

I ENCONTRO REGIONAL DE MATEMÁTICA APLICADA E  
COMPUTACIONAL - UFSC BLUMENAU

UMA INTRODUÇÃO AOS MÉTODOS  
MATEMÁTICOS EM FINANÇAS

PROF. VINÍCIUS ALBANI

*Universidade Federal de Santa Catarina*

2019

BLUMENAU - SC



# Conteúdo

<b>Introdução</b>	<b>4</b>
<b>1 Teoria de Probabilidade</b>	<b>7</b>
1.1 Definições Básicas . . . . .	7
1.2 O Teorema de Limite Central e a Lei dos Grandes Números	22
1.3 Esperança Condicional . . . . .	26
1.4 Martingais e Movimento Browniano . . . . .	30
<b>2 Cálculo Estocástico</b>	<b>37</b>
2.1 A Integral de Itô . . . . .	37
2.2 A Fórmula de Itô . . . . .	42
2.3 Equações Diferenciais Estocásticas . . . . .	47
<b>3 Finanças em Tempo Contínuo</b>	<b>55</b>
3.1 Conceitos Básicos . . . . .	56
3.2 O Modelo Matemático . . . . .	59
3.3 O Modelo de Black-Scholes . . . . .	61
3.4 Generalizações de Black-Scholes . . . . .	67
3.4.1 O Modelo de Dupire . . . . .	67
3.4.2 O Modelo de Heston . . . . .	70

<b>4</b>	<b>Simulações Numéricas</b>	<b>75</b>
4.1	O Método Euler-Maruyama . . . . .	75
4.2	Métodos de Diferenças Finitas . . . . .	81
<b>5</b>	<b>Calibração de Parâmetros</b>	<b>91</b>
5.1	A Regularização de Tikhonov . . . . .	91
5.2	Problemas Inversos em Finanças . . . . .	94
5.2.1	Volatilidade Implícita . . . . .	95
5.2.2	Volatilidade Local . . . . .	97
	<b>Bibliografia</b>	<b>105</b>

# Introdução

Estas notas têm como objetivo apresentar às pessoas familiares ao Cálculo, ou com formação na área de Ciências Exatas e em Engenharia, algumas técnicas de Métodos Matemáticos em Finanças, também chamada de Finanças Quantitativas. Esta área de pesquisa tem fornecido uma imensa classe de técnicas computacionais e analíticas, com base sólida em teorias matemáticas, aos atores dos mercados (financeiro, de commodities e de títulos).

Métodos Matemáticos em Finanças ou Finanças Quantitativas formam uma classe de técnicas que procuram descrever a evolução de ativos e derivativos em diversos mercados, como o mercado financeiro, o de mercado de títulos e o mercado de commodities, em que, a natureza aleatória destas quantidades não pode ser negligenciada. Assim, a teoria de processos estocásticos funciona como ferramenta básica para a construção de modelos precisos para preços de ativos. No entanto, muitas outras áreas da matemática e da estatística fornecem ferramentas fundamentais para o desenvolvimento de análises precisas. Dentre elas, podemos citar as áreas de Otimização, Equações Diferenciais Parciais, Análise Numérica, Econometria, Teoria de Jogos, Teoria de Controle e etc.

Em Finanças Quantitativas existem diversos objetivos, como a cobertura de risco, a alocação de recursos, a precificação de derivativos,

e etc. Assim, dependendo do que se quer estudar, uma área terá maior predominância que a outra. Por exemplo, em análise de investimentos, técnicas de otimização são fundamentais, já na precificação de derivativos, ferramentas de equações diferenciais parciais e análise numérica são cruciais. Assim, fica evidente a interdisciplinaridade desta área de pesquisa que continua a ser muito ativa .

Com o intuito de ser o mais auto-contido possível, estas notas apresentarão alguns capítulos contendo revisões de assuntos básicos, como Teoria de Probabilidade, a construção do movimento Browniano, e Cálculo Estocástico. Também serão apresentadas técnicas para a solução numérica dos modelos aqui apresentados, bem como para a calibragem ou estimação dos seus parâmetros.

Vale alertar aos leitores que estão tendo um primeiro contato com as técnicas apresentadas nestas notas que a intuição cresce com a experiência. Assim, num primeiro momento, o leitor poderá se sentir desconfortável com a teoria, mas, com a devida prática, alcançará uma compreensão mais precisa no futuro.

# Capítulo 1

## Fatos Básicos em Teoria de Probabilidade

Neste capítulo serão revisados alguns conceitos básicos de teoria de probabilidade, como espaço amostral,  $\sigma$ -álgebra, medida de probabilidade, variável aleatória, valor esperado e etc. Esta é uma adaptação do Capítulo 2 de [10].

### Definições Básicas

Seja  $\Omega$  um certo conjunto. Se  $A$  é um subconjunto de  $\Omega$ , e tem-se que  $A \subset \Omega$ . Então complementar de  $A$ , denotado por  $A^c$ , é definido como  $\Omega - A$ , isto é, o conjunto de todos os elementos de  $\Omega$  que não estão em  $A$ . No que se segue, será necessário o conceito de  $\sigma$ -álgebra.

**Definição 1.1.** *Uma  $\sigma$ -álgebra é uma coleção  $\mathcal{F}$  de subconjuntos de  $\Omega$  satisfazendo as seguintes propriedades:*

1.  $\emptyset$  e  $\Omega$  pertencem a  $\mathcal{F}$ .
2. Se  $A \in \mathcal{F}$ , então o seu complementar também pertence a  $\mathcal{F}$ .

3. Se a sequência de conjuntos  $\{A_j\}_{j \in \mathbb{N}}$  está contida em  $\mathcal{F}$ , então a união e a interseção de todos os  $A_j$ 's também estão em  $\mathcal{F}$ . Ou seja,

$$\bigcup_{j \in \mathbb{N}} A_j \in \mathcal{F} \quad e \quad \bigcap_{j \in \mathbb{N}} A_j \in \mathcal{F}.$$

Considere agora o conceito de medida de probabilidade:

**Definição 1.2.** *Seja  $\mathcal{F}$  uma  $\sigma$ -álgebra de  $\Omega$ . Uma função  $P$  que associa a cada conjunto em  $\mathcal{F}$  um valor no intervalo  $[0, 1]$ , isto é,  $P : \mathcal{F} \rightarrow [0, 1]$ , é chamada de medida de probabilidade se satisfizer:*

1.  $P(\emptyset) = 0$  e  $P(\Omega) = 1$ .
2. Se a sequência de conjuntos  $\{A_j\}_{j \in \mathbb{N}}$  está contida em  $\mathcal{F}$ , então vale a seguinte desigualdade, também chamada de sub-aditividade:

$$P \left( \bigcup_{j \in \mathbb{N}} A_j \right) \leq \sum_{j \in \mathbb{N}} P(A_j).$$

3. Vale a igualdade se a sequência acima for formada por conjuntos dois a dois disjuntos. Ou seja, se valer  $A_j \cap A_k = \emptyset$  para todos  $j \neq k$ , então

$$P \left( \bigcup_{j \in \mathbb{N}} A_j \right) = \sum_{j \in \mathbb{N}} P(A_j).$$

Segue da definição acima que se  $A \subset B$ , então  $P(A) \leq P(B)$ .

A trinca  $(\Omega, \mathcal{F}, P)$ , formada por um conjunto  $\Omega$ , uma  $\sigma$ -álgebra  $\mathcal{F}$  de  $\Omega$  e uma medida de probabilidade  $P$  definida em  $\mathcal{F}$ , é chamada de espaço de probabilidade. Neste contexto, um subconjunto  $A \subset \Omega$  é chamado de evento e um ponto  $\omega \in \Omega$  é chamado de ponto amostral. Já o conjunto  $\Omega$  é chamado de espaço amostral.



Quando uma dada propriedade é verdadeira a menos de um evento de probabilidade zero, é dito que esta propriedade ocorre quase certamente, podendo ser abreviada como “q.c.”.

**Exemplo 1.1** (O Modelo Binomial de Um Período). *Considere um jogo de moedas justo, isto é, em que as chaces de se obter cara ou coroa são iguais. Assim, podemos definir o espaço amostral como*

$$\Omega = \{\text{cara}, \text{coroa}\}.$$

*Assim, temos a seguinte  $\sigma$ -álgebra de  $\Omega$*

$$\mathcal{F} = \{\emptyset, \{\text{cara}\}, \{\text{coroa}\}, \Omega\}$$

*e a medida de probabilidade  $P : \mathcal{F} \rightarrow [0, 1]$  assumindo os valores:*

$$P(\emptyset) = 0, \quad P(\Omega) = 1, \quad P(\{\text{cara}\}) = 1/2 \quad \text{e} \quad P(\{\text{coroa}\}) = 1/2.$$

**Exercício 1.1.** *Verifique que a trinca  $(\Omega, \mathcal{F}, P)$  definida no exemplo anterior é de fato um espaço de probabilidade. Isto é, verifique que  $\mathcal{F}$  é uma  $\sigma$ -álgebra e que  $P$  é uma medida de probabilidade.*

Quando o conjunto  $\Omega$  tem um número finito de elementos, a  $\sigma$ -álgebra  $\mathcal{F}$  pode ser escolhida como sendo o conjunto contendo todos os subconjuntos de  $\Omega$ , isto é, o conjunto das partes de  $\Omega$ . No entanto, se  $\Omega$  é infinito e não enumerável, não é possível contruir uma medida de probabilidade que associe a cada subconjunto de  $\Omega$  um valor em  $[0, 1]$ . Para maiores detalhes, veja [3].

**Exemplo 1.2.** *A menor  $\sigma$ -álgebra contendo todos os conjuntos abertos de  $\mathbb{R}^n$  é chamada de  $\sigma$ -álgebra de Borel, ou simplesmente borelianos. Em geral, ela é denotada por  $\mathcal{B}$ . Seja  $f : \mathbb{R}^n \rightarrow \mathbb{R}_+$  integrável em  $\mathbb{R}^n$  e*

tal que  $\int_{\mathbb{R}^n} f(x)dx = 1$ . Defina

$$P(B) = \int_B f(x)dx,$$

para todo  $B \in \mathcal{B}$ . Então  $P$  é uma medida de probabilidade definida em  $\mathcal{B}$  e a trinca  $(\mathbb{R}^n, \mathcal{B}, P)$  é um espaço de probabilidade. A função  $f$  é chamada de função de densidade de probabilidade da medida  $P$ .

**Exemplo 1.3.** No exemplo anterior, se  $n = 1$  e  $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ , tem-se que  $P$  é a medida de probabilidade normal padrão.

Em geral, tem-se um espaço de probabilidade  $(\Omega, \mathcal{F}, P)$  abstrato, isto é, desconhecido ou “não-observável” e uma variável aleatória, que é uma função que associa elementos amostrais de  $\Omega$ , a elementos amostrais de um espaço de probabilidade “observável” como  $(\mathbb{R}^n, \mathcal{B}, P)$ . De modo mais preciso, tem-se a seguinte definição:

**Definição 1.3.** Sejam  $(\Omega, \mathcal{F}, P)$  um espaço de probabilidade e uma função  $X : \Omega \rightarrow \mathbb{R}^n$ .  $X$  é chamada variável aleatória  $n$ -dimensional se, para cada  $B \in \mathcal{B}$ , tem-se que a pré-imagem de  $B$  por  $X$  está na  $\sigma$ -álgebra  $\mathcal{F}$ , isto é,  $X^{-1}(B) \in \mathcal{F}$ . Equivalentemente, diz-se que  $X$  é  $\mathcal{F}$ -mensurável.

Observe que, em geral, omite-se a dependência da variável aleatória  $X$  de  $\omega \in \Omega$ , ou seja, ao invés de se ter  $X(\omega)$ , escreve-se, simplesmente,  $X$ . Além disso, ao a (medida de) probabilidade de  $X$  pertencer a um dado conjunto  $B$  é escrita como  $P(X \in B)$  ao invés de  $P(X^{-1}(B))$ .

A seguir, serão apresentados alguns exemplos de variáveis aleatórias.

**Exemplo 1.4.** No modelo binomial de um período, representado pelo jogo de moedas, a função  $X : \Omega \rightarrow \mathbb{R}$

$$X(\omega) = \begin{cases} 0, & \text{se } \omega = \text{cara,} \\ 1, & \text{se } \omega = \text{coroa} \end{cases}$$

é uma variável aleatória.

**Exercício 1.2.** Verificar que a função do exemplo anterior é de fato uma variável aleatória.

**Exemplo 1.5.** A função indicadora do conjunto  $A \in \mathcal{F}$  definida como

$$\mathcal{X}_A(\omega) = \begin{cases} 0, & \text{se } \omega \notin A, \\ 1, & \text{se } \omega \in A, \end{cases}$$

é uma variável aleatória. Mais geralmente, se os conjuntos  $A_1, A_2, \dots, A_n$  são elementos de  $\mathcal{F}$ , tais que  $\cup_{j=1}^n A_j = \Omega$ , e  $a_1, a_2, \dots, a_n$  são números reais, então

$$X = \sum_{j=1}^n a_j \mathcal{X}_{A_j}$$

é uma variável aleatória chamada de função simples.

A definição a seguir introduz o conceito de  $\sigma$ -álgebra gerada por uma dada variável aleatória  $X$ .

**Definição 1.4.** Seja  $X : \Omega \rightarrow \mathbb{R}^n$  uma variável aleatória. Então, a  $\sigma$ -álgebra gerada por  $X$  é definida como

$$\mathcal{F}(X) = \{X^{-1}(B) : B \in \mathcal{B}\},$$

lembrando que  $\mathcal{B}$  é a  $\sigma$ -álgebra dos borelianos de  $\mathbb{R}^n$ .

**Exercício 1.3.** Verificar que a  $\sigma$ -álgebra gerada pela variável aleatória  $X$  é de fato uma  $\sigma$ -álgebra.

Observe que a  $\sigma$ -álgebra gerada pela variável aleatória  $X$  contém, em termos probabilísticos, todas as informações relevantes a respeito de  $X$ . Em finanças quantitativas, o conceito de informação disponível até um dado instante de tempo é representado através de uma  $\sigma$ -álgebra.

Assim, a medida que o tempo passa, constrói-se uma família “crescente” de  $\sigma$ -álgebras chamada de filtro ou filtração.

Outro conceito fundamental em teoria de probabilidade e em finanças quantitativas é o de processo estocástico. Processos estocásticos são famílias indexadas de variáveis aleatórias, em que esse índice pode ser discreto ou contínuo. Como o objetivo destas notas é introduzir o leitor ao mundo das finanças em tempo contínuo, será apresentado apenas o caso de índices definidos em  $\mathbb{R}_+$ , o qual representará o tempo.

**Definição 1.5.** *Uma coleção  $\{X(t) : t \in \mathbb{R}_+\}$  de variáveis aleatórias é chamada de processo estocástico. Ou seja, para cada  $t \geq 0$ ,  $X(t) : \Omega \rightarrow \mathbb{R}^n$  é uma variável aleatória. Para cada  $\omega \in \Omega$  o mapa  $t \mapsto X(t, \omega)$  é o correspondente caminho amostral de  $X$ .*

Sejam  $(\Omega, \mathcal{F}, P)$  um espaço de probabilidade e  $X = \sum_{j=1}^m a_j \mathcal{X}_{A_j}$  uma função simples aleatória e real. A integral de variável aleatória simples  $X$  é definida como

$$\int_{\Omega} X dP = \sum_{j=1}^m a_j P(A_j).$$

Se  $X$  é uma variável aleatória não-negativa, define-se a sua integral como sendo o supremo, isto é, a menor cota superior, de todas as integrais de variáveis aleatórias simples  $Y$  tais que  $Y(\omega) \leq X(\omega)$  para todo  $\omega \in \Omega$ . Ou seja,

$$\int_{\Omega} X dP = \sup_{Y \leq X, Y \text{ simples}} \int_{\Omega} Y dP.$$

Se  $X$  é uma variável aleatória qualquer, então,  $X$  pode ser decomposta como a diferença  $X = X^+ - X^-$ , em que  $X^+ = \max\{0, X\}$ , e  $X^- = \max\{0, -X\}$ . Como  $X^+$  e  $X^-$  são variáveis aleatórias não negativas,

tem-se que a integral de  $X$  é dada por:

$$\int_{\Omega} X dP = \int_{\Omega} X^+ dP - \int_{\Omega} X^- dP,$$

desde que ao menos uma das integrais no lado direito da equação acima seja finita.

Notem que a integral definida acima não é a integral de Riemann e sim, uma adaptação da integral de Lebesgue para medidas de probabilidade. O leitor interessa neste assunto pode aprender um pouco mais sobre essa magnífica teoria em [3].

Se  $X : \Omega \rightarrow \mathbb{R}^n$  é uma variável aleatória, então  $X = (X_1, \dots, X_n)$ , com  $X_j : \Omega \rightarrow \mathbb{R}$ ,  $j = 1, \dots, n$ . A integral de  $X$  é definida como o seguinte vetor:

$$\int_{\Omega} X dP = \left( \int_{\Omega} X_1 dP, \dots, \int_{\Omega} X_n dP \right).$$

No que se segue, serão as propriedades da integral de  $X$  serão assumidas sem prova. Para maiores detalhes, veja [3].

A integral da variável aleatória  $X$  sobre o espaço amostral  $\Omega$  é chamada de esperança, valor médio, média ou valor esperado de  $X$ , o qual é denotado por  $E(X)$ , ou seja,

$$E(X) = \int_{\Omega} X dP.$$

A variância de  $X$  é definida como

$$Var(X) = E(|X - E(X)|^2) = E(|X|^2) - |E(X)|^2.$$

Aqui,  $|X| = \sqrt{X_1^2 + \dots + X_n^2}$  representa a norma euclidiana.

O resultado a seguir, conhecido como desigualdade de Chebyshev, é de grande importância em probabilidade, apesar de sua simplicidade.

**Lema 1.1.** *Se  $X$  é uma variável aleatória e  $1 \leq p < \infty$ , então*

$$P(|X| \geq \lambda) \leq \frac{1}{\lambda^p} E(|X|^p),$$

para todo  $\lambda > 0$ .

*Demonstração.* Observe que

$$\begin{aligned} E(|X|^p) &= \int_{\Omega} |X|^p dP \geq \int_{\{|X| \geq \lambda\}} |X|^p dP \geq \int_{\{|X| \geq \lambda\}} \lambda^p dP \\ &= \lambda^p \int_{\Omega} \mathcal{X}_{\{|X| \geq \lambda\}} dP = \lambda^p P(|X| \geq \lambda). \end{aligned}$$

□

Sejam  $x, y \in \mathbb{R}^n$ , com  $x = (x_1, \dots, x_n)$  e  $y = (y_1, \dots, y_n)$ . Diz-se que  $x \leq y$  quando  $x_j \leq y_j$  para todo  $j = 1, \dots, n$ .

A seguir serão apresentadas algumas definições adicionais.

**Definição 1.6.** *Sejam  $(\Omega, \mathcal{F}, P)$  um espaço de probabilidade e  $X: \Omega \rightarrow \mathbb{R}^n$  uma variável aleatória.*

1. *A função de distribuição de probabilidade de  $X$ , ou simplesmente a distribuição de  $X$ , é a função  $F_X: \mathbb{R}^n \rightarrow [0, 1]$  definida por*

$$F_X(x) = P(X \leq x),$$

para todo  $x \in \mathbb{R}^n$ .

2. *Se  $X_1, \dots, X_m$  são variáveis aleatórias com valores em  $\mathbb{R}^n$ , então a sua função de distribuição conjunta  $F_{X_1, \dots, X_m}: (\mathbb{R}^n)^m \rightarrow [0, 1]$  é definida como:*

$$F_{X_1, \dots, X_m}(x^1, \dots, x^m) = P(X_1 \leq x^1, \dots, X_m \leq x^m),$$

para todo  $x_j \in \mathbb{R}^n$ ,  $j = 1, \dots, m$ .

3. Se existir uma função integrável em  $\mathbb{R}^n$   $f_X : \mathbb{R}^n \rightarrow \mathbb{R}$  tal que

$$F_X(x) = F(x_1, \dots, x_n) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} f_X(s_1, \dots, s_n) ds_1 \dots ds_n,$$

então  $f_X$  é chamada de função de densidade de probabilidade de  $X$ , ou simplesmente densidade de  $X$ .

Observe que, se  $X$  possuir uma densidade  $f_X : \mathbb{R}^n \rightarrow \mathbb{R}$ , então

$$P(X \in B) = \int_B f(x) dx$$

para todo  $B$  em  $\mathcal{B}$ .

Observe que no lado direito da equação acima tem-se uma integral sobre um subconjunto de  $\mathbb{R}^n$ . Assim, se  $f_X$  é conhecida, esta integral pode ser calculada explicitamente.

**Exemplo 1.6.** Se  $X \Omega \rightarrow \mathbb{R}$  possui como densidade a função

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{|x-m|^2}{2\sigma^2}},$$

para todo  $x \in \mathbb{R}$ , então  $X$  tem distribuição normal (gaussiana) com média  $m$  e variância  $\sigma^2$ . Neste caso, denota-se

$$X = {}^d N(m, \sigma^2).$$

**Exemplo 1.7.** Se  $X \Omega \rightarrow \mathbb{R}^n$  possui como densidade a função

$$f_X(x) = \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{\det \Sigma}} e^{-\frac{1}{2}(x-m)^T \Sigma^{-1}(x-m)},$$

para todo  $x \in \mathbb{R}^n$ , com  $m \in \mathbb{R}^n$  fixo e  $\Sigma$  uma matriz simétrica, então  $X$  tem distribuição normal (gaussiana) com média  $m$  e matriz covariância

$\Sigma$ . Neste caso, denota-se

$$X = {}^d N(m, \Sigma).$$

Observe que, se  $X : \Omega \rightarrow \mathbb{R}^n$  é uma variável aleatória, que tem como distribuição a função  $F_X$  e densidade  $f_X$ , então, se  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  é uma função e  $Y = g(X)$  é integrável, tem-se que

$$E(Y) = \int_{\mathbb{R}^n} g(x)f(x)dx.$$

Em particular,

$$E(X) = \int_{\mathbb{R}^n} xf(x)dx \quad \text{e} \quad \text{Var}(X) = \int_{\mathbb{R}^n} |x - E(X)|^2 dx.$$

**Exercício 1.4.** *Verifique que os fatos citados acima são verdadeiros para uma variável aleatória  $X : \Omega \rightarrow \mathbb{R}^n$  e uma função simples  $g : \mathbb{R}^n \rightarrow \mathbb{R}$ , definida como  $g(x) = \sum_{j=1}^m b_j \mathcal{X}_{B_j}(x)$ , com  $b_j \in \mathbb{R}$  e  $B_j \in \mathcal{B}$ .*

**Exercício 1.5.** *Se  $X : \Omega \rightarrow \mathbb{R}^n$  é tal que  $X = {}^d N(m, \sigma^2)$ , mostre que  $E(X) = m$  e  $\text{Var}(X) = \sigma^2$  usando as identidades acima.*

Um conceito fundamental em probabilidade é o de independência, para isto, é necessário introduzir outras notações. Sejam  $(\Omega, \mathcal{F}, P)$  um espaço de probabilidade e  $A, B \in \mathcal{F}$  dois eventos, com  $P(B) > 0$ . Se  $\omega \in \Omega$  é escolhido aleatoriamente, sabendo-se que  $\omega \in B$ , então, qual seria a probabilidade de  $\omega$  também pertencer a  $A$ ? Como  $\omega \in B$ , o conjunto  $B$  pode ser pensado como um novo espaço amostral e define-se o espaço de probabilidade  $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ , com  $\tilde{\Omega} = B$ ,  $\tilde{\mathcal{F}} = \{C \cap B : C \in \mathcal{F}\}$  e  $\tilde{P} = P/P(B)$ . Então, a probabilidade de  $\omega$  estar em  $A$  é

$$\tilde{P}(A \cap B) = \frac{P(A \cap B)}{P(B)}.$$



De maneira mais precisa, tem-se a definição a seguir:

**Definição 1.7** (Probabilidade Condicional). *Sejam  $(\Omega, \mathcal{F}, P)$  um espaço de probabilidade e  $A, B \in \mathcal{F}$  dois eventos, com  $P(B) > 0$ . Então, a probabilidade condicional de  $A$  dado  $B$  é dada por:*

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

Matematicamente, dizer que  $A$  é independente de  $B$  é equivalente a se ter  $P(A|B) = P(A)$ . Intuitivamente, isto significa que o evento  $B$  ter ocorrido é irrelevante para determinar a probabilidade de  $A$  ter ocorrido. Assim, se  $A$  e  $B$  são independentes, tem-se que:

$$P(A) = P(A|B) = \frac{P(A \cap B)}{P(B)},$$

ou seja,  $P(A \cap B) = P(A)P(B)$ .

**Exercício 1.6.** *Mostre que, se  $A$  e  $B$  são eventos independentes, então os eventos  $A^c$  e  $B$  são independentes e  $A^c$  e  $B^c$  também são independentes.*

**Definição 1.8.** *Seja  $\{A_j\}_{j \in \mathbb{N}}$  uma sequência de eventos em  $\mathcal{F}$ . Estes eventos são ditos independentes se, para quaisquer escolha de índices  $1 \leq k_1 < k_2 < \dots < k_m$ , tem-se que*

$$P(A_{k_1} \cap A_{k_2} \cap \dots \cap A_{k_m}) = P(A_{k_1})P(A_{k_2}) \dots P(A_{k_m}).$$

Deste modo, é possível estender este conceito para  $\sigma$ -álgebras e variáveis aleatórias.

**Definição 1.9.** *Seja  $\{\mathcal{F}_j\}_{j \in \mathbb{N}}$  uma sequência de  $\sigma$ -álgebras tal que  $\mathcal{F}_j \subset \mathcal{F}$ . Estas  $\sigma$ -álgebras são independentes se, para toda a escolha de índices  $1 \leq k_1 < k_2 < \dots < k_m$  e de toda escolha de eventos  $A_{k_l} \in \mathcal{F}_{k_l}$*

com  $l = 1, \dots, m$ , tem-se que

$$P(A_{k_1} \cap A_{k_2} \cap \dots \cap A_{k_m}) = P(A_{k_1})P(A_{k_2}) \cdots P(A_{k_m}).$$

**Definição 1.10.** *Seja  $\{X_j\}_{j \in \mathbb{N}}$  uma sequência de variáveis aleatórias  $X_j : \Omega \rightarrow \mathbb{R}^n$  com  $j \in \mathbb{N}$ . Diz-se que tais variáveis aleatórias são independentes se, para toda escolha de inteiros  $m \geq 2$  e de conjuntos borelianos  $B_1, \dots, B_m$  em  $\mathbb{R}^n$ , tem-se que*

$$\begin{aligned} P(X_1 \in B_1, X_2 \in B_2, \dots, X_m \in B_m) = \\ P(X_1 \in B_1)P(X_2 \in B_2) \cdots P(X_m \in B_m). \end{aligned}$$

Observe que dizer que a sequência de variáveis aleatórias  $\{X_j\}_{j \in \mathbb{N}}$  é independente é equivalente a dizer que a sequência de  $\sigma$ -álgebras geradas pelas variáveis aleatórias  $X_j$  com  $j \in \mathbb{N}$ , isto é,  $\{\mathcal{F}(X_j)\}_{j \in \mathbb{N}}$ , é independente. Observe ainda que  $P(X \in A, Y \in B) = P(X^{-1}(A) \cap Y^{-1}(B))$ .

A prova do lema a seguir pode ser encontrada em [6]

**Lema 1.2.** *Sejam  $X_1, X_2, \dots, X_{m+n}$  variáveis aleatórias independentes,  $X_j : \Omega \rightarrow \mathbb{R}^k$ ,  $j = 1, \dots, m+n$ . Considere as funções  $f : (\mathbb{R}^k)^n \rightarrow \mathbb{R}$  e  $g : (\mathbb{R}^k)^m \rightarrow \mathbb{R}$ . Então, as variáveis aleatórias  $Y$  e  $Z$ , definidas como*

$$Y = f(X_1, \dots, X_n) \quad e \quad Z = g(X_{n+1}, \dots, X_{m+n})$$

*são independentes.*

Em outras palavras, o lema anterior mostra que funções de variáveis aleatórias independentes e distintas também são variáveis aleatórias independentes. A seguir serão apresentados sem prova alguns resultados relativos a variáveis aleatórias independentes. Para as demonstrações, veja [10].

**Teorema 1.1.** *As variáveis aleatórias  $X_1, \dots, X_m$ , com  $X_j : \Omega \rightarrow \mathbb{R}^n$*

e  $j = 1, \dots, n$ , são independentes se e só se a função de distribuição conjunta for igual ao produto das funções de distribuição, isto é,

$$F_{X_1, \dots, X_m}(x_1, \dots, x_m) = F_{X_1}(x_1) \cdots F_{X_m}(x_m),$$

para todo  $x_j \in \mathbb{R}^n$ ,  $j = 1, \dots, m$ . Se as variáveis aleatórias possuírem densidade, então, a independência é equivalente a

$$f_{X_1, \dots, X_m}(x_1, \dots, x_m) = f_{X_1}(x_1) \cdots f_{X_m}(x_m),$$

para todo  $x_j \in \mathbb{R}^n$ ,  $j = 1, \dots, m$ .

**Exercício 1.7.** Prove a primeira parte do teorema anterior, para isso use a definição de independência e de função de distribuição de probabilidade.

**Teorema 1.2.** Sejam  $X_1, \dots, X_m$ , com  $X_j : \Omega \rightarrow \mathbb{R}$  e  $j = 1, \dots, n$ , variáveis aleatórias independentes.

1. Se estas variáveis aleatórias têm módulo integrável, isto é  $E(|X_j|) < \infty$  para todo  $j = 1, \dots, m$ , então  $E(|X_1 \cdots X_m|) < \infty$  e

$$E(X_1 \cdots X_m) = E(X_1) \cdots E(X_m).$$

2. Se a variância destas variáveis aleatórias é finita, ou seja  $\text{Var}(X_j) < \infty$  para todo  $j = 1, \dots, m$ , então

$$\text{Var}(X_1 + \cdots + X_m) = \text{Var}(X_1) + \cdots + \text{Var}(X_m).$$

**Exercício 1.8.** Prove o primeiro item teorema anterior para o caso de variáveis aleatórias com densidade. Prove também o segundo item, usando argumentos de indução finita e o primeiro item para o caso geral.

Outro conceito fundamental em probabilidade é o Lema de Borel-Cantelli, mas antes de apresentá-lo, é necessário introduzir o conceito de evento infinitamente comum.

**Definição 1.11.** *Seja  $\{A_j\}_{j \in \mathbb{N}}$  uma seqüência de eventos em  $\mathcal{F}$ . Então o evento*

$$\bigcap_{n \in \mathbb{N}} \bigcup_{j \geq n} A_j = \{\omega \in \Omega : \omega \text{ pertence a infinitos eventos } A_n\}$$

é chamado de “ $A_n$  infinitamente comum” ou simplesmente “ $A_n$  i.c.”.

**Lema 1.3** (Borel-Cantelli). *Se uma seqüência de eventos  $\{A_j\}_{j \in \mathbb{N}}$  em  $\mathcal{F}$  é tal que  $\sum_{j \in \mathbb{N}} P(A_j) < \infty$ , então  $P(A_n \text{ i.c.}) = 0$ .*

*Demonstração.* Usando a definição de  $A_n$  i.c., tem-se que

$$P(A_n \text{ i.c.}) = P\left(\bigcap_{n \in \mathbb{N}} \bigcup_{j \geq n} A_j\right) \leq P\left(\bigcup_{j \geq n} A_j\right) \leq \sum_{j \geq n} P(A_j).$$

Como as desigualdades acima valem para todo  $n \in \mathbb{N}$  e  $\sum_{j \in \mathbb{N}} P(A_j) < \infty$ , tomando o limite quando  $n \rightarrow \infty$ , tem-se o resultado.  $\square$

Uma seqüência de variáveis aleatórias  $\{X_j\}_{j \in \mathbb{N}}$  definida num espaço de probabilidade é dita convergir em probabilidade para uma variável aleatória  $X$  quando, dado  $\varepsilon > 0$  tem-se que

$$\lim_{j \rightarrow \infty} P(|X_j - X|) = 0.$$

**Teorema 1.3.** *Se  $X_j \rightarrow X$  em probabilidade, então existe uma subseqüência  $\{X_{j_l}\}_{l \in \mathbb{N}}$  tal que  $X_{j_l}(\omega) \rightarrow X(\omega)$  para quase todo  $\omega$ , ou seja, a probabilidade do conjunto de pontos de  $\omega$  em que não ocorre a convergência é nula.*

*Demonstração.* Para cada  $l \in \mathbb{N}$  é possível selecionar  $j_l \in \mathbb{N}$  tal que

$$P\left(|X_{j_l} - X| > \frac{1}{l}\right) \leq \frac{1}{l^2}.$$

Considere os eventos  $A_l = [|X_{j_l} - X| > \frac{1}{l}]$ . Como  $\sum_{l \in \mathbb{N}} \frac{1}{l^2} < \infty$ , o Lema de Borel-Cantelli implica que  $P(A_l \text{ i.c.}) = 0$ . Portanto, para quase todo  $\omega \in \Omega$ ,  $|X_{j_l}(\omega) - X(\omega)| \leq \frac{1}{l}$  desde que  $j \geq J$ , para algum índice  $J = J(\omega)$ .  $\square$

Outra classe de ferramentas bastante útil em probabilidade é a das funções características.

**Definição 1.12.** *Seja  $X : \Omega \rightarrow \mathbb{R}^n$  uma variável aleatória. Então a função característica de  $X$ , denotada por  $\phi_X$ , é definida por*

$$\phi_X(\lambda) = E\left(e^{i\lambda^T X}\right),$$

para todo  $\lambda \in \mathbb{R}^n$ .

**Exercício 1.9.** *Mostre que se  $X \stackrel{d}{=} N(m\sigma^2)$ , então*

$$\phi_X(\lambda) = e^{im\lambda - \frac{\lambda^2\sigma^2}{2}},$$

para todo  $\lambda \in \mathbb{R}$ .

O resultado a seguir apresenta uma série de propriedades das funções características, para a sua prova, veja [10] e [6].

**Lema 1.4.** *1. Se  $X_1, \dots, X_m$  são variáveis aleatórias independentes com valores em  $\mathbb{R}^n$ , então para todo  $\lambda \in \mathbb{R}^n$*

$$\phi_{X_1 + \dots + X_m}(\lambda) = \phi_{X_1}(\lambda) \cdots \phi_{X_m}(\lambda).$$

2. Se  $X$  é uma variável aleatória com valores em  $\mathbb{R}$ , então, a  $k$ -ésima derivada de  $\phi_X$ , com  $k \in \mathbb{N}$ , satisfaz:

$$\phi_X^{(k)}(0) = i^k E(X^k).$$

3. Se  $X$  e  $Y$  são variáveis aleatórias com valores em  $\mathbb{R}^n$  e suas funções características são iguais, isto é,  $\phi_X(\lambda) = \phi_Y(\lambda)$  para todo  $\lambda \in \mathbb{R}^n$ , então,  $X$  e  $Y$  possuem a mesma distribuição de probabilidade, ou seja,  $F_X(x) = F_Y(x)$  para todo  $x \in \mathbb{R}^n$ .

Quando duas variáveis aleatórias  $X$  e  $Y$  possuem a mesma distribuição, diz-se que elas são iguais em distribuição e escreve-se  $X \stackrel{d}{=} Y$ . Observe que ao dizer que duas variáveis aleatórias são iguais em distribuição não se está dizendo que elas são iguais. Elas podem ter a mesma distribuição e ser independentes.

## O Teorema de Limite Central e a Lei dos Grandes Números

Outros fatos de suma importância em teoria da probabilidade com impactos cruciais em finanças são as leis dos grandes números e o teorema do limite central.

**Definição 1.13.** *Seja  $(\Omega, \mathcal{F}, P)$  um espaço de probabilidade e considere uma sequência de variáveis aleatórias  $\{X_j\}_{j \in \mathbb{N}}$  definida neste espaço tal que*

$$X_j : \Omega \rightarrow \mathbb{R}^n,$$

para  $j \in \mathbb{N}$ . Diz-se que esta é uma sequência de variáveis aleatórias identicamente distribuídas se as funções de distribuição de todas as

## 1.2. O TEOREMA DE LIMITE CENTRAL E A LEI DOS GRANDES NÚMEROS

variáveis aleatórias da sequência coincidirem, ou seja,

$$F_{X_1}(x) = F_{X_2}(x) = \cdots = F_{X_j}(x) = \cdots ,$$

para todo  $x \in \mathbb{R}^n$ .

Caso esta sequência de variáveis aleatórias seja independente, diz-se que  $\{X_j\}_{j \in \mathbb{N}}$  é uma sequência de variáveis aleatórias independentes e identicamente distribuídas e escreve-se que  $\{X_j\}_{j \in \mathbb{N}}$  é “i.i.d.”.

O resultado a seguir é chamado Lei dos Grandes Números *Forte* e sua prova pode ser encontrada em [10]. Também existe uma Lei dos Grandes Números *Fraca*, mas esta não será apresentada nestas notas.

**Teorema 1.4** (Lei dos Grandes Números *Forte*). *Seja  $(\Omega, \mathcal{F}, P)$  um espaço de probabilidade e uma sequência de variáveis aleatórias i.i.d.  $\{X_j\}_{j \in \mathbb{N}}$  definida neste espaço e tal que  $X_j \Omega \rightarrow \mathbb{R}^n$  e  $|E(X_j)| < \infty$ ,  $j \in \mathbb{N}$ . Denote  $m = E(X_j)$ , com  $j \in \mathbb{N}$ . Então,*

$$P \left( \lim_{j \rightarrow \infty} \frac{X_1 + X_2 + \cdots + X_j}{j} = m \right) = 1.$$

Observe que este teorema diz que para quase todo  $\omega \in \Omega$  vale o limite

$$\frac{X_1 + X_2 + \cdots + X_j}{j} \rightarrow m \quad \text{quando } j \rightarrow \infty.$$

**Lema 1.5.** *Seja  $(\Omega, \mathcal{F}, P)$  um espaço de probabilidade e uma sequência de variáveis aleatórias i.i.d.  $\{X_j\}_{j \in \mathbb{N}}$  definida neste espaço e tal que  $X_j \Omega \rightarrow \mathbb{R}$ ,*

$$P(X_j = 0) = p \quad \text{e} \quad P(X_j = 1) = q,$$

para todo  $j \in \mathbb{N}$ , com  $p + q = 1$  e  $p, q \geq 0$ . Então,

$$E(X_1 + \cdots + X_m) = mp \quad \text{e} \quad \text{Var}(X_1 + \cdots + X_m) = mpq,$$

para todo  $m \in \mathbb{N}$ .

*Demonstração.* Como  $p + q = 1$  e  $P(\Omega) = 1$ , segue que, para cada  $j \in \mathbb{N}$ , o evento  $A = [X_j \neq 0] \cap [X_j \neq 1]$  tem probabilidade nula. Como  $\Omega = A \cup [X_j = 0] \cup [X_j = 1]$  é uma união disjunta (por que?), tem-se que

$$\begin{aligned} E(X_j) &= \int_{\Omega} X_j dP = \int_A X_j dP + \int_{[X_j=0]} X_j dP + \int_{[X_j=1]} X_j dP \\ &= 0 + 0 + P(X_j = 1) = p, \end{aligned}$$

para cada  $j \in \mathbb{N}$ . Assim, pela linearidade do valor esperado  $E(X_1 + \dots + X_m) = mp$ .

De modo similar, para cada  $j \in \mathbb{N}$ ,

$$\begin{aligned} \text{Var}(X_j) &= \int_{\Omega} (X_j - p)^2 dP \\ &= \int_A (X_j - p)^2 dP + \int_{[X_j=0]} (X_j - p)^2 dP + \int_{[X_j=1]} (X_j - p)^2 dP \\ &= 0 + (0 - p)^2 P(X_j = 0) + (p - p)^2 P(X_j = 1) = pq. \end{aligned}$$

Pela independência da sequência das variáveis aleatórias, tem-se que

$$\text{Var}(X_1 + \dots + X_m) = \text{Var}(X_1) + \dots + \text{Var}(X_m)$$

e o resultado segue da estimativa acima.  $\square$

O teorema a seguir é uma versão simplificada do Teorema do Limite Central e estabelece que as somas parciais dos elementos da sequência de variáveis aleatórias do lema anterior, apropriadamente re-centralizadas e normalizadas, tendem, convergem em distribuição a uma variável aleatória normal, com média zero e variância igual a 1. Para a prova de resultado a seguir, veja [10].



## 1.2. O TEOREMA DE LIMITE CENTRAL E A LEI DOS GRANDES NÚMEROS

**Teorema 1.5** (Laplace-De Moivre). *Seja  $(\Omega, \mathcal{F}, P)$  um espaço de probabilidade e uma seqüência de variáveis aleatórias i.i.d.  $\{X_j\}_{j \in \mathbb{N}}$  definida neste espaço e tal que  $X_j \Omega \rightarrow \mathbb{R}$ ,*

$$P(X_j = 0) = p \quad e \quad P(X_j = 1) = q,$$

*para todo  $j \in \mathbb{N}$ , com  $p + q = 1$  e  $p, q \geq 0$ . Defina as somas parciais*

$$S_k = X_1 + \cdots + X_k,$$

*com  $k \in \mathbb{N}$ . Então, para todos  $-\infty < a < b < +\infty$ , tem-se que*

$$\lim_{k \rightarrow \infty} P\left(a \leq \frac{S_k - kp}{\sqrt{kpq}} \leq b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx.$$

Intuitivamente, o teorema anterior implica que o experimento de se jogar a mesma moeda repetidamente, no limite, pode ser modelado por uma variável aleatória.

O resultado a seguir, conhecido com Teorema do Limite Central, generaliza o teorema anterior para variáveis aleatórias com qualquer distribuição. Para um esboço de sua demonstração, veja [10].

**Teorema 1.6** (do Limite Central). *Seja  $(\Omega, \mathcal{F}, P)$  um espaço de probabilidade e uma seqüência de variáveis aleatórias i.i.d.  $\{X_j\}_{j \in \mathbb{N}}$  definida neste espaço e tal que  $X_j \Omega \rightarrow \mathbb{R}$ ,*

$$E(X_j) = m \quad e \quad \text{Var}(X_j) = \sigma^2 > 0, \quad \text{para todo } j \in \mathbb{N}.$$

*Defina as somas parciais*

$$S_k = X_1 + \cdots + X_k,$$

com  $k \in \mathbb{N}$ . Então, para todos  $-\infty < a < b < +\infty$ , tem-se que

$$\lim_{k \rightarrow \infty} P \left( a \leq \frac{S_k - km}{\sqrt{k}\sigma} \leq b \right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx.$$

## Esperança Condicional

Outros conceitos fundamentais em finanças quantitativas são o de esperança condicional e de martingal, pois ajudam a descrever matematicamente o fato de não ser possível prever o futuro, isto é, o futuro é incerto, mas que a informação disponível até certo momento pode ajudar a reduzir o grau de incerteza.

O conceito de probabilidade de um evento  $A$  condicionado à ocorrência do evento  $B$  foi introduzido há pouco. Como deveria ser definido o valor esperado de uma variável aleatória  $X$  condicionado a um evento  $B$ , isto é  $E(X|B)$ ? Considerando o espaço de probabilidade  $(B, \tilde{\mathcal{F}}, \tilde{P})$ , em que  $\tilde{F} = \{C \cap B : C \in \mathcal{F}\}$  e  $\tilde{P} = P/P(B)$ , se  $P(B) > 0$ , então

$$E(X|B) = \text{é o valor médio de } X \text{ em } B = \frac{1}{P(B)} \int_B X dP.$$

Outra pergunta que aparece neste contexto é, como definir o valor esperado de uma variável aleatória  $X$  condicionado a uma outra variável aleatória  $Y$ , ou seja,  $E(X|Y)$ ? Ou seja, dado  $\omega \in \Omega$  escolhido ao acaso, sabendo-se o valor de  $Y(\omega)$ , qual seria a melhor estimativa para o valor de  $X(\omega)$ ?

O exemplo a seguir apresentado em [10] procura ilustrar de modo intuitivo este conceito.

**Exemplo 1.8.** *Considere um espaço de probabilidade  $(\Omega, \mathcal{F}, P)$  e os eventos disjuntos  $A_1, \dots, A_m$  em  $\mathcal{F}$  e com probabilidade positiva, tais que  $\cup_{j=1}^m A_j = \Omega$ . Dados os números reais  $a_1, \dots, a_m$ , defina a variável*

*aleatória simples com valores reais*

$$Y = \sum_{j=1}^m a_j \mathcal{X}_{A_j}.$$

Assim,  $Y(\omega) = a_j$  se e só se  $a_j \in A_j$ , para  $j = 1, \dots, m$ .

Dada uma outra variável aleatória  $X$  definida neste mesmo espaço de probabilidade, qual seria a melhor estimativa para  $X$  dado  $Y$ ? Observe que, ao se saber o valor de  $Y(\omega)$ , sabe-se a qual evento  $A_j$   $\omega$  pertence. Assim, a melhor estimativa para  $X$  ao se esta informação é dada pelas médias de  $X$  em cada evento  $A_j$ , ou seja, tem-se que

$$E(X|Y) = \begin{cases} \frac{1}{P(A_1)} \int_{A_1} X dP & \text{em } A_1 \\ \frac{1}{P(A_2)} \int_{A_2} X dP & \text{em } A_2 \\ \vdots & \vdots \\ \frac{1}{P(A_m)} \int_{A_m} X dP & \text{em } A_m. \end{cases}$$

A esperança condicional de  $X$  dada uma outra variável aleatória  $Y$  é a melhor estimativa de  $X$  ao se “conhecer”  $Y$ . Observe que no exemplo acima  $E(X|Y)$  é uma variável aleatória, isto é, depende de  $\omega \in \Omega$ . Além disso, a  $\sigma$ -álgebra gerada por  $E(X|Y)$  está contida na  $\sigma$ -álgebra gerada por  $Y$   $\mathcal{F}(Y)$ , isto é,  $E(X|Y)$  é  $\mathcal{F}(Y)$ -mensurável. Observe ainda que  $\int_A X dP = \int_A E(X|Y) dP$ , para todo  $A \in \mathcal{F}(Y)$ .

**Exercício 1.10.** *Determine a  $\sigma$ -álgebra  $\mathcal{F}(Y)$  e mostre que ela e a  $\sigma$ -álgebra gerada por  $E(X|Y)$  coincidem. Mostre ainda que  $\int_A X dP = \int_A E(X|Y) dP$ , para todo  $A \in \mathcal{F}(Y)$ .*

Considere agora o caso geral:

**Definição 1.14.** *Sejam  $(\Omega, \mathcal{F}, P)$  um espaço de probabilidade e  $Y: \Omega \rightarrow \mathbb{R}^n$  uma variável aleatória. Então, a esperança condicional  $E(X|Y)$*

de uma variável aleatória integrável  $X$  dada  $Y$  é qualquer variável aleatória  $\mathcal{F}(Y)$ -mensurável que satisfaça:

$$\int_A X dP = \int_A E(X|Y) dP$$

para todo evento  $A \in \mathcal{F}(Y)$ .

Assim, para a definição da esperança condicional, o que importa é a  $\sigma$ -álgebra  $\mathcal{F}(Y)$  e não a variável aleatória  $Y$ . Isto motiva a definição a seguir:

**Definição 1.15.** *Sejam  $(\Omega, \mathcal{F}, P)$  um espaço de probabilidade e  $\mathcal{G}$  uma  $\sigma$ -álgebra contida em  $\mathcal{F}$ . Se  $X : \Omega \rightarrow \mathbb{R}^n$  é uma variável aleatória integrável, então a esperança condicional de  $X$  com respeito a  $\mathcal{G}$ , denotada por  $E(X|\mathcal{G})$  é qualquer variável aleatória definida em  $(\Omega, \mathcal{F}, P)$  satisfazendo:*

1.  $E(X|\mathcal{G})$  é  $\mathcal{G}$ -mensurável.
2.  $\int_A X dP = \int_A E(X|\mathcal{G}) dP$  para todo  $A$  em  $\mathcal{G}$ .

O teorema a seguir garante a existência e a unicidade de esperanças condicionais como a da definição anterior.

**Teorema 1.7.** *Sejam  $(\Omega, \mathcal{F}, P)$  um espaço de probabilidade e  $X : \Omega \rightarrow \mathbb{R}^n$  uma variável aleatória. Então, para cada  $\sigma$ -álgebra  $\mathcal{G} \subset \mathcal{F}$ , existe a esperança condicional  $E(X|\mathcal{G})$ . A unicidade ocorre a menos de conjuntos  $\mathcal{G}$ -mensuráveis com probabilidade zero.*

A esperança condicional  $E(X|\mathcal{G})$  pode ser interpretada como sendo a melhor estimativa de  $X$  dada a informação contida na  $\sigma$ -álgebra  $\mathcal{G}$ . A primeira condição na definição de esperança condicional garante que essa estimativa seja construída de acordo com a informação disponível

em  $\mathcal{G}$ . Já a segunda condição exige que essa estimativa seja consistente com  $X$ .

Observe ainda que valem as seguintes propriedades:

1.  $E(X|Y) = E(X|\mathcal{F}(Y))$ .
2.  $E(E(X|\mathcal{G})) = E(X)$ .
3.  $E(X|\mathcal{G}) = E(X)$  quando  $\mathcal{G} = \{\emptyset, \Omega\}$ , isto é,  $\mathcal{G}$  é a  $\sigma$ -álgebra trivial.

**Exercício 1.11.** *Prove as propriedades acima.*

A seguir são apresentadas propriedades adicionais da esperança condicional.

**Teorema 1.8.** *Sejam  $(\Omega, \mathcal{F}, P)$  um espaço de probabilidade e  $\mathcal{G}$  uma  $\sigma$ -álgebra contida em  $\mathcal{F}$ .*

1. *Se  $X$  é  $\mathcal{G}$ -mensurável, então  $E(X|\mathcal{G}) = X$  q.c.*
2. *Se  $a, b \in \mathbb{R}$ , então  $E(aX + bY|\mathcal{G}) = aE(X|\mathcal{G}) + bE(Y|\mathcal{G})$  q.c.*
3. *Se  $X$  é independente de  $\mathcal{G}$ , então  $E(X|\mathcal{G}) = E(X)$  q.c.*
4. *Se  $\mathcal{V} \subset \mathcal{G}$ , então*

$$E(X|\mathcal{V}) = E(E(X|\mathcal{G})|\mathcal{V}) = E(E(X|\mathcal{V})|\mathcal{G})$$

5. *Se  $X$  é  $\mathcal{G}$ -mensurável e  $XY$  é integrável, então  $E(XY|\mathcal{G}) = XE(X|\mathcal{G})$  q.c.*
6. *A desigualdade  $X \leq Y$  q.c. implica que  $E(X|\mathcal{G}) \leq E(Y|\mathcal{G})$  q.c.*

Para a prova deste resultado, veja [10], no entanto, o leitor interessado pode provar estas propriedades como exercício.

## Martingais e Movimento Browniano

Para explorar a intuição a respeito de processos martingais, considere o seguinte exemplo:

**Exemplo 1.9.** *Sejam  $(\Omega, \mathcal{F}, P)$  um espaço de probabilidade e  $\{Y_j\}_{j \in \mathbb{N}}$  uma sequência de variáveis aleatórias com valores reais e independentes, tal que  $E(X_j) = 0$  para todo  $j \in \mathbb{N}$ . Defina as somas parciais para cada  $m \in \mathbb{N}$ ,*

$$S_m = X_1 + \cdots + X_m.$$

*Pergunta-se, qual é a melhor estimativa para  $S_{m+k}$  dados os valores das  $m$  primeiras somas parciais  $S_1, S_2, \dots, S_m$ ? A resposta é:*

$$\begin{aligned} E(S_{m+k} | S_1, \dots, S_m) &= E(S_m | S_1, \dots, S_m) + E(Y_{m+1} + \cdots + Y_{m+k} | S_1, \dots, S_m) \\ &= S_m + E(Y_{m+1} + \cdots + Y_{m+k}) = S_m + 0 = S_m. \end{aligned}$$

*Em outras palavras, a melhor estimativa para  $S_{m+k}$ , dada a informação disponível até o momento  $m$ , é exatamente  $S_m$ .*

Para a definição de martingais em tempo contínuo é necessária a definição de história ou histórico de um processo estocástico.

**Definição 1.16.** *Seja  $\{X(t) : t \geq 0\}$  um processo estocástico real, isto é  $X_t : \Omega \rightarrow \mathbb{R}$ . Então,*

$$\mathcal{F}(t) = \mathcal{F}(X(s) : s \in [0, t])$$

*é a  $\sigma$ -álgebra gerada pelas variáveis aleatórias  $X(s)$ , para  $0 \leq s \leq t$ , a qual é chamada de história do processo até o tempo  $t$  (incluindo  $t$ ).*

**Definição 1.17.** *Seja  $\{X(t) : t \geq 0\}$  um processo estocástico tal que*

$$E(|X(t)|) < \infty$$

para todo  $t \geq 0$ .

1. Se  $X(s) = E(X(t)|\mathcal{F}(s))$  q.c. para todo  $t \geq s \geq 0$ , então o processo estocástico  $X(\cdot)$  é chamado de martingal.
2. Se  $X(s) \leq E(X(t)|\mathcal{F}(s))$  q.c. para todo  $t \geq s \geq 0$ , então o processo estocástico  $X(\cdot)$  é chamado de sub-martingal.
3. Se  $X(s) \geq E(X(t)|\mathcal{F}(s))$  q.c. para todo  $t \geq s \geq 0$ , então o processo estocástico  $X(\cdot)$  é chamado de super-martingal.

**Lema 1.6.** *Suponha que  $X(\cdot)$  seja um martingal real e que  $\Phi : \mathbb{R} \rightarrow \mathbb{R}$  seja uma função convexa, isto é, para todo  $x, y \in \mathbb{R}$  e  $\alpha \in [0, 1]$ , tem-se que*

$$\Phi(\alpha x + (1 - \alpha)y) \leq \alpha\Phi(x) + (1 - \alpha)\Phi(y).$$

Então, se  $E(|\Phi(t)|) < \infty$  para todo  $t \geq 0$ ,  $\Phi(X(\cdot))$  é um sub-martingal.

**Exercício 1.12.** *Seja  $(\Omega, \mathcal{F}, P)$  um espaço de probabilidade. Prove o lema anterior usando a desigualdade de Jensen condicional: se  $X$  é uma variável aleatória real e integrável,  $\mathcal{G}$  é uma  $\sigma$ -álgebra contida em  $\mathcal{F}$  e  $\Phi : \mathbb{R} \rightarrow \mathbb{R}$  é uma função convexa, então*

$$\Phi(E(X|\mathcal{G})) \leq E(\Phi(X)|\mathcal{G}).$$

*S*

O teorema a seguir apresenta uma importante generalização da desigualdade de Chebyshev para o caso de martingais e sub-martingais.

**Teorema 1.9.** *Seja  $X(\cdot)$  um processo estocástico real cujo os caminhos são contínuos q.c., ou seja, para quase todo  $\omega \in \Omega$ , o mapa  $t \mapsto X(t, \omega)$  é contínuo em  $t$ .*

1. Se  $X(\cdot)$  é um sub-martingal, então

$$P\left(\max_{0 \leq s \leq t} X(s) \geq \lambda\right) \leq \frac{1}{\lambda} E(\max\{0, X(t)\})$$

para todo  $\lambda > 0$  e todo  $t \geq 0$ .

2. Se  $X(\cdot)$  é um martingal e  $1 < p < \infty$ , então

$$E\left(\max_{0 \leq s \leq t} |X(s)|^p\right) \leq \left(\frac{p}{p-1}\right)^p E(|X(t)|^p)$$

para todo  $t \geq 0$ .

Para uma ideia da prova deste teorema, veja [10].

A seguir será apresentado um processo estocástico muito especial, conhecido como movimento browniano, ou processo de Wiener. Este é um processo estocástico, geralmente denotado pela letra  $B$  ou  $W$ , tal que, para cada  $t > 0$ ,  $W(t)$  tem distribuição gaussiana. Ele é amplamente usado na modelagem de preços de ativos financeiros, tendo aparecido neste contexto pela primeira vez no trabalho seminal de Louis Bachelier em 1900.

**Definição 1.18.** *Seja  $(\Omega, \mathcal{F}, P)$  um espaço de probabilidade. Um processo estocástico com valores reais  $W(\cdot)$  é chamado movimento browniano se satisfizer as seguintes condições:*

1.  $W(0) = 0$  q.c.
2.  $W(t) - W(s) \stackrel{d}{=} N(0, t - s)$  para todo  $t \geq s \geq 0$ .
3. Para toda escolha de índices  $0 < t_1 < t_2 < \dots < t_m$ , as variáveis aleatórias  $W(t_1), W(t_2) - W(t_1), \dots, W(t_m) - W(t_{m-1})$  são independentes, isto é, os incrementos do movimento Browniano são independentes.



Observe que  $E(W(t)) = 0$  e  $E(W^2(t)) = t$  para todo  $t \geq 0$ .

**Exercício 1.13.** *Mostre que o movimento browniano é um martingal.*

**Exercício 1.14.** *Calcule  $E(W(t)W(s))$  para  $0 < s \leq t$ .*

Observe que, definindo a função

$$g(x, t|y) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{(x-y)^2}{2t}}$$

então, para qualquer escolha de índices  $0 < t_1 < \dots < t_m$ , e de números reais  $a_j \leq b_j$ ,  $j = 1, \dots, m$ , vale que

$$P(a_1 \leq W(t_1) \leq b_1, a_2 \leq W(t_2) \leq b_2, \dots, a_m \leq W(t_m) \leq b_m) = \int_{a_1}^{b_1} \dots \int_{a_m}^{b_m} g(x_1, t_1|0)g(x_2, t_2-t_1|x_1) \dots g(x_m, t_m-t_{m-1}|x_{m-1})dx_m \dots dx_1.$$

O teorema a seguir apresenta uma importante propriedade sobre o movimento Browniano que será bastante útil no que se verá no contexto de finanças.

**Teorema 1.10.** *Seja  $W(\cdot)$  um movimento browniano unidimensional definido no espaço de probabilidade  $(\Omega, \mathcal{F}, P)$ . Então, para todo  $m \in \mathbb{N}$ , para toda escolha de índices  $0 = t_0 < t_1 < \dots < t_m$  e para cada função  $f: \mathbb{R}^m \rightarrow \mathbb{R}$ , tem-se que*

$$E(f(W(t_1), W(t_2), \dots, W(t_m))) = \int_{\mathbb{R}} \dots \int_{\mathbb{R}} f(x_1, \dots, x_m)g(x_1, t_1|0)g(x_2, t_2 - t_1|x_1) \dots g(x_m, t_m - t_{m-1}|x_{m-1})dx_m \dots dx_1.$$

A versão  $n$ -dimensional do movimento browniano será apresentada a seguir.

**Definição 1.19.** *Seja  $(\Omega, \mathcal{F}, P)$  um espaço de probabilidade. Um processo estocástico com valores em  $\mathbb{R}^n$*

$$W(\cdot) = (W_1(\cdot), \dots, W_n(\cdot))$$

*é chamado de movimento browniano  $n$ -dimensional se satisfizer as seguintes condições:*

1. *Para cada  $j = 1, \dots, n$ ,  $W_j(\cdot)$  é um movimento browniano unidimensional.*
2. *As  $\sigma$ -álgebras  $\mathcal{F}_j = \mathcal{F}(W_j : t \geq 0)$ , para  $j = 1, \dots, n$ , são independentes.*

De maneira análoga ao caso unidimensional, tem-se o seguinte resultado para um movimento browniano  $n$ -dimensional:

**Teorema 1.11.** *Se  $W(\cdot)$  é um movimento browniano  $n$ -dimensional, então  $W(t) = {}^d N(0, tI)$  para cada  $t \geq 0$  e*

$$P(W(t) \in A) = \frac{1}{(2\pi t)^{\frac{n}{2}}} \int_A e^{-\frac{|x|^2}{2t}} dx$$

*Além disso, para todo  $m \in \mathbb{N}$ , para toda escolha de índices  $0 = t_0 < t_1 < \dots < t_m$  e para cada função  $f: \mathbb{R}^m \rightarrow \mathbb{R}$ , tem-se que*

$$\begin{aligned} E(f(W(t_1), W(t_2), \dots, W(t_m))) &= \\ &\int_{\mathbb{R}} \dots \int_{\mathbb{R}} f(x_1, \dots, x_m) g(x_1, t_1 | 0) \\ &g(x_2, t_2 - t_1 | x_1) \dots g(x_m, t_m - t_{m-1} | x_{m-1}) dx_m \dots dx_1, \end{aligned}$$

onde

$$g(x, t | y) = \frac{1}{(2\pi t)^{\frac{n}{2}}} e^{-\frac{|x-y|^2}{2t}}.$$

**Definição 1.20.** Um processo estocástico  $X(\cdot)$  e sua história chamado processo de Markov ou markoviano se satisfizer a seguinte propriedade:

$$P(X(t) \in B | \mathcal{F}(s)) = P(X(t) \in B | X(s)) \text{ q.c.}$$

para todo  $0 \leq s \leq t$  e todo conjunto boreliano  $B \in \mathcal{B}$ , em que  $\mathcal{F}(t) = \{\mathcal{F}(X(s) : 0 \leq s \leq t)\}$  é a história do processo  $X(\cdot)$ .

**Teorema 1.12.** Se  $W(\cdot)$  é um movimento browniano  $n$ -dimensional, então  $W(\cdot)$  é um processo de Markov e

$$P(W(t) \in B | W(s)) = \frac{1}{(2\pi(t-s))^{\frac{n}{2}}} \int_B e^{-\frac{|x-W(s)|^2}{2(t-s)}} dx \text{ q.c.,}$$

para todo  $0 \leq s < t$  e todo boreliano  $B$ .

Para as provas dos resultados relativos ao movimento browniano, veja [10].



## Capítulo 2

# Introdução ao Cálculo Estocástico

A integral e a fórmula de Itô são conceitos fundamentais em finanças quantitativas, pois é através deles que se pode definir as equações diferenciais estocásticas, que, por sua vez, são usadas para modelar preços de ações e outros ativos nos mercados financeiro, de commodities e de títulos. A pedra angular da construção de todas estas técnicas é o movimento browniano, que foi definido no final do capítulo anterior. Este capítulo é uma adaptação dos capítulos 4 e 5 de [10].

### A Integral de Itô

Para definir a integral de Itô, primeiramente, é necessário introduzir algumas notações.

**Definição 2.1.** *Sejam  $(\Omega, \mathcal{F}, P)$  um espaço de probabilidade e  $W(\cdot)$  um movimento browniano unidimensional.*

1. *A  $\sigma$ -álgebra  $\mathcal{W}(t) = \mathcal{F}(W(s) : 0 \leq s \leq t)$  é chamada de história do movimento browniano até o tempo  $t$ .*

2. Uma família  $\mathcal{G}(\cdot)$  de  $\sigma$ -álgebras contida em  $\mathcal{F}$  é chamada não antecipativa e de filtro ou filtração se satisfizer:

(a)  $\mathcal{G}(s) \subset \mathcal{G}(t)$  para todo  $0 \leq s \leq t$ ,

(b)  $\mathcal{W}(t) \subset \mathcal{G}(t)$  para todo  $t \geq 0$ ,

(c)  $\mathcal{G}(t)$  é independente de  $\mathcal{W}^+(t) = \mathcal{F}(W(s) - W(t) : s \geq t)$ , para todo  $t \geq 0$ .

3. Um processo estocástico  $G(\cdot)$  é não antecipativo se, para cada  $t \geq 0$ ,  $G(t)$  for  $\mathcal{G}(t)$ -mensurável, ou seja, a pré-imagem de todo conjunto boreliano por  $G(t, \cdot)$  pertence a  $\mathcal{G}(t)$ .

4. O conjunto  $\mathbb{L}^2(0, T)$  é o espaço de todos os processos estocásticos unidimensionais  $G(\cdot)$  progressivamente mensuráveis, isto é, mensuráveis nas variáveis  $\omega$  e  $t$  e não antecipativos, tais que

$$E \left( \int_0^T G^2 dt \right) < \infty.$$

De posse dos conceitos apresentados acima, o próximo passo é definir a integral de Itô em  $(0, T)$  para um processo estocástico  $G(\cdot) \in \mathbb{L}^2(0, t)$ .

Primeiramente, considere um processo escada  $G(\cdot) \in \mathbb{L}^2(0, t)$ , isto é, existe uma partição do intervalo  $[0, T]$ , denotada por

$$P = \{0 = t_0 < t_1 < \dots < t_m = T\},$$

em que  $G(t) = G_k$  para  $t \in [t_k, t_{k+1})$  e  $G_k$  é  $\mathcal{G}(t_k)$  mensurável.

Assim, a integral de Itô em  $(0, T)$  para este processo escada  $G(\cdot)$  é dada por

$$\int_0^T G dW = \sum_{k=0}^{m-1} G_k (W(t_{k+1}) - W(t_k))$$

**Exercício 2.1.** *Mostre que o processo escada acima é não antecipativo.*

Para estender a definição de integral de Itô para processos mais gerais, serão usadas técnicas semelhantes às da definição da integral de Riemann. No entanto, ao invés de se ter os diâmetros dos subintervalos da partição na soma de Riemann, tem-se as diferenças  $W(t_{k+1}) - W(t_k)$ . Portanto, é necessário provar que essas somas parciais convergem quando o diâmetro da partição tende a zero. Notem que isto não é nem um pouco óbvio devido às propriedades dos caminhos do movimento browniano.

A demonstração do lema a seguir é muito longa e por isso, não será apresentada aqui. Ela pode ser encontrada em [10].

**Lema 2.1.** (*Varição Quadrática*) *Dado um intervalo  $[a, b] \subset [0, T]$ , seja  $\{P_m\}_{m \in \mathbb{N}}$  uma sequência de partições de  $[a, b]$ , tal que o diâmetro de  $P_m$  satisfaz o limite  $\lim_{m \rightarrow \infty} |P_m| = 0$  e  $P_m = \{a = t_0 < t_1^m < \dots < t_{k_m}^m = b\}$ . Então,*

$$\lim_{m \rightarrow \infty} \sum_{j=0}^{k_m} (W(t_{j+1}^m) - W(t_j^m))^2 = b - a.$$

A seguir são apresentadas algumas propriedades da integral de Itô de funções escada em  $\mathbb{L}^2(0, T)$ .

**Lema 2.2** (*Propriedades da Integral de Itô para Processos Escada*). *Sejam  $a, b \in \mathbb{R}$  constantes quaisquer e  $G(\cdot), H(\cdot) \in \mathbb{L}^2(0, T)$  processos escada. Então, valem as seguintes propriedades:*

1.  $\int_0^T (aG(t) + bH(t))dW(t) = a \int_0^T G(t)dW(t) + b \int_0^T H(t)dW(t),$
2.  $E \left( \int_0^T G(t)dW(t) \right) = 0,$

$$3. E \left( \left( \int_0^T G(t) dW(t) \right)^2 \right) = E \left( \int_0^T G^2(t) dt \right).$$

Observe que, apara cada processo  $G(\cdot) \in \mathbb{L}^2(0, T)$ , existe uma sequência  $\{G_m(\cdot)\}_{m \in \mathbb{N}}$  de processos escada contida em  $\mathbb{L}^2(0, T)$  atis que

$$\lim_{m \rightarrow \infty} E \left( \int_0^T |G(t) - G_m(t)|^2 dt \right) = 0$$

e assim

$$\int_0^T G(t) dW(t) = \lim_{m \rightarrow \infty} \int_0^T G_m(t) dW(t).$$

Assim, para qualquer processo  $G(\cdot) \in \mathbb{L}^2(0, T)$ , sua integral de Itô é o limite de integrais de Itô de processos escada que convergem para  $G(\cdot)$  com respeito a métrica de  $\mathbb{L}^2(0, T)$ .

Uma propriedade importante de integrais de Itô é a identidade a seguir:

$$E \left( \int_0^T G(t) dW(t) \int_0^T H(t) dW(t) \right) = E \left( \int_0^T G(t) H(t) dt \right)$$

ara quaisquer processos  $G(\cdot)$  e  $H(\cdot)$  em  $\mathbb{L}^2(0, T)$ .

Para cada processos  $G(\cdot) \in \mathbb{L}^2(0, T)$  e todo  $t \in [0, T]$  define-se o processo estocástico integral de Itô:

$$I(t) = \int_0^t G(t) dW(t).$$

O processo  $I(t)$  é um martingal e possui uma versão com caminhos contínuos.

Considere agora a versão muultidimensional da integral de Itô.

**Definição 2.2.** *Seja  $W = (W_1(\cdot), \dots, W_m(\cdot))$  um movimento browniano  $m$ -dimensional. Uma família não-antecipativa de  $\sigma$ -álgebras  $\mathcal{G}(\cdot)$  satisfaz as seguintes propriedades:*



1.  $\mathcal{G}(s) \subset \mathcal{G}(t)$  para todos  $0 \leq s \leq t$ .
2.  $\mathcal{W}(t) \subset \mathcal{G}(t)$  para todo  $t \geq 0$ , em que  $\mathcal{W}(t) = \mathcal{F}(W(s) : 0 \leq s \leq t)$ ,  $t \geq 0$ , é a história do movimento browniano  $m$ -dimensional  $W(\cdot)$ .
3.  $\mathcal{G}(t)$  é independente de  $\mathcal{W}^+(t) = \mathcal{F}(W(s) - W(t) : t \leq s < \infty)$ .

Antes de definir a integral propriamente dita, é necessário definir os espaços onde os integrandos estão definidos..

Um processo estocástico  $G(\cdot) = (G_{ij}(\cdot))$  com valores no espaço de matrizes  $\mathbb{R}^{n \times m}$  pertence ao conjunto  $\mathbb{L}_{n \times m}^2(0, T)$ , se, para cada  $i = 1, \dots, n$  e  $j = 1, \dots, m$ ,  $G_{ij}(\cdot) \in \mathbb{L}^2(0, T)$ .

Um processo estocástico  $F(\cdot) = (F_1(\cdot), \dots, F_n(\cdot))$  com valores em  $\mathbb{R}^n$  pertence ao conjunto  $\mathbb{L}_n^1(0, T)$ , se, para cada  $i = 1, \dots, n$ ,  $F_i(\cdot) \in \mathbb{L}^1(0, T)$ .

**Definição 2.3.** Se  $G(\cdot)$  pertence a  $\mathbb{L}_{n \times m}^2(0, T)$ , então, a integral

$$\int_0^T G(t) dW(t)$$

é uma variável aleatória  $n$ -dimensional, cujo a  $i$ -ésima componente é dada por:

$$\sum_{j=1}^m \int_0^T G_{ij}(t) dW_j(t), \text{ para } i = 1, \dots, n.$$

Assim, tem-se as seguintes identidades:

$$E \left( \int_0^T G(t) dW(t) \right) = 0,$$

e

$$E \left( \left| \int_0^T G(t) dW(t) \right|^2 \right) = E \left( \int_0^T |G(t)|^2 dt \right),$$

em que  $|G(t)|^2 = \sum_{i=1}^n \sum_{j=1}^m |G_{ij}(t)|^2$ .

## A Fórmula de Itô

Um conceito muito importante na teoria do cálculo estocástico é o da fórmula de Itô e das diferenciais estocásticas.

Diz-se que um processo estocástico  $F(\cdot)$  está no espaço  $\mathbb{L}^1(0, T)$  se  $F(\cdot)$  satisfizer

$$E \left( \int_0^T |F(t)| dt \right) < \infty$$

Assim, pode-se definir a diferencial estocástica de um processo  $X(\cdot)$ .

**Definição 2.4** (Diferencial Estocástica). *Suponha que  $X(\cdot)$  seja um processo estocástico unidimensional que satisfaça a equação:*

$$X(r) = X(s) + \int_s^r F(t) dt + \int_s^r G(t) dW(t)$$

para algum  $F(\cdot) \in \mathbb{L}^1(0, T)$  e algum  $G(\cdot) \in \mathbb{L}^2(0, T)$  e todos  $0 \leq s \leq r \leq T$ . Então, diz-se que o processo  $X(\cdot)$  possui a diferencial estocástica

$$dX(t) = F(t)dt + G(t)dW(t)$$

para  $0 \leq t \leq T$ .

Note que a diferencial estocástica é apenas uma representação mais simples da equação integral do processo  $X(\cdot)$ . A partir da definição acima, é possível apresentar a regra de Itô, que seria uma versão para processos de Itô da regra da cadeia do cálculo diferencial.

**Teorema 2.1** (Fórmula de Itô). *Suponha que o processo estocástico unidimensional  $X(\cdot)$  possua a diferencial estocástica*

$$dX(t) = F(t)dt + G(t)dW(t),$$

para  $F(\cdot) \in \mathbb{L}^1(0, T)$ ,  $G(\cdot) \in \mathbb{L}^2(0, T)$  e  $t \in [0, T]$ . Seja  $u : \mathbb{R} \times [0, T] \rightarrow$

$\mathbb{R}$  uma função contínua tal que as derivadas parciais  $\frac{\partial u}{\partial t}$ ,  $\frac{\partial u}{\partial x}$  e  $\frac{\partial^2 u}{\partial x^2}$  existem e são contínuas, isto é,  $u$  é de classe  $C^{2,1}$ . Defina o processo estocástico  $Y(\cdot)$  por

$$Y(t) = u(X(t), t),$$

para todo  $t \in [0, T]$ . Então, o processo  $Y(\cdot)$  possui a diferencial estocástica

$$\begin{aligned} dY(t) &= \frac{\partial u}{\partial t} dt + \frac{\partial u}{\partial x} dX(t) + \frac{1}{2} \frac{\partial^2 u}{\partial x^2} G^2(t) dt \\ &= \left( \frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} F(t) + \frac{1}{2} \frac{\partial^2 u}{\partial x^2} G^2(t) \right) dt + \frac{\partial u}{\partial x} G(t) dW(t). \end{aligned}$$

Observe que, por simplicidade, na diferencial de  $Y(\cdot)$  foi omitida a dependência das derivadas de  $u$  com respeito ao par  $(X(t), t)$ . Além disso, para todos  $0 \leq s \leq r \leq T$ , tem-se que a diferencial de  $Y(\cdot)$  é equivalente a:

$$\begin{aligned} Y(r) - Y(s) &= u(X(r), r) - u(X(s), s) \\ &= \int_s^r \frac{\partial u}{\partial t}(X(t), t) + \frac{\partial u}{\partial x}(X(t), t)F(t) + \frac{1}{2} \frac{\partial^2 u}{\partial x^2}(X(t), t)G^2(t) dt \\ &\quad + \int_s^r \frac{\partial u}{\partial x}(X(t), t)G(t)dW(t) \end{aligned}$$

quase certamente. Observe ainda que, como os caminhos do processo  $X(t) = X(0) + \int_0^t F(s)ds + \int_0^t G(s)dW(s)$  são contínuos q.c., então, os processos  $\frac{\partial u}{\partial t}(X(\cdot), \cdot)$ ,  $\frac{\partial u}{\partial x}(X(\cdot), \cdot)$  e  $\frac{\partial^2 u}{\partial x^2}(X(\cdot), \cdot)$  são contínuos q.c., daí, as integrais acima na formulação de  $Y(\cdot)$  estão bem definidas.

A seguir são apresentadas alguns exemplos de aplicações da fórmula de Itô.

**Exemplo 2.1.** *Seja  $X(\cdot) = W(\cdot)$  e  $u(x, t) = x^m$ ,  $m \in \mathbb{N}$  fixo. Então, a diferencial de  $X(\cdot)$  é  $dX(t) = dW(t)$ , para  $t \in [0, T]$  e, as derivadas*

parciais de  $u$  são:  $\frac{\partial u}{\partial t} = 0$ ,  $\frac{\partial u}{\partial x} = mx^{m-1}$  e  $\frac{\partial^2 u}{\partial x^2} = m(m-1)x^{m-2}$ .  
Assim,  $Y(\cdot) = W^m(\cdot)$  possui a seguinte diferencial estocástica:

$$d(W^m(t)) = mW^{m-1}(t)dW(t) + \frac{1}{2}m(m-1)W^{m-2}(t)dt.$$

Então, quando  $m = 2$ , tem-se que

$$d(W^2(t)) = 2W(t)dW(t) + dt$$

**Exercício 2.2.** Encontre a diferencial estocástica do processo  $Y(t) = tW(t)$ . Observe que, neste caso  $u(x, t) = tx$  e  $X(t) = W(t)$ .

**Exemplo 2.2.** Seja  $X(t) = W(t)$ ,  $u(x, t) = e^{\lambda x - \frac{\lambda^2}{2}t}$ ,  $F(t) = 0$  e  $G(t) = 0$ . Então,

$$\begin{aligned} d\left(e^{\lambda W(t) - \frac{\lambda^2}{2}t}\right) &= \left(-\frac{\lambda^2}{2}e^{\lambda W(t) - \frac{\lambda^2}{2}t} + \frac{\lambda^2}{2}e^{\lambda W(t) - \frac{\lambda^2}{2}t}\right)dt + \lambda e^{\lambda W(t) - \frac{\lambda^2}{2}t}dW(t) \\ &= \lambda e^{\lambda W(t) - \frac{\lambda^2}{2}t}dW(t). \end{aligned}$$

Assim,  $Y(t) = e^{\lambda W(t) - \frac{\lambda^2}{2}t}dW(t)$  satisfaz:

$$dY(t) = \lambda e^{\lambda W(t) - \frac{\lambda^2}{2}t}dW(t) \quad e \quad Y(0) = 1.$$

Em cálculo estocástico a expressão  $e^{\lambda W(t) - \frac{\lambda^2}{2}t}dW(t)$  tem o mesmo papel de  $e^{\lambda t}$  no cálculo diferencial.

**Teorema 2.2** (Regra do Produto de Itô). *Suponha que  $X_1(\cdot)$  e  $X_2(\cdot)$  sejam dois processos estocásticos que possuem as seguintes diferenciais estocásticas:*

$$dX_1(t) = F_1(t)dt + G_1(t)dW(t) \quad e \quad dX_2(t) = F_2(t)dt + G_2(t)dW(t),$$

com  $F_1(\cdot), F_2(\cdot) \in \mathbb{L}^1(0, T)$  e  $G_1(\cdot), G_2(\cdot) \in \mathbb{L}^2(0, T)$ . Então, o processo dado pelo produto de  $X_1(\cdot)$  e  $X_2(\cdot)$  possui a seguinte diferencial estocástica:

$$d(X_1 X_2)(t) = X_2(t) dX_1(t) + X_1(t) dX_2(t) + G_1(t) G_2(t) dt.$$

Observe que o termo  $G_1 G_2$  na identidade do teorema acima é o termo de correção de Itô e quando tem-se que  $G_1(t) = G_2(t) = 0$  para todo  $t$ , então, tem-se as regras da derivada do produto e da integração por partes, ambas provenientes do cálculo diferencial e integral.

A seguir será apresentada a fórmula de Itô multidimensional, que é fundamental para avaliar funções que dependam de vários processos estocásticos de Itô definidos sobre o mesmo movimento Browniano.

**Teorema 2.3** (Fórmula de Itô Generalizada). *Suponha que  $X_j(\cdot)$ , para  $j = 1, \dots, m$ , sejam processos estocásticos (de Itô) que possuem as seguintes diferenciais estocásticas:*

$$dX_j(t) = F_j(t) dt + G_j(t) dW(t),$$

com  $F_j(\cdot) \in \mathbb{L}^1(0, T)$  e  $G_j(\cdot) \in \mathbb{L}^2(0, T)$ , para  $j = 1, \dots, m$ . Seja  $u : \mathbb{R}^m \times [0, T] \rightarrow \mathbb{R}$  uma função contínua, tal que as suas derivadas parciais  $\frac{\partial u}{\partial t}$ ,  $\frac{\partial u}{\partial x_j}$  e  $\frac{\partial^2 u}{\partial x_i \partial x_j}$ , com  $i, j = 1, \dots, m$ , existem e são contínuas. Então,

$$d(u(X_1, \dots, X_n, t)) = \frac{\partial u}{\partial t} dt + \sum_{j=1}^m \frac{\partial u}{\partial x_j} dX_j(t) + \frac{1}{2} \sum_{i,j=1}^m \frac{\partial^2 u}{\partial x_i \partial x_j} G_i(t) G_j(t) dt.$$

**Exercício 2.3.** Use a fórmula de Itô generalizada para provar a fórmula do produto de Itô.

A fórmula de Itô para processos multidimensionais será apresentada no que se segue.

**Definição 2.5.** *Seja  $X(\cdot) = (X_1(\cdot), \dots, X_n(\cdot))$  um processo estocástico  $n$ -dimensional, tal que*

$$X(r) = X(s) + \int_s^r F(t)dt + \int_s^r G(t)dW(t)$$

*para algum processo  $n$ -dimensional  $F(\cdot)$  em  $\mathbb{L}_n^1(0, T)$ , algum processo matricial  $G(\cdot)$  em  $\mathbb{L}_{n \times m}^2(0, T)$  e todos  $0 \leq s \leq t \leq T$ . Assim, diz-se que  $X(\cdot)$  possui a diferencial estocástica*

$$dX(t) = F(t)dt + G(t)dW(t).$$

*Isto significa que, para cada  $i = 1, \dots, n$ , tem-se*

$$dX_i(t) = F_i(t)dt + \sum_{j=1}^m G_{ij}(t)dW_j(t).$$

**Teorema 2.4** (Fórmula de Itô  $n$ -Dimensional). *Suponha que o processo estocástico  $n$ -dimensional  $X(\cdot)$  possua a diferencial estocástica*

$$dX(t) = F(t)dt + G(t)dW(t),$$

*como na definição acima. Seja  $u : \mathbb{R}^n \times [0, T] \rightarrow \mathbb{R}$  uma função contínua tal que as suas derivadas parciais  $\frac{\partial u}{\partial t}$ ,  $\frac{\partial u}{\partial x_i}$ ,  $i = 1, \dots, n$  e  $\frac{\partial^2 u}{\partial x_i \partial x_j}$ ,  $i, j = 1, \dots, n$  existem e são contínuas. Então, o processo estocástico  $Y(\cdot) = u(X(\cdot), \cdot)$  possui a diferencial estocástica*

$$\begin{aligned} dY(t) = du(X(t), t) &= \frac{\partial u}{\partial t} dt + \sum_{i=1}^n \frac{\partial u}{\partial x_i} dX_i(t) \\ &+ \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2 u}{\partial x_i \partial x_j} \sum_{l=1}^m G_{il}(t) G_{jl}(t) dt. \end{aligned}$$

Note que na diferencial estocástica acima o argumento  $(X(t), t)$  das derivadas de  $u$  foi suprimido, por simplicidade.

Para uma ideia da prova do teorema acima, veja [10].

**Exercício 2.4.** Use a fórmula de Itô  $n$ -dimensional para mostrar que a diferencial estocástica do produto de dois movimentos brownianos unidimensionais e independentes  $W$  e  $\widetilde{W}$  é dada por

$$d(W\widetilde{W}) = Wd\widetilde{W} + \widetilde{W}dW.$$

## Equações Diferenciais Estocásticas

Sejam  $W(\cdot)$  um movimento browniano  $m$ -dimensional e  $X_0$  uma variável aleatória  $n$ -dimensional, com  $X_0$  independente de  $W(\cdot)$ . No que se segue será considerada a seguinte família de  $\sigma$ -álgebras:

$$\mathcal{G}(t) = \mathcal{F}(X_0, W(s) : 0 \leq s \leq t), \text{ para } t \geq 0.$$

Ou seja, para cada  $t \geq 0$ ,  $\mathcal{G}(t)$  é a  $\sigma$ -álgebra gerada por  $X_0$  e pela história do movimento browniano  $W(\cdot)$  até o tempo  $t$ .

Para  $T > 0$  fixo, considere as funções determinísticas (não-aleatórias)

$$b : \mathbb{R}^n \times [0, T] \rightarrow \mathbb{R}^n \text{ e } B : \mathbb{R}^n \times [0, T] \rightarrow \mathbb{R}^{n \times m}.$$

Neste caso,  $b = (b_1, \dots, b_n)$  e  $B = (b_{ij})$ ,  $i = 1, \dots, n$  e  $j = 1, \dots, m$ .

**Definição 2.6.** Diz-se que um processo estocástico  $n$ -dimensional  $X(\cdot)$  é uma solução da equação diferencial estocástica (de Itô)

$$\begin{cases} dX(t) = b(X(t), t)dt + B(X(t), t)dW(t), \\ X(0) = X_0, \end{cases}$$

para  $0 \leq t \leq T$ , quando

1.  $X(\cdot)$  for progressivamente mensurável com respeito a família de  $\sigma$ -álgebras  $\mathcal{G}(\cdot)$ .
2.  $b(X(\cdot), \cdot)$  é um processo estocástico em  $\mathbb{L}_n^1(0, T)$ .
3.  $B(X(\cdot), \cdot)$  é um processo estocástico em  $\mathbb{L}_{n \times m}^2(0, T)$ .
4.  $X(t) = X_0 + \int_0^t b(X(s), s)ds + \int_0^t B(X(s), s)dW(s)$  q.c., para todo  $t \in [0, T]$ .

A seguir serão apresentados alguns exemplos de equações diferenciais estocásticas.

**Exemplo 2.3.** *Seja  $m = n = 1$  e suponha que  $g[0, T] \rightarrow \mathbb{R}$  seja uma função contínua, então a única solução da equação diferencial estocástica*

$$\begin{cases} dX(t) = g(t)X(t)dW(t), \\ X(0) = 1 \end{cases}$$

é  $X(t) = \exp\left(-\frac{1}{2}\int_0^t g^2(s)ds + \int_0^t g(s)dW(s)\right)$ , para  $t \in [0, T]$ .

**Exercício 2.5.** *Verifique a afirmação do exemplo anterior. Para isto, aplique a fórmula de Itô ao processo  $Y(t) = u(X(t), t)$ , com*

$$X(t) = -\frac{1}{2}\int_0^t g^2(s)ds + \int_0^t g(s)dW(s)$$

e  $u(x, T) = \exp(x)$ .

O próximo exemplo é similar ao anterior

**Exemplo 2.4.** *Seja  $m = n = 1$  e suponha que  $f, g[0, T] \rightarrow \mathbb{R}$  sejam funções contínuas. Então a única solução da equação diferencial estocástica*

$$\begin{cases} dX(t) = f(t)X(t)dt + g(t)X(t)dW(t), \\ X(0) = 1 \end{cases}$$



é  $X(t) = \exp\left(\int_0^t [f(s) - \frac{1}{2}g^2(s)]ds + \int_0^t g(s)dW(s)\right)$ , para  $t \in [0, T]$ .

**Exercício 2.6.** Use argumentos análogos ao do exercício anterior para mostra a afirmação do exemplo acima.

Existe uma outra definição de solução de uma equação diferencial estocástica, mas, por simplicidade, apenas este caso será tratado aqui. A seguir um teorema de existência e unicidade será apresentado, para sua prova, veja [10].

**Teorema 2.5** (Existência e Unicidade de Solução). *Suponha que  $b : \mathbb{R}^n \times [0, T] \rightarrow \mathbb{R}^n$  e  $B : \mathbb{R}^n \times [0, T] \rightarrow \mathbb{R}^{n \times m}$  sejam funções contínuas e que satisfaçam as condições abaixo:*

1. Para todo  $t \in [0, T]$ ,  $b(\cdot, t)$  e  $B(\cdot, t)$  são funções uniformemente Lipschitz-contínuas com constante  $L$ . Ou seja,

$$|b(x, t) - b(y, t)| \leq L|x - y| \text{ e } |B(x, t) - B(y, t)| \leq L|x - y|,$$

para todo  $t \in [0, T]$  e todos  $x, y \in \mathbb{R}^n$ .

2. Ambas as funções têm crescimento (ou decrescimento) linear. Isto é,

$$|b(x, t)| \leq L(1 + |x|) \text{ e } |B(x, t)| \leq L(1 + |x|).$$

Seja  $X_0$  uma variável aleatória  $n$ -dimensional tal que  $E(|X_0|^2) < \infty$  e  $X_0$  é independente de  $\mathcal{W}^+(0)$ , em que  $W(\cdot)$  é um movimento browniano  $m$ -dimensional dado. Então, existe um único processo estocástico  $X(\cdot)$  no espaço  $\mathbb{L}_n^2(0, T)$  que é solução da equação diferencial estocástica

$$\begin{cases} dX(t) = b(X(t), t)dt + B(X(t), t)dW(t), & t \in [0, T], \\ X(0) = X_0. \end{cases}$$

Observe que por “única solução” da equação diferenciais estocástica, entende-se que, se  $X(\cdot)$  e  $\tilde{X}(\cdot)$  são dois processos em  $\mathbb{L}_n^2(0, T)$ , com caminhos amostrais contínuos q.c., e ambos são soluções da mesma equação diferencial estocástica, então

$$P(X(t) = \tilde{X}(t) \text{ para todo } t \in [0, T]) = 1.$$

Uma equação diferencial estocástica é linear, se os coeficientes  $b$  e  $B$  assumirem as formas:

$$b(x, t) = c(t) + D(t)x \text{ e } B(x, t) = E(t) + F(t)x,$$

com  $c : [0, T] \rightarrow \mathbb{R}^n$ ,  $D : [0, T] \rightarrow \mathbb{R}^{n \times n}$ ,  $E : [0, T] \rightarrow \mathbb{R}^{n \times m}$  e  $F : [0, T] \rightarrow L(\mathbb{R}^n, \mathbb{R}^{n \times m})$ , em que  $L(\mathbb{R}^n, \mathbb{R}^{n \times m})$  é o espaço dos operadores lineares limitados que associam elementos de  $\mathbb{R}^n$  a elementos de  $\mathbb{R}^{n \times m}$ .

Uma equação diferencial estocástica linear é homogênea quando os coeficientes satisfizerem  $c \equiv E \equiv 0$ .

A seguir são apresentadas algumas técnicas de solução de equações diferenciais estocásticas.

**Exemplo 2.5.** *Considere a equação diferencial estocástica linear e unidimensional*

$$\begin{cases} dX(t) = d(t)X(t)dt + f(t)X(t)dW(t), & t \in [0, T], \\ X(0) = X_0. \end{cases}$$

Suponha que  $X(t)$  seja dado pelo produto  $X(t) = X_1(t)X_2(t)$ , onde

$$\begin{cases} dX_1(t) = f(t)X_1(t)dW(t), & t \in [0, T], \\ X_1(0) = X_0, \end{cases}$$

e

$$\begin{cases} dX_2(t) = A(t)dt + B(t)dW(t), & t \in [0, T], \\ X(0) = 1, \end{cases}$$

com  $A$  e  $B$  funções ainda não definidas. Então, aplicando a fórmula de Itô, tem-se que

$$\begin{aligned} dX(t) &= d(X_1 X_2)(t) \\ &= X_1(t)dX_2(t) + X_2(t)dX_1(t) + f(t)X_1(t)B(t)dt \\ &= f(t)X_1(t)X_2(t)dt + [X_1(t)dX_2(t) + f(t)X_1(t)B(t)dt] \\ &= f(t)X(t)dW(t) + X_1(t)[dX_2(t) + f(t)B(t)dt]. \end{aligned}$$

Escolha  $A$  e  $B$  tais que

$$dX_2(t) + f(t)B(t)dt = d(t)X_2(t)dt,$$

ou seja,

$$A(t)dt + B(t)dW(t) + f(t)B(t)dt = d(t)X_2(t)dt.$$

Assim,  $B \equiv 0$  e  $A(t) = d(t)X_2(t)$ . Isto implica que  $dX_2(t) = d(t)X_2(t)dt$ . Neste caso,  $X_2(\cdot)$  não é aleatório e, necessariamente,  $X_2(t) = e^{\int_0^t d(s)ds}$ .

Como  $X_1(t) = X_0 \exp\left(\int_0^t f(s)dW(s) - \frac{1}{2} \int_0^t f^2(s)ds\right)$ , tem-se que,

$$X(t) = X_1(t)X_2(t) = X_0 \exp\left(\int_0^t f(s)dW(s) + \int_0^t \left[d(s) - \frac{1}{2}f^2(s)\right] ds\right).$$

**Exemplo 2.6.** Considere seguinte equação diferencial estocástica unidimensional:

$$\begin{cases} dX(t) = [c(t) + d(t)X(t)]dt + [e(t) + f(t)X(t)]dW(t), & t \in [0, T], \\ X(0) = X_0. \end{cases}$$

Como acima, assuma que o processo  $X(\cdot)$  possa ser representado pelo produto  $X(t) = X_1(t)X_2(t)$ , onde

$$\begin{cases} dX_1(t) = d(t)X_1(t)dt + f(t)X_1(t)dW(t), & t \in [0, T], \\ X_1(0) = 1, \end{cases}$$

e

$$\begin{cases} dX_2(t) = A(t)dt + B(t)dW(t), & t \in [0, T], \\ X(0) = X_0, \end{cases}$$

com  $A$  e  $B$  funções ainda não definidas. Então, aplicando a fórmula de Itô, tem-se que

$$\begin{aligned} dX(t) &= d(X_1X_2)(t) \\ &= X_1(t)dX_2(t) + X_2(t)dX_1(t) + f(t)X_1(t)B(t)dt \\ &= d(t)X_1(t)X_2(t)dt + f(t)X_1(t)X_2(t)dt + [X_1(t)dX_2(t) + f(t)X_1(t)B(t)dt] \\ &= d(t)X(t)dt + f(t)X(t)dW(t) + X_1(t)[A(t)dt + B(t)dW(t) + f(t)B(t)dt]. \end{aligned}$$

Para que a equação acima e a equação para  $X(\cdot)$  sejam iguais, é necessário que

$$X_1(t)[A(t)dt + B(t)dW(t) + f(t)B(t)dt] = c(t)dt + e(t)dW(t).$$

Isto implica que

$$A(t) = \frac{1}{X_1(t)}[c(t) - f(t)e(t)] \text{ e } B(t) = \frac{e(t)}{X_1(t)}.$$

Observe que

$$X_1(t) = \exp\left(\int_0^t f(s)dW(s) + \int_0^t \left[d(s) - \frac{1}{2} \int_0^s f^2(u)du\right] ds\right),$$

então  $X_1(t) > 0$  q.c., para todo  $t \in [0, T]$ .

Consequentemente, tem-se que

$$X_2(t) = X_0 + \int_0^t \frac{1}{X_1(t)} [c(s) - f(s)e(s)] dt + \int_0^t \frac{e(s)}{X_1(s)} dW(s).$$

Ao aplicar-se as fórmulas para  $X_1(\cdot)$  e  $X_2(\cdot)$  obtidas acima e usando o fato de que  $X(\cdot) = X_1(\cdot)X_2(\cdot)$ , é possível encontrar a fórmula para  $X(\cdot)$ .

**Exercício 2.7.** Encontre a fórmula para  $X(\cdot)$  no exemplo acima.



## Capítulo 3

# Introdução às Finanças em Tempo Contínuo

Neste capítulo serão introduzidos alguns modelos matemáticos para mercados financeiros. Mais precisamente, modelos que descrevem a dinâmica de ativos e derivativos, como as opções. Isto é especialmente relevante quando se deseja fazer análises de risco, de alocação de recursos ou precificar contratos derivativos.

Observe que a dinâmica do preço de ações, como as da Petrobras, da Vale S.A. ou do Itaú-Unibanco, são naturalmente aleatórias, podendo haver mudanças súbitas e bastante relevantes. Diante desta realidade, como se proteger de grandes perdas? Uma possibilidade é através da venda ou da compra de opções e contratos derivativos, os quais têm como objetivo a redução da exposição a certas fontes de risco. Assim, uma pergunta fundamental seria como encontrar o preço justo de tais instrumentos financeiros?

## Conceitos Básicos

Um conceito fundamental em finanças é o de arbitragem. Como definido em [20], uma estratégia de investimento que começa sem capital, tem probabilidade zero de perda de dinheiro e possui uma probabilidade positiva de se apurar lucro é uma estratégia de arbitragem.

Para ilustrar melhor este conceito, será apresentado um exemplo.

**Exemplo 3.1.** *Considere uma roleta que paga 2 : 1 quando sai vermelho e nada se sair preto.*

*As probabilidades dos resultados são:*

*Vermelho: 70% e Preto: 30%.*

*ogando muitas vezes, para cada \$1,00 investido, espera-se receber:*

$$2 \times 0,7 + 0 \times 0,3 = \$1,40.$$

*Assuma que um jogador venda por \$60,00 um bilhete que paga \$100,00 se sair vermelho e nada se sair preto. Assim, pergunta-se, este bilhete está caro ou barato? Procedendo como acima, o preço do bilhete deve ser \$70,00.*

*Antes de responder a pergunta, considere as seguintes situações:*

- 1. O jogador fica com o dinheiro guardado. Se se sair vermelho, ele tem um prejuízo de \$40,00, mas se sair preto, ele tem um lucro de \$60,00.*
- 2. O jogador aposta \$60,00 na roleta. Nesse caso, se sair vermelho, ele paga os \$100,00 ao comprador do bilhete e tem um lucro de \$20,00. Caso saia preto, ele fica com nada, mas também não leva prejuízo.*



3. Se o jogador apostar \$50,00 na roleta, não importa o que aconteça, ele terá um lucro de \$10,00.

*Então, pode-se dizer que o bilhete está caro, pois leva a arbitragem. Nas duas últimas possibilidades, a probabilidade de prejuízo é nula, e a probabilidade de se obter lucro é positiva.*

No exemplo acima, qual seria o preço justo do bilhete? Seria \$50,00, pois corresponde às probabilidades 50% – 50% para vermelho e preto. Ou seja,

$$\$100,00 \times 0,5 + \$0 \times 0,5 = \$50,00.$$

Observe ainda que, nas três situações do exemplo acima, se o preço for \$50,00, não haverá lucro.

Uma pergunta que surge naturalmente deste exemplo é, por que a probabilidade 50% – 50% levaria ao preço correto do bilhete?

Esta medida de probabilidade que leva ao preço correto de contratos derivativos é conhecida como probabilidade neutra ao risco. Para que ela exista e seja única, em geral, assume-se a hipótese de mercados eficientes, isto é, que não há oportunidades de arbitragem e, que o mercado é completo, ou seja, que todas as fontes de incerteza estão contidas no modelo matemático em consideração.

Assim, sob tais condições, é possível garantir a unicidade do preço justo de derivativos. O resultado que garante tal afirmação é o Teorema de Girsanov. Para o seu enunciado e sua demonstração, veja [16, Capítulo 3]. Observe que, em alguns casos, a hipótese de completude dos mercados pode ser enfraquecida, levando ao caso em que existe mais de um preço justo para tais contratos. Isso ocorre, por exemplo, em modelos para a dinâmica de ativos que consideram vários fatores.

No que se segue, ambas as hipóteses, de completude e de eficiência dos mercados, serão assumidas como verdadeiras.

Outro conceito de suma importância em finanças é o de *hedging*, isto

é, o de estratégias de investimento que têm como objetivo a redução da exposição a fontes de risco. Assim, constrói-se uma carteira, onde alocam-se partes do capital em ações e contratos derivativos. Vale ressaltar que, neste contexto assume-se por simplicidade que não há custos de transação e que é possível comprar e vender (a descoberto) infinitésimos de ações e derivativos.

Existem diversos tipos contratos derivativos os mais simples e mais negociados são as opções de compra e venda europeias e americanas.

1. **Opção de Compra Europeia:** é um contrato que dá o direito, mas não a obrigação, de comprar uma ação por um preço fixo, chamado de *strike*, numa data futura, que é o seu vencimento.
2. **Opção de Venda Europeia:** é um contrato similar à opção de compra, mas dá o direito de vender uma ação por um preço fixo no seu vencimento.
3. **Opções Americanas de Compra e de Venda:** similares às versões europeias mas podem ser exercidas em qualquer data entre a emissão destes contratos e os respectivos vencimentos. Devido a essa liberdade de exercício, tais contratos podem ser mais caros que as opções europeias.

Os preços destes contratos são funções dos preços das ações (ou ativos) subjacentes. Outro aspecto interessante é que opções europeias podem ser precificadas por equações diferenciais parciais, enquanto que opções americanas podem ser precificadas através de problemas de fronteira livre. Como dito acima, existem diversos outros tipos de opções, como opções asiáticas, opções *lookback*, opções reais, *spread options* e *basket options*.

## O Modelo Matemático

Tipicamente, a dinâmica do preço de um ativo ou ação é dada por um processo semi-martingal, isto é, se  $S_t$  denotar o preço no tempo  $t \geq 0$ , então, tem-se que

$$S_t = \text{Alguma Coisa}_t + \text{Martingal}_t.$$

Para a definição de martingais, veja o Capítulo 1. O Teorema da Representação Martingal, ver [16], estabelece que, sob certas hipóteses, martingais podem ser representados como processos de Itô, ou seja, possuem diferencial estocástica. Este é o cerne de finanças quantitativas em tempo contínuo.

Para tornar as ideias um pouco mais claras, considere a série temporal de preços de uma ação:

$$\{s_{t_0}, s_{t_1}, s_{t_2}, \dots, s_{t_n}\}, \quad \text{com} \quad \Delta t = t_i - t_{i-1}, \quad i = 1, 2, \dots, n.$$

Considere agora os retornos logarítmicos ou log-retornos:

$$y_i = \log(s_{t_i}/s_{t_{i-1}}).$$

Assuma, por simplicidade, que esta série de log-retornos possua as seguintes características:

1. independência,
2. mesma distribuição,
3. distribuição gaussiana.

Então, tem-se que

$$E(y) = \mu\Delta t = \frac{1}{n} \sum_{i=1}^n y_i \quad \text{e} \quad \text{Var}(y) = \sigma^2\Delta t = \frac{1}{n} \sum_{i=1}^n y_i^2 - (\mu\Delta t)^2$$

Assim, sob as hipóteses acima, o preço do ativo evolui satisfazendo a seguinte dinâmica.

$$\ln(S_t) = \ln(S_0) + \mu t + \text{“aleatoriedade”}$$

Tal aleatoriedade pode ser modelada da seguinte forma:

$$\sigma W_t,$$

em que  $W(t)$  é um movimento browniano.

Portanto, as ferramentas básicas para construir modelos para as dinâmicas de preços de ações e ativos em tempo contínuo são as equações diferenciais estocásticas e o Lema de Itô (ou fórmula de Itô).

De um modo geral, o preço de um ativo no tempo  $t$  é dado da seguinte forma:

**Definição 3.1.** *Seja  $(\Omega, \mathcal{F}, \tilde{P})$  um espaço de probabilidade com filtração  $\mathcal{G}(\cdot) = \{\mathcal{G}_t\}_{t \geq 0}$ . Então, o preço de um ativo no tempo  $t \geq 0$  é dado pela seguinte equação diferencial estocástica:*

$$\begin{cases} dS(t) = \mu(t)S(t)dt + \sigma(t)S(t)dW(t), \\ S(0) = S_0, \end{cases}$$

em que  $W_t$  é um movimento browniano,  $\mu_t \in \mathbb{L}^1(0, \infty)$ ,  $\sigma_t \in \mathbb{L}^2(0, \infty)$  e  $S_0$  é dado.

A filtração do espaço de probabilidade da definição pode ser a história do movimento browniano, a qual pode ser chamada de filtração

browniana. Veja [16]. Os parâmetros  $\mu(\cdot)$  e  $\sigma(\cdot)$  são conhecidos, respectivamente, como *drift* e *volatilidade*. Dependendo da escolha da medida de probabilidade do espaço, o drift pode ser substituído por uma taxa de juros livre de risco. Esta taxa é usada por uma conta bancária para remunerar os depósitos ou representa o custo de empréstimos.

## O Modelo de Black-Scholes

Introduzido em 1973, o modelo de Black-Scholes [5] representou um grande avanço na descrição matemática de fenômenos do mercado financeiro.

Seja  $(\Omega, \mathcal{F}, \tilde{P})$  um espaço de probabilidade com filtração  $\mathcal{G}(\cdot) = \{\mathcal{G}_t\}_{t \geq 0}$ . Assuma que o preço de uma ação ou ativo no tempo  $t \geq 0$  seja dado pela equação diferencial estocástica

$$dS_t = S_t(\mu dt + \sigma dW_t),$$

com  $W_t$  um movimento browniano,  $S_0$  dado,  $\mu \geq 0$  e  $\sigma \geq 0$  constantes. Neste contexto, como calcular o preço de contratos derivativos?

Primeiramente, assuma que:

1. não haja custos de transação,
2. que a ação não pague dividendos,
3. que não haja oportunidades de arbitragem,
4. que o preço da ação seja indefinidamente divisível, isto é, pode-se comprar/vender infinitésimos de uma ação,
5. que a venda a descoberto seja permitida,
6. que as negociações possam ocorrer a qualquer instante, ou seja, continuamente no tempo e,

7. que a taxa de juros livre de risco e a volatilidade sejam conhecidas e constantes.

Agora, considere o seguinte portfólio sem risco composto por uma conta bancária, representada por  $P(t)$ , uma ação, representada por  $S(t)$ , e uma opção de compra, cujo preço é representado por  $C(S(t), t)$ :

$$Y(t) = \phi(t)P(t) + \pi(t)S(t) - C(S(t), t),$$

em que  $P(t) = P(0) \exp(r \cdot t)$  representa um investimento numa conta bancária que remunera à taxa de juros livre de risco, contínua e anualizada de valor  $r > 0$ . Os parâmetros  $\phi$  e  $\pi$  representam as quantidades investidas em cada um dos ativos, ou seja, a conta bancária  $P(t)$  e a ação  $S(t)$ . Observe que o investidor vende a descoberto uma opção e investe o dinheiro obtido na transação na conta bancária e na ação, assim,  $\phi(0) + \pi(0) = 1$ . Como o portfólio é sem risco, o capital investido deve evoluir de acordo com a taxa de juros livre de risco  $r$ , ou seja, a dinâmica do portfólio é dada por

$$dY(t) = rY(t)dt.$$

Por outro lado, tem-se que

$$dY(t) = \phi(t)dP(t) + \pi(t)dS(t) + dC(S(t), t).$$

Aplicando a fórmula de Itô a  $C(S(t), t)$ , e assumindo que  $C$  é contínua

assim como as suas derivadas parciais  $\frac{\partial C}{\partial t}$ ,  $\frac{\partial C}{\partial S}$  e  $\frac{\partial^2 C}{\partial S^2}$ , tem-se:

$$\begin{aligned} dY(t) &= \phi(t)dP(t) + \pi(t)dS(t) + dC(S(t), t) \\ &= \phi(t)rP(t)dt + \pi(t)(\mu S(t)dt + \sigma S(t)dW(t)) \\ &- \left[ \frac{\partial C}{\partial t}(S(t), t)dt + \frac{\partial C}{\partial x}(S(t), t)dS(t) + \frac{1}{2} \frac{\partial^2 C}{\partial x^2}(S(t), t)\sigma^2 S(t)^2 dt \right] \\ &= \phi(t)rP(t)dt + \pi(t)(\mu S(t)dt + \sigma S(t)dW(t)) \\ &- \left[ \frac{\partial C}{\partial t}(S(t), t)dt + \frac{\partial C}{\partial x}(S(t), t)(\mu S(t)dt + \sigma S(t)dW(t)) \right. \\ &\quad \left. + \frac{1}{2} \frac{\partial^2 C}{\partial x^2}(S(t), t)\sigma^2 S(t)^2 dt \right]. \end{aligned}$$

Assumindo que  $\pi(t) = \frac{\partial C}{\partial x}(S(t), t)$ , e usando o fato de

$$Y(t) = \phi(t)P(t) + \pi(t)S(t) - C(S(t), t),$$

segue que

$$0 = rC(S(t), t) - rS(t) \frac{\partial C}{\partial x}(S(t), t) - \frac{\partial C}{\partial t}(S(t), t) - \frac{1}{2} \sigma^2 S(t)^2 \frac{\partial^2 C}{\partial x^2}(S(t), t).$$

Ou seja, chega-se a famosa equação diferencial parcial de Black-Scholes:

$$\frac{\partial C}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + r S \frac{\partial C}{\partial S} - r C = 0, \quad 0 < t < T, \quad S > 0.$$

Se  $C$  representar uma opção de compra com vencimento em  $T$  e strike  $K$ , tem-se que, para o tempo  $t = T$  a opção tem como payoff a função:

$$C(T, S) = \max\{0, S - K\}.$$

Isto estabelece a condição terminal para a equação de Black-Scholes, ou seja, o preço de uma opção de compra com strike  $K$  e vencimento  $t$ ,

para todo tempo  $t \in [0, T]$  e para qualquer preço da ação  $S \in (0, +\infty)$  é a solução do seguinte problema de valor terminal e de contorno:

$$\left\{ \begin{array}{l} \frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + rS \frac{\partial C}{\partial S} - rC = 0, \quad 0 < t < T, \quad S > 0, \\ C(0, t) = 0, \quad 0 < t < T, \\ C(S, t) \approx S, \quad \text{quando } S \rightarrow +\infty, \quad 0 < t < T, \\ C(S, T) = \max\{0, S - K\}, \quad S > 0. \end{array} \right. \quad (3.1)$$

Observe que na equação acima o argumento  $(S(t), t)$  foi omitido. Usando mudanças de variáveis e renormalizações apropriadas, prova-se que o problema de valor terminal acima é equivalente a um problema de inicial associado a equação do calor

$$\frac{\partial u}{\partial s}(x, s) = \frac{\partial^2 u}{\partial x^2}(x, s), \quad s \geq 0, \quad x \in \mathbb{R}.$$

A partir disso, é possível encontrar a solução analítica da equação de Black-Scholes para uma opção de compra europeia com vencimento em  $T$  e strike  $K$ . Ou seja, para todo  $t \in [0, T]$

$$C(t, S) = SN(d_1) - Ke^{-r(T-t)}N(d_2), \quad (3.2)$$

em que

$$N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{y^2}{2}} dy, \quad (3.3)$$

$$d_1(t, S) = \frac{\log(S/K) + (r + \sigma^2/2)(T - t)}{\sigma\sqrt{T - t}}, \quad (3.4)$$

e

$$d_2(t, S) = d_1 - \sigma\sqrt{T - t}. \quad (3.5)$$



Existe ainda uma outra forma de se chegar na equação de Black-Scholes sem precisar usar uma estratégia de investimento. Observe que, sob a hipótese de ausência de arbitragem, tem-se a existência da medida de probabilidade neutra ao risco, denotada por  $Q$ . Sob  $Q$ , o preço da ação  $S(t)$  satisfaz a seguinte dinâmica:

$$dS(t) = rS(t)dt + \sigma S(t)d\widetilde{W}(t),$$

em que  $\widetilde{W}(t)$  é um movimento browniano na medida  $Q$  ou um  $Q$ -movimento browniano. Neste caso, uma opção de compra europeia é dada pelo seguinte valor esperado:

$$C(S(t), t, K, T) = e^{-r(T-t)}\widetilde{E}(\max\{0, S(T) - K\} | \mathcal{G}(t)),$$

em que  $\widetilde{E}$  representa o valor esperado na medida  $Q$ .

De um modo mais geral, opções europeias com um payoff  $f(S(T))$  satisfazem:

$$V(S(t), t) = e^{-r(T-t)}\widetilde{E}[f(S(T)) | \mathcal{G}(t)].$$

Então, sob hipóteses apropriadas, uma opção europeia com vencimento em  $T > 0$  e payoff  $f(S(T))$  é a solução da seguinte equação diferencial parcial:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} - rV = 0, \quad 0 < t < T, \quad S > 0,$$

com a condição terminal abaixo:

$$V(S, T) = f(S).$$

Este resultado é uma consequência do Teorema de Feynmann-Kac. Para o seu enunciado, veja [16].

É possível mostrar que o problema de valor terminal e de contorno

em (3.1) é equivalente ao problema de valor inicial e de contorno do calor. Para ver isto, primeiramente, defina:

$$S = Ke^x, \quad t = T - \frac{\tau}{\frac{1}{2}\sigma^2}, \quad C(S, t) = Kv(x, \tau),$$

em que a função  $v$  é solução do seguinte problema de valor inicial e de contorno:

$$\left\{ \begin{array}{l} \frac{\partial v}{\partial \tau} = \frac{\partial^2 v}{\partial x^2} + (k-1)\frac{\partial v}{\partial x} - kv, \quad 0 < \tau < T, \quad x \in \mathbb{R}, \\ \lim_{x \rightarrow -\infty} v(x, \tau) = 0, \quad 0 < \tau < T, \\ v(x, \tau) \approx e^x, \quad \text{quando } x \rightarrow +\infty, \\ v(x, 0) = \max\{0, e^x - 1\}, \quad x \in \mathbb{R}, \end{array} \right.$$

em que  $k = \frac{r}{\frac{1}{2}\sigma^2}$ .

Agora, defina a função  $u(x, \tau)$  de modo que  $v(x, \tau) = e^{\alpha x + \beta \tau} u(x, \tau)$ , em que

$$\alpha = -\frac{1}{2}(k-1) \quad \text{e} \quad \beta = -\frac{1}{4}(k+1)^2.$$

Então,  $u$  é solução da equação do calor

$$\frac{\partial u}{\partial \tau} = \frac{\partial^2 u}{\partial x^2}, \quad 0 < \tau < T, \quad x \in \mathbb{R},$$

com a condição inicial

$$u(x, 0) = \max\{0, e^{\frac{1}{2}(k+1)x} - e^{\frac{1}{2}(k-1)x}\}, \quad x \in \mathbb{R}.$$

**Exercício 3.1.** *Determine as condições de contorno satisfeitas por  $u$ . Verifique que  $u$  e  $v$  de fato satisfazem as equações diferenciais parciais*

*e as condições iniciais e de contorno apresentadas acima.*

## Generalizações do Modelo de Black-Scholes

O modelo de Black-Scholes é inegavelmente o modelo mais conhecido em finanças quantitativas. No entanto, sua aplicação é bastante limitada, atualmente, devido a sua grande simplicidade. Por usar apenas dois parâmetros constantes, ele não consegue capturar de maneira satisfatória as incertezas do mercado. Assim, para corrigir este problema, diversas generalizações tem sido propostas, muitas delas usam parâmetros não-constantes, podendo ser determinísticos ou aleatórios. O grande problema é encontrar modelos que consigam ser sofisticados o suficiente para capturar de forma precisa as incertezas do mercado e que, ao mesmo tempo, sejam simples de ser avaliado computacionalmente. Modelos muito sofisticados podem ter uso bastante limitado, pois podem levar a problemas computacionais bastante relevantes. Outra questão relevante é a de se calibrar ou estimar os parâmetros de um modelo. Quanto mais sofisticado, mais difícil é encontrar parâmetros que descrevam bem a realidade.

Nesta seção serão tratadas duas generalizações bastante populares do modelo de Black-Scholes, a saber, os modelos de Dupire [8] e de Heston [11].

### O Modelo de Dupire

O modelo de Dupire [8] foi introduzido em 1994 por Bruno Dupire e assume que a volatilidade constante no modelo de Black-Scholes é substituída por uma função que depende do tempo e do preço da ação. Esse parâmetro é conhecido como volatilidade local e o modelo de Dupire é também chamado de Modelo da Volatilidade Local.

Seja  $(\Omega, \mathcal{F}, Q)$  um espaço de probabilidade com filtração  $\mathcal{G}(\cdot)$ . Neste

modelo, sob a probabilidade neutra ao risco  $Q$ , o preço do ativo satisfaz a seguinte equação diferencial estocástica:

$$\begin{cases} dS_t = (r - q) S(t) dt + \sigma(S(t), t) S(t) d\widetilde{W}(t), & t \geq 0 \\ S_0 \text{ é dado,} \end{cases}$$

em que  $\widetilde{W}(\cdot)$  é um  $Q$ -movimento browniano,  $\sigma(S(t), t)$  é uma função determinística do preço do ativo e do tempo,  $r$  é a taxa de juros livre de risco e  $q$  representa uma taxa de pagamento de dividendos da ação.

Novamente, opções de compra europeias são dadas pelo valor esperado na medida  $Q$ :

$$C(S(t), t, K, T) = \widetilde{\mathbb{E}}[e^{-r(T-t)} \max\{0, S_T - K\} | \mathcal{F}(t)].$$

Primeiramente, deve-se fixar o tempo  $t = 0$  e o preço da ação  $S(t) = S_0$ , deixando o vencimento  $T$  e o strike  $K$  variar. Além disso, ao se assumir que para cada  $T$ ,  $S(T)$  tenha uma densidade de probabilidade  $\varphi(S, T)$ , tem-se:

$$C(S_0, 0, K, T) = e^{-rT} \int_0^\infty \max\{0, S - K\} \varphi(S, T) dS.$$

A densidade de probabilidade  $\varphi(S, T)$  satisfaz a equação de Fokker-Planck e, após integrar por partes esta equação, entre outras manipulações algébricas, chega-se ao seguinte problema de valor inicial e de contorno (veja [11]):

$$\left\{ \begin{array}{l} \frac{\partial C}{\partial T} = \frac{1}{2}\sigma^2 K^2 \frac{\partial^2 C}{\partial K^2} - (r - q)K \frac{\partial C}{\partial K} - qC, \quad T > 0, \quad K \geq 0 \\ \lim_{K \rightarrow 0} C(K, T) = S_0, \quad T > 0, \\ \lim_{K \rightarrow +\infty} C(K, T) = 0, \quad T > 0, \\ C(K, 0) = \max\{0, S_0 - K\}, \quad K > 0. \end{array} \right.$$

Observe que o argumento  $(K, T)$  na equação foi omitido. Como o parâmetro  $\sigma$  na equação acima é uma função, o problema acima não possui uma solução analítica explícita, sendo necessário o uso de técnicas computacionais, como diferenças finitas ou elementos finitos, para encontrar a sua solução. Veja [1].

Fazendo as mudanças de variáveis

$$\tau = T \quad \text{e} \quad y = \log(K/S_0),$$

substituindo  $C$  por  $\tilde{C}(K, \tau) = e^{-r\tau}C(K, \tau)$ , com  $r$  a taxa de juros contínua livre de risco, e definindo:

$$u(y, \tau) = \tilde{C}(S_0 e^y, \tau) \quad \text{e} \quad a(y, \tau) = \frac{1}{2}\sigma^2(S_0 e^y, \tau),$$

tem-se agora o seguinte problema de valor de contorno e inicial:

$$\left\{ \begin{array}{l} -u_\tau + a(u_{yy} - u_y) - ru_y = 0, \quad \tau > 0, \quad y \in \mathbb{R}, \\ u(y, 0) = S_0(1 - e^y)^+, \quad y \in \mathbb{R}, \\ \lim_{y \rightarrow -\infty} u(y, \tau) = S_0, \quad \tau > 0, \\ \lim_{y \rightarrow +\infty} u(y, \tau) = 0, \quad \tau > 0. \end{array} \right. \quad (3.6)$$

Este possui uma expressão uma única solução, desde que o parâmetro de difusão  $a$  seja suficientemente regular, i.e., se  $a$  satisfizer, por exemplo, as seguintes condições:

1.  $c_1 < a(y, \tau) < c_2$ , para todo  $(y, \tau) \in D = [0, T_{\max}] \times \mathbb{R}$ , com  $0 < c_1 \leq c_2 < \infty$  constantes dadas.
2.  $a$  é contínua e duas vezes diferenciável, com derivadas parciais contínuas.
3. existe uma constante  $a_0$  tal que,  $a - a_0$  é de quadrado integrável em  $D$ , assim como as suas derivadas parciais.

Observe que, mesmo sob estas hipóteses, o problema acima não possui uma expressão para a sua solução, sendo necessário o uso de métodos numéricos para encontrar uma aproximação para a mesma. Este problema foi bastante estudado e diversos modelos de solução foram propostos. Veja, pro exemplo, [2] e referências aí contidas.

## O Modelo de Heston

Este modelo, introduzido em 1993 por Heston em [13], assume que a volatilidade constante do modelo de Black-Scholes é substituída por uma

volatilidade estocástica, cujo a dinâmica é governada por um processo de Itô com reversão à média.

Seja  $(\Omega, \mathcal{F}, Q)$  um espaço de probabilidade com filtração  $\mathcal{G}(\cdot)$ . Então, o preço de um ativo satisfaz o seguinte sistema de equações diferenciais estocásticas:

$$\left\{ \begin{array}{l} dS_t = \mu S(t) dt + \sqrt{v(t)} S_t dW(t), \quad t \geq 0 \\ dv(t) = -\lambda(v(t) - \bar{v})dt + \eta \sqrt{v(t)} dZ(t), \\ S_0, v_0 \quad \text{são dados,} \\ d\langle W(t), Z(t) \rangle = \rho dt \end{array} \right.$$

Novamente, o preço de uma opção de compra europeia com vencimento em  $T$  e strike  $K$  é dado por:

$$C(S(t), t, K, T) = \tilde{\mathbb{E}}(e^{-r(T-t)} \max\{0, S_T - K\} | \mathcal{F}(t)).$$

É possível usar, neste contexto, um argumento similar ao aplicado ao modelo de Black-Scholes para se encontrar uma equação diferencial parcial de precificação de opções europeias. Veja [11].

Ao considerar o preço de uma opção de compra europeia como função do preço do ativo  $S$  da volatilidade  $v$  e do tempo  $t$ , tem-se um problema de valor terminal, com a equação diferencial parcial:

$$\begin{aligned} \frac{\partial C}{\partial t} + \frac{1}{2} v S^2 \frac{\partial^2 C}{\partial S^2} + \rho \nu v S \frac{\partial^2 C}{\partial v \partial S} + \frac{1}{2} \eta^2 v \frac{\partial^2 C}{\partial v^2} + r S \frac{\partial C}{\partial S} - r C \\ = \lambda(v - \bar{v}) \frac{\partial C}{\partial v}, \quad S \geq 0, \quad v \geq 0, \quad t \in [0, T], \end{aligned}$$

e a condição terminal

$$C(S, v, T) = \max\{0, S - K\}, \quad S \geq 0, \quad v \geq 0.$$

Existe uma fórmula para a solução analítica deste problema de valor final, para apresentá-la, primeiramente observe que, ao se fazer as mudanças de variáveis

$$\tau = T - t \quad \text{e} \quad x = \log(S_t/K),$$

a função  $V(x, v, \tau) = C(Ke^x, v, T - \tau)$  satisfaz a equação diferencial parcial

$$\begin{aligned} -\frac{\partial V}{\partial \tau} + \frac{1}{2}v\frac{\partial^2 V}{\partial x^2} - \frac{1}{2}v\frac{\partial V}{\partial x} + \rho v v\frac{\partial^2 V}{\partial v \partial x} + \frac{1}{2}\eta^2 v\frac{\partial^2 V}{\partial v^2} + r\frac{\partial V}{\partial x} - rV \\ = \lambda(v - \bar{v})\frac{\partial C}{\partial v}, \quad x \in \mathbb{R}, \quad v \geq 0, \quad \tau \in [0, T], \end{aligned}$$

e a condição inicial

$$V(x, v, 0) = K \max\{0, e^x - 1\}, \quad x \in \mathbb{R}, \quad v \geq 0.$$

Assim, assumamos que  $V$  assumamos a seguinte forma:

$$V(x, v, \tau) = K [e^x P_1(x, v, \tau) - P_0(x, v, \tau)].$$

Para  $j = 0, 1$ , tem-se que

$$P_j(x, v, \tau) = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \operatorname{Re} \left[ \frac{\exp(C_j(u, \tau)\bar{v} + D_j(u, \tau)v + iux)}{iu} \right] du$$

Defina os parâmetros:

$$\alpha_j = -\frac{u^2}{2} - \frac{i u}{2} + i j u, \quad \beta_j = \lambda - j \rho \eta - i \rho \eta u, \quad \gamma = \frac{\eta^2}{2}.$$

$$r_{\pm, j} = \frac{\beta_j \pm d_j}{\eta^2}, \quad \text{com } d_j = \sqrt{\beta_j^2 - 4\alpha_j \gamma}, \quad \text{e } g = \frac{r_{-, j}}{r_{+, j}}.$$

Usando os parâmetros acima, tem-se que os termos na equação de  $V$



são:

$$C_j(u, \tau) = \lambda \left[ r_{-,j} \tau - \frac{2}{\eta^2} \log \left( \frac{1 - g_j e^{-d_j \tau}}{1 - g_j} \right) \right]$$

e

$$D_j(u, \tau) = r_{-,j} \frac{1 - e^{-d_j \tau}}{1 - g_j e^{-d_j \tau}}.$$

Essa fórmula, apesar de ser complexa, é bastante útil para a identificação dos parâmetros do modelo de Heston a partir de preços de opções de compra europeias. Para uma implementação numérica destas fórmulas, veja a dissertação [17].



## Capítulo 4

# Simulações Numéricas

Neste capítulo serão apresentadas técnicas computacionais para a simulação e a solução numérica de equações diferenciais estocásticas e equações diferenciais parciais.

### O Método Euler-Maruyama para Equações Diferenciais Estocásticas

O método de Euler-Maruyama foi desenvolvido para encontrar a solução de equações diferenciais estocásticas. Sua principal vantagem é a simplicidade de implementação. Este tipo de método constrói uma sequência de realizações de aproximações dos caminhos da solução de uma equação diferencial estocástica. Uma boa referência para estes métodos é o artigo [14].

Considere um processo estocástico  $X(\cdot)$ , definido no espaço de probabilidade  $(\Omega, \mathcal{F}, P)$  com filtração  $\mathcal{G}(\cdot)$ , e que possua diferencial estocástica

$$dX(t) = f(X(t))dt + g(X(t))dW(t), \quad t \in [0, T],$$

onde  $W(\cdot)$  é um movimento browniano neste mesmo espaço, e com  $X(0) = X_0$  dado.

Seja  $\Delta t = T/L$  para algum inteiro positivo  $L$ , e  $\tau_j = j\Delta t$ , uma discretização com tamanho de passo  $\Delta t$ , do intervalo  $[0, T]$ . Então, o método de Euler-Maruyama (EM) assume a seguinte forma:

$$X_{j+1} = X_j + f(X_j)\Delta t + g(X_j)(W(\tau_{j+1}) - W(\tau_j)), \quad j = 0, 1, \dots, L-1.$$

Observe que o método acima é uma aproximação das seguintes equações:

$$X(\tau_{j+1}) = X(\tau_j) + \int_{\tau_j}^{\tau_{j+1}} f(X(t))dt + \int_{\tau_j}^{\tau_{j+1}} g(X(t))dW(t), \quad j = 0, 1, \dots, L-1.$$

Assumindo que  $f$  e  $g$  sejam funções reais determinísticas (não-aleatórias), tem-se que, pela definição de movimento browniano, o termo  $W(\tau_{j+1}) - W(\tau_j)$  tem distribuição gaussiana com média zero e variância igual a  $\Delta t$ , isto é,

$$W(\tau_{j+1}) - W(\tau_j) = {}^d N(0, \Delta t).$$

Portanto, para cada  $j = 0, 1, \dots, L-1$ , os valores de  $X_{j+1}$  do método EM são completamente determinados por  $X_j$  e por realizações de uma variável aleatória com distribuição  $N(0, \Delta t)$ .

**Exemplo 4.1.** *Considere o modelo de Black-Scholes para o preço de uma ação:*

$$dS(t) = S(t)(\mu dt + \sigma dW(t)), \quad t \in [0, T],$$

com  $S(0) = S_0$ . A solução desta equação diferencial estocástica é dada por

$$S(t) = S_0 \exp \left( \left( \mu - \frac{1}{2}\sigma^2 \right) t + \sigma W(t) \right)$$

ou

$$S(\tau_{j+1}) = S(\tau_j) \exp \left( \left( \mu - \frac{1}{2} \sigma^2 \right) \Delta t + \sigma (W(\tau_{j+1}) - W(\tau_j)) \right)$$

Neste caso, é possível comparar as soluções analítica e a dada pelo método de Euler-Maruyama, que é dada por

$$S(\tau_{j+1}) = S(\tau_j) + S(\tau_j) (\mu \Delta t + \sigma (W(\tau_{j+1}) - W(\tau_j)))$$

Para este Assuma que  $\mu = 0.03$ ,  $\sigma = 0.4$ ,  $S_0 = 1$  e  $T = 1$ . Para a discretização do método EM, assuma que  $L = 100$ . A comparação entre um caminho dado pela solução analítica (segunda formulação) e o dado pelo EM pode ser vista na figura 4.1. Observe que a mesma realização de  $W(\tau_{j+1}) - W(\tau_j)$  foi usada para calcular  $S_{j+1}$  e  $S(\tau_{j+1})$ .

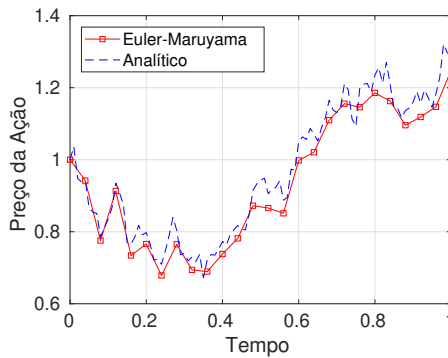


Figura 4.1: Comparação entre um caminho dado pela solução analítica do preço de uma ação dado pelo modelo de Black-Scholes e a solução dada pelo modelo de Euler-Maruyama.

**Exercício 4.1.** Construa uma função numérica que calcule as soluções analítica e numérica do exemplo anterior.

**Exemplo 4.2.** *Considere agora o modelo de Heston:*

$$\left\{ \begin{array}{l} dS_t = \mu S(t) dt + \sqrt{v(t)} S_t dW(t), \quad t \geq 0 \\ dv(t) = -\lambda(v(t) - \bar{v})dt + \eta\sqrt{v(t)}dZ(t), \\ S_0, v_0 \quad \text{são dados,} \\ d\langle W(t), Z(t) \rangle = \rho dt \end{array} \right.$$

*Este modelo não possui uma fórmula para a solução analítica. A solução numérica dada pelo modelo de Euler-Maruyama é:*

$$S_{j+1} = S_j + \mu S_j \Delta t + \sqrt{v_j} S_j (W(\tau_{j+1}) - W(\tau_j))$$

$$v_{j+1} = v_j + \lambda(v_j - \bar{v})\Delta t + \eta\sqrt{v_j}(Z(\tau_{j+1}) - Z(\tau_j))$$

*Para construir movimentos brownianos com correlação  $\rho$ , considere dois movimentos brownianos independentes  $W_1(\cdot)$  e  $W_2(\cdot)$ . Assim,  $W(\cdot) = W_1(\cdot)$  e*

$$Z(t) = \rho W_1(t) + \sqrt{1 - \rho^2} W_2(t), \quad \text{para todo } t \geq 0.$$

*Para este exemplo, assuma que  $\mu = 0.03$ ,  $v_0 = 0.4$ ,  $\eta = 0.4$ ,  $\lambda = 0.1$ ,  $\bar{v} = 0.4$ ,  $S_0 = 1$ ,  $\rho = 0.5$  e  $T = 1$ . Para a discretização do método EM, assuma que  $L = 100$ .*

*Exemplos de caminhos de  $S(\cdot)$  e de  $v(\cdot)$  gerados pelo método de Euler-Maruyama podem ser vistos na figura 4.2.*

**Exercício 4.2.** *Novamente, construa uma função numérica que calcule as soluções analítica e numérica do exemplo anterior.*

Uma das possibilidades em se usar este tipo de técnica é o de aproximar valores esperados por médias amostrais. Por exemplo, ao se

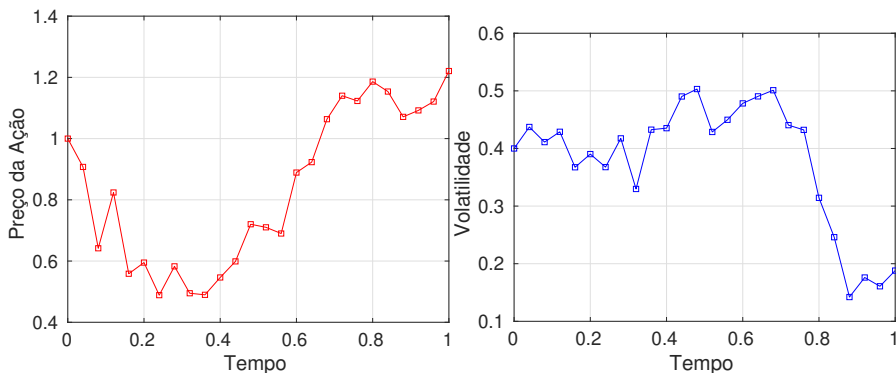


Figura 4.2: Exemplos de caminhos de  $S(\cdot)$  (esquerda) e de  $v(\cdot)$  (direita) do modelo de Heston gerados pelo método de Euler-Maruyama.

gerar um grande número de caminhos, o preço de uma opção de compra europeia no tempo zero, que é dado pelo valor esperado, na medida neutra ao risco,

$$C(S_0, 0, K, T) = e^{-rT} E(\max\{0, S_T - K\} | S(0) = S_0),$$

pode ser aproximado por

$$e^{-rT} E(\max\{0, S_T - K\} | S(0) = S_0) \approx e^{-rT} \frac{1}{N} \sum_{j=1}^N \max\{0, S_T^j - K\},$$

em que  $S_T^j$ ,  $j = 1, \dots, N$  são realizações da variável aleatória  $S_T$  que podem ser obtidas usando o método de Euler-Maruyama.

**Exemplo 4.3.** *Considere novamente o modelo de Black-Scholes. Use os parâmetros do exemplo acima sobre a aplicação do método de Euler-Maruyama ao mesmo modelo. Usando a fórmula de Black-Scholes, tem-se que, o preço de uma opção de compra europeia, sob o modelo de Black-Scholes, com vencimento em 1 ano e strike \$0.70 é \$0.3496.*

*Usando os mesmos parâmetros do exemplo anterior para obter a*

*solução numérica da equação diferencial estocástica usando método de Euler-Maruyama, são geradas 10.000 realizações de  $S_T$  e tem-se que*

$$e^{-rT} \frac{1}{N} \sum_{j=1}^N \max\{0, S_T^j - K\} = \$0.3521.$$

*Observe que o resultado não é exato, mas a diferença é muito pequena e representa um erro relativo de 0.0073, menor que 1%.*

**Exemplo 4.4.** *Repetindo o experimento do exemplo anterior, mas agora, usando o modelo de Heston, assumindo que o vencimento é em 1 ano e o strike é \$0.70, tem-se que o preço da opção de compra europeia usando a aproximação do valor esperado é \$0.3967, já o preço dado pela fórmula é \$0.3983. Novamente, a diferença é muito pequena, apresentando um erro relativo de 0.004, menor que 1%.*

Esta aproximação do valor esperado é chamada de integração de Monte Carlo. Existem maneiras bem mais sofisticadas de se realizar este tipo de aproximação. No entanto, esta funciona bem. Os métodos de Monte Carlo, que usam amostragens de variáveis aleatórias são muito usados na prática, principalmente quando não existem meios determinísticos de se encontrar preços de derivativos. Em geral, estes métodos são usados com modelos mais sofisticados que o de Black-Scholes para a precificação de derivativos e opções mais complexos, como opções exóticas e dependentes de caminho.

A grande dificuldade de se usar métodos de Monte Carlo é o alto custo computacional, pois a geração de um grande número cenários, dependendo do modelo, pode ser proibitiva.

Para exemplos de códigos em Matlab para a solução numérica de equações diferenciais estocásticas lineares usando o método de Euler-Maruyama, veja o artigo [14]. Uma implementação em Matlab da fórmula de precificação de opções dada pelo o modelo de Heston, usando



a transformada rápida de Fourier, pode ser encontrada em [17].

**Exercício 4.3.** *Refazer os experimentos dos exemplos anteriores. O leitor pode se basear nos códigos em Matlab apresentados nas referências [14, 17].*

## Métodos de Diferenças Finitas

Os métodos de diferenças finitas formam uma classe de técnicas para encontrar soluções numéricas de equações diferenciais parciais. Existem diversos outros métodos, como o de elementos finitos e o de volumes finitos. No entanto, devido a sua simplicidade, versatilidade e precisão, apenas alguns esquemas de diferenças finitas serão abordados neste texto. O leitor interessado pode saber mais sobre estes métodos em [21].

No que se segue serão apresentados alguns exemplos de esquemas de diferenças finitas para a equação do calor unidimensional. Observe que estes métodos podem ser generalizados facilmente para outras equações.

Observe que, a partir de certas mudanças de variáveis, a equação de Black-Scholes pode ser transformada na equação do calor

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}.$$

O preço de uma opção de compra europeia  $C(K, T)$  pode ser obtido a partir de  $u$  através da seguinte expressão:

$$C(K, T) = K^{\frac{1}{2}(1+k)} S_0^{\frac{1}{2}(1-k)} e^{\frac{1}{8}(k+1)^2 \sigma^2 (T-t)} u \left( \log(S_0/K), \frac{1}{2} \sigma^2 (T-t) \right),$$

em que  $k = \frac{r}{\frac{1}{2} \sigma^2}$ . Para mais detalhes, veja [25].

A ideia central dos esquemas de diferenças finitas é a aproximação das derivadas parciais por diferenças entre valores da função  $u$  numa

malha, em que a distância entre esses pontos é pequena.

Por exemplo, se  $f : \mathbb{R} \rightarrow \mathbb{R}$  é uma função três vezes continuamente diferenciável, tem-se que a expansão em série de Taylor de  $f$  em torno de  $x$  é dada por

$$f(x + \Delta x) = f(x) + f'(x)\Delta x + \frac{f''(x)}{2}(\Delta x)^2 + O((\Delta x)^3),$$

em que o  $O((\Delta x)^3)$  representa os termos de ordem maior ou igual  $(\Delta x)^3$ . Portanto,

$$\frac{f(x + \Delta x) - f(x)}{\Delta x} = f'(x) + \frac{f''(x)}{2}(\Delta x) + O((\Delta x)^2),$$

ou seja,

$$f'(x) = \frac{f(x + \Delta x) - f(x)}{\Delta x} + O(\Delta x).$$

Neste caso específico, no método de diferenças finitas, substitui-se a derivada de  $f$  pela aproximação de diferenças finitas no lado direito e omite-se os termos de ordem maior ou igual a  $\Delta x$ . Isto é, considera-se que

$$f'(x) \approx \frac{f(x + \Delta x) - f(x)}{\Delta x}.$$

Esta aproximação em particular é chamada de diferença avançada. Se, ao invés de se considerar  $x + \delta x$ , fosse usado  $x - \Delta x$ , ter-se-ia a aproximação por diferenças finitas atrasada, ou seja,

$$f'(x) \approx \frac{f(x - \Delta x) - f(x)}{\Delta x},$$

onde a expansão em série de Taylor de  $f$ , nesse caso é dada por:

$$f(x - \Delta x) = f(x) - f'(x)\Delta x + \frac{f''(x)}{2}(\Delta x)^2 + O((\Delta x)^3),$$

Ambas as aproximações por diferenças finitas acima são de primeira

ordem, isto é, os termos remanescentes da série de Taylor que aparecem na aproximação da derivada são  $O(\Delta x)$ .

Observe agora caso conhecido como diferenças centradas, isto é,

$$\frac{f(x + \Delta x) - f(x - \Delta x)}{2\Delta x} = f'(x) + O((\Delta x)^2).$$

Neste caso, a aproximação

$$f'(x) \approx \frac{f(x + \Delta x) - f(x - \Delta x)}{2\Delta x},$$

é de segunda ordem. No entanto, este caso não é muito considerado na prática devido a questões de estabilidade numérica. O que é mais comum de se fazer é usar, ao invés de  $\Delta x$ , usar  $\Delta x/2$ , que leva à aproximação

$$f'(x) \approx \frac{f(x + \Delta x/2) - f(x - \Delta x/2)}{\Delta x},$$

a qual também é de segunda ordem e aparece no esquema conhecido como de Crank-Nicolson.

Para encontrar uma aproximação por diferenças finitas para a segunda derivada de  $f$ , primeiramente, soma-se as expansões de  $f$  para  $x + \Delta x$  e  $x - \Delta x$  em torno de  $x$ , respectivamente, e obtém-se

$$f(x + \Delta x) - f(x - \Delta x) = 2f(x) + f''(x)(\Delta x)^2 + O((\Delta x)^4).$$

Assim, subtraindo  $2f(x)$  e dividindo por  $(\Delta x)^2$  em ambos os lados da equação acima tem-se que

$$f''(x) = \frac{f(x + \Delta x) - 2f(x) + f(x - \Delta x)}{(\Delta x)^2} + O((\Delta x)^2).$$

Ou seja,

$$f''(x) \approx \frac{f(x + \Delta x) - 2f(x) + f(x - \Delta x)}{(\Delta x)^2}.$$

e obtém-se uma aproximação de segunda ordem para a segunda derivada de  $f$ , que é chamada diferença centrada de segunda ordem.

Assim, para a função  $u(x, \tau)$  solução da equação do calor, tem-se que ela satisfaz:

Como os esquemas de diferenças finitas são obtidos com respeito aos pontos  $x$ ,  $x + \Delta x$  e  $x - \Delta x$ , pode-se considerar uma discretização de um intervalo  $[a, b]$ , com tamanho de passo  $\Delta x = (b - a)/N$ , denotada por  $x_i = a + i\Delta x$ ,  $i = 0, 1, \dots, N$ .

De modo similar, pode-se considerar uma discretização, ou malha (uniforme) de diferenças finitas, do retângulo  $[a, b] \times [c, d]$ , com os tamanhos de passo  $\Delta x = (b - a)/I$  e  $\Delta \tau = (d - c)/J$ , e denotada por

$$(x_i, \tau_j) = (a + i\Delta x, c + j\Delta \tau), \quad i = 0, 1, \dots, I, \quad j = 0, 1, \dots, J.$$

Assim, os valores de uma função  $u : [a, b] \times [c, d] \rightarrow \mathbb{R}$  nos pontos da malha de diferenças finitas são denotados por:

$$u_i^j = u(a + i\Delta x, c + j\Delta \tau), \quad i = 0, 1, \dots, I, \quad j = 0, 1, \dots, J.$$

Considere novamente o problema de valor de contorno e inicial de-

finido pela equação do calor:

$$\left\{ \begin{array}{l} \frac{\partial u}{\partial \tau} = \frac{\partial^2 u}{\partial x^2}, \quad (x, \tau) \in [a, b] \times [0, T_{\max}], \\ u(x, 0) = f(x), \quad x \in [a, b] \\ u(a, \tau) = g(\tau), \quad \tau \in [0, T_{\max}], \\ u(b, \tau) = h(\tau), \quad \tau \in [0, T_{\max}]. \end{array} \right. \quad (4.1)$$

No caso da equação do calor associada ao problema de valor inicial e de contorno de Black-Scholes, o intervalo  $[a, b]$  é substituído por  $\mathbb{R}$ . No entanto, ao se considerar a solução numérica deste problema, é necessário truncá-lo para valores grandes de  $-a$  e  $b$ . Neste caso, pode-se considerar  $-a = b = 5$ . Assim, assume-se que as condições de contorno no infinito são satisfeitas nestes pontos de truncamento.

**O Esquema FTCS** A equação diferencial parcial em (4.1) pode ser aproximada por diferenças finitas, por exemplo, usando diferenças avançadas para a derivada temporal e diferenças centradas de segunda ordem para a segunda derivada espacial. Ou seja, a solução de (4.1) satisfaz a equação:

$$\frac{u_i^{j+1} - u_i^j}{\Delta \tau} + O(\Delta \tau) = \frac{u_{i+1}^j - 2u_i^j + u_{i-1}^j}{(\Delta x)^2} + O((\Delta x)^2).$$

Portanto, ignorando os termos  $O(\Delta \tau)$  e  $O((\Delta x)^2)$ , tem-se o esquema de diferenças finitas avançado no tempo e centrado no espaço (Forward Time Central Space - FTCS):

$$\frac{u_i^{j+1} - u_i^j}{\Delta \tau} = \frac{u_{i+1}^j - 2u_i^j + u_{i-1}^j}{(\Delta x)^2}.$$

Ao se definir o parâmetro  $\beta = \Delta\tau/(\Delta x)^2$ , este esquema pode ser reescrito como:

$$u_i^{j+1} = \beta u_{i+1}^j + (1 - 2\beta)u_i^j + \beta u_{i-1}^j.$$

Este é um método explícito, pois,  $u_i^{j+1}$  está completamente determinado a partir de  $u_{i-1}^j$ ,  $u_i^j$  e  $u_{i+1}^j$ .

As condições de contorno e inicial são dadas por:

$$u_i^0 = f(x_i), \quad i = 0, 1, \dots, I, \quad u_I^j = g(\tau_j) \text{ e } u_1^j = h(\tau_j), \quad j = 0, 1, \dots, J.$$

Observe que, para o método funcionar bem, isto é, ser estável, é necessário que o parâmetro  $\beta$  esteja no intervalo  $(0, 0.5]$ . Ou seja,  $\Delta\tau$  e  $\Delta y$  devem ser escolhidos de modo que esta condição seja satisfeita.

**Exercício 4.4.** *Construa uma função numérica que calcule a solução de diferenças finitas do esquema FTCS, usando como parâmetros de entrada os coeficientes do problema de valor inicial e de contorno (4.1) e as respectivas condições de contorno. Use, por exemplo,  $f(x) = \max(0, e^x - 1)$ ,  $g(\tau) = 1$  e  $h(\tau) = 0$ .*

**Exercício 4.5.** *Use a função construída no exercício anterior para calcular o preço de uma opção de compra europeia no modelo de Black-Scholes, com os seguintes dados:  $\sigma = 0.4$ ,  $r = 0.03$ ,  $\tau = T - t = 0.1$  (ano),  $K = \$0.70$  e  $S_0 = \$1.00$ . Compare o resultado com o obtido usando a fórmula de Black-Scholes.*

**O Esquema BTCS** Agora, aproxime a equação diferencial em (4.1) por diferenças finitas usando diferenças atrasadas para a derivada temporal e diferenças centradas de segunda ordem para a segunda derivada espacial:

$$\frac{u_i^{j+1} - u_i^j}{\Delta\tau} + O(\Delta\tau) = \frac{u_{i+1}^{j+1} - 2u_i^{j+1} + u_{i-1}^{j+1}}{(\Delta x)^2} + O((\Delta x)^2).$$

Novamente, ignorando os termos  $O(\Delta\tau)$  e  $O((\Delta x)^2)$ , tem-se o esquema de diferenças finitas atrasado no tempo e centrado no espaço (Backward Time Central Space - BTCS):

$$\frac{u_i^{j+1} - u_i^j}{\Delta\tau} = \frac{u_{i+1}^{j+1} - 2u_i^{j+1} + u_{i-1}^{j+1}}{(\Delta x)^2}.$$

Ele pode ser reescrito como:

$$-\beta u_{i-1}^{j+1} + (1 + 2\beta)u_i^{j+1} - \beta u_{i+1}^{j+1} = u_i^j.$$

Este é um método implícito, pois, para se obter  $u_{i-1}^{j+1}$ ,  $u_i^{j+1}$  e  $u_{i+1}^{j+1}$  a partir de  $u_i^j$ , é necessário resolver um sistema linear. Ou seja, é necessário resolver recursivamente os seguinte sistemas lineares:

$$Mu^{j+1} = b^j, \quad j = 0, \dots, J,$$

em que

$$u^{j+1} = (u_1^{j+1}, \dots, u_{I-1}^{j+1})^T, \quad j = -1, 0, 1, \dots, J$$

$$b^j = u^j + \beta(u_0^{j+1}, 0, \dots, 0, u_I^{j+1})^T, \quad j = 0, 1, \dots, J$$

e

$$M = \begin{bmatrix} 1 + 2\beta & -\beta & 0 & \cdots & 0 \\ -\beta & 1 + 2\beta & -\beta & \cdots & 0 \\ 0 & -\beta & \ddots & \ddots & \\ \vdots & & \ddots & \ddots & -\beta \\ 0 & 0 & & -\beta & 1 + 2\beta \end{bmatrix}.$$

Estes sistemas lineares podem ser resolvidos de diversas formas, por exemplo, usando fatoração LU, o método de Gauss-Seidel, o método de sobre-relaxações sucessivas (SOR) e etc. Para mais opções, veja o livro [12].

**Exercício 4.6.** *Construa uma função numérica que calcule a solução de diferenças finitas do esquema BTCS, usando como parâmetros de entrada os coeficientes do problema de valor inicial e de contorno (4.1) e as respectivas condições de contorno. Use, por exemplo,  $f(x) = \max 0, e^x - 1$ ,  $g(\tau) = 1$  e  $h(\tau) = 0$ .*

**Exercício 4.7.** *Use a função construída no exercício anterior para calcular o preço de uma opção de compra europeia no modelo de Black-Scholes, com os seguintes dados:  $\sigma = 0.4$ ,  $r = 0.03$ ,  $\tau = T - t = 0.1$  (ano),  $K = \$0.70$  e  $S_0 = \$1.00$ . Compare o resultado com o obtido usando a fórmula de Black-Scholes.*

**O Método de Crank-Nicolson** O esquema de Crank-Nicolson possui duas vantagens muito importantes em esquemas de diferenças finitas, ele é de segunda ordem no espaço e no tempo e possui uma região de estabilidade maior que o esquema FTCS. Intuitivamente, ele é uma média entre os esquemas FTCS e BTCS, ou seja, a equação diferencial em (4.1) é aproximada por:

$$\begin{aligned} & \frac{u_i^{j+1} - u_i^j}{\Delta\tau} + O((\Delta\tau)^2) \\ &= \frac{1}{2} \left( \frac{u_{i+1}^{j+1} - 2u_i^{j+1} + u_{i-1}^{j+1}}{(\Delta x)^2} + \frac{u_{i+1}^j - 2u_i^j + u_{i-1}^j}{(\Delta x)^2} \right) + O((\Delta x)^2). \end{aligned}$$

Ignorando os termos  $O(\Delta\tau)$  e  $O((\Delta x)^2)$ , tem-se:

$$\frac{u_i^{j+1} - u_i^j}{\Delta\tau} = \frac{1}{2} \left( \frac{u_{i+1}^{j+1} - 2u_i^{j+1} + u_{i-1}^{j+1}}{(\Delta x)^2} + \frac{u_{i+1}^j - 2u_i^j + u_{i-1}^j}{(\Delta x)^2} \right).$$

Ou usando o parâmetro  $\beta$ , tem-se:

$$u_i^{j+1} - \frac{1}{2}\beta(u_{i-1}^{j+1} - 2u_i^{j+1} + u_{i+1}^{j+1}) = u_i^j + \frac{1}{2}\beta(u_{i-1}^j - 2u_i^j + u_{i+1}^j)$$



Definindo-se a variável auxiliar

$$Z_i^j = (1 - \beta)u_i^j + \frac{1}{2}\beta(u_{i-1}^j + u_{i+1}^j),$$

tem-se que

$$(1 + \beta)u_i^{j+1} - \frac{1}{2}\beta(u_{i-1}^{j+1} + u_{i+1}^{j+1}) = Z_i^j$$

Assim, para encontrar a solução de diferenças finitas dada pelo esquema de Crank-Nicolson, basta encontrar o vetor  $Z^j$  e resolver um sistema linear para encontrar  $u^{j+1}$ . Ou seja, deve-se resolver:

$$Mu^{j+1} = b^j, \quad j = 0, 1, \dots, J,$$

em que

$$b^j = Z^j + \frac{1}{2}(u_0^{j+1}, 0, \dots, 0, u_I^{j+1})^T, \quad j = 0, 1, \dots, J,$$

$$Z^j = (Z_1^j, \dots, Z_{I-1}^j)^T, \quad j = 0, 1, \dots, J$$

$$u^{j+1} = (u_1^{j+1}, \dots, u_{I-1}^{j+1}), \quad j = -1, 0, 1, \dots, J.$$

e

$$M = \begin{bmatrix} 1 + \beta & -\frac{1}{2}\beta & 0 & \cdots & 0 \\ -\frac{1}{2}\beta & 1 + \beta & -\frac{1}{2}\beta & \cdots & 0 \\ 0 & -\frac{1}{2}\beta & \ddots & \ddots & \\ \vdots & & \ddots & \ddots & -\frac{1}{2}\beta \\ 0 & 0 & & -\frac{1}{2}\beta & 1 + \beta \end{bmatrix}.$$

Observe que  $Z^j = Nu^j$ , em que

$$N = \begin{bmatrix} 1 - \beta & \frac{1}{2}\beta & 0 & \cdots & 0 \\ \frac{1}{2}\beta & 1 - \beta & \frac{1}{2}\beta & \cdots & 0 \\ 0 & \frac{1}{2}\beta & \ddots & \ddots & \\ \vdots & & \ddots & \ddots & \frac{1}{2}\beta \\ 0 & 0 & & \frac{1}{2}\beta & 1 - \beta \end{bmatrix}.$$

**Exercício 4.8.** *Construa uma função numérica que calcule a solução de diferenças finitas do esquema de Crank-Nicolson, usando como parâmetros de entrada os coeficientes do problema de valor inicial e de contorno (4.1) e as respectivas condições de contorno. Use, por exemplo,  $f(x) = \max 0, e^x - 1$ ,  $g(\tau) = 1$  e  $h(\tau) = 0$ .*

**Exercício 4.9.** *Use a função construída no exercício anterior para calcular o preço de uma opção de compra europeia no modelo de Black-Scholes, com os seguintes dados:  $\sigma = 0.4$ ,  $r = 0.03$ ,  $\tau = T - t = 0.1$ (ano),  $K = \$0.70$  e  $S_0 = \$1.00$ . Compare o resultado com o obtido usando a fórmula de Black-Scholes.*

## Capítulo 5

# Técnicas de Calibração de Parâmetros

### A Regularização de Tikhonov

Considere um fenômeno físico/biológico/financeiro genérico que possa ser descrito por uma função que associa um conjunto de parâmetros em grandezas observáveis, ou seja,

$$F : \mathcal{D}(F) \subset X \longrightarrow Y,$$

em que  $X$  e  $Y$  são espaços vetoriais (normados) e  $\mathcal{D}(F)$  é o domínio de  $F$ .

Assim, dado  $y$  na imagem de  $F$ , queremos encontrar um elemento  $x^\dagger \in \mathcal{D}(F)$  que seja solução da equação

$$F(x) = y. \tag{5.1}$$

A equação (5.1) pode não ser linear e define um *problema inverso*. Quando a função  $F$  é linear e inversível, então  $x^\dagger$  é simplesmente

$x^\dagger = A^{-1}y$ . No entanto, esta não é uma hipótese realista, pois, mesmo que  $X$  e  $Y$  sejam finito-dimensionais e  $A$  uma matriz,  $A$  poderia não ser bem condicionada, isto é, estar muito próxima de ser singular, e assim, pequenas perturbações na medição de  $y$  poderiam levar a soluções completamente diferentes e distantes de  $x^\dagger$ . Além disso, em geral não é possível observar  $y$  exatamente, mas apenas uma versão com ruído  $y^\delta$ , que se relaciona com  $y$ , através da seguinte estimativa:

$$\|y - y^\delta\|_Y \leq \delta. \quad (5.2)$$

Deste modo, não é possível resolver (5.1) diretamente, pois  $y^\delta$  pode nem estar na imagem de  $F$ .

Assim, ao invés de procurar a solução de (5.1), procura-se a solução de um problema aproximado e mais regular, isto é, aplica-se alguma técnica de regularização. Uma das técnicas mais conhecidas é a regularização do tipo Tikhonov. Mais precisamente, substitui-se o problema (5.1) pelo problema de minimização

$$\min\{\|F(x) - y^\delta\|_Y^p + \alpha f(x)\} \text{ sujeito a } x \in \mathcal{D}(F), \quad (5.3)$$

em que  $p > 0$  e  $\alpha > 0$  são constantes apropriadas e  $f : \mathcal{D}(f) \subset X \rightarrow \mathbb{R}_+$  é uma função convexa.

Dependendo das propriedades de  $F$ , talvez seja necessário assumir hipóteses adicionais sobre a função  $f$ , de modo a garantir a existência e a estabilidade de soluções do problema de minimização em (5.3).

O exemplo a seguir apresenta a solução do problema inverso em (5.1) via regularização de Tikhonov para o caso de  $F$  ser linear e o termo de penalização ser a norma de  $X$  ao quadrado. Esse caso é a regularização de Tikhonov clássica.

**Exemplo 5.1.** *Suponha que  $p = 2$ ,  $X$  e  $Y$  sejam espaços de Hilbert e  $F$  seja linear e limitado, i.e.,  $F(x) = Ax$ . Além disso, seja  $f(x) = \|x\|_X^2$ .*

Assim, o problema (5.3) se torna

$$\min\{\|F(x) - y^\delta\|_Y^2 + \alpha\|x\|_X^2\} \text{ sujeito a } x \in X,$$

i.e., um problema de minimização quadrática irrestrita, cujo a solução é dada por

$$x_\alpha^\delta = (A^*A + \alpha Id)^{-1}A^*y^\delta,$$

onde  $Id : Y \rightarrow Y$  é o operador identidade.

Existem ainda diversas outras técnicas de regularização, aqui foi apresentada apenas a de Tikhonov, por simplicidade e por ser uma das técnicas mais usadas em diversas aplicações. O leitor deve observar que, o funcional de Tikhonov em (5.3) é basicamente um problema de mínimos quadrados, quando  $p = 2$  combinado com uma penalização. O termo de penalização é que torna o problema solúvel, no entanto, a escolha do parâmetro  $\alpha$ , chamado de parâmetro de regularização, é fundamental para a construção de soluções confiáveis. Se  $\alpha$  for pequeno demais, então, as soluções podem começar a ficar instáveis, produzindo ruído. Já, se  $\alpha$  for grande demais, começa-se a se introduzir uma tendência que pode comprometer a qualidade da solução. Portanto, uma técnica bastante conhecida de escolha de  $\alpha$  é o princípio da discrepância de Morozov, em que, escolhe-se  $\alpha$ , de modo que

$$\|F(x) - y^\delta\|_Y \approx \tau\delta.$$

De modo mais prático, toma-se o maior  $\alpha$  tal que

$$\|F(x) - y^\delta\|_Y \leq \tau\delta,$$

com  $\tau > 1$  uma constante dada.

Regularização de Problemas Inversos tem sido tema de pesquisa

bastante ativa nas últimas décadas, com aplicações em diversas áreas do conhecimento humano. Uma boa introdução ao assunto pode ser encontrada em [4]. Outras referências introdutórias, que exploram aspectos computacionais destas técnicas são [22, 24]. Para quem quiser saber mais sobre a teoria, recomenda-se, primeiramente o livro [9] e posteriormente os livros [15, 19]. Existem ainda revistas especializadas sobre o tema

Para encontrar aproximações dos minimizadores do funcional de Tikhonov em (5.3), em geral, aplica-se técnicas de otimização não-linear irrestrita, como aquelas inspiradas no método do gradiente. Uma boa referência para estes métodos é o livro [18].

Para ilustrar estas técnicas, denote o funcional de Tikhonov a ser minimizado por:

$$\mathcal{F}(x) = \|F(x) - y^\delta\|^p + \alpha f(x)$$

Assuma que  $X$  e  $Y$  são finito-dimensionais,  $p > 1$  e as funções  $F$  e  $f$  são diferenciáveis, então, pode-se considerar o gradiente de  $\mathcal{F}$ . Os métodos do tipo gradiente são baseados em iterações dadas por

$$x^{k+1} = x^k - \eta_k \nabla \mathcal{F}(x^k),$$

em que  $\eta_k$  deve ser escolhido de maneira apropriada. Veja [18]. Para um chute inicial  $x^0$ , a iteração acima é repetida até que um número máximo de iterações é atingido ou uma tolerância é satisfeita.

## Problemas Inversos em Finanças

Um dos maiores problemas em finanças quantitativas é o de calibração de parâmetros. Em geral, tem-se o problema direto bem definido, isto é, as funções que associam parâmetros aos preços são contínuas. No

entanto, o problema inverso é mal posto, ou seja, as inversas destas funções não existem ou não estão bem definidas. Assim, alguma técnica de regularização deve ser aplicada.

Estes problemas inversos, em geral, podem ser divididos em duas classes principais:

1. Calibração a partir de séries temporais de preços e log-retornos. Neste caso, estes problemas estão relacionados às análises de risco e de alocação de recursos, precisando de técnicas em tempo real. As principais técnicas de calibração são Máxima Verossimilhança e Filtro de Kalman. Os modelos usados nesta classe de problemas são paramétricos e as principais referências são da área de Econometria, veja, por exemplo, [23].
2. Calibração usando dados implícitos, ou seja, dados de preços de opções de compra e venda europeias. Neste caso, os problemas estão relacionados à precificação de derivativos e outros instrumentos financeiros complexos. Os problemas inversos associados necessitam de técnicas de regularização. Veja, por exemplo, [1].

Nestas notas, apenas o segundo caso será tratado.

## Cálculo da Volatilidade Implícita

Considere o modelo de Black-Scholes e a fórmula para o preço de uma opção de compra europeia dada nas equações (3.2)-(3.5).

Ao se substituir os valores do preço do ativo  $S_0$ , do strike  $K$ , da taxa de juros livre de risco  $r$ , do tempo ao vencimento  $\tau = T - t$  e da volatilidade  $\sigma$  pelos valores:

$$S_0 = \$ 1.00, K = \$ 0.7, r = 0.01 \text{ (1\%)}, \tau = 0.5 \text{ (anos)} \text{ e } \sigma = 0.40,$$

chega-se ao valor de, aproximadamente,

$$C(S_0, K, r, \tau, \sigma) = \$0.3147.$$

Agora, se o valor de  $C$ ,  $S_0$ ,  $r$  e  $\tau$  são conhecidos, como encontrar o valor de  $\sigma$ ? Ou seja, para um preço de opção de compra  $\tilde{C}$  dado, qual o valor da volatilidade de Black-Scholes associado? Apesar de sua simplicidade, este é um problema inverso bem conhecido e pode ser facilmente resolvido através de métodos para a solução de equações não lineares, como o método da secante ou o método de Newton. Veja, por exemplo, o livro de análise numérica [7].

Assim, para aplicar, por exemplo, o método da secante, é suficiente construir uma função numérica que calcule o valor da opção de compra Europeia como função da volatilidade, mantendo os outros parâmetros fixados.

**Exercício 5.1.** *Construa uma função na linguagem de sua preferência (R, Python, Matlab, C, C++, etc.) que calcule o preço de uma opção de compra europeia no modelo de Black-Scholes que tenha como parâmetro de entrada o preço do ativo  $S_0$ , o strike  $K$ , a taxa de juros contínua e livre de risco  $r$ , o tempo ao vencimento  $\tau$  e a volatilidade  $\sigma$ .*

**Exercício 5.2.** *Usando a função do exercício anterior, aplique o método da secante para encontrar a volatilidade implícita, isto é, de Black-Scholes, dados os valores da opção de compra Europeia  $C$ , o preço do ativo  $S_0$ , o strike  $K$ , a taxa de juros contínua e livre de risco  $r$  e o tempo ao vencimento  $\tau$ . Teste o seu código com os seguintes dados:*

1.  $S_0 = \$1.00$ ,  $K = \$0.7$ ,  $r = 0.01$  (1%),  $\tau = 0.5$  (anos) e  $C = 0.3147$ .
2.  $S_0 = \$1.00$ ,  $K = \$0.7$ ,  $r = 0.01$  (1%),  $\tau = 0.5$  (anos) e  $C = 0.3268$ .



3.  $S_0 = \$100$ ,  $K = \$120$ ,  $r = 0.03$  (1%),  $\tau = 0.1$  (anos) e  $C = 0.1176$ .

A volatilidade implícita ainda é bastante usada no mercado financeiro. Devido a simplicidade do modelo de Black-Scholes, existem muitas fórmulas fechadas para o cálculo de instrumentos financeiros, como opções e derivativos, que usam esta quantidade como parâmetro de cálculo. Além disso, esta grandeza, por ser “adimensional”, permite uma melhor avaliação do risco associado a determinado ativo que está incorporado aos preços de opções de compra e venda. Este fenômeno é conhecido como “smile”, ou seja, para cada vencimento, toma-se a volatilidade implícita como função dos strikes. De modo bastante simplificado, quanto mais acentuado for o “smile” maior o risco associado ao ativo subjacente. Além disso, a assimetria deste smile também traduz uma maior ou menor impressão de risco de queda ou subida do ativo no período até o vencimento das opções.

Na toolbox de finanças quantitativas do Matlab existem funções que calculam preços de opções de compra e venda europeias e a volatilidade implícita, sob o modelo de Black-Scholes. Se o leitor tiver acesso a uma licença do Matlab que inclua esta toolbox poderá usar as referidas funções para comparar os resultados obtidos nos exercícios acima.

O leitor talvez tenha observado que nesta subseção não foi necessário o uso de técnicas de regularização. Isso se deve ao fato de a função que associa a volatilidade ao preço de opção ser inversível para  $\sigma > 0$ .

### Calibragem da Superfície de Volatilidade Local

Seja  $\tilde{C}$  uma superfície de preços de opções de compra europeias. Isto é, para um tempo  $t$  e um preço de uma ação  $S_t$  fixados, considere os preços de opções para todos os vencimentos e todos os strikes disponíveis.

Assuma que esses preços sejam dados pela equação de Dupire. Então, a superfície de volatilidade local correspondente, denotada por  $\sigma_{TRUE}^2$ , é solução da seguinte equação:

$$\tilde{C} = C(\sigma_{TRUE}^2). \quad (5.4)$$

Observe que  $C(\sigma_{TRUE}^2)$  é a solução da equação de Dupire.

Infelizmente, apenas um conjunto finito e esparsos de dados ruidosos, denotado por  $C^\delta$ , está disponível. Como acima,  $\tilde{C}$  e  $C^\delta$  se relacionam através da inequação

$$\|\tilde{C} - C^\delta\| \leq \delta,$$

com  $\delta > 0$  o nível de ruído.

Como o problema inverso é mal posto, aplicando-se regularização do tipo Tikhonov o problema inverso em (5.4) é substituído pela solução do seguinte problema de minimização irrestrita:

$$\operatorname{argmin} \left\{ \|C(\sigma^2) - C^\delta\|^2 + \alpha f(\sigma^2) : \sigma^2 \in Q \right\},$$

em que  $Q$  é o conjunto de superfícies de volatilidade local apropriadas.

A solução numérica deste problema se dá da seguinte forma:

1. A equação de Dupire é resolvida pelo esquema de Crank-Nicolson.
2. A minimização do funcional de Tikhonov é resolvida por um método do tipo gradiente.
3. As iterações terminam sempre que se atinge a tolerância:

$$\frac{\|C(\sigma_k^2) - C^\delta\|}{\|C^\delta\|} < \text{tol},$$

em que  $\text{tol} = 0.01$ , tipicamente.

Estes passos serão descritos de modo mais detalhado abaixo e têm como referência básica o artigo [2].

**Solução Numérica da Equação de Dupire:** or simplicidade, assumamos que a taxa de juros livre de risco seja nula. Primeiramente, aproximamos o problema de valor inicial e de contorno de Dupire, dado pela equação (3.6) no domínio  $D = [0, \tau_{\max}] \times [-5, 5]$ . As condições de contorno são impostas em  $y = \pm 5$ , respectivamente. Sejam  $J, L \in \mathbb{N}$  fixados e considere a discretização do conjunto  $D$  com a notação  $\tau_i = i\Delta\tau$ ,  $i = 0, 1, 2, \dots, I$ , e  $y_j = j\Delta y$ ,  $j = -J, -J + 1, \dots, 0, 1, \dots, J$ . Denote a solução da equação de Dupire por  $v_j^i := v(\tau_i, y_j)$ , e a volatilidade local por  $a_j^i := a(\tau_i, y_j)$ . Defina ainda os parâmetros  $\beta := \Delta\tau/\Delta y$  e  $\eta = \Delta\tau/\Delta y^2$ . A equação de Dupire em (3.6) pode então ser discretizada pelo esquema de Crank-Nicolson como

$$\begin{aligned} v_j^i - \frac{1}{2}\eta a_j^i(v_{j+1}^i - 2v_j^i + v_{j-1}^i) + \frac{1}{4}\beta a_j^i(v_{j+1}^i - v_{j-1}^i) = \\ v_j^{i-1} + \frac{1}{2}\eta a_j^{i-1}(v_{j+1}^{i-1} - 2v_j^{i-1} + v_{j-1}^{i-1}) - \frac{1}{4}\beta a_j^{i-1}(v_{j+1}^{i-1} - v_{j-1}^{i-1}) \end{aligned} \quad (5.5)$$

O esquema (5.5) pode ser escrito em forma matricial da seguinte forma:

$$[Id + M(a^i)]v^i + b(a^i) = [Id - M(a^{i-1})]v^{i-1} - b(a^{i-1}), \quad i = 1, \dots, I, \quad (5.6)$$

em que

$$\begin{aligned}
 a^i &= (a_{-J+1}^i, \dots, a_0^i, \dots, a_{J-1}^i)^T, \\
 v^i &= (v_{-J+1}^i, \dots, v_0^i, \dots, v_{J-1}^i)^T, \\
 [M(a^i)]_{j,j} &= \eta a_j^i, \quad j = -J+1, \dots, 0, J-1, \\
 [M(a^i)]_{j,j+1} &= \frac{1}{2} \left( \frac{1}{2} \beta - \eta \right) a_j^i, \quad j = -J+1, \dots, 0, J-1, \\
 [M(a^i)]_{j,j-1} &= -\frac{1}{2} \left( \frac{1}{2} \beta + \eta \right) a_j^i \quad j = -J+1, \dots, 0, J-1, \\
 b(a^i) &= \left( -\frac{1}{2} \left( \frac{1}{2} \beta + \eta \right) a_{-J}^i, 0, \dots, 0 \right)^T.
 \end{aligned}$$

Maiores detalhes sobre esta solução numérica podem ser encontrados no artigo [2].

**Exercício 5.3.** *Construa uma função numérica que calcule a solução do problema de valor inicial e de contorno (3.6) usando o esquema de Crank-Nicolson (5.5), tendo como parâmetros de entrada  $S_0$ , o número de pontos espaciais  $2J+1$ , o número de pontos temporais  $I+1$ , a matriz de volatilidades  $a = (a_{ij}, i = 0, \dots, I, j = -J, \dots, -1, 0, 1, \dots, J)$ , e o tempo máximo  $T_{\max}$ .*

**A Minimização do Funcional de Tikhonov:** Resolvendo o problema de minimização por um método do tipo gradiente e denotando a  $k$ -ésima iterada por  $a^k$ , avalia-se:

$$a^{k+1} = a^k - \lambda_k \nabla \mathcal{F}(a^k), \quad (5.7)$$

até que o critério de parada mencionado acima seja satisfeito, para uma tolerância pré-definida. A superfície de volatilidade local  $a$  resultante é tomada como a solução regularizada, caso a tolerância seja satisfeita. No caso contrário, diminui-se o parâmetro de regularização

$\alpha$  e se iniciam novamente as iterações.

Observe que a derivada do funcional de Tikhonov é dada por:

$$\nabla \mathcal{F}(a) = \nabla J^\delta(a) + \alpha \nabla f_{a_0}(a),$$

em que

$$J^\delta(a) := \|v(a) - v^\delta\|_Y^2 = \sum_{i=1}^{I-1} \|v^i(a) - v^{\delta,i}\|^2. \quad (5.8)$$

O passo  $\lambda_k$  no método gradiente em (5.7) é dado por

$$\lambda_k = \min \left( 2.5, \frac{J^\delta(a^k)}{2\|\nabla J^\delta(a^k)\|^2} \right). \quad (5.9)$$

Esta fórmula controla o tamanho do passo, levando em conta que se  $\|\nabla J^\delta(a^k)\|$  é pequeno demais e o método pode se tornar instável.

Para calcular o gradiente do termo  $J^\delta(a)$ , é necessário considerar o estado adjunto da derivada do operador direto com respeito a  $a$ . Para facilitar a notação, a dependência em  $a$  será omitida sempre que for possível. Considere o estado adjunto  $u$ , o qual é solução da seguinte equação adjunta:

$$[Id + M(a^{i+1})^T]u^i = [Id - M(a^{i+1})^T]u^{i+1} + 2 \left( P_l^* P_l \hat{v}^i - P_l^* \hat{v}^{\delta,i} \right),$$

$$i = I - 1, \dots, 0, \quad (5.10)$$

com  $u^I = 0$ .

**Exercício 5.4.** *Construa uma função numérica que calcule a solução do problema de adjunto acima usando o esquema de Crank-Nicolson (5.5), tendo como parâmetros de entrada  $S_0$ , o número de pontos espaciais  $2J+1$ , o número de pontos temporais  $I+1$ , a matriz de volatilidades  $a = (a_{ij}, i = 0, \dots, I, j = -J, \dots, -1, 0, 1, \dots, J)$ , o tempo máximo  $T_{\max}$ , a solução numérica do problema de Dupire (3.6) encontrada no*

exercício anterior e os dados de opções de compra  $v^\delta$ .

Sejam  $a : b$  o produto de dois vetores  $a$  e  $b$  que dão o vetor com entradas  $a_j b_j$ ,  $j = -J + 1, \dots, 0, \dots, J - 1$ , e  $\partial_{y,m}$  e  $\partial_{yy,m}$  as matrizes de diferenças centradas de primeira e segunda ordem, isto é,

$$\partial_{y,m} = \frac{1}{2\Delta y} \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ -1 & 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & -1 & 0 & 1 \\ 0 & \cdots & 0 & -1 & 0 \end{bmatrix}$$

e

$$\partial_{yy,m} = \frac{1}{(\Delta y)^2} \begin{bmatrix} 2 & 1 & 0 & \cdots & 0 \\ -1 & 2 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & -1 & 2 & 1 \\ 0 & \cdots & 0 & -1 & 2 \end{bmatrix}$$

Assim, para cada  $i = 1, \dots, I - 1$ , o gradiente do termo  $J^\delta(a)$  é, aproximadamente, dado por

$$\nabla J^\delta(a)^i \approx -\Delta\tau[\partial_{yy,m} - \partial_{y,m}]v^i : u^i + [\partial_{yy,m} - \partial_{y,m}]v^i : u^{i+1}, \quad (5.11)$$

com  $v$  solução da equação de Dupire e  $u$  solução do estado adjunto.

**Exercício 5.5.** *Construa uma função numérica que calcule o gradiente do termo  $J^\delta(a)$ , usando como parâmetros de entrada as soluções do problema de Dupire (3.6) e do estado adjunto, calculados nos exercícios anteriores.*

O termo de penalização pode definido como:

$$f_{a_0}(a) = \|a - a_0\|^2 + 10\|\partial_{y,n}a\|^2 + \|\partial_{\tau,n}a\|^2,$$

em que,

$$\|a\|^2 = \sum_{i=0}^I \sum_{j=-J}^J a_{ij}^2$$

e as matrizes de diferenças avançadas são dadas por

$$\partial_{y,n} = \frac{1}{\Delta y} M \quad \text{e} \quad \partial_{\tau,n} = \frac{1}{\Delta \tau} M$$

$$M = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -1 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & -1 & 1 \\ 0 & \cdots & 0 & -1 & \end{bmatrix}$$

O gradiente de  $f_{a_0}$  é dado por:

$$\nabla f_{a_0}(a) = 2(a - a_0) + 20\partial_{y,n}^T \partial_{y,n} a + 2\partial_{\tau,n}^T \partial_{\tau,n} a.$$

Agora, considere a seguinte superfície de volatilidade local sintética:

$$\sigma(\tau, y) = \begin{cases} \frac{2}{5} - \frac{4}{25} e^{-\tau/2} \cos\left(\frac{4\pi y}{5}\right), & \text{se } -2/5 \leq y \leq 2/5 \\ 2/5, & \text{caso contrário.} \end{cases} \quad (5.12)$$

Use a função para o cálculo de preços de opções de compra com base no modelo de Dupire para gerar dados sintéticos, considerando os seguintes dados:

1. Use a superfície de volatilidade local (5.12).
2. Use  $S_0 = 1$ ,  $r = 0$ ,  $T_{\max} = 0.5$  e os tamanhos de passo  $\Delta\tau = 0.005$  e  $\Delta y = 0.025$ . Com base nos tamanhos de passo o leitor deve ser capaz de encontrar os valores de  $J$  e  $I$ .

Com base nos dados acima e na solução calculada, considere os preços de opções para os tempos ao vencimento  $\tau_i = i \cdot 0.1$ ,  $i = 1, 2, \dots, 5$ , e strikes  $y_j = j \cdot 0.05$ ,  $j = -10, -9, \dots, 0, 1, \dots, 10$ .

Ao denotar os dados sem ruído por  $v = v(a)$ , pode-se adicionar ruído gaussiano da seguinte forma:

$$v^\delta(\tau, y) = v(\tau, y) (1 + \delta\eta). \quad (5.13)$$

em que  $\eta$  é uma variável aleatória normal padrão. Pode-se usar ainda, por exemplo,  $\delta = 0.01$

Considere ainda os dados,  $a_0 \equiv 0.101$  e  $\alpha = 10^{-4}$ .

Após implementar o método do gradiente para identificar a volatilidade local a partir dos preços gerados acima, inicia-se as iterações com a volatilidade local constante  $a^0 \equiv 0.101$ . Como só existem valores de preços de opções para alguns pontos da malha, considera-se a volatilidade local como incógnita nestes pontos apenas, para os demais pontos usando interpolação linear com respeito a  $y$  e a  $\tau$  e assumindo que

$$a(\tau, y) = \begin{cases} a(\tau, -0.5) & \text{if } \tau > 0.1, y \leq -0.5, \text{ (deep in the money)} \\ a(\tau, 0.5) & \text{if } \tau > 0.1, y \geq 0.5, \text{ (deep out of the money)} \\ a(0.1, y) & \text{if } \tau \leq 0.1. \end{cases}$$

Na figura 5.1 está apresentada uma comparação entre a volatilidade verdadeira, usada para gerar os dados e a volatilidade reconstruída, usando o método do gradiente. Observe que a reconstrução não é perfeita, pois os dados são esparsos e estão com ruído, ou seja, a informação é incompleta e possui ruído, como acontece na prática. Outras funções convexas podem ser usadas como penalização. Veja a introdução de [2] para algumas referências.



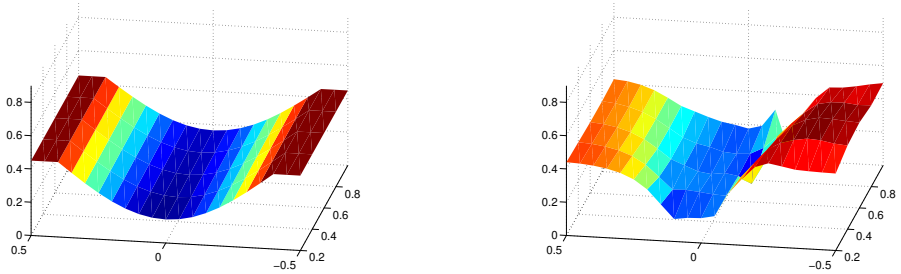


Figura 5.1: Comparação entre a superfície de volatilidade local verdadeira (esquerda) e a volatilidade reconstruída usando o método do gradiente (direita).

**Exercício 5.6.** *Agora, tente refazer os passos acima e construir um código que calibre a superfície de volatilidade local a partir de um conjunto de preços de opções de compra europeias.*



# Bibliografia

- [1] Y. Achdou and O. Pironneau, *Computational Methods for Option Pricing*, Frontiers in Applied Mathematics, SIAM, 2005.
- [2] V. Albani, U. Ascher, and J. Zubelli, *Local Volatility Models in Commodity Markets and Online Calibration*, Journal of Computational Finance **21** (2018), no. 5, 63–95.
- [3] R. G. Bartle, *The elements of integration and Lebesgue measure*, vol. 27, Wiley Online Library, 1995.
- [4] Johann Baumeister and Antonio Leitao, *Topics in inverse problems*, (2005).
- [5] F. Black and M. Scholes, *The Pricing of Options and Corporate Liabilities*, J. Polit. Econ. **81** (1973), no. 3, 637–654.
- [6] L. Breiman, *Probability*, Addison-Wesley, 1968.
- [7] S. Conte and C. De Boor, *Elementary Numerical Analysis: An Algorithmic Approach*, Mcgraw-hill, 1972.
- [8] B. Dupire, *Pricing with a smile*, Risk Magazine **7** (1994), 18–20.
- [9] H. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems*, Mathematics and its Applications, vol. 375, Kluwer Academic Publishers Group, Dordrecht, 1996.

- [10] Lawrence C Evans, *An introduction to stochastic differential equations*, vol. 82, American Mathematical Soc., 2012.
- [11] Jim Gatheral, *The Volatility Surface: A Practitioner's Guide*, Wiley Finance, John Wiley & Sons, 2006.
- [12] G. H. Golub and C. F. Van Loan, *Matrix computations*, third ed., Johns Hopkins University Press, 1996.
- [13] Steven L Heston, *A closed-form solution for options with stochastic volatility with applications to bond and currency options*, The review of financial studies **6** (1993), no. 2, 327–343.
- [14] Desmond Higham, *An Algorithmic Introduction to Numerical Simulation of Stochastic Differential Equations*, SIAM Rev. **43** (2001), no. 3, 525–546.
- [15] B. Kaltenbacher, A. Neubauer, and O. Scherzer, *Iterative Regularization Methods for Nonlinear ill-Posed Problems*, Walter de Gruyter, 2008.
- [16] Ralph Korn and Elke Korn, *Option Price and Portfolio Optimization: Modern Methods of Mathematical Finance*, Graduate Studies in Mathematics, vol. 31, AMS, 2001.
- [17] Nimalin Moodley, *The Heston Model: A Practical Approach with Matlab Code*, Master's thesis, University of the Witwatersrand, 2005.
- [18] J. Nocedal and S. Wright, *Numerical Optimization*, Springer Series in Operations Research and Financial Engineering, Springer, 2006.
- [19] Otmar Scherzer, Markus Grasmair, Harald Grossauer, Markus Haltmeier, and Frank Lenzen, *Variational Methods in Imaging*,

- Applied Mathematical Sciences, vol. 167, Springer, New York, 2008.
- [20] S. Shreve, *Stochastic calculus for finance*, vol. I, Springer, 2004.
- [21] John C Strikwerda, *Finite difference schemes and partial differential equations*, vol. 88, Siam, 2004.
- [22] A. Tarantola, *Inverse problem theory and methods for model parameter estimation*, vol. 89, SIAM, 2005.
- [23] Ruey S Tsay, *Analysis of financial time series*, vol. 543, John wiley & sons, 2005.
- [24] C.R. Vogel, *Computational methods for inverse problems*, vol. 23, SIAM, 2002.
- [25] Paul Wilmott, Sam Howinson, and Jeff Dewynne, *The Mathematics of Financial Derivatives: A Student Introduction*, Cambridge University Press, 1995.