i}{\partial \tilde{D}_{i+1}} = \frac{1}{p_i}.$$

Já para a parte contínua, temos que:

$$\begin{aligned} \frac{\partial^2 H^c}{\partial \tilde{D}_i^2}(\tilde{D}) &= \frac{\partial}{\partial \tilde{D}_i} [-c_{i-1}(\beta_{i-1}) + \beta_{i-1}K_i + c_i\beta_i - \beta_iK_i] \\ &= c'_{i-1}\beta_i - 1 \frac{\partial \beta_{i-1}}{\partial \tilde{D}_i} + \frac{\partial \beta_{i-1}}{\partial \tilde{D}_i} K_i + c'_i\beta_i \frac{\partial \beta_i}{\partial \tilde{D}_i} - \frac{\partial \beta_i}{\partial \tilde{D}_i} K_i \\ &= (K_i - \bar{K}_{i-1}) \frac{\partial \beta_{i-1}}{\partial \tilde{D}_i} + (\bar{K}_i - K_i) \frac{\partial \beta_i}{\partial \tilde{D}_i} \\ &= -\frac{(K_i - \bar{K}_{i-1})^2}{p_{i-1}c''_{i-1}(\beta_{i-1})} - \frac{(K_i - \bar{K}_i)^2}{p_i c''_i(\beta_i)} \end{aligned} \tag{3.2.7}$$

e se $i < n$, temos que

$$\begin{aligned}
\frac{\partial^2 H^c}{\partial \tilde{D}_i \partial \tilde{D}_{i+1}}(\tilde{D}) &= \frac{\partial}{\partial \tilde{D}_{i+1}} [c_i(\beta_i) - \beta_i K_i] \\
&= (c'_i(\beta_i) - K_i) \frac{\partial \beta_i}{\partial \tilde{D}_{i+1}} \\
&= -\frac{(\bar{K}_i - K_i)(K_i - \bar{K}_{i+1})}{p_i c''_i(\beta_i)}.
\end{aligned} \tag{3.2.8}$$

Cabe ressaltar que nas passagens anteriores nós utilizamos as relações

$$\begin{aligned}
\frac{\partial \beta_{i-1}}{\partial \tilde{D}_i} &= -\frac{(K_i - \bar{K}_{i-1})}{p_{i-1} c''_{i-1}(\beta_{i-1})}, \\
\frac{\partial \beta_i}{\partial \tilde{D}_i} &= \frac{(K_i - \bar{K}_i)}{p_i c''_i(\beta_i)} \quad e \\
\frac{\partial \beta_i}{\partial \tilde{D}_{i+1}} &= -\frac{(K_{i+1} - \bar{K}_i)}{p_i c''_i(\beta_i)},
\end{aligned}$$

que são obtidas derivando a equação $c'_{i-1}(\beta_{i-1}) = \bar{K}_{i-1}$ em termos de \tilde{D}_i e a equação $c'_i(\beta_i) = \bar{K}_i$ em termos de \tilde{D}_i e \tilde{D}_{i+1} . \square

Teorema 3.4. *A função $H : \Omega \rightarrow \mathbb{R}$ é estritamente côncava.*

Demonstração. Seja $\tilde{D} \in \Omega$. É suficiente mostrar que $H''(\tilde{D})$ é estritamente negativa definida, ou seja, $\langle H''(\tilde{D})d, d \rangle < 0$, se $d \neq 0$. Para tal, provaremos que $[H^d]''(\tilde{D})$ é estritamente negativa definida e que $[H^c]''(\tilde{D})$ é negativa definida, ou seja, $\langle H''(\tilde{D})d, d \rangle \leq 0$, se $d \neq 0$.

Para cada $i = 0, 1, \dots, n$, seja p_i , \bar{K}_i e β_i dadas pelas equações (3.1.4), (3.1.6) e (3.1.7), respectivamente.

Lembrando que se uma matriz $A = (a_{i,j})$ de ordem n é simétrica, então vale que

$$\langle Ad, d \rangle = \sum_{i=1}^n a_{i,i} d_i^2 + 2 \sum_{i=1}^{n-1} a_{i,i+1} d_i d_{i+1} \quad \forall d = (d_1, \dots, d_n) \in \mathbb{R}^n.$$

Aplicando essa última relação para $A = [H^d]''(\tilde{D})$, com $d \neq 0$, obtemos a seguinte estimativa:

$$\begin{aligned}
\langle [H^d]''(\tilde{D})d, d \rangle &= -\sum_{i=1}^n \left(\frac{1}{p_{i-1}} + \frac{1}{p_i} \right) d_i^2 + 2 \sum_{i=1}^{n-1} \frac{d_i d_{i+1}}{p_i} \\
&= -\frac{d_1^2}{p_0} - \sum_{i=1}^{n-1} \frac{d_i^2 + d_{i+1}^2}{p_i} - \frac{d_n^2}{p_n} + 2 \sum_{i=1}^{n-1} \frac{d_i d_{i+1}}{p_i} \\
&= -\frac{d_1^2}{p_0} - \sum_{i=1}^{n-1} \frac{(d_i - d_{i+1})^2}{p_i} - \frac{d_n^2}{p_n} < 0.
\end{aligned} \tag{3.2.9}$$

Agora, usando o mesmo procedimento para $A = [H^c]''(\tilde{D})$, temos que

$$\begin{aligned}
\langle [H^c]''(\tilde{D})d, d \rangle &= - \sum_{i=1}^n \left[\frac{(K_i - \bar{K}_{i-1})^2}{p_{i-1}c''_{i-1}(\beta_{i-1})} + \frac{(K_i - \bar{K}_i)^2}{p_i c''_i(\beta_i)} \right] d_i^2 \\
&\quad - 2 \sum_{i=1}^{n-1} \frac{(\bar{K}_i - K_i)(K_{i+1} - \bar{K}_i)d_i d_{i+1}}{p_i c''_i(\beta_i)} \\
&= - \frac{(K_1 - \bar{K}_0)^2 d_1^2}{p_0 c''_0(\beta_0)} \\
&\quad - \sum_{i=1}^{n-1} \frac{[(\bar{K}_i - K_i)d_i + (K_{i+1} - \bar{K}_i)d_{i+1}]^2}{p_i c''_i(\beta_i)} \\
&\quad - \frac{(K_n - \bar{K}_n)^2 d_n^2}{p_n c''_n(\beta_n)} \leq 0. \tag{3.2.10}
\end{aligned}$$

Portanto, das duas últimas inequações e do Teorema 3.1, podemos dizer que

$$\langle [H]''(\tilde{D})d, d \rangle = \langle [H^c]''(\tilde{D})d, d \rangle + \langle [H^d]''(\tilde{D})d, d \rangle < 0, \quad \forall d \neq 0$$

□

O corolário a seguir pode ser visto como um resumo de tudo que foi feito neste capítulo. Para encontrar a DME só é necessário saber os preços das opções de compra em cada strike, pois o vetor \tilde{D} , que possui os preços das opções digitais, é ponto de máximo da função entropia H .

Corolário 2. *São equivalentes:*

- (i) $g_{\tilde{D}} \in \mathcal{G}$ é uma DME.
- (ii) $\tilde{D} \in \Omega$ é um ponto crítico de H .
- (iii) $g_{\tilde{D}} \in \mathcal{G}$ é contínua.

Demonstração. Já foi visto que $(i) \Rightarrow (ii)$, pois ao encontrar a densidade de máxima entropia a função H está no seu ponto crítico. Em [12] mostra-se que $(ii) \Leftrightarrow (iii)$ e, finalmente, o último teorema mostra que $(iii) \Rightarrow (i)$. □

Observe que o objetivo do capítulo foi cumprido, ou seja, é possível obter a DME somente com os preços das opções de compra. Isto ocorre porque o preço das opções digitais são encontrados a partir do ponto crítico de $H : \Omega \rightarrow \mathbb{R}$. Assim, estamos prontos para elaborar um algoritmo capaz de encontrar somente a partir dos preços das opções de compra. Este será o tema do próximo capítulo.

Capítulo 4

O Algoritmo

O corolário do Teorema 3.4 mostrou que, se $\tilde{D} \in \Omega$ é uma raiz de H' , então $g_{\tilde{D}}$ é uma DME. Como a hessiana de H é conhecida analiticamente, o método de Newton-Raphson é uma possível escolha de algoritmo para encontrar a raiz de H' numericamente e, conseqüentemente, a DME.

A escolha do método de Newton-Raphson possui algumas restrições. Porém, no problema em questão essas restrições são eliminadas se utilizarmos uma outra versão conhecida como Método de Newton Modificado. Tal método é apresentado em [3], página 487.

Em geral há dois tipos de problema no método de Newton: i) encontrar a raiz errada no caso de haver mais de uma e ii) alto custo computacional pelo fato da hessiana possuir n^2 entradas. Porém, conforme já foi visto, nenhum dos dois problemas ocorrem no método em questão porque H só possui um ponto crítico e sua hessiana é tridiagonal e simétrica, reduzindo o número de entradas para $2n - 1$

Em cada passo do método de Newton para encontrar a raiz $d \in \mathbb{R}^n$ de H' deve-se resolver o sistema linear $H''(\tilde{D})d = H'(\tilde{D})$, onde $\tilde{D} \in \Omega$ é o palpite corrente de solução. Portanto o método falharia se $H''(\tilde{D})$ não fosse inversível. Porém, conforme demonstrado no Teorema 3.4, H é fortemente côncava, ou seja, $\langle H''(\tilde{D})d, d \rangle < 0$, se $d \neq 0$. Isso garante que $H''(\tilde{D})$ é inversível, segundo [10].

Após toda a base teórica apresentada e condirações feitas, podemos resumir o método prático de obtenção da DME nos seguintes passos:

(Passo 1) Coleta dos dados de mercado:

Para um determinado ativo e com uma dada maturidade T , coletamos os preços de mercado para a opção de compra correspondente a Strikes dados.

(Passo 2) Validação dos dados:

Testamos, para cada preço da opção, a condição de ausência de arbitragem. Ou seja:

$$-\frac{\tilde{C}_i - \tilde{C}_{i-1}}{K_i - K_{i-1}} > \tilde{D}_i > -\frac{\tilde{C}_{i+1} - \tilde{C}_i}{K_{i+1} - K_i}$$

Onde \tilde{D}_i é calculado usando-se a derivada discreta centralizada de \tilde{C}_i . Assim sendo, usamos:

$$\tilde{D}_i = -\frac{\tilde{C}_{i+1} - \tilde{C}_{i-1}}{K_{i+1} - K_{i-1}}$$

Valores que não satisfazem a condição de não arbitragem são retirados e novos vetores de Strikes e de preços da opção são criados somente com os dados válidos.

(Passo 3) Construção de $H''(\tilde{D})$ e $H'(\tilde{D})$:

A partir de \tilde{D} gerado pela derivada discreta centralizada de \tilde{C}_i , que será o "chute inicial" no método de Newton Rapson, construímos $H''(\tilde{D})$ e $H'(\tilde{D})$ com as fórmulas dadas pelos Teoremas 3.2 e 3.3.

(Passo 4) Utilização do método de Newton Rapson presente em [3] para encontrar o ponto crítico de $H'(\tilde{D})$: Para obter \tilde{D} ótimo resolvemos, em cada etapa do método de Newton, o sistema

$$H''(\tilde{D})d = H'(\tilde{D}).$$

Onde \tilde{D} é valor corrente da opção digital e d é o valor de incremento de \tilde{D} . Desta forma, a cada passo nós fazemos $\tilde{D}_i = \tilde{D}_i + 0.5d_i$. Além disso, o valor de controle de condição de parada, chamado de λ , é dado pela equação $H'(\tilde{D})^T d$.

(Passo 5) Obtenção de cada α_i e β_i :

Com o \tilde{D} ótimo obtemos cada α_i e β_i resolvendo numericamente as equações

$$\alpha_i = p_i e^{-c_i(\beta_i)} \quad e$$

$$c'_i(\beta_i) = \bar{K}_i$$

Capítulo 5

Exemplos Numéricos

Dois tipos de testes foram feitos: um com dados sintéticos gerados a partir do modelo de Black e Scholes e outro com dados reais retirados do Boletim Diário da BMEF Bovespa [2] de opções de compra sobre o preço de Dollar.

A ideia é comparar os resultados obtidos com preços gerados pelo modelo de Black e Scholes, chamados aqui de preços sintéticos, com os resultados obtidos utilizando os preços reais retirados do mercado a partir de [2]. Para tal, usaremos os strikes retirados de [2] para gerar preços de Black e Scholes.

Todos os cálculos estão associados à expressão da DME $g(x)$ deduzida no Teorema 2.2. Mais precisamente, dado $x \in [K_i, K_{i+1}[$, então a DME é dada por

$$g(x) = \alpha_i e^{\beta_i x},$$

em que cada α_i e β_i são dados pelas equações (2.1.2) e (2.1.3).

Com o exposto acima em mente encontramos a DME com o algoritmo do Capítulo 4. Além disso, baseado em [17], também faremos uma estimativa da incerteza associada à calibragem dos parâmetros que formam a DME. Ou seja, cada α_i e β_i .

Assim, a partir da expressão da DME nós calculamos (usando Matlab e com expressões dadas em [11] e [12]):

(i) Volatilidade Implícita

Função *blsimpv*($F, K, r, T, C(K)$). Onde F é valor corrente do ativo, K um strike arbitrário escolhido, r a taxa Selic, T o vencimento em anos e $C(K)$ o preço da opção de compra retirada do mercado em K .

(ii) Preço da opção $\tilde{C}(K)$ para qualquer strike K (interpolação ou extrapolação)

Conforme [11], o preço descontado da opção de compra em algum strike arbitrário $K \in [K_i, K_{i+1}[$ é dado por:

$$\tilde{C}(K) = \begin{cases} \tilde{C}_i - (K - K_i) \left(\tilde{D}_i + \frac{\alpha_i}{\beta_i} e^{\beta_i K_i} \right) + \frac{\alpha_i}{\beta_i^2} (e^{\beta_i K_i} - e^{\beta_i K}) & \beta_i \neq 0 \\ \tilde{C}_i - (K - K_i) \tilde{D}_i + \frac{\alpha_i}{2} (K - K_i)^2 & \beta_i = 0 \end{cases}$$

(iii) Derivada de $\tilde{C}(K)$ em termos de K ou Opção Digital (binária)

Este é um parâmetro essencial no teste de ausência de arbitragem. Novamente, segundo [11], temos que:

$$\tilde{C}'(K) = -\tilde{D}(K) = \begin{cases} -\left(\tilde{D}_i - \frac{\alpha_i}{\beta_i} (e^{\beta_i K_i} - e^{\beta_i K}) \right) & \beta_i \neq 0 \\ -\tilde{D}_i + \alpha_i (K - K_i) & \beta_i = 0 \end{cases}$$

(iv) Segunda derivada de $\tilde{C}(K)$ em termos de K

Derivando em termos de K as expressões do item anterior, obteremos:

$$\tilde{C}''(K) = \begin{cases} \alpha_i e^{\beta_i K} & \beta_i \neq 0 \\ \alpha_i & \beta_i = 0 \end{cases}$$

(v) Função Cumulativa de $g(x)$

Em um dado $x = K$, teremos:

$$G(K) = \begin{cases} 1 - \tilde{D}_i + \frac{\alpha_i}{\beta_i} (e^{\beta_i K_i} - e^{\beta_i K}) & \beta_i \neq 0 \\ 1 - \tilde{D}_i + \alpha_i (K - K_i) & \beta_i = 0 \end{cases}$$

Além disso, cada item citado acima será calculado para diferentes strikes e diferentes maturidades para que sejam gerados gráficos de *volatilidade implícita*, *derivada primeira de $C(K)$* e *função densidade de probabilidade*.

5.1 Testes

Para os testes com valores sintéticos, os seguintes parâmetros foram usados:

(a) $\sigma = 0.25$

(b) Taxa Selic

$$r = 0.11$$

(c) Valor da cotação do Dólar em 8 de maio de 2014

$$F = R\$2.219$$

(d) Vetor de Maturidades

Os vencimentos usados (em anos) foram $T = \frac{1}{365}[23 \ 53 \ 83]$. Cabe ressaltar que estes valores foram retirados de [2].

(e) Vetores de Stikes retirados do mercado

Novamente, a partir de valores retirados de [2], usamos os seguintes vetores de strike (em reais):

$$K1 = [1.975 \ 2.000 \ 2.025 \ 2.200 \ 2.225 \ 2.250 \ 2.275]$$

$$K2 = [2.300 \ 2.425 \ 2.500]$$

$$K3 = [2.275 \ 2.400]$$

Cabe ressaltar que, com exceção do valor $\sigma = 0.25$, todos os demais valores também serão usados nos testes com valores reais.

5.1.1 Valores Sintéticos

Para este tópico a função "*entropiasintetica(K, T, F, strike)*", criada em Matlab e cujo código se encontra no Apêndice, foi criada para cálculo da DME, dos preços extrapolados para diversos strikes e da volatilidade implícita. Com isto, geramos, para cada item, gráficos comparando os valores extrapolados com os valores iniciais vindos dos strikes originais de [2]. Para gerar os gráficos, usamos a função:

"*Plotagem(strikeextrapolado, strikeinicial, CKextrapolado, CKinicial, volextrapolado, volinicial)*", cujo código se encontra no apêndice.

5.1.2 Valores Reais

Agora, duas funções criadas em Matlab serão utilizadas: "testearbitragem(C,K,T)" e "entropiareal(C,K,T,F,strike)". Assim, no uso de ambas as funções, além dos parâmetros citados no início desta Seção, também usaremos os seguintes vetores iniciais de preços de opções de compra (em reais):

$$C1 = [0.2370 \ 0.0370 \ 0.0229 \ 0.0134 \ 0.0078]$$

$$C2 = [0.0169 \ 0.0026 \ 0.006]$$

$$C3 = [0.0400 \ 0.0120]$$

Continuando, a primeira função tem por objetivo fazer o teste de condição de ausência de arbitragem observado em (3.1.1). Mais precisamente, vimos nos Capítulos 2 e 3 que a hipótese de ausência de arbitragem é fundamental para a existência da DME deduzida. Além disso, na Seção 3.1, vimos que a ausência de arbitragem ocorre quando

$$-\frac{\tilde{C}_i - \tilde{C}_{i-1}}{K_i - K_{i-1}} > \mathbb{E}^g[I_{S(T) > K_i}] > -\frac{\tilde{C}_{i+1} - \tilde{C}_i}{K_{i+1} - K_i} \quad \forall i = 1, 2, \dots, n.$$

Assim sendo, para estimar o valor de $\mathbb{E}^g[I_{S(T) > K_i}]$, que é o preço da opção digital ou a derivada de $\tilde{C}(K)$ em termos de K , usamos a derivada numérica centralizada presente na página 298 de [5], isto é,

$$\mathbb{E}^g[I_{S(T) > K_i}] = \tilde{C}'(K_i) = -\frac{\tilde{C}_{i+1} - \tilde{C}_{i-1}}{K_{i+1} - K_{i-1}}.$$

Com isso, se em algum strike K_i a condição (3.1.1) não for satisfeita, então este strike e seu respectivo valor da opção de compra são excluídos.

A segunda função, "entropiareal(C,K,T,F,strike)", usa somente os strikes com preços satisfazendo (3.1.1), que foram testados pela primeira função, para gerar a DME, preços extrapolados de opção de compra e volatilidade implícita. Com estes valores, geramos gráficos comparando os valores extrapolados com os valores iniciais vindos dos strikes de [2]. Mais uma vez todos os códigos das funções usadas estão anexadas no Apêndice.

5.2 Análise de Incerteza dos Parâmetros

Nesta seção faremos uma análise da incerteza dos parâmetros encontrados, ou seja, cada α_i e β_i . Para tal, nos basearemos em [17].

Ao usar os preços de opção de compra retirados do mercado temos um erro associado aos dados. Além disso, limitações numéricas também estão presentes quando utilizamos o algoritmo no Matlab. Em [17] há a apresentação de um método para estimativa de erros associados a experimentos nas mais diversas áreas da ciência.

É claro que este tipo de análise é mais rica quando um tratamento estatístico pode ser feito. Isto ocorre quando a calibração ou experimento pode ser repetido para diversas medições ou retiradas de dados. Além disso, quanto maior o número de medições, mais confiabilidade podemos obter. Entretanto, neste trabalho a fonte de dados (preços das opções de compra para determinados strikes) não nos permite este tratamento estatístico mais completo.

Porém, ainda é possível fazer uma estimativa da incerteza associada a calibração dos parâmetros e este é o objetivo desta seção.

5.2.1 Estimação da Incerteza

O objetivo central do trabalho, que é obter a densidade de probabilidade dos preços de um ativo pela maximização da entropia, pode ser reformulado da seguinte maneira:

Ponto de partida Preços de opções de compra retirado do mercado.

Função objetivo Determinar a DME é equivalente a determinar cada α_i e β_i .

Ou seja, podemos pensar que a função objetivo (cada α_i e β_i) depende de C ou, mais precisamente, de \tilde{C} .

Com o fato exposto acima em mente e baseado em [17], quando queremos avaliar a incerteza associada a calibração de uma função objetivo f que depende da medição de uma variável x (poderia ser mais de uma variável), definimos a incerteza (ou erro) como a diferença entre o valor verdadeiro e o calibrado.

Desta forma, segundo [17], temos que

$$f_{calibrado} - f_{real} = \frac{\partial f}{\partial x} (x_{medido} - x_{real}) \quad (5.2.1)$$

Assim sendo, $\forall i = 1, 2, \dots, n$, estimaremos a incerteza das funções $\alpha_i(\tilde{C})$ e $\beta_i(\tilde{C})$ usando (5.2.1):

$$\alpha_{calibrado} - \alpha_{real} = \frac{\partial \alpha}{\partial \tilde{C}} (\tilde{C}_{medido} - \tilde{C}_{real})$$

e

$$\beta_{calibrado} - \beta_{real} = \frac{\partial \beta}{\partial \tilde{C}} (\tilde{C}_{medido} - \tilde{C}_{real})$$

5.2.2 Resultados Práticos da Análise de Incerteza

Agora chegamos ao objetivo central desta seção, que é a análise da incerteza propriamente dita. Para tal, em cada α_i e β_i o erro depende do cálculo da derivada parcial da função objetivo em relação ao preço da opção de compra e do erro na medição deste preço.

Assim sendo, para cada i , faremos

$$\frac{\partial \alpha_i}{\partial \tilde{C}} = \frac{\alpha_{i+1} - \alpha_i}{\tilde{C}_{i+1} - \tilde{C}_i}$$

$$\frac{\partial \beta_i}{\partial \tilde{C}} = \frac{\beta_{i+1} - \beta_i}{\tilde{C}_{i+1} - \tilde{C}_i}$$

onde assumiremos que $\alpha_{n+1} = \beta_{n+1} = 0$.

Além disso, para estimar $\tilde{C}_{medido} - \tilde{C}_{real}$ usaremos uma estimativa do bid-ask spread, que é a diferença entre maior preço oferecido pelo comprador da opção e o menor preço disponível para a venda da mesma. A partir de [2], o bid-ask spread será dado pela média das diferenças dos últimos valores oferecidos de compra e venda da opção.

Os valores do bid-ask spread médio (em reais) para cada vencimento foi 0,00003, 0,00016 e 0,00017, respectivamente. A partir desses valores e da equação (5.2.1) um código em matlab (que está presente no apêndice) foi criado para encontrar o erro percentual de cada α_i e β_i .

Desta forma, obtivemos os seguintes erros percentuais:

Primeiro vencimento, $T = 23/365$ anos Erro percentualEm cada α

$$0,0210 \quad 8,0787 * 10^{67} \quad 0,0040 \quad 2,1956 \quad 7,4464 \quad 3,2966 * 10^{12}$$

Em cada β

$$0,0750 \quad 0,2085 \quad 0,0001 \quad 0,0001 \quad 7,4464 \quad 5,3462$$

Segundo vencimento, $T = 53/365$ anos Erro percentualEm cada α

$$0,0482 \quad 2,8284 \quad 5,2445 * 10^{11}$$

Em cada β

$$0,1527 \quad 2,8284 \quad 3,7071$$

Terceiro vencimento, $T = 83/365$ anos Erro percentualEm cada α

$$578,1334 \quad 4,5215 \quad 4,5215 * 10^{12}$$

Em cada β

$$0,1006 \quad 7,8393 \quad 18,2917$$

Observamos pelos resultados que os erros percentuais para os valores de β são aceitáveis. Com exceção do último valor de β para o terceiro vencimento, cujo erro foi de cerca de 18%, todos os demais ficaram abaixo de 8%.

Já nos erros percentuais de α , alguns valores perderam significância. Porém, em todos esses casos os valores correspondentes de α eram muito pequenos e os erros tão altos indicam que os parâmetros podem assumir o valor nulo.

Assim sendo, segundo [5], valores nulos (ou próximos disso) do parâmetro causam problemas no cálculo do erro percentual $\frac{\alpha_{calibrado} - \alpha_{real}}{\alpha_{real}}$. É o que [5] chama de função mal condicionada, algo que causa esta perda de significância do parâmetro (ou instabilidade).

5.3 Gráficos

5.3.1 Teste de Arbitragem

A seguir, para cada vetor $C1, C2, C3, K1, K2$ e $K3$, construiu-se o gráfico de $-\tilde{D}(K)$ antes e depois do teste (isso é equivalente a testar a condição (3.1.1)). Ou seja, no mesmo gráfico há a curva com os valores iniciais (antes do teste) e finais (depois do teste e já com a retirada dos valores fora da condição).

O indício de ausência de arbitragem, que é consequência de (3.1.1), é uma derivada de $\tilde{C}(k)$ estritamente decrescente. Como fizemos o gráfico de $-\tilde{D}(K)$, então devemos ver gráficos estritamente crescentes. Assim sendo, ficará claro que nos dois primeiros vencimentos alguns strikes tiveram de ser retirados, mas que no terceiro isso não foi necessário.

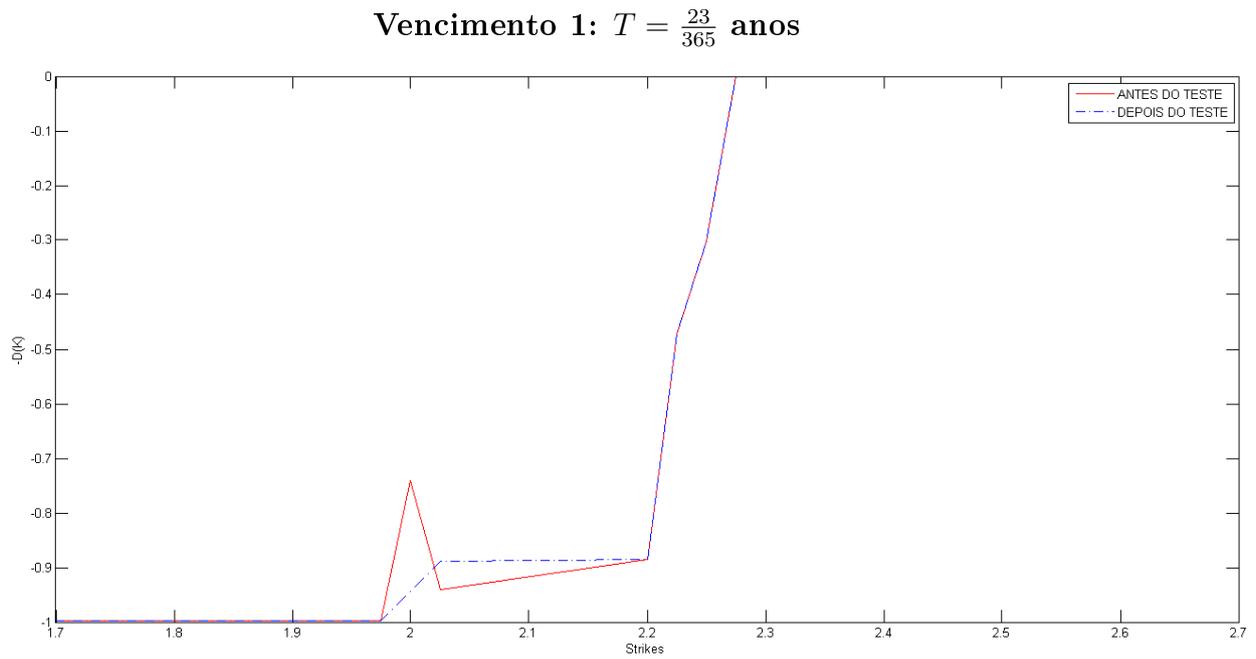


Figura 5.1: Primeiro teste de ausência de arbitragem

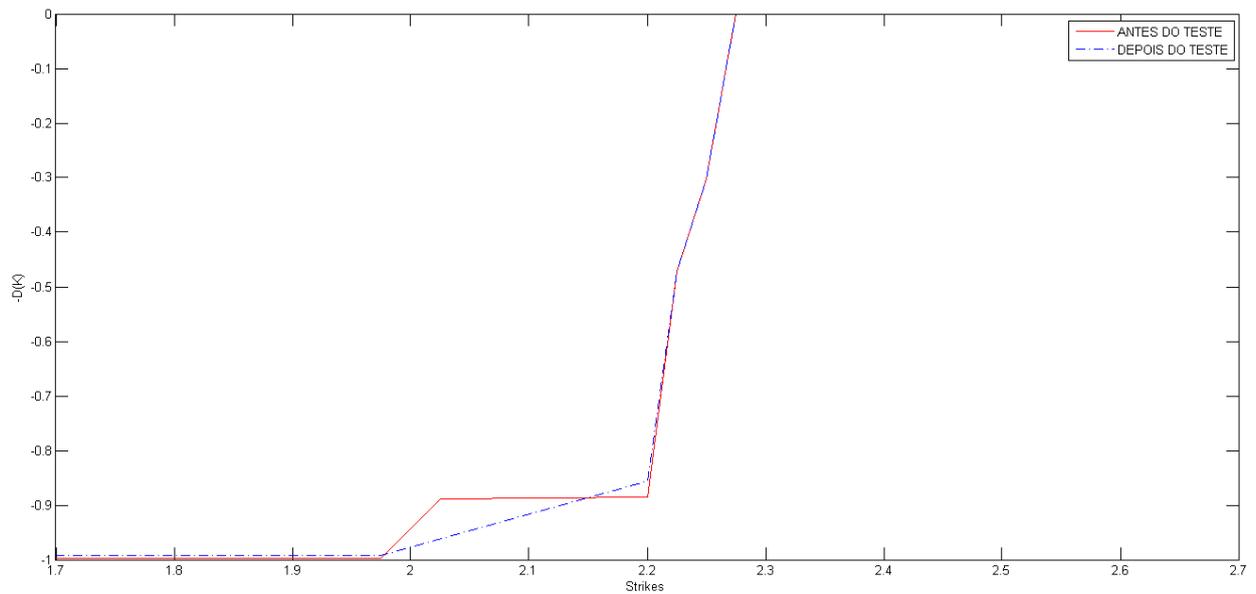


Figura 5.2: Segundo teste de ausência de arbitragem

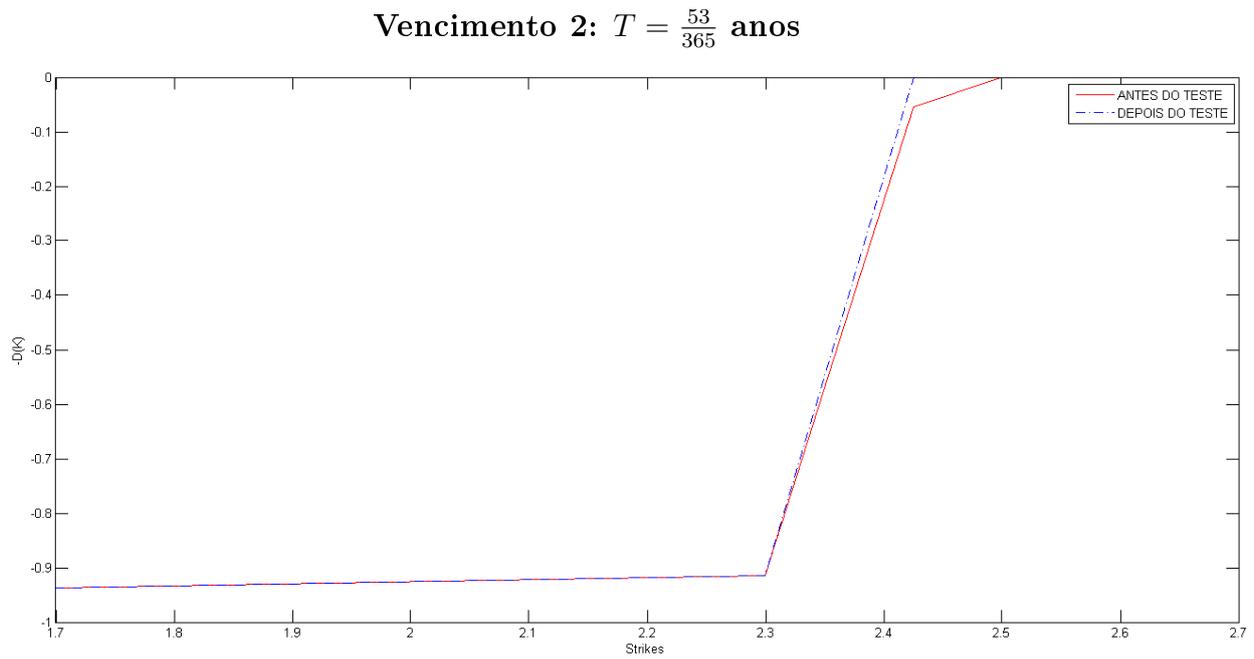


Figura 5.3: Terceiro teste de ausência de arbitragem

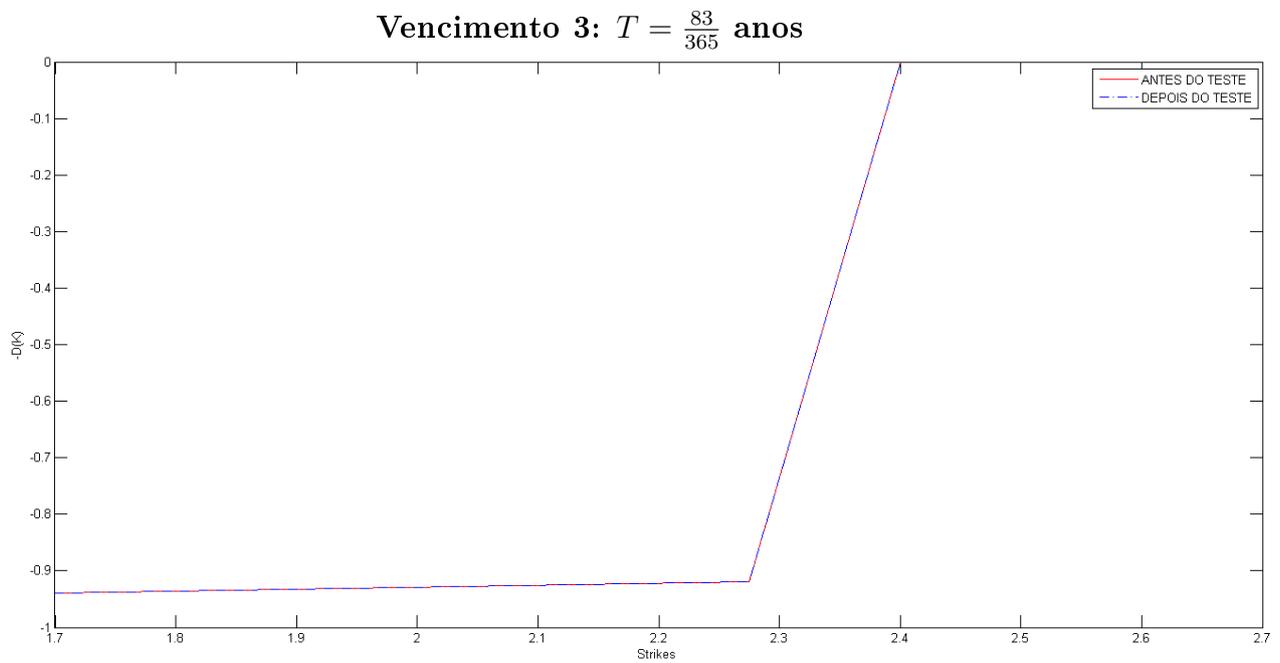


Figura 5.4: Quarto teste de ausência de arbitragem

5.3.2 Gráfico da Densidade

Após os testes feitos na Seção anterior, gerou-se a função densidade com os valores que satisfazem (3.1.1). A seguir, gráficos para cada maturidade foram gerados comparando os resultados com valores reais e sintéticos.

Nestes gráficos, para qualquer uma das maturidades, percebe-se uma disparidade razoável entre os valores para dados reais e sintéticos.

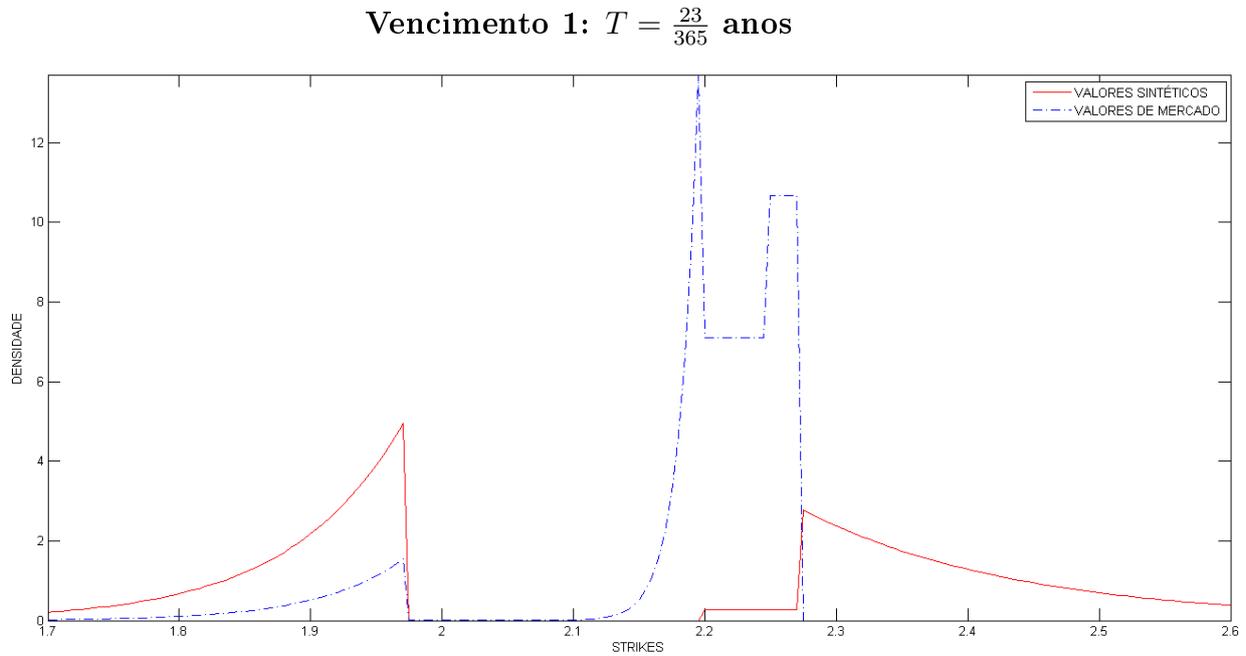


Figura 5.5: Primeiro gráfico da DME

Vencimento 2: $T = \frac{53}{365}$ anos

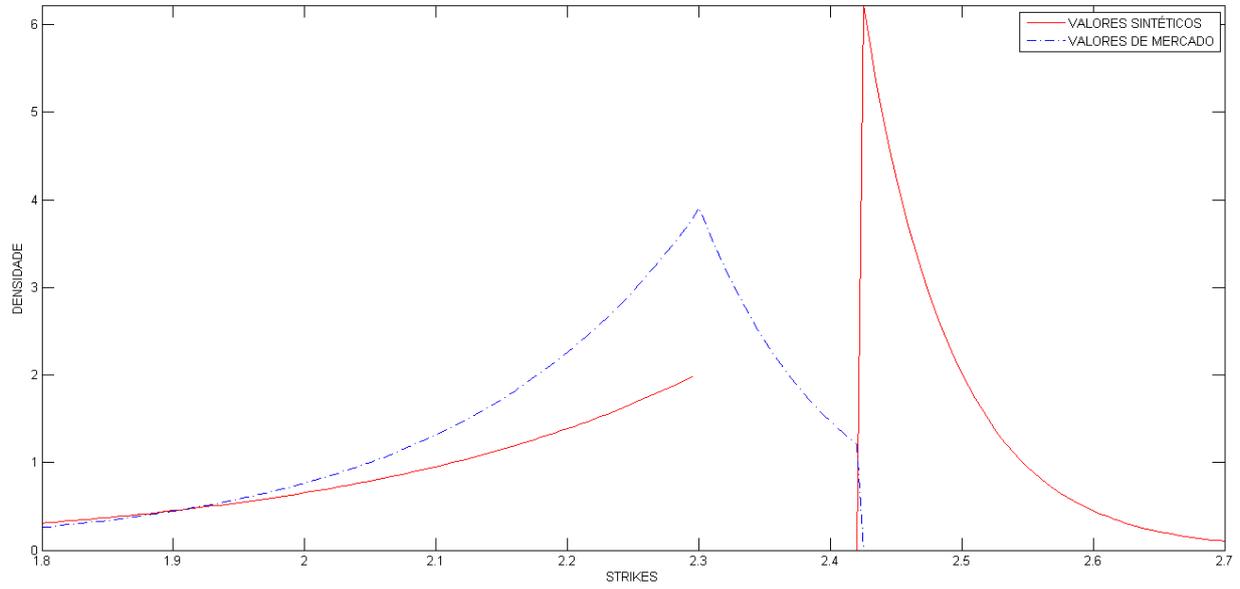


Figura 5.6: Segundo gráfico da DME

Vencimento 3: $T = \frac{83}{365}$ anos

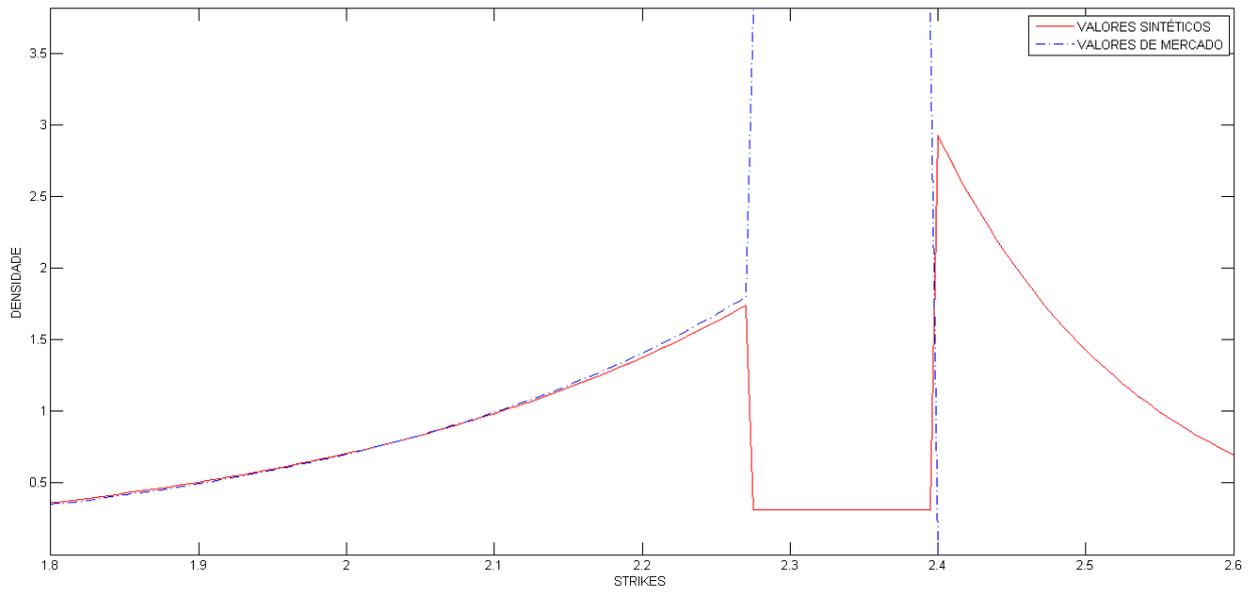


Figura 5.7: Terceiro gráfico da DME

5.3.3 Preços Extrapolados das Opções de Compra para Valores Sintéticos

Com a DME encontrada no item anterior, calculou-se o preço da opção de compra para uma faixa de strikes extrapolados. Assim sendo, em cada caso, construímos gráficos comparando os valores extrapolados com os valores iniciais de Black e Scholes calculados nos strikes originais vindos de [2].

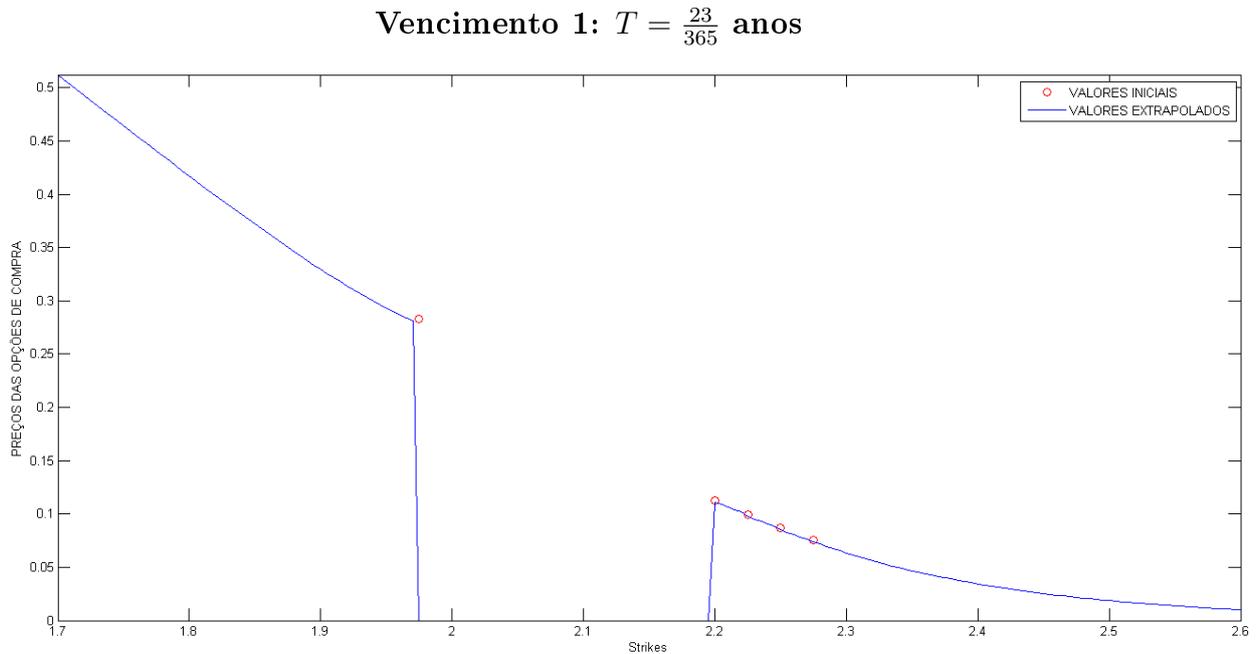


Figura 5.8: Preços sintéticos - maturidade 1

Vencimento 2: $T = \frac{53}{365}$ anos

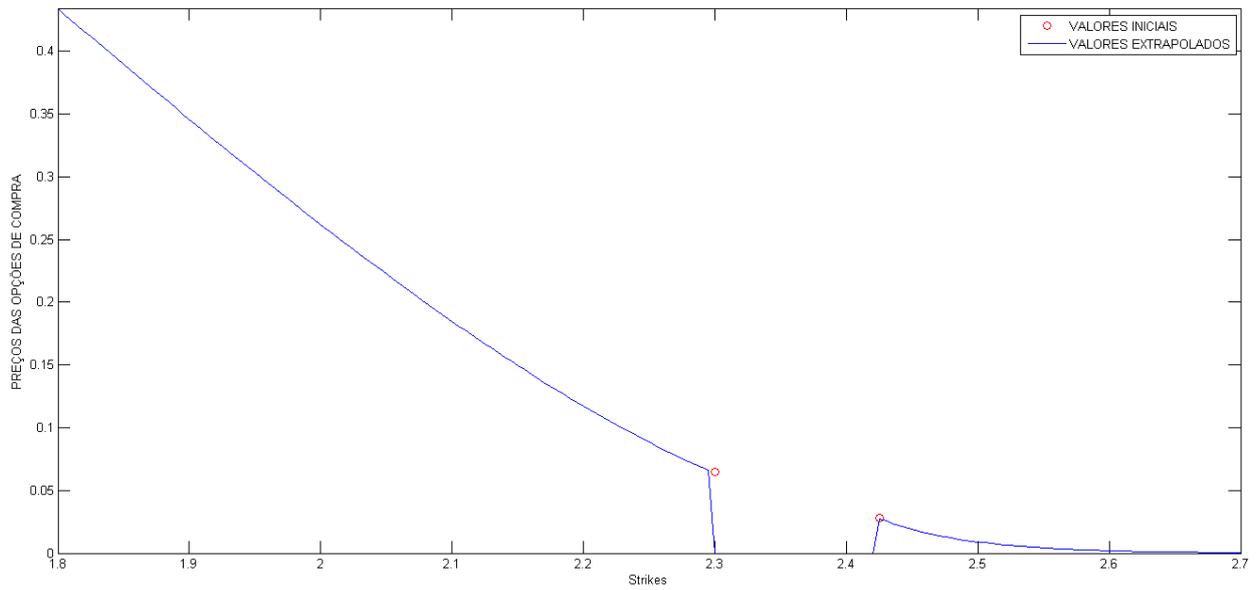


Figura 5.9: Preços sintéticos - maturidade 2

Vencimento 3: $T = \frac{83}{365}$ anos

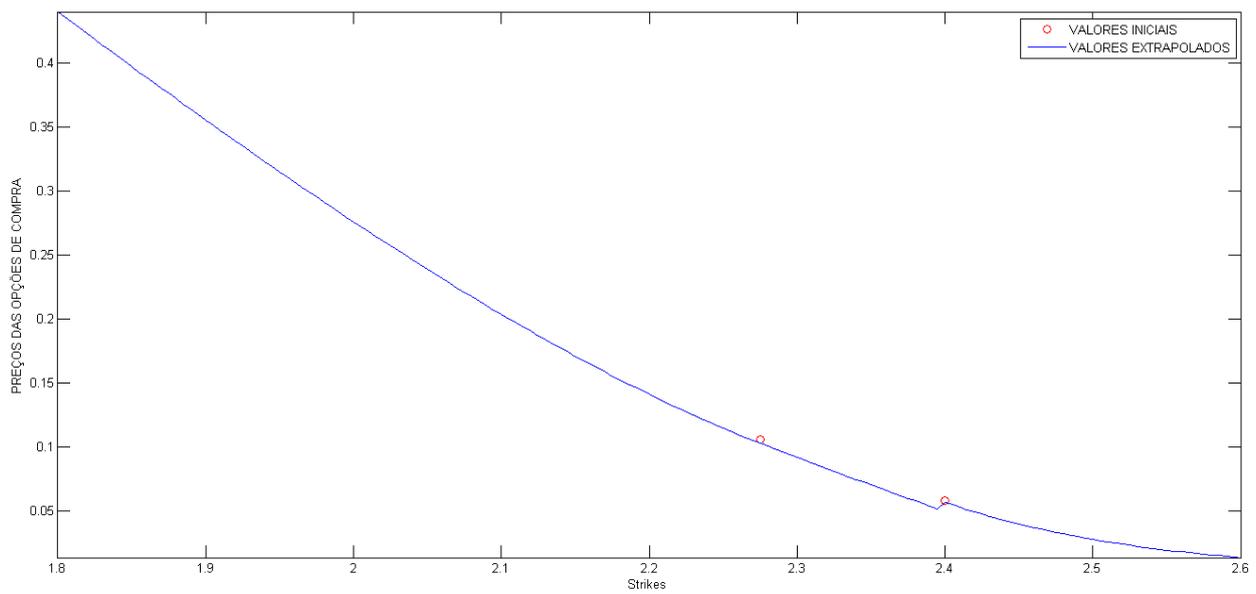


Figura 5.10: Preços sintéticos - maturidade 3

5.3.4 Preços Extrapolados das Opções de Compra para Valores Reais

Com a DME encontrada no item anterior, calculou-se o preço da opção de compra para uma faixa de strikes extrapolados. Assim sendo, em cada caso, construímos gráficos comparando os valores extrapolados com os valores iniciais dos strikes originais vindos de [2].

Vencimento 1: $T = \frac{23}{365}$ anos

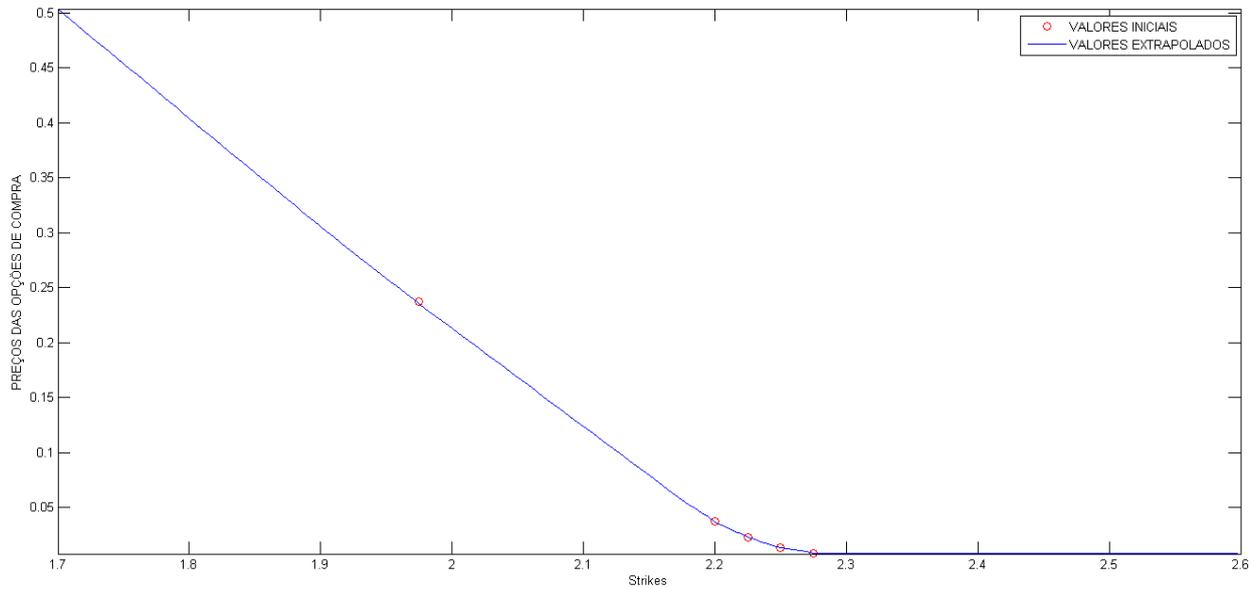


Figura 5.11: Preços reais - maturidade 1

Vencimento 2: $T = \frac{53}{365}$ anos

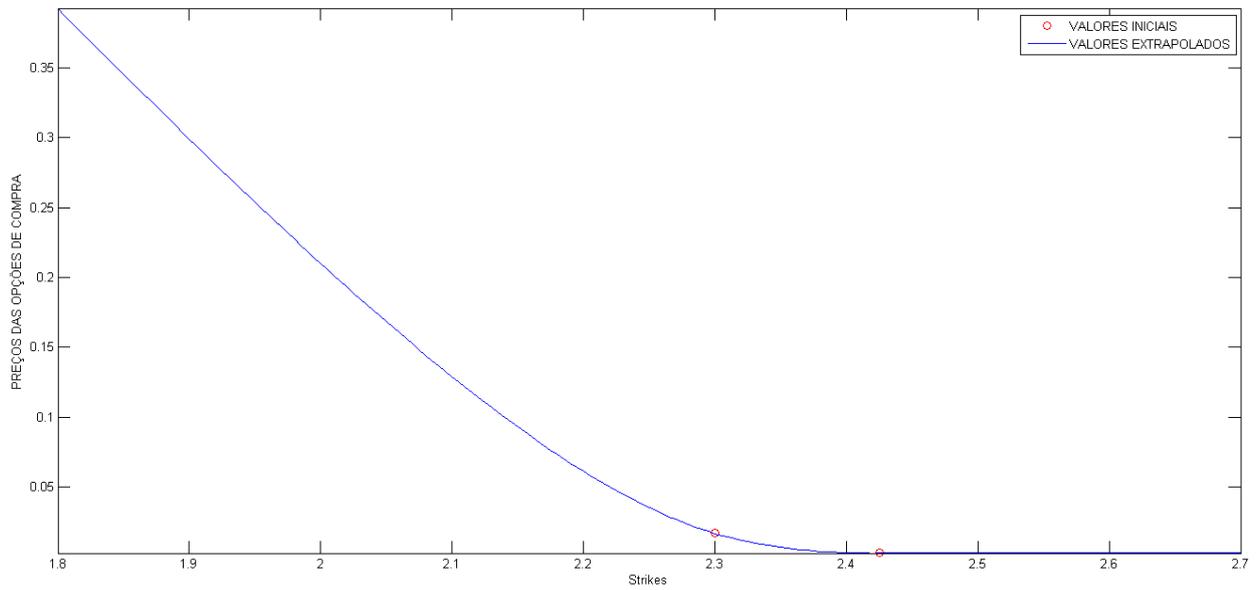


Figura 5.12: Preços reais - maturidade 2

Vencimento 3: $T = \frac{83}{365}$ anos

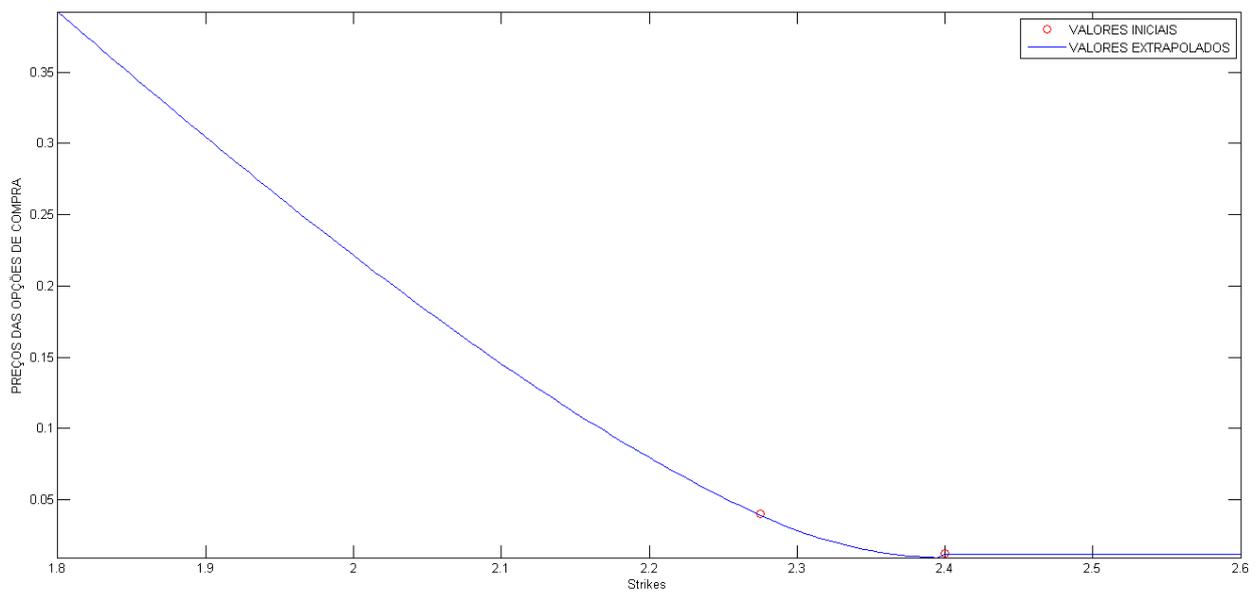


Figura 5.13: Preços reais - maturidade 3

5.3.5 Volatilidade Implícita para Valores Sintéticos

Agora, após encontrar a DME, usou-se a função $blsimpv(F, K, r, T, C(K))$ do Matlab, onde F é valor corrente do ativo, K um strike arbitrário escolhido, r a taxa Selic, T o vencimento em anos e $C(K)$ o preço da opção de compra retirada do mercado em K . Assim, gráficos para os dados sintéticos foram gerados comparando os valores extrapolados com os de Black e Scholes vindos dos strikes de [2].

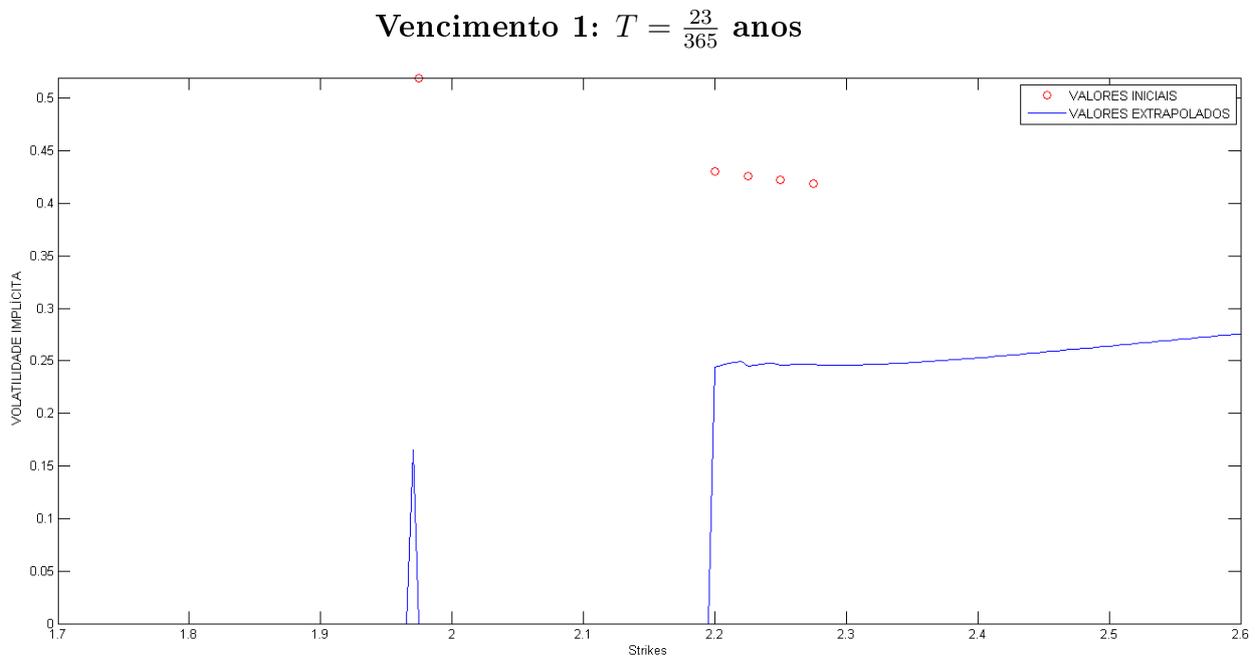


Figura 5.14: Volatilidade implícita sintética - vencimento 1

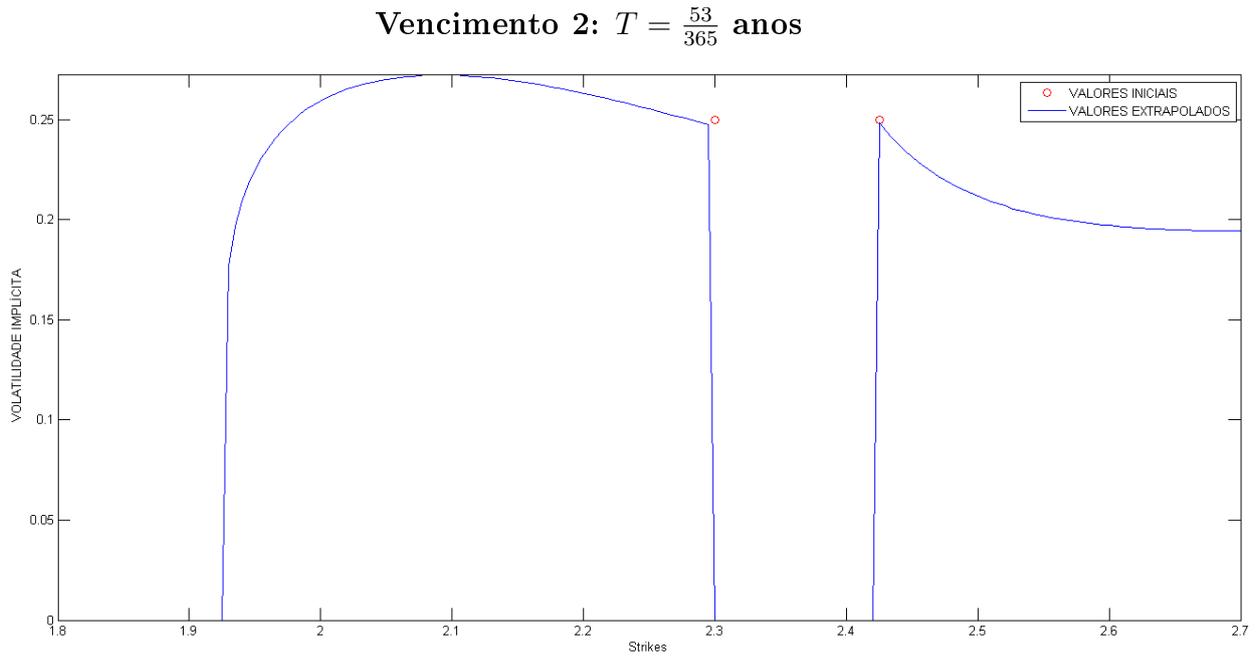


Figura 5.15: Volatilidade implícita sintética - vencimento 2

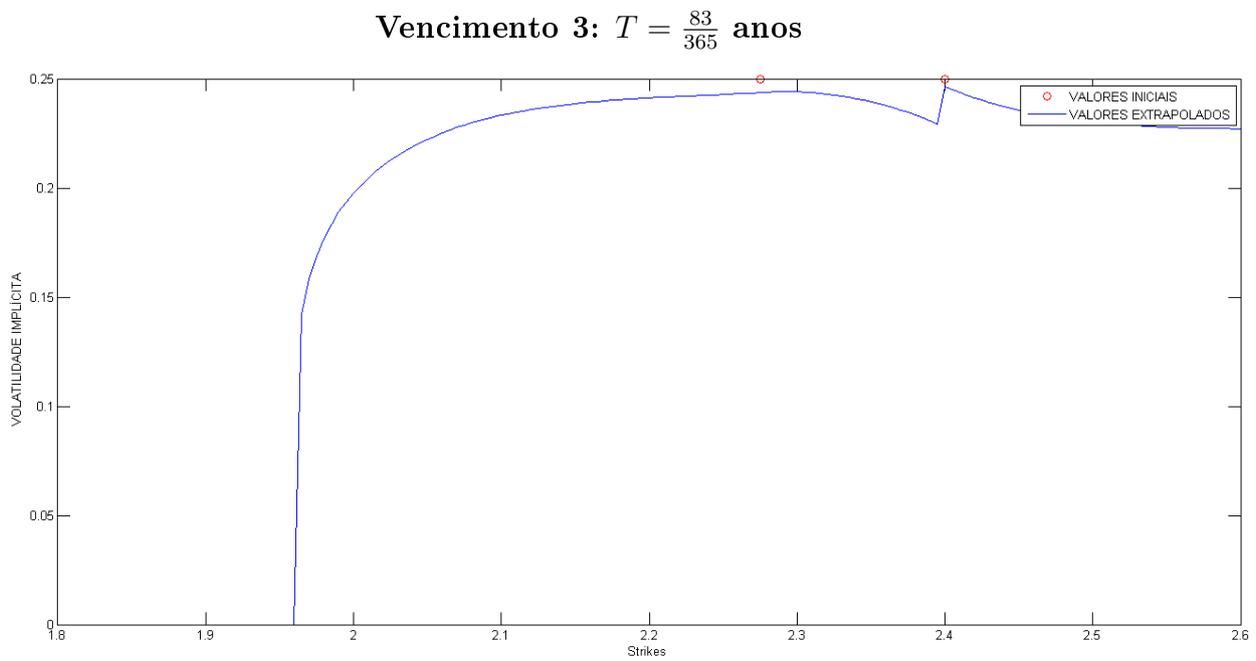


Figura 5.16: Volatilidade implícita sintética - vencimento 3

5.3.6 Volatilidade Implícita para Valores Reais

Finalmente, mais uma vez após encontrar a DME, usou-se a função $blsimpv(F, K, r, T, C(K))$ do Matlab, onde F é valor corrente do ativo, K um strike arbitrário escolhido, r a taxa Selic, T o vencimento em anos e $C(K)$ o preço da opção de compra retirada do mercado em K . Assim, gráficos para os dados reais foram gerados comparando os valores extrapolados com os de [2].

Vencimento 1: $T = \frac{23}{365}$ anos

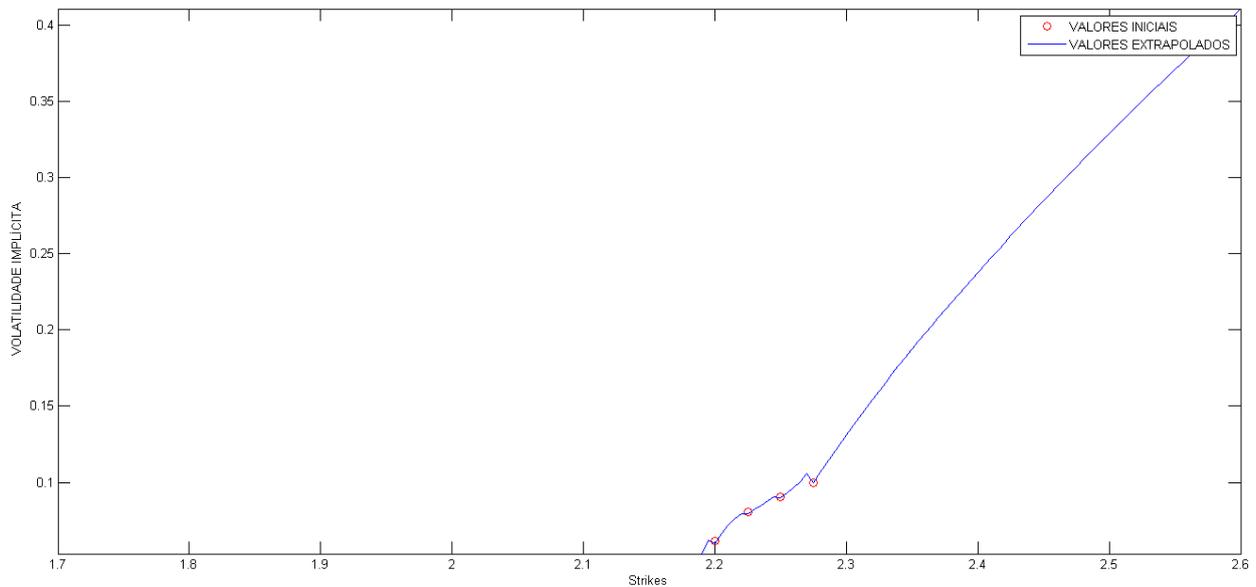


Figura 5.17: Volatilidade implícita real - vencimento 1

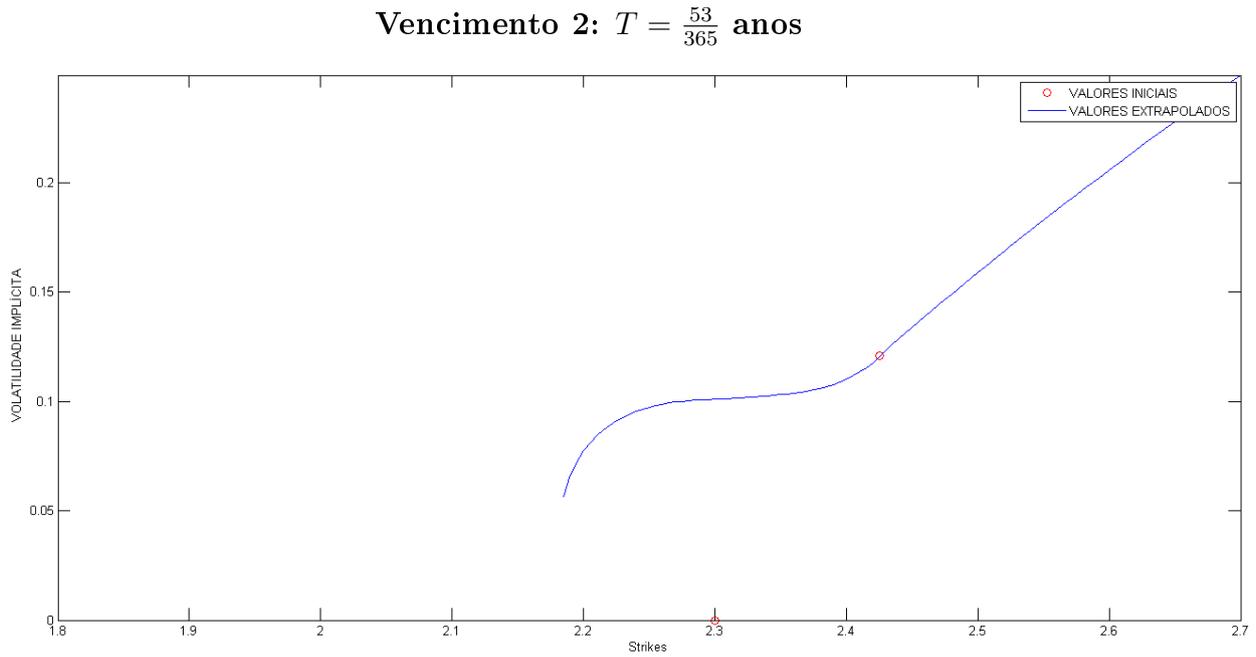


Figura 5.18: Volatilidade implícita real - vencimento 2

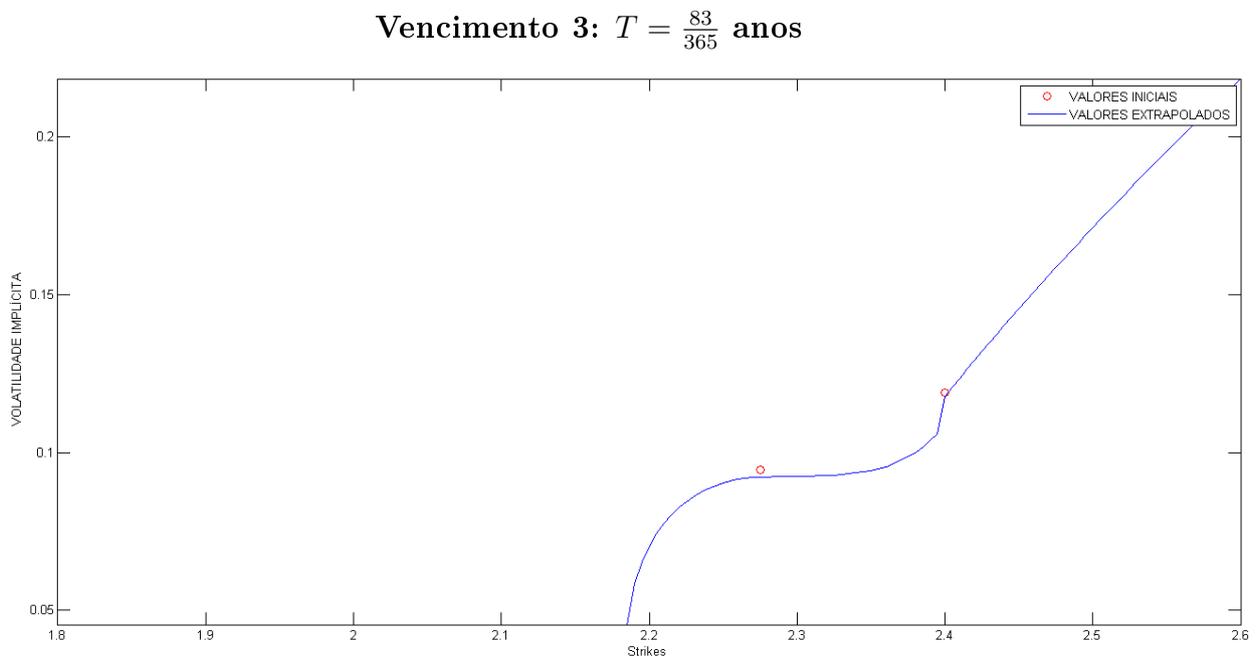


Figura 5.19: Volatilidade implícita real - vencimento 3

Capítulo 6

Conclusão

O principal objetivo do trabalho é encontrar, a partir de preços de opções de compra praticados no mercado, para uma dada maturidade, uma função distribuição de probabilidade do ativo associado. Mais precisamente, encontramos uma função densidade de probabilidade em $[0, \infty[$ que atende, para uma dada maturidade T , aos preços de opções de compra com strikes estritamente crescente $K_0 = 0, K_1, \dots, K_n, K_{n+1} = \infty$.

A ferramenta usada para encontrar essa distribuição é a Entropia de Shannon. Segundo [11], a entropia mede quanto a a distribuição é não enviesada. Logo, ao encontrar a densidade de máxima entropia, estamos encontrando a função menos enviesada que está de acordo com os dados extraídos do mercado.

Para tal, com o Teorema 1.1 mostramos que encontrar essa densidade em $[0, \infty[$ é equivalente a encontrar densidades nos intervalos $[K_i, K_{i+1}[$. Além disso, com auxílio do Teorema de Csiszár 2.1, chegamos ao Teorema 2.2, que é um dos resultados centrais do trabalho. Ou seja, vimos que para cada intervalo $[K_i, K_{i+1}[$ a expressão da densidade que maximiza a entropia é:

$$g(x) = \alpha_i e^{\beta_i x}, \quad \forall x \in [K_i, K_{i+1}[,$$

A partir disso, o objetivo passou a ser encontrar um algoritmo para obter cada α_i e β_i . Inicialmente, a densidade de máxima entropia (DME) dependia dos valores das chamadas opções digitais. Isso era uma limitação considerável do método, pois o preço de opções digitais não é algo fácil de se obter no mercado. Porém, com ajuda dos Teoremas 3.1, 3.2, 3.3 e 3.4 foi possível escrever a Entropia como função das opções digitais e, com isso, achar o ponto crítico desta função. Tal ponto crítico, além de sempre existir, é de máximo. Portanto, os preços das opções digitais não precisam ser extraídos do mercado. Na verdade, os únicos dados extraídos do mercado são os preços das opções de compra nos strikes K_i .

Cabe ressaltar mais uma vez que a densidade é uma informação muito valiosa. Com ela, não somente o preço da opção de compra pode ser calculado para qualquer strike, mas também o preço de outros derivativos mais complexos. Sabendo a expressão do payoff, o valor livre de arbitragem do derivativo pode ser calculado com a expressão (equação (1.0.1)):

$$C(K, T) = \mathbb{E}^g [e^{-rT} f(S(T))] = \int_0^\infty e^{-rT} f(s) g(s) ds,$$

onde $f(S(T))$ é o payoff da opção e g é função densidade de probabilidade associada a medida neutra ao risco.

Além disso, também podemos encontrar a volatilidade implícita para qualquer strike em $[0, \infty[$. Em modelos como o Black e Scholes temos que assumir que a volatilidade é constante, mas conforme vimos nos gráficos da Seção 5.3.6, isso não é a realidade do mercado. Assim,

encontrar essa DME fornece mais esta vantagem. Obter a volatilidade do ativo, parâmetro muito valioso no mercado.

Por fim, vimos que o método possui bastante consistência numérica e que métodos já consagrados como o de Newton Raphson, podem ser usados. Portanto, a utilidade do trabalho se justifica pela consistência do algoritmo aliada a importância da função densidade para a manipulação de derivativos em geral.

Apêndice A

Definições Importantes

A.1 Martingais

Um conceito fundamental na precificação de opções é o de Martingais. Em geral, assume-se que o preço de um dado ativo, como função do tempo, descreve um processo estocástico semi-martingal. Assim sendo, vamos definir os conceitos de filtração, processo estocástico, martingal, sub-martingal e super-martingal. As definições podem ser encontradas em [9].

Seja $(\Omega, \mathcal{F}, \mathcal{P})$ um espaço de probabilidade completo com espaço amostral Ω , σ -álgebra \mathcal{F} e medida de probabilidade P .

Definição A.1 (Filtração). *Seja $\{\mathcal{F}_t\}_{t \in I}$ uma família de sub- σ -álgebras de \mathcal{F} , I um conjunto ordenado de índices com $\mathcal{F}_s \subset \mathcal{F}_t$ para $s < t$, s e $t \in I$. Tal família é chamada de Filtração.*

A σ -álgebra \mathcal{F}_t normalmente modela os eventos observados no tempo t . Com isso, a variável aleatória X_t é \mathcal{F}_t -mesnsurável e, portanto, podemos determinar seu valor da informação dada no tempo t .

Definição A.2 (Processo Estocástico). *Um conjunto $\{(X_t, \mathcal{F}_t)\}$ composto de de uma filtração $\{\mathcal{F}_t\}_{t \in I}$ e de uma família em \mathbb{R}^n de variáveis aleatórias $\{X_t\}_{t \in I}$, com $\{X_t\}_{t \in I}$ \mathcal{F}_t -mesnsurável é chamado de *Proceso Estocástico com filtração \mathcal{F}_t .**

Definição A.3 (Martingais). *Um processo real $\{(X_t, \mathcal{F}_t)\}$ com $E | X_t | < \infty$ para todo $t \in I$, onde I é um conjunto ordenado de índices, é chamado:*

Martingal *quando para todo $s, t \in I$ com $s \leq t$ nós temos $E(X_t | \mathcal{F}_s) = X_s$, quase certamente,*

Super-Martingal *quando para todo $s, t \in I$ com $s \leq t$ nós temos $E(X_t | \mathcal{F}_s) \leq X_s$, quase certamente,*

Super-Martingal *quando para todo $s, t \in I$ com $s \leq t$ nós temos $E(X_t | \mathcal{F}_s) \geq X_s$ quase certamente.*

Uma boa noção intuitiva do que são os martingais, que pode ser vista em [15] na página 57, diz que dada a informação no processo presente, representado por \mathcal{F}_s , se tivermos que intuir sobre o estado futuro X_t , com $t \geq s$, na média, os valores do processo permanecem no mesmo estado. Na prática essa ideia é bem simples e a usamos com frequência na modelagem financeira como no caso do fechamento da cotação de uma moeda estrangeira. Salvo o caso de informações privilegiadas, o valor esperado da cotação no dia seguinte é o próprio valor do fechamento no dia anterior.

A.2 Multiplicadores de Lagrange

Na Seção 2.2 usamos multiplicadores de Lagrange para encontrar as equações que fornecem cada α_i e β_i . Assim sendo, nesta seção este conceito será apresentado.

Porém, um outro conceito, que será apresentado antes, muito importante relacionado à teoria de Multiplicadores de Lagrange é o de Derivada de Fréchet. No que segue, $(X, \|\cdot\|_X)$ e $(Y, \|\cdot\|_Y)$ denotam espaços de Banach, $U \subset X$ um conjunto aberto, $f : U \rightarrow Y$ uma aplicação e $\mathcal{L}(X, Y)$ o espaço dos operadores lineares contínuos. Além disso, usaremos a notação $r(h) = o(\|h\|_X)$ de uma aplicação $r : X \rightarrow Y$ se, e somente se,

$$\lim_{h \rightarrow 0} \frac{\|r(h)\|_Y}{\|h\|_X} = 0.$$

Segundo [1], temos a seguinte definição:

Definição A.4 (Derivada de Fréchet). *Seja x um ponto do conjunto aberto $U \subset X$. Uma aplicação $f : U \rightarrow Y$ é diferenciável à Fréchet em $x \in U$ se existe um operador linear $A \in \mathcal{L}(X, Y)$ tal que*

$$f(x+h) - f(x) - Ah = o(\|h\|).$$

O operador A é chamado de derivada de Fréchet da aplicação f em x e denotado por $Df(x)$ ou $f'(x)$. Se $f : U \rightarrow Y$ é diferenciável em todo ponto de U , então $Df : U \rightarrow \mathcal{L}(X, Y)$ é chamado de derivada à Fréchet de f .

Agora, vamos estabelecer o conceito de multiplicadores de Lagrange presente em [1]. No que segue, sejam X um espaço de Banach, $F \in \mathcal{C}^1(X, \mathbb{R})$ e um conjunto de vínculo:

$$S := \{v \in X; F(v) = 0\}.$$

Definição A.5 (multiplicadores de Lagrange). *Suponhamos que para todo $u \in S$, temos que $F'(u) \neq 0$. Se $J \in \mathcal{C}^1(X, \mathbb{R})$ (ou também sobre uma vizinhança de S ou \mathcal{C}^1 sobre S), dizemos que $c \in \mathbb{R}$ é valor crítico de J sobre S se existe $u \in S$ e $\lambda \in \mathbb{R}$ tais que*

$$J(u) = c \quad e \quad J'(u) = \lambda f'(u).$$

O ponto u é chamado de ponto crítico de J sobre S e o número real λ é chamado de multiplicador de Lagrange para o valor crítico c (ou para o ponto crítico u).

Assim sendo, em [1] é enunciada uma proposição que garante a existência do multiplicador de Lagrange λ . Na demonstração deste fato, usa-se o Teorema da Função Implícita presente em [18].

Apêndice B

O Conceito de Entropia de Shannon

No presente capítulo, que se baseia em [16], apresentaremos algumas das noções principais que construíram a definição de entropia utilizada na obtenção da distribuição de probabilidade dos preços de ativos financeiros.

A entropia de Shannon possui origem na teoria da informação [6]. A partir de um "gerador de informação", que pode ser um simples telégrafo, uma função ou até mesmo um processo estocástico (que é o nosso caso, pois estamos considerando que o preço do ativo, como função do tempo, satisfaz um processo estocástico), vemos em [16] que a entropia é construída de forma a medir o quanto de informação que pode ser produzida pelo sistema. Desta forma, ao maximizar a entropia, estamos maximizando a capacidade do sistema de gerar informação.

Com o exposto acima em mente, podemos agora entender com mais clareza o que foi exposto na introdução deste trabalho. Lá, a obtenção da distribuição de probabilidade via máxima entropia foi vista como uma maneira de obter uma função densidade que concordava com os dados retirados do mercado (preço das opções de compra) e que, ao mesmo tempo, era a distribuição menos enviesada possível.

De fato, se a entropia consegue medir o quanto de informação o sistema consegue gerar, então ao maximizar a entropia nós estamos encontrando uma distribuição de probabilidade mais "imparcial" possível, pois o sistema que gera informação (onde a informação, no caso, são preços do ativo) está em uma situação de "geração máxima de dados". Ou seja, estamos na situação menos enviesada possível, pois não há preços privilegiados pelo sistema.

Assim sendo, vamos agora apresentar o início da construção do conceito de entropia feito por Shannon para um caso particular de sistema de informação. Tais conceitos estão presentes no artigo do próprio Shannon em [16], além de [6].

B.1 Sistemas de Comunicação

No geral, segundo o exposto em [16], um sistema de comunicação é composto por cinco elementos. São eles:

Gerador de Informação É o produtor das mensagens e, no nosso caso, é o elemento de principal interesse. Essas mensagens podem ser de vários tipos, tais como uma sequência de letras produzidas por um telégrafo, um função do tempo como na telefonia, uma função do tempo e de outras variáveis como ocorre nas televisões ou até mesmo um processo estocástico.

Transmissor Algo capaz de operar a mensagem de maneira a produzir uma espécie de sinal que poderá ser transmitido.

Canal É o meio utilizado para transmissão da mensagem. Pode ser uma antena ou um cabo coaxial.

Receptor Algo capaz de fazer a operação inversa ao do transmissor, ou seja, transformar o sinal transmitido na mensagem original.

Destinatário É simplesmente o ponto final ou pessoa que originalmente deveria receber a mensagem.

Neste trabalho de finanças, conforme citado no início do capítulo, o elemento principal do sistema é o Gerador de Informação. Em outras palavras, uma mensagem particular é apenas uma dentre outras de um conjunto de possibilidades. Desta forma, o sistema deve ser capaz de operar qualquer umas das mensagens do universo de possibilidades a qualquer tempo. Ou seja, estamos interessados na capacidade do sistema de comunicação.

Considerando que o número de mensagens é finito, então este próprio valor ou qualquer função monótona desse número pode ser uma medida de quanta informação é produzida pelo sistema, dado que uma mensagem é escolhida ao acaso, sendo todas igualmente prováveis. Neste caso, se utilizarmos uma função monótona para fazer essa medida, a mais conveniente é a função logaritmo. Isso ocorrerá mesmo em casos mais complexos, que é o presente neste trabalho, pois as mensagens (preços) variam em conjunto contínuo e não são igualmente prováveis de ocorrer.

Vamos agora nos concentrar na capacidade do Gerador de Informação e na construção natural do conceito de entropia.

B.2 Geradores Discretos e o Conceito de Entropia

Vimos que, no caso em que o conjunto de possíveis mensagens é finito e as possibilidades são equiprováveis, uma maneira de medir a capacidade do sistema é usar a função logaritmo. A grande questão agora é estudar o efeito de uma estatística no sistema de encadeamento das mensagens. Por exemplo, imaginando que as mensagens são letras relacionadas a um texto em alguma determinada língua, certas letras possuirão uma frequência maior do que outras. Em um texto em português temos maior incidência de letras "a" do que de letras "x".

Repare que é exatamente esse o foco do nosso problema. Nossa "mensagem" são os preços do ativo e nosso interesse é justamente determinar como a escolha dessas mensagens se comporta e isso é traduzido pela distribuição de probabilidade.

Voltando ao caso discreto e finito, vamos designar por p_1, p_2, \dots, p_n as probabilidades de acontecimento de cada evento. Desta forma, como podemos medir o grau ou a quantidade de informação produzida pelo processo? Será possível achar uma medida de quanta incerteza está envolvida na seleção de algum evento?

Caso essa medida exista, digamos $H(p_1, p_2, \dots, p_n)$, é interessante que ela satisfaça as seguintes propriedades:

- i H deve ser contínua em p_i
- ii Caso os eventos sejam equiprováveis, H deve ser estritamente crescente em n , pois teremos um universo maior de escolha
- iii Caso a escolha seja quebrada em escolhas sucessivas, o valor original de H deve ser ponderado pelos valores sucessivos de H .

Assim sendo, em [16], temos o seguinte teorema:

Teorema B.1. *A única função que satisfaz as três condições acima é da forma*

$$H = -K \sum_{i=1}^n p_i \ln p_i \quad (\text{B.2.1})$$

Onde K é uma constante positiva.

Demonstração. Seja $H(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}) = A(n)$. Da terceira condição requerida nós podemos decompor a escolha vinda de s^m possibilidades equiprováveis em uma série de m escolhas equiprováveis vindas de s e obter

$$A(s^m) = mA(s).$$

Analogamente, $A(t^n) = nA(t)$. Nós podemos escolher n arbitrariamente grande e achar m que satisfaça

$$s^m < t^n < s^{m+1}$$

Assim, aplicando logaritmos e dividindo tudo por $n \ln s$,

$$\frac{m}{n} \leq \frac{\ln t}{\ln s} \leq \frac{m}{n} + \frac{1}{n}$$

ou, equivalentemente,

$$\left| \frac{m}{n} - \frac{\ln t}{\ln s} \right| < \epsilon$$

onde ϵ é arbitrariamente pequeno. Agora, da monotocidade da $A(n)$,

$$A(s^m) \leq A(t^n) \leq A(s^{m+1})$$

$$mA(s) \leq nA(t) \leq (m+1)A(s)$$

Portanto, dividindo por $nA(s)$,

$$\frac{m}{n} \leq \frac{A(t)}{A(s)} \leq \frac{m}{n} + \frac{1}{n}$$

ou

$$\left| \frac{m}{n} - \frac{A(t)}{A(s)} \right| < \epsilon$$

$$\left| \frac{A(t)}{A(s)} - \frac{\ln t}{\ln s} \right| < 2\epsilon$$

com $A(t) = K \ln t$, sendo K uma constante positiva para satisfazer a segunda condição.

Agora, suponha que tenhamos uma escolha a partir de n possibilidades com probabilidades racionais $p_i = \frac{n_i}{\sum n_i}$ onde cada n_i é inteiro. Nós podemos parar uma escolha a partir de $\sum n_i$ possibilidades em uma escolha vinda de n possibilidades com probabilidades p_1, p_2, \dots, p_n e então, se a i -ésima possibilidade for a escolhida, efetuamos uma escolha a partir de n_i com probabilidades equiprováveis. Assim, da terceira condição, nós igualamos as escolhas totais a partir de $\sum n_i$ calculadas pelos dois métodos

$$K \ln \sum n_i = H(p_1, p_2, \dots, p_n) + K \sum p_i \ln n_i$$

Portanto

$$H = K \left[\sum p_i \left(\ln \sum n_i \right) - \sum p_i \ln n_i \right]$$

$$H = -K \sum p_i \ln \left(\frac{n_i}{\sum n_i} \right) = -K \sum p_i \ln p_i$$

Se as probabilidades não forem racionais, então podemos aproximá-las por números racionais e, usando a hipótese de continuidade, a fórmula obtida continua válida. \square

Ainda em [16], algumas propriedades interessantes de H são apresentadas, entre elas

1. H é nulo se, e somente se, tivermos algum p_i igual a 1 com os demais nulos. Caso contrário, H será positivo. Repare que isso é um resultado totalmente coerente, pois H só é nulo quando temos certeza do que vai acontecer.
2. Para um dado n e supondo que todos os eventos são equiprováveis, o valor máximo de H é $\ln n$.
3. Chamando de $H_x(y)$ a entropia do modelo y dado que sabemos o que ocorreu no modelo x , então vale

$$H_x(y) \leq H(y)$$

Ou seja, a medida de incerteza de y pode diminuir quando sabemos o que ocorreu com x .

Ainda em [16], uma extensão natural para o caso contínuo é feita, mas o importante é que as principais propriedades e a ideia central do conceito de entropia exposto aqui, se preserva. Ou seja, a entropia definida como

$$H(g) := - \int_0^{\infty} g(x) \ln g(x) dx$$

consegue medir o quanto de informação o sistema consegue gerar e, portanto, quando achamos g que maximiza H , estamos encontrando a densidade de probabilidade menos enviesada possível.

Apêndice C

Códigos de Matlab

Código i *Script para a obtenção da DME com valores sintéticos.*

```
function [volimpv, Alpha, Beta, CK, DK, CKlinha, CK2linha, densidade, cumulativa] ...
    = entropiasintetica(K, T, F, strike)
%definindo os parâmetros
r=0.11; %taxa Selic no Brasil
T=T/365;
sigma=0.25;
tamanho=size(K);
outrotamanho=size(strike);
m=outrotamanho(2);
n=tamanho(2)-1;
C=[1:n+2];
K=[K 1e10];
lambda = 1; %usado nos testes para encontrar o D mo
D=[0:n+1];
D(1)=1;
D(n+2)=0;
%gerando os preços das calls
for i=[1:n+2]
    C(i)=blsprice(F, K(i), r, T, sigma)
end
C=exp(-r*T)*C;
%gerando o preço da opção digital (chute inicial)
for i=[2:n+1]
    D(i)=normcdf((log(F/K(i))-0.5*sigma^2*T)/sigma*sqrt(T));
end
D=exp(-r*T)*D;
%encontrando o preço da opção digital com Newton
while lambda/2 > 1e-3
    p=[0:n+1]; %iniciando pi
    Kmed = [0:n+1]; %iniciando Kmed
    for i = [1:n+1] %formando pi e Kmed
        p(i)=D(i)-D(i+1);
        Kmed(i)=((C(i)+K(i)*D(i))-(C(i+1)+K(i+1)*D(i+1)))/p(i);
    end
end
```

```

for i = [1:n]
    FF=@(x)((K(i+1)*exp(x*K(i+1))-K(i)*exp(x*K(i)))/...
        (exp(x*K(i+1))-exp(x*K(i)))-1/x-Kmed(i));
    beta(i)=fsolve(FF,0.005);
end
%beta da borda
beta(n+1)=1/(K(n+1)-...
    ((C(n+1)+K(n+1)*D(n+1))-C(n+2))/(D(n+1)-D(n+2)));

alpha=[0:n];
for i = [1:n+1]
    alpha(i)=beta(i)*...
        ((D(i)-D(i+1))/(exp(beta(i)*K(i+1))-exp(beta(i)*K(i))));
end

Hdois = zeros(n,n);
cdois = [0:n];

for i = [1:n]
    cdois(i)= 1/beta(i)^2-...
        (((K(i+1)-K(i))^2)*exp(beta(i)*(K(i)+K(i+1))))/...
        (exp(beta(i+1)*K(i+1))-exp(beta(i)*K(i)))^2;
end
cdois(n+1)=1/beta(n+1)^2;

for i = [1:n]
    Hdois(i,i)=-1/p(i)-1/p(i+1)-...
        ((K(i+1)-Kmed(i))^2)/(p(i)*cdois(i))-...
        ((K(i+1)-Kmed(i+1))^2)/p(i+1)*cdois(i+1);
end

for i = [1:n-1]
    Hdois(i,i+1)=1/p(i+1)-...
        (Kmed(i+1)-K(i+1))*(K(i+2)-Kmed(i+1))/p(i+1)*cdois(i+1);
    Hdois(i+1,i)=1/p(i+1)-...
        ((Kmed(i+1)-K(i+1))*(K(i+2)-Kmed(i+1)))/p(i+1)*cdois(i+1);
end

Hhum=zeros(n,1);

for i = [1:n]
    Hhum(i)=log(alpha(i)*exp(beta(i)*K(i+1)))...
        -log(alpha(i+1)*exp(beta(i+1)*K(i+1)));
end

deltaD = -Hdois\Hhum;
lambda = Hhum'*deltaD;

```

```

for i=[1:n]
    D(i+1)=D(i+1)+ 0.5*deltaD(i);
end
end
Dsol=D;
%criando a fun que gera os betas
beta=[0:n];
%gerando os betas (exceto o ltimo)
for i=[1:n]
    FFF=@(x)((K(i+1)*exp(x*K(i+1))-K(i)*exp(x*K(i)))/(exp(x*K(i+1))-...
        exp(x*K(i)))-1/x-((C(i)+K(i)*D(i))-(C(i+1)+K(i+1)*D(i+1)))/...
        (D(i)-D(i+1)));
    beta(i)=fsolve(FFF,0.005);
end
%beta da borda
beta(n+1)=1/(K(n+1)-((C(n+1)+K(n+1)*D(n+1))-C(n+2))/(D(n+1)-D(n+2)));

%gerando o alpha
alpha=[0:n];
for i=[1:n+1]
    alpha(i)=beta(i)*((D(i)-D(i+1))/(exp(beta(i)*K(i+1))-...
        exp(beta(i)*K(i))));
end
Alpha=alpha;
Beta=beta;
%achando a volatilidade implicita
for s=[1:m]
    contador = 0;
    j=1;
    while contador == 0
        if (K(j)<=strike(s))&&(strike(s)<K(j+1))&&(beta(j)~=0)
            DK(s)=Dsol(j)-...
                (alpha(j)/beta(j))*(exp(beta(j)*strike(s))-exp(beta(j)*K(j)));
            CK(s)=C(j)-(strike(s)-K(j))*(Dsol(j)+...
                (alpha(j)/beta(j))*exp(beta(j)*K(j)))+...
                (alpha(j)/beta(j)^2)*(exp(beta(j)*strike(s))-...
                exp(beta(j)*K(j)));
            CKlinha(s)=-(Dsol(j)+(alpha(j)/beta(j))*exp(beta(j)*K(j)))+...
                (alpha(j)/beta(j))*exp(beta(j)*strike(s));
            CK2linha(s)=alpha(j)*exp(beta(i)*strike(s));
            densidade(s)=alpha(j)*exp(beta(j)*strike(s));
            cumulativa(s)=1-Dsol(j)+(alpha(j)/beta(j))*(exp(beta(j)...
                *strike(s))-exp(beta(j)*K(j)));
            contador=contador+1;
        end
    end
    if (K(j)<=strike(s))&&(strike(s)<K(j+1))&&(beta(j)==0)
        DK(s)=Dsol(j)-alpha(j)*(strike(s)-K(j));
    end
end

```

```

    CK(s)=C(j)-(strike(s)-K(j))*Dsol(j)+(alpha(j)/2)*...
        (strike(s)-K(j))^2;
    CKlinha(s)=-Dsol(j)+alpha(j)*(strike(s)-K(j));
    CK2linha(s)=alpha(j);
    densidade(s)=alpha(j)*exp(beta(j)*strike(s));
    cumulativa(s)=1-Dsol(j)+alpha(j)*(strike(s)-K(j));
    contador=contador+1;
end
j=j+1;
end
volimpv(s) = blsimpv(F, strike(s), r, T, CK(s));
end
end

```

Código ii *Script que testa a condição de não arbitragem dos valores reais retirados do mercado.*

```

function [novoC, novoK]=testearbitragem(Cplot, Kplot, T)

%definindo os parametros
r=0.11; %Taxa Selic do Brasil
T=T/365;
sigma=0.25;
tamanho=size(Kplot);
n=tamanho(2)-1;
C=[Cplot 0];
K=[Kplot 1e10];
D=[0:n];
D(1)=1;
F=2.219;
%digital de B&S usada para teste de no arbitragem
for i=[2:n+1]
D(i)=-(C(i+1)-C(i-1))/(K(i+1)-K(i-1));
end
deltateste(1)=0;
for i=[2:n+2]
    deltatete(i)=-(C(i)-C(i-1))./(K(i)-K(i-1));
end
%teste de no arbitragem
ll = 2;
novoC(1)=F;
novoK(1)=0;
for i=[2:n+1]
    if (deltateste(i)>D(i))&& (D(i)>deltateste(i+1))
        novoC(ll)=C(i);
        novoK(ll)=K(i);
        ll = ll+1;
    end
end
end

```

```

novotamanho=size(novoK);
m=novotamanho(2)-1;
novoCplot=[novoC 0];
novoKplot=[novoK 1e10];
novoD=[0:m];
novoD(1)=1;
%digital de B&S usada para teste de no arbitragem
for i=[2:m+1]
novoD(i)=-(novoCplot(i+1)-novoCplot(i-1))/(novoKplot(i+1)-novoKplot(i-1));
end

plot(Kplot,-D,'-r',novoK,-novoD,'-b')
xlabel('Strikes');
ylabel('-D(K)');
legend('ANTES DO TESTE','DEPOIS DO TESTE');
axis([1.7 2.7 -1 0]);
end

```

Código iii *Script para a obtenção da DME com valores reais.*

```

function [volimpv,Alpha,Beta,CK,DK,CKlinha,CK2linha,densidade,cumulativa]...
=entropiareal(C,K,T,F,strike)
%definindo os parmetros
r=0.11;%Taxa Selic do Brasil
T=T/365;
tamanho=size(K);
n=tamanho(2)-1;
outrotamanho=size(strike);
m=outrotamanho(2);
C=exp(-r*T)*[C 0];
K=[K 1e10];
lambda = 1; %usado nos testes para encontrar o D mo
D=[0:n+1];
D(1)=1;
D(n+2)=0;
CK=[1:m];
DK=[1:m];
CKlinha=[1:m];
CK2linha=[1:m];
densidade=[1:m];
cumulativa=[1:m];
volimpv=[1:m];
%gerando o preo da op digital (chute inicial)
for i=[2:n+1]
D(i)=-(C(i+1)-C(i-1))/(K(i+1)-K(i-1));
end
%encontrando o preo mo da op digital com Newton
while lambda/2 > 1e-3

```

```

p=[0:n+1]; %iniciando pi
Kmed = [0:n+1]; %iniciando Kmed
for i = [1:n+1] %formando pi e Kmed
    p(i)=D(i)-D(i+1);
    Kmed(i)=((C(i)+K(i)*D(i))-(C(i+1)+K(i+1)*D(i+1)))/p(i);
end
for i=[1:n]
    FF=@(x)((K(i+1)*exp(x*K(i+1))-K(i)*exp(x*K(i)))/...
        (exp(x*K(i+1))-exp(x*K(i)))-1/x-Kmed(i));
    beta(i)=fsolve(FF,0.005);
end
%beta da borda
beta(n+1)=1/(K(n+1)-...
    ((C(n+1)+K(n+1)*D(n+1))-C(n+2))/(D(n+1)-D(n+2)));

alpha=[0:n];
for i=[1:n+1]
    alpha(i)=beta(i)*...
        ((D(i)-D(i+1))/(exp(beta(i)*K(i+1))-exp(beta(i)*K(i))));
end

Hdois = zeros(n,n);
cdois = [0:n];

for i = [1:n]
    cdois(i)= 1/beta(i)^2-...
        (((K(i+1)-K(i))^2)*exp(beta(i)*((K(i)+K(i+1)))))/...
        (exp(beta(i+1)*K(i+1))-exp(beta(i)*K(i)))^2);
end
cdois(n+1)=1/beta(n+1)^2;

for i = [1:n]
    Hdois(i,i)=-1/p(i)-1/p(i+1)-...
        ((K(i+1)-Kmed(i))^2)/(p(i)*cdois(i))-...
        ((K(i+1)-Kmed(i+1))^2)/p(i+1)*cdois(i+1);
end

for i = [1:n-1]
    Hdois(i,i+1)=1/p(i+1)-...
        (Kmed(i+1)-K(i+1))*(K(i+2)-Kmed(i+1))/p(i+1)*cdois(i+1);
    Hdois(i+1,i)=1/p(i+1)-...
        ((Kmed(i+1)-K(i+1))*(K(i+2)-Kmed(i+1)))/p(i+1)*cdois(i+1);
end

Hhum=zeros(n,1);

for i = [1:n]
    Hhum(i)=log(alpha(i)*exp(beta(i)*K(i+1)))...

```

```

        -log(alpha(i+1)*exp(beta(i+1)*K(i+1)));
end

deltaD = -Hdois\Hhum;
lambda = Hhum'*deltaD;

for i=[1:n]
    D(i+1)=D(i+1)+ 0.5*deltaD(i);
end
end

Dsol=D;

%criando a fun  que gera os betas
beta=[0:n];
%gerando os beta (exceto o ltimo)
for i=[1:n]
    FFF=@(x)((K(i+1)*exp(x*K(i+1))-K(i)*exp(x*K(i)))/(exp(x*K(i+1))-...
        exp(x*K(i)))-1/x-((C(i)+K(i)*D(i))-(C(i+1)+K(i+1)*D(i+1)))/...
        (D(i)-D(i+1)));
    beta(i)=fsolve(FFF,0.005);
end
%beta da borda
beta(n+1)=1/(K(n+1)-((C(n+1)+K(n+1)*D(n+1))-C(n+2))/(D(n+1)-D(n+2)));

%gerando o alpha

alpha=[0:n];
for i=[1:n+1]
    alpha(i)=beta(i)*((D(i)-D(i+1))/(exp(beta(i)*K(i+1))-...
        exp(beta(i)*K(i))));
end
Alpha=alpha;
Beta=beta;
%achando a volatilidade implicita
for s=[1:m]
    contador = 0;
    j=1;
    while contador == 0
        if (K(j)<=strike(s))&&(strike(s)<K(j+1))&&(beta(j)~=0)
            DK(s)=Dsol(j) - ...
                (alpha(j)/beta(j))*(exp(beta(j)*strike(s))-exp(beta(j)*K(j)));
            CK(s)=C(j)-(strike(s)-K(j))*(Dsol(j)+...
                (alpha(j)/beta(j)*exp(beta(j)*K(j)))+...
                (alpha(j)/beta(j)^2)*(exp(beta(j)*strike(s))-...
                exp(beta(j)*K(j)));
            CKlinha(s)=-(Dsol(j)+(alpha(j)/beta(j)*exp(beta(j)*K(j)))+...
                (alpha(j)/beta(j))*exp(beta(j)*strike(s)));
        end
        j=j+1;
        contador = CK(s);
    end
end

```

```

    CK2linha(s)=alpha(j)*exp(beta(i)*strike(s));
    densidade(s)=alpha(j)*exp(beta(j)*strike(s));
    cumulativa(s)=1-Dsol(j)+(alpha(j)/beta(j))*(exp(beta(j)*...
        strike(s))-exp(beta(j)*K(j)));
    contador=contador+1;
end
if (K(j)<=strike(s))&&(strike(s)<K(j+1))&&(beta(j)==0)
    DK(s)=Dsol(j)-alpha(j)*(strike(s)-K(j));
    CK(s)=C(j)-(strike(s)-K(j))*Dsol(j)+(alpha(j)/2)*...
        (strike(s)-K(j))^2;
    CKlinha(s)=-Dsol(j)+alpha(j)*(strike(s)-K(j));
    CK2linha(s)=alpha(j);
    densidade(s)=alpha(j)*exp(beta(j)*strike(s));
    cumulativa(s)=1-Dsol(j)+alpha(j)*(strike(s)-K(j));
    contador=contador+1;
end
j=j+1;
end
volimpv(s) = blsimpv(F, strike(s), r, T, CK(s));
end
end

```

Código iv *Script para a Plotagem dos gráficos da DME, Preços Extrapolados das Opções e Volatilidade Implícita.*

```

function Plotagem(strike , greal , gsint , CKreal , CKsint , volreal , volsint)

tamanho=size(strike);
plot(strike , gsint , '-r' , strike , greal , '-.b')
xlabel('Strikes');
ylabel('Densidade');
legend('VALORES SINTETICOS' , 'VALORES DE MERCADO');
axis([strike(1) strike(tamanho(2)) min(min(greal , gsint)) ...
    max(max(greal , gsint))] );

plot(strike , CKsint , '-r' , strike , CKreal , '-.b')
xlabel('Strikes');
ylabel('PREÇOS (EXTRAPOLADOS) DAS OPCOES DE COMPRA');
legend('VALORES SINTETICOS' , 'VALORES DE MERCADO');
axis([strike(1) strike(tamanho(2)) min(min(CKreal , CKsint)) ...
    max(max(CKreal , CKsint))] );
grid

plot(strike , volsint , '-r' , strike , volreal , '-.b')
xlabel('Strikes');
ylabel('VOLATILIDADE IMPLICITA');
legend('VALORES SINTETICOS' , 'VALORES DE MERCADO');

```

```
axis ([ strike (1) strike (tamanho (2)) min (min (volreal , volsint)) ...
      max (max (volreal , volsint))]);
grid
end
```

Código v *Script para estimativa dos erros percentuais de cada α e β .*

```
function [erroBeta ,erroAlpha ,PercentualBeta ,PercentualAlpha]= ...
    erro (C,BD, Alpha ,Beta)

tamanho=size (Beta);
n=tamanho (2);
erroBeta =[1:n];
erroAlpha =[1:n];
PercentualAlpha =[1:n];
PercentualBeta =[1:n];
for i =[1:n-1]
    erroBeta (i)=BD*(Beta (i+1)-Beta (i))/(C(i+1)-C(i));
    erroAlpha (i)=BD*(Alpha (i+1)-Alpha (i))/(C(i+1)-C(i));
    PercentualBeta (i)=100*erroBeta (i)/Beta (i);
    PercentualAlpha (i)=100*erroAlpha (i)/Alpha (i);
end
erroBeta (n)=BD*(Beta (n)/C(n));
erroAlpha (n)=BD*(Beta (n)/C(n));
PercentualBeta (n)=100*erroBeta (n)/Beta (n);
PercentualAlpha (n)=100*erroAlpha (n)/Alpha (n);

end
```


Referências Bibliográficas

- [1] Rodrigo A. O. Arruda, *Teoria do Pontos Críticos via Minimização*, Universidade Federal da Paraíba (2004).
- [2] BM&F Bovespa, *Boletim diário: Segmento BM&F*, 2014.
- [3] Stephen Boyd and Lieven Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.
- [4] Peter W. Buchen and Michael Kelly, *The Maximum Entropy Distribution of an Asset Inferred from Option Prices*, Journal of Financial and Quantitative Analysis **31** (1996), no. 1, 143–159.
- [5] Samuel D. Conte and Carl De Boor, *Elementary Numerical Analysis*, McGraw-Hill, 1972.
- [6] Thomas Cover and Joy Thomas, *Elements of Information Theory*, John Wiley & Sons, 2006.
- [7] Imre Csiszar, *I-Divergence Geometry of Probability Distributions and Minimization Problems*, The Annals of Probability **3** (1975), 146–158.
- [8] John C. Hull, *Options, Futures and Other Derivatives*, Prentice Hall, 2005.
- [9] Ralph Korn and Elke Korn, *Option Pricing and Portfolio Optimization*, vol. 31, AMS, 2000.
- [10] Elon L. Lima, *álgebra linear*, IMPA, 2008.
- [11] Cassio Neri and Lorenz Schneider, *Maximum Entropy Distributions Inferred from Options Portfolios*, Finance and Stochastics **16** (2012), no. 2, 293–318.
- [12] ———, *A Family of Maximum Entropy Densities Matching Call Option Prices*, Applied Mathematical Finance **20** (2013), no. 6, 548–577.
- [13] Raombanarivo Dina Ramilijaona, *Problemas Inversos em Engenharia Financeira: Regularização com Critério de Entropia.*, Universidade Estadual do Rio de Janeiro (2013).
- [14] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, 1970.
- [15] Paulo R. C. Ruffino, *Uma Iniciação aos Sistemas Dinâmicos Estocásticos*, Publicações Matemáticas, IMPA, 2009.
- [16] C. E. Shannon, *A Mathematical Theory of Communication*, The Bell System Technical Journal (1948).

- [17] M. B. M.-L. Stern, Fred Muste, *Summary of Experimental Uncertainty Assessment Methodology with Example*, IIHR Technical Report No 406 (1999).
- [18] Kosaku Yosida, *Functional Analysis*, Springer-Verlag, Heidelberg, 1995.