FUNDAÇÃO GETULIO VARGAS
EPGE
Escola de Pós-Graduação em Economia
Ensaios Econômicos

Escola de Pós-Graduação em Economia da Fundação Getulio Vargas

N° 351 ISSN 0104-8910

Introdução à Integração Estocástica (Revisado em Julho de 1999)

Paulo Klinger Monteiro

Agosto de 1999

URL: http://hdl.handle.net/10438/412

Klinger Monteiro, Paulo
Introdução à Integração Estocástica (Revisado em Julho de 1999)/ Paulo Klinger Monteiro - Rio de Janeiro : FGV,EPGE, 2010
(Ensaios Econômicos; 351)

Inclui bibliografia.

CDD-330

Introdução à integração estocástica.

Paulo Klinger Monteiro[1]

Revisão: Julho de 1999

[1] Agradeço os comentários de Flávio Menezes(ANU), Clóvis de Faro(EPGE-FGV), Marco Bonomo(EPGE-FGV) e Renato Flores(EPGE-FGV).

Contents

1 Notação. **3**

2 Espaços de probabilidades. **5**

3 Independência. **14**

4 Processos estocásticos. **15**

5 Martingalas e esperança condicional. **16**

6 Integral estocástica. **19**

7 A fórmula de Black-Scholes **30**

8 A fórmula de Feynman-Kac. **31**

9 Integração estocástica em várias dimensões. **32**

10 A fórmula de Girsanov. **40**

10.1 Princípio do máximo para equações elípticas. 42

11 Exercícios. **46**

12 Tabela de símbolos **53**

Prefácio.

A integração estocástica é a ferramenta básica para o estudo do apreçamento de ativos derivados[1] nos modelos de finanças de tempo contínuo. A fórmula de Black e Scholes é o exemplo mais conhecido. Os movimentos de preços de ações, são frequentemente modelados – tanto teóricamente quanto empíricamente – como seguindo uma equação diferencial estocástica. O livro texto de D. Duffie, "Dynamic asset pricing theory", usa livremente conceitos como o teorema de Girsanov e a fórmula de Feynman-Kac.

Um conhecimento básico da integração estocástica é cada vez mais necessário para quem quer acompanhar a literatura moderna em finanças.

Esta introdução à integração estocástica é dirigida para alunos de doutourado e no final de mestrado. Um conhecimento sólido[2] de continuidade, limites e facilidade de operar com a notação de conjuntos é fundamental para a compreensão do texto que se segue. Um conhecimento básico de integral de Lebesgue é recomendável. No entanto incluí no texto as definições básicas e os resultados fundamentais da teoria da integral de Lebesgue usados no texto.

[1] Opções por exemplo.

[2] No nível do livro de R. Bartle, "The elements of real analysis", John Wiley & Sons.

1 Notação.

A notação usada é na sua maior parte padrão.

Números.

Denoto por $\mathbb{R}$ o conjunto dos números reais, por $\mathbb{Q}$ o conjunto dos números racionais e $\mathbb{N}$ o conjunto dos números naturais $\{1, 2, \ldots\}$. Frequentemente precisamos considerar o conjunto dos números reais acrescido dos pontos ∞ e $-\infty$. Definimos $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$. E $[-\infty, \infty] = \bar{\mathbb{R}}$. As definições de soma, multiplicação e de ordem são extendidas a $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ de modo natural:

1. $x + \infty = \infty$ se $x \in \mathbb{R} \cup \{\infty\}$;
2. $x - \infty = x + (-\infty) = -\infty$ se $x \in \mathbb{R} \cup \{-\infty\}$;
3. A diferença $\infty - \infty$ não é definida;
4. $x \cdot \infty = \infty$ se $x \in (0, \infty]$ e $x \cdot \infty = -\infty$ se $x \in (-\infty, 0)$;
5. $-\infty < x < \infty$ para todo $x \in \mathbb{R}$.

O módulo de um número real x é denotado $|x|$. A parte positiva de x é denotada $x^+ = \max\{x, 0\}$ e a parte negativa $x^- = \max\{-x, 0\}$. É claro que $|x| = x^+ + x^-$.

Funções e sequências.

Se f é uma função entre os conjuntos A e B escrevemos $f : A \to B$. Se A e B são conjuntos de números reais dizemos que a função f é estritamente crescente se $a < b$ implicar $f(a) < f(b)$. E f é crescente (ou às vezes dita não-decrescente) se $a < b$ implicar $f(a) \leq f(b)$. Uma sequência de números reais $x : \mathbb{N} \to \mathbb{R}$ é denotada por $(x_n)_n$. Se a sequência tiver limite este será denotado por $\lim_{n\to\infty} x_n$. Se $(x_n)_n$ é uma sequência e $k : \mathbb{N} \to \mathbb{N}$ é uma função estritamente crescente podemos formar a nova sequência $(x_{k(n)})_n$ cujo n-ésimo termo é $x_{k(n)}$. Se uma sequência tem limite toda subsequência tem o mesmo limite. Se $(x_n)_n$ é uma sequência sempre existe $\bar{x} = \limsup_{n\to\infty} x_n \in \bar{\mathbb{R}}$. O limite superior de $(x_n)_n$ é caracterizado pelas propriedades:

1. Existe subsequência $(x_{k(n)})_n$ tal que $\lim_{n\to\infty} x_{k(n)} = \bar{x}$;
2. Se a subsequência $(x_{k(n)})_n$ converge então $\lim_{n\to\infty} x_{k(n)} \leq \bar{x}$.

De maneira análoga definimos $\liminf_{n\to\infty} x_n$, o menor limite de uma subsequência de $(x_n)_n$. Uma sequência é convergente se e somente se $\liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n \in \mathbb{R}$.

Conjuntos.

Se A e B são conjuntos denoto por:

1. $A \cup B$ a união de A e B;
2. $A \setminus B$ a diferença de A por B.
3. Se $A \subset \Omega$, $A^c = \Omega \setminus A$ denota o complementar de A em relação à Ω.
4. $A \cap B$ a intersecção de A e B.
5. $A \times B$ o produto cartesiano de A e B
6. $\mathcal{P}(\Omega)$ o conjunto dos subconjuntos de Ω.
7. $A \Delta B$ a diferença simétrica dos conjuntos A e B.

Se $A_n, n \in \mathbb{N}$ é uma sequência de conjuntos denoto por:

1. $A_n \uparrow A$ se $A_n \subset A_{n+1}$ para todo n e $A = \cup_{n=1}^{\infty} A_n$.
2. $A_n \downarrow A$ se $A_{n+1} \subset A_n$ para todo n e $A = \cap_{n=1}^{\infty} A_n$.
3. $\limsup A_n = \cap_{k=1}^{\infty} \cup_{p=k}^{\infty} A_p$ e $\liminf A_n = \cup_{k=1}^{\infty} \cap_{p=k}^{\infty} A_p$.

Se I é um conjunto não vazio de indices e $\mathcal{X} = \{X_i; i \in I\}$ é uma família de conjuntos, a união da família $\mathcal{X}$ é denotada $\cup_{i \in I} X_i = \{x; \text{existe } i \in I, x \in X_i\}$. A intersecção é denotada $\cap_{i \in I} X_i = \{x; x \in X_i, \text{para todo } i \in I\}$. O produto cartesiano de $\mathcal{X}$ é denotado por

$$\Pi_{i \in I} X_i = \{f : I \to \cup_{i \in I} X_i; f(i) \in X_i \text{ para todo } i \in I\}.$$

Se $A \subset \Omega$ é um conjunto definimos a função indicadora de A, $\chi_A : \Omega \to \{0, 1\}$ por

$$\chi_A(\omega) = \begin{cases} 1 & \text{se} \quad \omega \in A, \\ 0 & \text{se} \quad \omega \notin A. \end{cases}$$

Classes de diferenciabilidade.

Suponhamos $U \subset \mathbb{R}^n$ aberto e $f = (f_1, \ldots, f_m) : U \to \mathbb{R}^m$ diferenciável. Dizemos que f é de classe C^1 se a derivada f' for uma função contínua. A função f é C^1 se e somente existem e são contínuas as derivadas parciais $\frac{\partial f_i}{\partial x_j} : U \to \mathbb{R}$, $1 \leq i \leq m$, $1 \leq j \leq n$. O operador Laplaciano de uma função $f : \mathbb{R}^n \to \mathbb{R}$ é denotado $\Delta f := \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i^2}$.

Norma no $\mathbb{R}^n$ e em espaços mais gerais.

Seja V um espaço vetorial real. Uma norma em V é uma função $|\cdot| : V \to \mathbb{R}$ com as seguintes propriedades:

1. $|v| = 0$ implica $v = 0$;
2. $|v + w| \leq |v| + |w|$ para todo $v, w \in V$.(desigualdade triangular);
3. $|rv| = |r||v|$ se $r \in \mathbb{R}$ e $v \in V$.

Um espaço vetorial munido de uma norma é chamado de espaço normado.

Exemplo 1 *No $\mathbb{R}^n$ podemos definir variás normas. A norma da soma é definida por $|x|_S = |x_1|+\ldots+|x_n|$ se $x \in \mathbb{R}^n$. A norma do máximo é definida por $|x|_M = \max\{|x_1|, \ldots, |x_n|\}$. E finalmente a norma mais utilizada no $\mathbb{R}^n$ é a norma euclidiana: $|x| = \sqrt{\sum_{l=1}^n x_l^2}$. Se $\mathcal{M}$ é o espaço vetorial das matrizes $m \times n$ definimos a norma da matriz $A = (a_{ij})_{i\leq m, j\leq n}$ por $|A| = \sqrt{\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2}$.*

Mais tarde veremos outros exemplos de normas como a norma do espaço L^p.

Espaço de Banach.

Seja $(V, |\cdot|)$ um espaço normado.

Definição 1 *1. Uma sequência $(x_n)_n$ em V converge para $x \in V$ se para todo $\epsilon > 0$ existe $N \in \mathbb{N}$ tal que para todo índice $n \geq N$ vale que $|x_n - x| < \epsilon$.*

2. Uma sequência $(x_n)_n$ em V é de Cauchy se para todo $\epsilon > 0$ existe um inteiro N tal que se $m \geq N$ e $n \geq N$ então $|x_n - x_m| < \epsilon$.

Se uma sequência converge o seu limite é único. Toda sequência convergente é de Cauchy. A recíproca entretanto não é verdadeira.

Definição 2 (Espaço de Banach) *Um espaço normado V é um espaço de Banach se toda sequência de Cauchy em V for convergente.*

Os exemplos mais importantes de espaços de Banach são os espaços L^p.

2 Espaços de probabilidades.

Nesta seção apresento de maneira condensada as definições e resultados da teoria da integral de Lebesgue, necessários para o entendimento da integração estocástica. O livro do Robert G. Bartle, "The elements of integration" é uma ótima referência para quem nunca estudou teoria da medida. Os resultados e as definições nesta seção foram sempre que possível particularizados para espaços de probabilidade.

Definição 3 *Se Ω é um conjunto, uma álgebra em Ω é uma família $\mathcal{A} \subset \mathcal{P}(\Omega)$ que satisfaz às seguintes propriedades:*

1. $\Omega \in \mathcal{A}$;
2. $A \in \mathcal{A}$ *implica* $A^c \in \mathcal{A}$;
3. *Se* $\{A_1, \ldots, A_N\} \subset \mathcal{A}$ *então* $\cup_{n=1}^{N} A_n \in \mathcal{A}$.

Se além disso, $\mathcal{A}$ for fechada para uniões enumeráveis de elementos de $\mathcal{A}$ dizemos que $\mathcal{A}$ é uma σ-álgebra.

Observação 1 *Se A é uma álgebra então é imediato de se verificar que se $A, B \in \mathcal{A}$ então $A \cap B, A \setminus B = A \cap B^c \in \mathcal{A}$. Uma σ-álgebra é fechada por intersecção enumeável.*

Um par ordenado $(\Omega, \mathcal{A})$ formado pelo conjunto Ω e a σ-álgebra $\mathcal{A}$ é chamado de espaço mensurável. Os elementos de $\mathcal{A}$ são os conjuntos mensuráveis ou eventos. O conjunto Ω é também chamado de espaço amostral.

Exemplo 2 *O leitor fácilmente verificará os exemplos a seguir de espaços mensuráveis.*

1. $(\Omega, \mathcal{P}(\Omega))$;
2. $(X, \mathcal{X})$ *sendo X um conjunto não enumerável e $\mathcal{X}$ a σ-álgebra formada pelos subconjuntos enumeráveis de X e os seus complementos.*
3. $(\Omega, \cap_{t \in T} \mathcal{A}_t)$ *sendo $\mathcal{A}_t$ uma σ-álgebra para todo $t \in T$. Ou seja a intersecção de σ-álgebra é uma σ-álgebra.*

Proposição 1 (A σ-álgebra gerada por um conjunto.) *Dado um subconjunto $\mathbf{U} \subset \mathcal{P}(\Omega)$ existe $\sigma(\mathbf{U})$, a menor σ-álgebra de Ω que contém $\mathbf{U}$.*

Demonstração: Seja $\Gamma = \{\mathcal{C}; \mathcal{C} \supset \mathbf{U}$ é uma σ-álgebra de $\Omega\}$. Então $\Gamma \neq \emptyset$ pois $\mathcal{P}(\Omega) \in \Gamma$. Portanto $\sigma(\mathbf{U}) := \cap_{\mathcal{C} \in \Gamma} \mathcal{C}$ é uma σ-álgebra por ser intersecção de σ-álgebras. E é a menor σ-álgebra que contém $\mathbf{U}$. QED

Definição 4 *A σ-álgebra de Borel no $\mathbb{R}^n$, $\mathcal{B}_n$, é a menor σ-álgebra do $\mathbb{R}^n$ que contem $\Pi_{i=1}^{n}(a_i, b_i)$, sendo $(a_i, b_i) \subset \mathbb{R}$ um intervalo aberto.*

Definição 5 *Uma função $f : \Omega \to \bar{\mathbb{R}}$ é mensurável se :*

1. $f^{-1}(\infty) \in \mathcal{A}$ *e* $f^{-1}(-\infty) \in \mathcal{A}$;
2. $f^{-1}(B) \in \mathcal{A}$ *para todo* $B \in \mathcal{B} = \mathcal{B}_1$.

Se $f : \Omega \to \mathbb{R}$ é mensurável dizemos que f é uma variável aleatória. Para verificar que uma função é mensurável basta verificar para conjuntos B num conjunto que gere a σ-álgebra dos boreleanos. Portanto para verificar que uma função real f é mensurável basta verificar que $f^{-1}(B) \in \mathcal{A}$ para $B = (-\infty, a)$, $a \in \mathbb{Q}$. Mais geralmente $f : \Omega \to \bar{\mathbb{R}}$ é mensurável se $f^{-1}([-\infty, z)) \in \mathcal{A}$ para todo z racional.

Proposição 2 *A soma de funções mensuráveis é mensurável. O produto de funções mensuráveis é mensurável. Se f é mensurável então $|f|$ é mensurável.*

Em geral toda função contínua é mensurável. Toda função monótona é mensurável. A função indicadora, χ_E é mensurável se e somente se $E \in \mathcal{A}$.

Proposição 3 *Se $(f_n)_n$ é uma sequência de funções mensuráveis então:*

1. $\inf_n f_n$ *e* $\sup_n f_n$ *são mensuráveis.*
2. $\liminf_n f_n$ *e* $\limsup_n f_n$ *são mensuráveis.*
3. *Se existe* $\lim_{n\to\infty} f_n(\omega)$ *para todo* $\omega \in \Omega$ *então* $f(\omega) = \lim_{n\to\infty} f_n(\omega)$ *é mensurável.*

Definição 6 *Uma função com valores reais é simples se sua imagem é um conjunto finito. Se $f : \Omega \to \mathbb{R}$ é simples e mensurável então f pode ser escrita na forma: $f = \sum_{j=1}^n a_j \chi_{E_j}$, $E_j \in \mathcal{A}$.*

Na representação de uma função simples podemos supor que os a_j são distintos e que os conjuntos E_j são dois a dois disjuntos. Nesse caso a representação de f é única.

O lema a seguir é fundamental para a definição da integral de Lebesgue.

Lema 1 *Seja $f : \Omega \to \mathbb{R}_+$ mensurável. Existe então uma sequência g_n de variáveis aleatórias tais que:*

1. $0 \le g_n(\omega) \le g_{n+1}(\omega)$ *para todo* $\omega \in \Omega$ *e* $n \ge 1$*;*
2. $f(\omega) = \lim_n g_n(\omega)$ *para todo* $\omega \in \Omega$*;*
3. g_n *é uma função simples.*

Demonstração: Para $n \in \mathbb{N}$ e $1 \le k \le n2^n$ seja

$$E_{kn} = \left\{\omega \in \Omega; \frac{k}{2^n} \le f(\omega) < \frac{k+1}{2^n}\right\},$$

para $k < n2^n$ e $E_{(n2^n)n} = \{\omega \in \Omega; f(\omega) \ge n\}$. Os conjuntos E_{kn} são dois a dois disjuntos, $\cup_k E_{kn} = \Omega$. Definamos

$$g_n(\omega) = \sum_{k=1}^{n2^n} \frac{k}{2^n} \chi_{E_{kn}}(\omega).$$

É imediato de se verificar que g_n satisfaz as propriedades desejadas.

Definição 7 (Probabilidade.) *Uma (medida de)probabilidade é uma função* $P : \mathcal{A} \to [0,1]$ *tal que*

1. $P(\emptyset) = 0$;
2. *P é contavelmente aditiva: Se $\{E_n; n \in \mathbb{N}\}$ é uma família de eventos, dois a dois disjuntos então $P(\cup_n E_n) = \sum_{n=1}^{\infty} P(E_n)$.*

Proposição 4 *Se P é uma medida de probabilidade então:*

1. $P(\Omega) = 1$;
2. $P(A^c) = 1 - P(A)$ *para todo* $A \in \mathcal{A}$;
3. *Se* $A \subset B$ *então* $P(A) \leq P(B)$;
4. *Se* $A_n \uparrow A$ *então* $P(A) = \lim_n P(A_n) = \sup_n P(A_n)$.
5. *Se* $A_n \downarrow A$ *então* $P(A) = \lim_n P(A_n) = \inf_n P(A_n)$.
6. $P(\liminf A_n) \leq \liminf P(A_n) \leq \limsup P(A_n) \leq P(\limsup A_n)$.
7. $P(\cup_n A_n) \leq \sum_{n=1}^{\infty} P(A_n)$.

Exemplo 3
1. *Suponhamos $(\Omega, \mathcal{A}) = ([0,1], \mathcal{B}[0,1])$. Existe uma única probabilidade $\mu : \mathcal{B} \to [0,1]$ tal que $\mu((a,b)) = b - a$ para todo intervalo $(a,b) \subset [0,1]$.*
2. *Suponhamos que $f : [0,1] \to \mathbb{R}$ é contínua e crescente tal que $f(0) = 0$ e $f(1) = 1$. Existe então uma única probabilidade λ_f definida nos boreleanos tal que $\lambda_f((a,b)) = f(b) - f(a)$ para todo $a < b$. A probabilidade λ_f é chamada de medida de Borel-Stieltjes gerada pela f.*

Um espaço de probabilidade é um terno $(\Omega, \mathcal{A}, P)$, sendo $\mathcal{A}$ uma σ-álgebra e P uma probabilidade em Ω. Dizemos que uma propriedade vale quase certamente em Ω se o conjunto N de pontos de Ω para os quais ela não é valida está contido num subconjunto $M \in \mathcal{A}$ tal que $P(M) = 0$. Assim por exemplo dizemos que duas funções mensuráveis $f, g : \Omega \to \mathbb{R}$ são iguais quase certamente se $P(\{\omega \in \Omega; f(\omega) \neq g(\omega)\}) = 0$. Se $(\Omega, \mathcal{A}, \mu)$ é um espaço de probabilidade, podemos extender a σ-álgebra $\mathcal{A}$ de um modo natural:

Definição 8 (Completamento de um espaço de probabilidade.) *Definamos*

$$\tilde{\mathcal{A}} = \{B \subset \Omega;\ \textit{existe } C \in \mathcal{A} \textit{ tal que } B \Delta C \subset N \in \mathcal{A},\ P(N) = 0\}.$$

E definamos a probabilidade $\tilde{P} : \tilde{\mathcal{A}} \to [0,1]$ por $\tilde{P}(B) = P(C)$ se $C \in A$ e $B \Delta C \subset N$, $P(N) = 0$.

É fácil de se ver que o completamento de $\tilde{\mathcal{A}}$ é $\tilde{\mathcal{A}}$.

Definição 9 (Integral de uma função simples.) *Seja $f : \Omega \to \mathbb{R}$ uma função simples. Então a integral de f com relação à probabilidade μ é definida por:*

$$\int f d\mu = \sum_{j=1}^{n} a_j \mu(E_j).$$

Definição 10 (A integral de Lebesgue) *Se $f : \Omega \to \mathbb{R}_+$ é mensurável definimos a integral de f por*

$$\int f d\mu = \int_\Omega f(\omega) d\mu(\omega) = \sup \left\{ \int \phi d\mu; 0 \leq \phi \leq f, \phi \text{ função simples} \right\}.$$

Se $A \in \mathcal{A}$ define-se a integral em A por $\int_A f d\mu = \int \chi_A f d\mu$.

Teorema 1 (Teorema da convergência monótona.) *Se $(f_n)_n$ é uma sequência monótona crescente de funções mensuráveis $f_n : \Omega \to \mathbb{R}_+$ que converge para f então*

$$\int f d\mu = \lim_{n \to \infty} \int f_n d\mu.$$

Teorema 2 (Lema de Fatou.) *Seja $(f_n)_n$ uma sequência de funções mensuráveis não-negativas. Então*

$$\int \liminf_{n \to \infty} f_n d\mu \leq \liminf_{n \to \infty} \int f_n d\mu.$$

Demonstração: Seja $g_m = \inf\{f_p; p \geq m\}$. Portanto (g_m) é uma sequência crescente de funções mensuráveis. Como

$$\int g_m d\mu \leq \int f_n d\mu \text{ se } m \leq n,$$

temos que

$$\int g_m d\mu \leq \liminf_{n \to \infty} \int f_n d\mu.$$

Portanto o teorema da convergência monótona implica que

$$\int \liminf_{n \to \infty} f_n d\mu = \lim_{m \to \infty} \int g_m d\mu \leq \liminf_{n \to \infty} \int f_n d\mu.$$

QED

Uma função mensurável f, é integrável se $\int |f| d\mu < \infty$. É claro então que a parte positiva de f, $f^+ = \max\{f, 0\}$ e a parte negativa de f, $f^- = \max\{-f, 0\}$ são integráveis. Se f é integrável definimos a integral de f por

$$\int f d\mu = \int f^+ d\mu - \int f^- d\mu.$$

Proposição 5 *Se f, g são integráveis e $r \in \mathbb{R}$ vale que:*

1.
$$\int (rf + g)\, d\mu = r \int f d\mu + \int g d\mu.$$

2. *Se $f \geq 0$ então $\int f d\mu \geq 0$.*

3.
$$\left| \int f d\mu \right| \leq \int |f| d\mu.$$

Teorema 3 (Teorema da convergência dominada.) *Seja $(f_n)_n$ uma sequência de funções mensuráveis que converge quase certamente para uma função f. Se existe g integrável tal que $|f_n| \leq g$ para todo n então f é integrável e*

$$\lim_{n\to\infty} \int f_n d\mu = \int f d\mu.$$

Se $\mathcal{F}$ é uma família de variáveis aleatórias existe $\tilde{\mathcal{A}} = \sigma(\mathcal{F})$, a menor σ-álgebra que torna toda função $f \in \mathcal{F}$ mensurável. Com efeito, a família

$$\Gamma = \{\mathcal{C}; \mathcal{C} \text{ é } \sigma\text{-álgebra de } \Omega \text{ e toda } f \in \mathcal{F} \text{ é } \mathcal{C} \text{ mensurável}\},$$

é não vazia, pois $\mathcal{A} \in \Gamma$. E portanto $\tilde{\mathcal{A}} = \cap_{\mathcal{C}\in\Gamma}\mathcal{C} \subset \mathcal{A}$ é uma σ-álgebra, pois é uma intersecção de σ-álgebras.

Espaços L^p.

Inicialmente definimos o espaço $L = \{f : \Omega \to \mathbb{R}; f \text{ é } \mu \text{ integrável}\}$. Dizemos que $f, g \in L$ são equivalentes se $\mu(\{\omega \in \Omega; f(\omega) \neq g(\omega)\}) = 0$. A classe de equivalência de $f \in L$ é denotada $[f]$. Definimos então $L^1(\Omega, \mathcal{A}, \mu) = \{[f]; f \in L\}$. Se definirmos $|[f]|_1 = \int |f| d\mu$ obtemos um norma e portanto L^1 é um espaço normado. Mais do que isso vale o

Teorema 4 *O espaço $L^1(\Omega, \mathcal{A}, \mu)$ é um espaço de Banach.*

Uma generalização muito importante dos espaço L^1 são os espaços L^p, $1 \leq p \leq \infty$. Suponhamos inicialmente $1 \leq p < \infty$. O espaço $L^p = L^p(\Omega, \mathcal{A}, \mu)$ é o conjunto das classes de equivalência $[f]$ sendo que $\int |f|^p d\mu < \infty$. A função $|f|_p = \left(\int |f|^p d\mu\right)^{\frac{1}{p}}$ é uma norma no espaço L^p que é de Banach nessa norma.

Teorema 5 (Desigualdade de Hölder.) *Suponhamos que $f \in L^p$ e $g \in L^q$ sendo que $1 < p$, $1/p + 1/q = 1$. Então $fg \in L^1$ e $|fg|_1 \leq |f|_p |g|_q$.*

No caso $p = 2 = q$ a desigualdade de Hölder se reduz à desigualdade de Cauchy-Schwarz.

Teorema 6 (Desigualdade de Minkowski.) *Se f e g são mensurável então* $|f+g|_p \leq |f|_p + |g|_p$.

Tambem se usa muito o espaço L^∞ mas a definição da norma é um pouco diferente. Dizemos que uma função mensurável f, é essencialmente limitada se existe $M > 0$ tal que $\mu\{\omega \in \Omega; |f(\omega)| > M\} = 0$. O supremo essencial de f é então definido por

$$|f|_\infty = \inf\{M > 0; \mu\{\omega \in \Omega; |f(\omega)| > M\} = 0\}.$$

A função $|\cdot|_\infty$ é uma norma e o espaço L^∞ é o conjunto das classes de equivalências das funções essencialmente limitadas munido da norma $|\cdot|_\infty$. Este espaço é de Banach.

Tipos de convergência.

Os conceitos de convergência nos espaços de probabilidade são além dos usuais de convergência uniforme e de convergência pontual, os conceitos de convergência em probabilidade, quase certa e em L^p. Sejam $f_n : \Omega \to \mathbb{R}$, $n \geq 1$ e $f : \Omega \to \mathbb{R}$ funções mensuráveis.

a. A sequência $(f_n)_n$ converge uniformemente para f se para todo $\epsilon > 0$ existe um inteiro n_0 tal que para todo inteiro $n > n_0$ temos $|f_n(\omega) - f(\omega)| < \epsilon$ para todo $\omega \in \Omega$.

b. A sequência $(f_n)_n$ converge pontualmente para f se para todo $\omega \in \Omega$, $\lim_{n\to\infty} f_n(\omega) = f(\omega)$.

c. A sequência $(f_n)_n$ converge quase certamente para f se convergir pontualmente para f exceto num conjunto mensurável de probabilidade nula.

d. f_n converge para f em probabilidade se para todo $\epsilon > 0$,

$$\lim_{n\to\infty} \mu\left(\{\omega \in \Omega; |f_n(\omega) - f(\omega)| > \epsilon\}\right) = 0.$$

O limite em probabilidade, se houver, é único quase certamente. Mais precisamente se f e g são limites em probabilidade de f_n então $f = g$ quase certamente. Num espaço de probabilidade a convergência quase certa implica a convergência em probabilidade. A recíproca é parcialmente verdade: a convergência em probabilidade implica convergência em probabilidade numa subsequência.

Proposição 6 *1. Se $X_n \xrightarrow{P} X$ então existe uma subsequência $(X_{k_n})_n$ tal que $\lim_{n\to\infty} X_{k_n} = X$ quase certamente;*

2. Se $\lim_{n\to\infty} X_n = X$ quase certamente, então $X_n \xrightarrow{P} X$.

Convergência em L^p.

Uma sequência $([f_n])_n$ em L^p converge para $[f]$ se convergir na norma de L^p, isto é, se $\lim_{n\to\infty} |f_n - f|_p = 0$. As vezes dizemos que $[f_n]$ converge para $[f]$ na média de ordem p. A partir de agora serei menos cuidadoso na representação de elementos de L^p escrevendo f_n quando devería escrever $[f_n]$. Esta impresição é muito conveniente e não causa nenhuma dificuldade sendo universalmente utilizada.

Teorema 7 *Se (f_n) é uma sequência em L^p que converge uniformemente para f então $f \in L^p$ e (f_n) converge para f em L^p.*

Teorema 8 *Suponhamos que $f_n \in L^p$ e $f_n \to f$ quase certamente. Se existe $g \in L^p$ tal que para todo n, $|f_n| \leq g$ quase certamente então $f \in L^p$ e $f_n \to f$ em L^p.*

Demonstração: Como f_n converge quase certamente para f temos que $|f| \leq g$ quase certamente também. Logo $f \in L^p$. Portanto para todo n temos que $|f_n - f|^p \leq 2^p g^p$, quase certamente. Logo pelo teorema da convergência dominada de Lebesgue, $\lim_{n\to\infty} \int |f_n - f|^p d\mu = \int \lim_{n\to\infty} |f_n - f|^p d\mu = 0$. QED

Definição 11 (Sequência de Cauchy em probabilidade) *A sequência (f_n) é de Cauchy em probabilidade se*

$$\lim_{m,n\to\infty} \mu\left(\omega \in \Omega; |f_n(\omega) - f_m(\omega)| \geq \epsilon\right) = 0$$

para todo $\epsilon > 0$.

Teorema 9 *Seja (f_n) uma sequência Cauchy em probabilidade. Então existe uma subsequência $(f_{k_n})_n$ que converge quase certamente e em probabilidade para uma certa função mensurável f.*

Corolário 1 *Toda sequência de Cauchy em probabilidade, é convergente em probabilidade. O limite é quase certamente único.*

Teorema 10 *Seja (f_n) uma sequência de funções em L^p que converge em probabilidade para f e seja $g \in L^p$ tal que $|f_n| \leq g$, quase certamente. Então $f \in L^p$ e (f_n) converge para f em L^p.*

Os resultados mais sofisticados de integração estocástica necessitam de teoremas mais gerais do que o teorema da convergência dominada de Lebesgue. O teorema de Vitali é a base destes resultados.

Teorema 11 (Teorema de convergência de Vitali.) *Suponhamos $1 \leq p < \infty$. E seja a sequência (f_n) em L^p. Então $f_n \to f$ em L^p se e somente se valem (i) e (ii) abaixo.*

(i) *(f_n) converge para f em probabilidade;*

(ii) *Para todo $\epsilon > 0$ existe $\delta(\epsilon) > 0$ tal que se $B \in \mathcal{A}$ e $\mu(B) < \delta(\epsilon)$ então*

$$\int_B |f_n|^p d\mu < \epsilon^p, \forall n \in \mathbb{N}.$$

Igualdade de probabilidades.

Em geral é bem dificil de se verificar que duas probabilidades coincidem. O teorema de extensão de Hahn mostra que basta verificar que as probabilidades coincidem numa álgebra.

Teorema 12 (Teorema de extensão de Hahn.) *Suponhamos que μ é uma medida de probabilidade numa álgebra $\hat{\mathcal{A}}$. Então μ possui uma única extensão à $\sigma(\hat{\mathcal{A}})$. Em outras palavras, se duas medidas de probabilidade coincidem numa álgebra, elas coincidem na σ-álgebra gerada pela álgebra.*

Portanto por exemplo existe uma única medida em $[0,1]$ tal que $\mu([c,d]) = d-c$ para todo $0 \leq c \leq d \leq 1$.

Medida produto e os teoremas de Fubbini e Tonelli.

Definição 12 *Seja $(A, \mathcal{A})$ e $(B, \mathcal{B})$ dois espaços mensuráveis. Os subconjuntos de $A \times B$ da forma $X \times Y$, $X \in \mathcal{A}$, $Y \in \mathcal{B}$ são chamados de retângulos (ou retângulos mensuráveis).*

Seja $C = A \times B$ e C_0 o conjunto das uniões finitas de retângulos. É fácil de se checar que toda união finita de retângulos pode ser escrita como uma união finita disjunta de retângulos mensuráveis.

Lema 2 *A coleção C_0 é uma álgebra.*

Definição 13 *A σ-álgebra produto de $(A, \mathcal{A})$ e $(B, \mathcal{B})$, denotada $\mathcal{A} \otimes \mathcal{B}$ é a menor σ-álgebra de $A \times B$ que contem C_0, isto é, é a σ-álgebra gerada pelos retângulos.*

Teorema 13 *Suponhamos que $(A, \mathcal{A}, P)$ e $(B, \mathcal{B}, Q)$ são espaços de probabilidade. Existe então uma única medida de probabilidade π em $(A \times B, \mathcal{A} \otimes \mathcal{B})$ tal que $\pi(X \times Y) = P(X)Q(Y)$ para todo retângulo $X \times Y$.*

Definição 14 *Se $C \subset A \times B$ e $a \in A$ então a seção em a de C é o conjunto*

$$C_a = \{b \in B; (a,b) \in C\}.$$

Similarmente definimos a seção em b de C:

$$C_b = \{a \in A; (a,b) \in C\}.$$

Se $f : A \times B \to \mathbb{R}$ e $a \in A$ definimos a seção em a de f por $f_a(b) = f(a,b)$ e a seção em $b \in B$ de f por $f_b(a) = f(a,b)$.

Lema 3 **(a)** *Se $C \in \mathcal{A} \times \mathcal{B}$ então cada seção de C é mensurável.*

(b) *Se f é mensurável de $A \times B$ em $\mathbb{R}$, então cada seção de f é mensurável.*

Lema 4 *Seja $(A, \mathcal{A}, P)$ e $(B, \mathcal{B}, Q)$ espaços de probabilidade. Se $E \in \mathcal{A} \otimes \mathcal{B}$ então*

$$f(a) = Q(E_a), g(b) = P(E_b)$$

são mensuráveis e

$$\int_A f dP = \pi(E) = \int_B g dQ.$$

Teorema 14 (Teorema de Tonelli.) *Seja $(A, \mathcal{A}, P)$ e $(B, \mathcal{B}, Q)$ espaços de probabilidade. E seja $f : A \times B \to \mathbb{R}_+$ mensurável. Então*

$$\int_A \left(\int_B f(a,b) dQ(b) \right) dP(a) = \int_{A \times B} f d\pi = \int_B \left(\int_A f(a,b) dP(a) \right) dQ(b).$$

Teorema 15 (Teorema de Fubini.) *Se $\int_{A \times B} |f| d\pi < \infty$ então*

$$\int_A \left(\int_B f(a,b) dQ(b) \right) dP(a) = \int_{A \times B} f d\pi = \int_B \left(\int_A f(a,b) dP(a) \right) dQ(b).$$

Observação 2 *Note que se $f \geq 0$ o teorema de Tonelli é mais geral que o teorema de Fubini pois dispensa a hipótese de integrabilidade. O teorema de Fubini demonstra-se fácilmente decompondo-se $f = f^+ - f^-$.*

3 Independência.

Dizemos que dois eventos A e B são independentes se $P(A \cap B) = P(A)P(B)$. Um conjunto finito de eventos $\{A_1, \ldots, A_k\}$ é independente se para todo $1 \leq i_1 < i_2 < \ldots < i_t \leq k$ temos $P(A_{i_1} \cap \cdots \cap A_{i_t}) = P(A_{i_1}) \cdots P(A_{i_t})$. Uma família, $X_1, \ldots, X_n$ de variáveis aleatórias do espaço $(\Omega, \mathcal{A}, P)$ é independente se para todos boreleanos $B_1, \ldots, B_n$ de $\mathbb{R}$ temos

$$P(\cap_{i=1}^n \{\omega \in \Omega; X^i(\omega) \in B_i\}) = \Pi_{i=1}^n P(\{\omega \in \Omega; X^i(\omega) \in B_i\}).$$

Mais geralmente a família $\mathcal{A}_1, \ldots, \mathcal{A}_n$ de σ-álgebras de Ω é independente se para todo $A_i \in \mathcal{A}_i, 1 \leq i \leq n$,

$$P(\cap_{i=1}^n A_i) = \Pi_{i=1}^n P(A_i).$$

Note que $\{A_1, \ldots, A_k\}$ é independente se e somente se as σ-álgebras, $\mathcal{A}_i = \{A_i, A_i^c, \emptyset, \Omega\}$, $1 \leq i \leq k$ são independentes. A partir de agora usaremos com alguma frequência a notação $E[X]$ no lugar de $\int f(\omega) d\mu(\omega)$. Se X, Y

são variáveis aleatórias integráveis e independentes, então XY é integrável e $E[XY] = E[X]E[Y]$. No lema a seguir notemos que

$$\{\omega \in \Omega; \omega \in A_n \text{ i.v.}\} := \{\omega \in \Omega; \omega \in A_n \text{ infinitas vezes}\} = \limsup A_n.$$

Muito usado é o seguinte resultado:

Lema 5 (Borel-Cantelli) *Seja $(A_n)_n$ uma família de conjuntos mensuráveis. Então:*

$$\sum_{n=1}^{\infty} P(A_n) < \infty \textit{ implica } P(A_n \textit{ i.v.}) = 0.$$

Demonstração. Note que

$$P(A_n \text{ i.v.}) = P(\cap_{k=1}^{\infty} \cup_{n=k}^{\infty} A_n) \leq P(\cup_{n=k}^{\infty} A_n) \leq \sum_{n=k}^{\infty} P(A_n).$$

Portanto $P(A_n \text{ i.v.}) \leq \lim_{k\to\infty} \sum_{n=k}^{\infty} P(A_n) = 0$. QED

O lema de Borel-Cantelli tem uma recíproca parcial que inclui o caso dos eventos A_n serem dois a dois independentes:

Lema 6 *Suponhamos que $\sum_{n=1}^{\infty} P(A_n) = \infty$. Se $P(A_i \cap A_j) \leq P(A_i)P(A_j)$ sempre que $i \neq j$ então $P(A_n \textit{ i.v.}) = 1$.*

4 Processos estocásticos.

A distribuição de uma variável aleatória X é a função $F = F_X : \mathbb{R} \to \mathbb{R}$ definida por $F(x) = P(\{\omega \in \Omega; X(\omega) \leq x\})$. Temos que F é não decrescente, $F(-\infty) = 0$ e $F(\infty) = 1$ e que F é contínua à direita. Se $X^2 \in L^1(P)$ podemos definir a variância de X por $E[(X - E[X])] = \int_{\Omega} (X - EX)^2 dP$. Assim se X e Y são independentes, então $Var(X + Y) = Var(X) + Var(Y)$. A variável aleatória X tem distribuição normal de média μ e variância $\sigma^2 > 0$ se

$$F_X(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt.$$

Definição 15 *Uma filtração em $(\Omega, \mathcal{A}, P)$ é uma família crescente de σ-álgebras, $(\mathcal{A}_t)_t \subset \mathcal{A}$ tal que:*

1. *$\mathcal{A}_t = \cap_{s>t} \mathcal{A}_s$ para todo $t \geq 0$;*
2. *$\mathcal{A}_0$ contém os subconjuntos nulos de $\mathcal{A}$.*

Definição 16 *Um processo estocástico, $\{X_t\}_{t\geq 0}$, é uma família de variáveis aleatórias $X_t : \Omega \to \mathbb{R}$. Um processo estocástico é adaptado se X_t é $\mathcal{A}_t$ mensurável para todo t.*

Um movimento Browniano de média μ e variância σ^2 (também chamado de processo de Wiener), é um processo estocástico, $\{B_t\}_{t\geq 0}$, tal que:

1. $B_0 = 0$, quase certamente;
2. Para todos $0 \leq t_0 < t_1 < \ldots < t_n$, $n \geq 1$, $B_{t_1} - B_{t_0}, \ldots, B_{t_n} - B_{t_{n-1}}$ são independentes;
3. Se $0 \leq s < t$, $B_t - B_s$ tem distribuição normal de média $(t-s)\mu$ e variância $(t-s)\sigma^2$;
4. A trajetória, $t \to B_t(\omega)$ é contínua quase certamente em $\omega \in \Omega$.

Se $\mu = 0$ e $\sigma^2 = 1$ dizemos que o movimento Browniano é normalizado.

5 Martingalas e esperança condicional.

Esperança condicional.

Suponhamos que $X \in L^1(P)$ e $\mathcal{B}$ é uma σ-álgebra de Ω contida em $\mathcal{A}$. Existe uma função $E(X|\mathcal{B}) : \Omega \to \mathbb{R}$, $\mathcal{B}$ mensurável, tal que para todo $B \in \mathcal{B}$,

$$\int_B E(X|\mathcal{B})(\omega)dP(\omega) = \int_B X(\omega)dP(\omega).$$

A função $E(X|\mathcal{B})$ chama-se a esperança condicional de X dada a σ-álgebra $\mathcal{B}$. Valem as seguintes propriedades:

1. $E(X|\mathcal{B})$ é única a menos de um subconjunto $B \in \mathcal{B}$ de probabilidade nula.
2. $E(aX + bY|\mathcal{B})(\omega) = aE(X|\mathcal{B})(\omega) + bE(X|\mathcal{B})(\omega)$ para quase todo $\omega \in \Omega$.
3. Se $X \geq 0$ então $E(X|\mathcal{B}) \geq 0$.
4. $E(1|\mathcal{B}) = 1$.
5. Se X é $\mathcal{B}$ mensurável, então $E(X|\mathcal{B}) = X$ quase certamente. Mais geralmente $E(XY|\mathcal{B}) = XE(Y|\mathcal{B})$.
6. Se X é independente de $\mathcal{B}$ e $X \in L^1$ então $E(X|\mathcal{B}) = EX$.
7. Se $\phi : \mathbb{R} \to \mathbb{R}$ é convexa, $E[\phi(X)|\mathcal{B}] \geq \phi(E[X|\mathcal{B}])$.

Tempo de Parada.

O conceito de tempo de parada é fundamental.

Definição 17 *Seja* $(\Omega, \mathcal{A}, (\mathcal{A}_t)_t, P)$ *um espaço de probabilidades com uma filtração. Uma função* $\tau : \Omega \to [0, \infty]$ *é um tempo de parada se para todo* $t \geq 0$, $\{\omega \in \Omega; \tau(\omega) \leq t\} \in \mathcal{A}_t$.

O exemplo a seguir ilustra a importância do conceito de tempo de parada.

Exemplo 4 *Suponha* $(X_t)_{t\geq 0}$ *um processo estocástico contínuo e adaptado. Seja* $E = (-\infty, b]$*. A primeira vez,* $\tau_b(\omega)$*, que* $X_t(\omega)$ *é maior que* b*, é um tempo de parada. De fato, por definição* $\tau_b(\omega) = \inf\{t > 0; X_t(\omega) > b\}$*. Logo* $\{\omega \in \Omega; \tau_b(\omega) \leq t\} = \cap_{n=1}^{\infty}\{\omega \in \Omega; \tau_b(\omega) < t + 1/n\} = \cap_{n=1}^{\infty}\{\omega \in \Omega; \text{existe } s < t + 1/n, X_s(\omega) > b\} =$
$\cap_{n=1}^{\infty} \cup_{s\in\mathbb{Q}, s<t+1/n} \{\omega \in \Omega; X_s(\omega) > b\} \in \cap_{n=1}^{\infty}\mathcal{A}_{t+1/n} = \mathcal{A}_t$.

Proposição 7 *Se* $\sigma, \tau, \tau_n, n \geq 1$ *são tempos de parada então* $\max\{\sigma, \tau\}$, $\min\{\sigma, \tau\}$, $\sup_n \tau_n$ *e* $\limsup_{n\to\infty} \tau_n$ *são tempos de parada.*

A demontração dessa proposição não é dificil e será omitida.

Martingalas.

Um processo estocástico adaptado, $(X_t)_{t\geq 0}$, é uma martingala com relação à filtração $(\mathcal{A}_t)_{t\geq 0}$, se

1. i
2. $X_t \in L^1(P)$ para todo t;
3. $E(X_t|\mathcal{A}_s) = X_s$ para todo $t \geq s$.

Se no lugar da igualdade em (3) tivermos $E(X_t|\mathcal{A}_s) \geq X_s$ para todo $t \geq s$ dizemos que $(X_t)_t$ é uma sub-martingala.

As martingalas satisfazem a uma desigualdade muito útil:

Proposição 8 (A desigualdade das martingalas.) *Suponhamos que* (X_n) *seja uma martingala com relação à filtração* $\mathcal{A}_n = \sigma\{X_1, \ldots, X_n\}$*. Isto é,* $E(|X_n|) < \infty$ *e* $E(X_{n+1}|X_1, \ldots, X_n) = X_n$, $n \in \mathbb{N}$*. Então*

1. $P(\max_{1\leq k\leq n} |X_k| > \epsilon) \leq \frac{1}{\epsilon}E|X_n|$ *para todo* $\epsilon > 0, n \in \mathbb{N}$.
2. *Se além disso,* $E|X_n|^p < \infty$ *para algum* $p \geq 1$ *e para todo* n*, então* $P(\max_{1\leq k\leq n} |X_k| > \epsilon) \leq \frac{1}{\epsilon^p}E|X_n|^p$.

Demonstração: Seja $A_k = \cap_{j<k}[|X_j| \leq \lambda] \cap [|X_k| > \lambda]$ para k entre 1 e n e $A = [\max_{1\leq k\leq n}|X_k| > \lambda]$. Temos que A é uma união disjunta dos A_k. Portanto

$$\lambda P(A) = \sum_{k=1}^{n} \lambda P(A_k) \leq \sum_{k=1}^{n} E(|X_k|\chi_{A_k}).$$

Como A é mensurável com relação à $X_1, \ldots, X_k$,

$$E|X_n| \geq \sum_{k=1}^{n} E(|X_n|\chi_{A_k}) = \sum_{k=1}^{n} EE(|X_n|\chi_{A_k}|X_1, \ldots, X_k) =$$

$$\sum_{k=1}^{n} E[\chi_{A_k} E(|X_n||X_1, \ldots, X_k)] \geq \sum_{k=1}^{n} E[\chi_{A_k}|E(X_n|X_1, \ldots, X_k)|] = \sum_{k=1}^{n} E(\chi_{A_k}|X_k|).$$

Assim terminamos a demonstração da primeira parte. Para demonstrar a segunda parte usamos a primeira parte e que $|X_n|^p$ é uma sub-martingala para $p \geq 1$.

A proposição anterior possui uma generalização muito importante para processos estocásticos separáveis.

Definição 18 *Um processo estocástico, $(X_t)_{t\geq 0}$, é separável se existe um conjunto enumerável denso em $[0,\infty)$, $\{t_j\}_{j=1}^{\infty}$, e um subconjunto N de Ω de medida nula tal que para todo J, intervalo aberto de $\mathbb{R}_+$,*

$$\cap_{t\in J}\{\omega \in \Omega \setminus N; X(t,\omega) \in F\} = \cap_{n:t_n\in J}\{\omega \in \Omega \setminus N; X(t_n,\omega) \in F\}$$

para todo subconjunto fechado $F \subset \mathbb{R}$.

Notemos que todo processo estocástico cujas trajetórias, $t \to X_t(\omega)$ são contínuas quase certamente, é separável. A generalização prometida é a seguinte:

Proposição 9 *Se $(X_t)_{t\geq 0}$ é uma martingala separável, então*

$$P(\max_{0\leq s\leq t}|X_s| > \epsilon) \leq \frac{1}{\epsilon}E|X_t|.$$

Além disso, se $E|X_s|^p < \infty$ para algum $p \geq 1$ e para todo $s \leq t$, então

$$P(\max_{0\leq s\leq t}|X_s| > \epsilon) \leq \frac{1}{\epsilon^p}E|X_t|^p.$$

Não demonstrarei esta proposição.

6 Integral estocástica.

Seja $(B_t)_{t\geq 0}$, um movimento Browniano normalizado, no espaço de probabilidades $(\Omega, \mathcal{A}, P)$. Seja $(\mathcal{A}_t)_t$ uma filtração de $\mathcal{A}$ tal que B_s é $\mathcal{A}_t$ mensurável se $s \leq t$ e tal que $\sigma(B_{s+t} - B_t; s \geq 0)$, a menor σ-álgebra que torna $\{B_{s+t} - B_t; s \geq 0\}$ mensuráveis, é independente de $\mathcal{A}_t$ para todo t.

Inicialmente definiremos a integral estocástica para funções degrau. Suponha então que $f : \mathbb{R}_+ \times \Omega \to \mathbb{R}$ e que existam $0 = t_0 < t_1 < \ldots < t_{n+1} = T$ tais que $f(t, \omega) = f_i(\omega)$ se $t_i \leq t < t_{i+1}, 0 \leq i \leq n, \omega \in \Omega$, sendo f_i $\mathcal{A}_{t_i}$ mensurável. Definimos então para $\omega \in \Omega$,

$$\int_0^T f(t) dB_t(\omega) = \sum_{i=0}^n f_i(\omega)(B_{t_{i+1}}(\omega) - B_{t_i}(\omega)).$$

A variável aleatória $\int_0^T f(t) dB_t$ acima definida é a integral estocástica da função degrau f com relação ao movimento Browniano $(B_t)_t$.

A classe de funções degrau é pequena demais. Nosso objetivo é extender a integral acima ao fecho em probabilidade das funções degrau.

É facil de se verificar que a integral estocástica é linear nas funções degrau.

Lema 7 *Se f é uma função degrau tal que $f(t) \in L^2(P)$ para todo t, então*

$$E \int_0^T f(t) dB_t = 0 \text{ e } E \left(\int_0^T f(t) dB_t \right)^2 = E \int_0^T f^2(t) dt.$$

Demonstração: Para demonstrarmos a primeira igualdade notemos primeiramente que $E|f_i| \leq E\left(1 + |f_i|^2\right) = 1 + E|f_i|^2 < \infty$. Também vale que $E|B_{t_{i+1}} - B_{t_i}| < \infty$. Logo usando as propriedades (5) e (6) da esperança condicional e o item (3) da definição do movimento Browniano:

$$E \int_0^T f(t) dB_t = E \left(\sum_{i=0}^n f_i \cdot (B_{t_{i+1}} - B_{t_i}) \right) = \sum_{i=0}^n E\left(f_i \cdot (B_{t_{i+1}} - B_{t_i})\right) =$$

$$\sum_{i=0}^n E\left(E\left[f_i \cdot (B_{t_{i+1}} - B_{t_i}) | \mathcal{A}_{t_i}\right]\right) = \sum_{i=0}^n E(f_i) \cdot E(B_{t_{i+1}} - B_{t_i}) = 0.$$

Demonstremos agora a segunda igualdade. Vale a seguinte identidade:

$$\left(\int_0^T f(t) dB_t \right)^2 = \left(\sum_{i=0}^n f_i (B_{t_{i+1}} - B_{t_i}) \right)^2 =$$

$$2 \sum_{0 \leq i < j \leq n} f_i f_j (B_{t_{i+1}} - B_{t_i})(B_{t_{j+1}} - B_{t_j}) + \sum_{i=0}^n f_i^2 (B_{t_{i+1}} - B_{t_i})^2.$$

Precisamos verificar que cada parcela é integrável. Temos que $Ef_i^2(B_{t_{i+1}} - B_{t_i})^2 < \infty$ por independência. Logo

$$2E|f_i f_j(B_{t_{i+1}} - B_{t_i})(B_{t_{j+1}} - B_{t_j})| \leq E|f_i(B_{t_{i+1}} - B_{t_i})|^2 + E|f_j(B_{t_{j+1}} - B_{t_j})|^2 < \infty.$$

Podemos então calcular: $E\left(\int_0^T f(t)dB_t\right)^2 =$

$$\sum_{0\leq i<j\leq n} 2E\left[f_i f_j(B_{t_{i+1}} - B_{t_i})(B_{t_{j+1}} - B_{t_j})\right] + \sum_{i=0}^{n} E\left[f_i^2(B_{t_{i+1}} - B_{t_i})^2\right] =$$

$$= \sum_{0\leq i<j\leq n} 2E\left[f_i f_j(B_{t_{i+1}} - B_{t_i})\right] E[(B_{t_{j+1}} - B_{t_j})] + \sum_{i=0}^{n} E\left[f_i^2(B_{t_{i+1}} - B_{t_i})^2\right] =$$

$$= \sum_{i=0}^{n} E\left[f_i^2 E(B_{t_{i+1}} - B_{t_i})^2\right] = \sum_{i=0}^{n} E[f_i^2(t_{i+1} - t_i)] = E\int_0^T f^2(t)dt.$$

Definição 19 *Um processo estocástico definido em* $(\Omega, \mathcal{A}, P), (f_t)_t$ *é não antecipativo com relação à filtração* $(\mathcal{A}_t)_t$ *se*

1. $(f_t)_t$ *for separável*
2. $(t, \omega) \to f(t, \omega)$ *é mensurável com relação à* $\sigma(\mathcal{B}^1 \times \mathcal{A})$.
3. $(f_t)_t$ *é adaptado:* f_t *é* $\mathcal{A}_t$ *mensurável para todo* t.

Definimos então $L^2_\Omega = L^2_\Omega[0,T]$ como o conjunto dos processos estocásticos não-antecipativos $(f_t)_t$ tais que:

$$\text{para todo } T > 0,\ P(\{\omega; \int_0^T |f(t,\omega)|^2 dt < \infty\}) = 1.$$

E definimos $M^2_\Omega = \{f \in L^2_\Omega; \int_\Omega \int_0^T |f(t,\omega)|^2 dt\, dP(\omega) < \infty \text{ para todo } T > 0\}$. O próximo lema é fundamental para passarmos ao limite em probabilidades.

Lema 8 *Para toda função degrau* $f \in L^2_\Omega$ *e para todo* $\epsilon > 0, \delta > 0$ *temos*

$$P\left(\left|\int_0^T f(t)dB_t\right| > \epsilon\right) \leq P\left(\int_0^T f^2(t)dt > \delta\right) + \delta/\epsilon^2.$$

Demonstração: Seja

$$f_\delta(t) = \begin{cases} f(t) & \text{se} \quad t_k \leq t < t_{k+1} \text{ e } \sum_{i=0}^k f^2(t_i)(t_{i+1} - t_i) \leq \delta, \\ 0 & \text{se} \quad t_k \leq t < t_{k+1} \text{ e } \sum_{i=0}^k f^2(t_i)(t_{i+1} - t_i) > \delta. \end{cases}$$

Como $P(|\int_0^T f(t)dB_t| > \epsilon) \leq P(|\int_0^T f_\delta(t)dB_t| > \epsilon) + P(\int_0^T f(t)^2(\omega)dt > \delta)$ pela desigualdade de Tchebichev, (i.e. $P(X \geq \delta) \leq (EX^2)/\delta^2$) obtemos

$$P\left(\omega; \left|\int_0^T f(t)(\omega)dB_t\right| > \epsilon\right) \leq \frac{E\left(\int_0^T f_\delta(t)dB_t(\omega)\right)^2}{\epsilon^2} + P\left(\omega; \int_0^T f^2(t)(\omega)dt > \delta\right).$$

Finalmente notemos que f_δ é uma função degrau pois $\{\omega; \sum_{i=0}^k f^2(t_i)(t_{i+1} - t_i) \leq \delta\}$ é $\mathcal{A}_{t_k}$ mensurável. Portanto $E(\int_0^T f_\delta(t)dB_t)^2 = E\int_0^T f_\delta^2(t)dt \leq \delta$ o que termina a demonstração.

A proposição a seguir é a base para a extensão da integral de Itô a L^2_Ω.

Proposição 10 *Para toda $f \in L^2_\Omega$ e $T > 0$ existe uma sequência $f_n \in L^2_\Omega$ de funções degrau tais que*

$$\int_0^T |f_n(t) - f(t)|^2 \, dt \stackrel{q.c.}{\to} 0.$$

Além disso se $f \in M^2_\Omega$, $E\left[\int_0^T |f_n(t) - f(t)|^2 \, dt\right] \to 0$.

Não demonstrarei essa proposição pois sua demonstração é relativamente longa e não necessitaremos dela nada além da existência de tal sequência. Se necessário podemos supor, no lugar de funções degrau, que as funções f_n são contínuas quase certamente.

Passamos a extender a integral de Itô para L^2_Ω. Seja então $f \in L^2_\Omega$ e $f_n \in L^2_\Omega$ uma sequência de funções degrau dadas pela proposição anterior. A sequência $(\int_0^T f_n(t)dB_t)_n$ é de Cauchy em probabilidades pois se $\epsilon > 0$,

$$P\left(\omega \in \Omega; \left|\int_0^T (f_n(t,\omega) - f_m(t,\omega))dB_t(\omega)\right| > \epsilon\right) \leq$$

$$P\left(\omega; \int_0^T |f_n(t) - f_m(t)|^2 \, dt > \delta\right) + \delta/\epsilon^2.$$

Existe então pelo lema (1), $\int_0^T f(t)dB_t$, o limite em probabilidade de $\int_0^T f_n(t)dB_t$. Este limite é único e independe da sequência de funções degrau utilizadas. Chamamos $\int_0^T f(t)dB_t$ de integral estocástica (ou integral de Itô) de f com relação ao movimento Browniano $(B_t)_t$. O teorema a seguir é imediato e não será demonstrado.

Teorema 16 *Se $f_1, f_2 \in L^2_\Omega$ e $r_1, r_2 \in \mathbb{R}$ então $r_1 f_1 + r_2 f_2 \in L^2_\Omega$ e*

$$\int_0^T (r_1 f_1(t) + r_2 f_2(t))dB_t = r_1 \int_0^T f_1(t)dB_t + r_2 \int_0^T f_2(t)dB_t.$$

O próximo teorema nos permite calcular a média e a variância da integral de Itô.

Teorema 17 *Se $f \in M^2_\Omega$ então*

$$E\left(\int_0^T f(t)dB_t\right) = 0 \text{ e } E\left(\int_0^T f(t)dB_t\right)^2 = E\int_0^T f^2(t)dt.$$

Demonstração: Sejam $f_n \in L^2_\Omega$, $n \in \mathbb{N}$ funções degrau, tais que $E\int_0^T (f_n(t) - f(t))^2 dt \to 0$. Vale $E\int_0^T f_n(t)dB_t = 0$ para todo n. Também $E(\int f_n(t)dB_t)^2 = E\int f_n^2(t)dt$. Temos então que $\lim_{n\to\infty} E(\int f_n(t)dB_t)^2 = E\int f^2(t)dt$. Como $E(\int f_n(t)dB_t - \int f_m(t)dB_t)^2 = E\int |f_n(t) - f_m(t)|^2 dt \to 0$ vale portanto

$$E\left(\int f_n(t)dB_t - \int f(t)dB_t\right)^2 \to 0,$$

e

$$E\left(\int f(t)dB_t\right)^2 = \lim_{n\to\infty} E\left(\int f_n(t)dB_t\right)^2 = E\int f^2(t)dt.$$

Passarei a considerar as propriedades da integral de Itô, $I_t = \int_0^t f(s)dB_s$, como função de $t \in [0,T]$. Os próximos lemas são para demonstrar que $(I_t)_t$ é uma martingala.

Lema 9 *Se $f \in L^2_\Omega[\alpha,T]$ e X é limitada, $\mathcal{A}_\alpha$ mensurável então $Xf \in L^2_\Omega[\alpha,T]$ e*

$$\int_\alpha^T Xf(t)dB_t = X \cdot \int_\alpha^T f(t)dB_t.$$

Demonstração: É claro que $Xf \in L^2_\Omega[\alpha,T]$. Se f é uma função degrau a relação acima segue da definição de integral. O caso geral segue por aproximação.

Teorema 18 *Se $f \in M^2_\Omega[\alpha,T]$ então*

$$E\left(\int_\alpha^T f(t)dB_t|\mathcal{A}_\alpha\right) = 0,$$

e

$$E\left(\left|\int_\alpha^T f(t)dB_t\right|^2 |\mathcal{A}_\alpha\right) = E\left(\int_\alpha^T f^2(t)dt|\mathcal{A}_\alpha\right) = \int_\alpha^T E[f^2(t)|\mathcal{A}_\alpha]dt.$$

Demonstração: Seja X limitada, $\mathcal{A}_\alpha$ mensurável. Então $Xf \in M^2_\Omega[\alpha,T]$ e $E(\int_\alpha^T Xf(t)dB_t) = 0$. Logo pelo lema anterior $E(X\int_\alpha^T f(t)dB_t) = 0$ e portanto $E\{XE[\int_\alpha^T f(t)dB_t|\mathcal{A}_\alpha]\} = 0$ para todo X demonstrando que $E\left[\int_\alpha^T f(t)dB_t|\mathcal{A}_\alpha\right] = 0$. A segunda parte do teorema tem demonstração análoga.

Teorema 19 *Se $f \in M^2_\Omega[0,T]$ então $I_t = \int_0^t f(s)dB_s$ é uma martingala.*

Demonstração: Sejam $0 \le t' < t \le T$. Então de $I_t = I_{t'} + \int_{t'}^t f(s)dB_s$ vem

$$E(I_t|\mathcal{A}_{t'}) = I_{t'} + E\left(\int_{t'}^t f(s)dB_s|\mathcal{A}_{t'}\right) = I_{t'}.$$

A integral de Itô como definida pode não ser contínua em função de t. Mas no entanto, sempre pode ser modificada, num conjunto de medida zero e ser então contínua quase certamente. Inicialmente precisamos da seguinte definição:

Definição 20 *Dois processos estocásticos, X e Y, são equivalentes se*

$$\text{para todo } t \ge 0,\ P(X(t) \ne Y(t)) = 0.$$

Nesse caso dizemos que Y é uma versão de X.

Teorema 20 *A integral de Itô possui uma versão contínua.*

Demonstração: Seja $f \in M^2_\Omega$ e $(f_n)_n$ uma sequência de funções degrau tais que

$$E\left(\int_0^T (f(t) - f_n(t))^2 dt\right) \to 0.$$

Notemos que $I_n(t) = \int_0^t f_n(s)dB_s$, $0 \le t \le T$ é contínua e $I_n(t) - I_m(t)$ é uma martingala. Portanto pela desigualdade das martingalas,

$$P\left(\sup_{0\le t\le T} |I_n(t) - I_m(t)| > \epsilon\right) \le \frac{1}{\epsilon^2} E\left(\left|\int_0^T (f_m(s) - f_n(s))dB_s\right|^2\right) =$$

$$\frac{1}{\epsilon^2} E\left(\int_0^T (f_m(s) - f_n(s))^2 ds\right) \to 0.$$

Para $\epsilon = 1/2^k$ existe n_k suficientemente grande para que

$$P\left(\sup_{0\le t\le T} |I_{n_k}(t) - I_m(t)| > 1/2^k\right) < 1/k^2 \text{ se } m \ge n_k.$$

Sem perda de generalidade n_k cresce com k. Logo

$$P\left(\sup_{0\le t\le T} |I_{n_k}(t) - I_{n_{k+1}}(t)| > 1/2^k\right) < 1/k^2,$$

e portanto, como $\sum_k 1/k^2 < \infty$, pelo lema de Borel-Cantelli obtemos

$$P\left(\sup_{0\le t\le T} |I_{n_k}(t) - I_{n_{k+1}}(t)| > 1/2^k \text{ infinitas vezes}\right) = 0.$$

Ou seja para quase todo ω, $|I_{n_k}(t)(\omega) - I_{n_{k+1}}(t)(\omega)| \leq 1/2^k$, $0 \leq t \leq T$. Logo com probabilidade um, $I_{n_k}(t)$ converge uniformemente para uma função contínua $J(t)$. Como $\int_0^t f_n(s)dB_s \to \int_0^t f(s)dB_s$ segue-se $J(t) = \int_0^t f(s)dB_s$ quase certamente. Suponhamos agora que $f \in L^2_\Omega[0,T]$ e definamos $f_N(t) = f(t)$ se $\int_0^t f^2(s)ds \leq N$ e $f_N(t) = 0$ caso contrário. É claro que $f_N \in M^2_\Omega[0,T]$. Seja $J_N(t) = \int_0^t f_N(s)dB_s$ uma versão contínua. E $\Omega_N = \{\omega; \int_0^T f^2(t)(\omega)dt < N\}$. Se $\omega \in \Omega_N$ e $M > N$, $f_N(t) = f_M(t)$ para $t \in [0,T]$. Temos então $J_M(t) = J_N(t)$ se $0 \leq t \leq T$. Logo $\tilde{J}(t) = \lim_{M\to\infty} J_M(t)$ é contínua em t quase certamente $\omega \in \Omega_N$. Como $P(\Omega_N) \to 1, (\tilde{J}(t))_t$ é um processo contínuo. E de $P(\int_0^t |f(s) - f_N(s)|^2 ds > 0) = P(\int_0^t f^2(s)ds > N) \to 0$ vem $J_M(t) \xrightarrow{P} \int_0^t f(s)dB_s = I_t$.

A partir de agora somente usaremos a versão contínua da integral de Itô. Vale também o teorema a seguir, demonstrado anteriormente para funções degrau.

Teorema 21 *Se* $f \in L^2_\Omega, \epsilon > 0, \delta > 0$ *então*

$$P\left(\max_{0\leq t\leq t}\left|\int_0^T f(t)dB_t\right| > \epsilon\right) \leq P\left(\int_0^T f^2(t)dt > \delta\right) + \delta/\epsilon^2.$$

Demonstração: Definamos $\chi_\delta(z) = \chi_{(-\infty,\delta]}(z)$ e $f_\delta(t) = f(t)\cdot\chi_\delta(\int_0^t f^2(s)ds)$. Vale então que

$$P\left(\sup_{0\leq t\leq T}\left|\int_0^t f(s)dB_s\right| > \epsilon\right) \leq$$

$$\leq P\left(\sup_{0\leq t\leq T}\left|\int_0^t f(s)dB_s - \int_0^t f_\delta(s)dB_s\right| > 0\right) + P\left(\sup_{0\leq t\leq T}\left|\int_0^t f_\delta(s)dB_s\right| > \epsilon\right).$$

Como $\int_0^t f_\delta(s)dB_s$ é uma martingala temos pela desigualdade das martingalas que:

$$P\left(\sup_{0\leq t\leq T}\left|\int_0^t f_\delta(s)dB_s\right| > \epsilon\right) \leq \frac{1}{\epsilon^2}E\left|\int_0^T f_\delta(s)dB_s\right|^2 = \frac{1}{\epsilon^2}E\int_0^T f_\delta(s)^2 ds \leq \frac{\delta}{\epsilon^2}.$$

Consideremos agora a primeira parcela. Se

$$\sup_{0\leq t\leq T}\left|\int_0^t f(s)(\omega)dB_s(\omega) - \int_0^t f_\delta(s)(\omega)dB_s(\omega)\right| > 0$$

existe $\tilde{t}$ tal que $\int_0^{\tilde{t}} f(s)(\omega)dB_s(\omega) \neq \int_0^{\tilde{t}} f_\delta(s)(\omega)dB_s(\omega)$. Logo pela definição de f_δ, $\int_0^{\tilde{t}} f^2(s)(\omega)ds > \delta$ e portanto $\int_0^T f^2(s)ds > \delta$. QED

O próximo teorema mostra que a integral estocástica possui uma importante propriedade de continuidade.

Teorema 22 *Se $f, f_n \in L^2_\Omega$ e $\int |f_n(t) - f(t)|^2 dt \xrightarrow{P} 0$ então $\int f_n(t)dB_t \xrightarrow{P} \int f(t)dB_t$.*

Demonstração: Seja $\epsilon > 0, \rho > 0$. Temos,

$$P\left(\left|\int_0^T (f_n(t) - f(t))dB_t\right| > \epsilon\right) \leq P\left(\int_0^T (f_n(t) - f(t))^2 dt > \epsilon^2\rho\right) + \rho.$$

Logo, por hipótese, a primeira parcela do membro direito da desigualdade anterior tende a zero. Fazendo $\rho \to 0$ terminamos a demonstração.

Algumas integrais fáceis de calcular.

Para treinar um pouco calculemos as seguintes integrais.

1. $\int_0^T dB_t = B_T$;
2. $\int_0^T B_t dB_t = \frac{1}{2}B_T^2 - T/2$;
3. $\int_0^T f(t)dB_t = -\int_0^T f'(t)B_t dt + f(T)B_T$ se $f : \mathbb{R}_+ \to \mathbb{R}$ é contínuamente diferenciável.

A integral em (1) é imediata. Para calcular (2) seja $f_n(t,\omega) = B_{Tk/n}(\omega)$ se $k/n \leq t/T < (k+1)/n, 0 \leq k \leq n-1$. É claro que f_n é uma função degrau. Logo

$$\int f_n(t)dB_t = \sum_{k=0}^{n-1} B_{\frac{Tk}{n}}\left(B_{\frac{T(k+1)}{n}} - B_{\frac{Tk}{n}}\right) =$$

$$-\frac{1}{2}\sum_{k=0}^{n-1}\left(B_{\frac{T(k+1)}{n}} - B_{\frac{Tk}{n}}\right)^2 + \frac{1}{2}\sum_{k=0}^{n-1}\left(B^2_{\frac{T(k+1)}{n}} - B^2_{\frac{Tk}{n}}\right) =$$

$$= -\frac{1}{2}\sum_{k=0}^{n-1}\left(B_{\frac{T(k+1)}{n}} - B_{\frac{Tk}{n}}\right)^2 + \frac{1}{2}B_T^2.$$

A variável aleatória, $Z_n = \sum_{k=0}^{n-1}(B_{\frac{T(k+1)}{n}} - B_{\frac{Tk}{n}})^2$, tem média T e variância $\frac{2T^2}{n} = \sum_{k=0}^{n-1}\frac{2T^2}{n^2} \to 0$. Portanto $Z_n \xrightarrow{q.c.} T$ e então $\int B_t dB_t = \frac{1}{2}B_T^2 - T/2$. Para calcular (3) seja $f_n(t,\omega) = f(k/n)$, $k/n \leq t/T < (k+1)/n$. Logo

$$\int f_n(t)dB_t = \sum_{k=0}^{n-1} f(Tk/n)(B_{\frac{T(k+1)}{n}} - B_{\frac{Tk}{n}}) = \sum_{k=0}^{n-1} f(Tk/n)B_{\frac{T(k+1)}{n}} - \sum_{k=0}^{n-1} f(Tk/n)B_{\frac{Tk}{n}} =$$

$$\sum_{k=0}^{n-1}\left(f(Tk/n) - f(\frac{T(k+1)}{n})\right)B_{\frac{T(k+1)}{n}} + \sum_{k=0}^{n-1}\left(f(\frac{T(k+1)}{n})B_{\frac{T(k+1)}{n}} - f(Tk/n)B_{\frac{Tk}{n}}\right) =$$

$$= -\sum_{k=0}^{n-1} \frac{Tf'(\zeta_{kn})}{n} B_{\frac{T(k+1)}{n}} + f(T)B_T,$$

sendo $\zeta_{kn} \in (Tk/n, T(k+1)/n)$. Como $t \to B_t(\omega)$ é quase certamente contínua obtemos:

$$\lim_{n\to\infty} \sum_{k=0}^{n-1} \frac{Tf'(\zeta_{kn})}{n} B_{\frac{T(k+1)}{n}}(\omega) = \int_0^T f'(t)B_t(\omega)dt \quad \text{quase certamente.}$$

Isto termina a verificação de (3).

A fórmula de Itô.

Para demonstrarmos a fórmula de Itô necessitamos do

Teorema 23 *Suponhamos que $a_i, b_i \in L^2_\Omega$, $i = 1, 2$ e que X_1, X_2 são processos estocásticos tais que se $0 \le t_1 \le t_2 \le T$,*

$$X_i(t_2) - X_i(t_1) = \int_{t_1}^{t_2} a_i(t)dt + \int_{t_1}^{t_2} b_i(t)dB_t \quad \textit{quase certamente.}$$

Então o processo estocástico $Y = X_1X_2$ é tal que

$$Y_{t_2} - Y_{t_1} = \int_{t_1}^{t_2} (X_1a_2(t) + X_2a_1(t))dt + \int_{t_1}^{t_2} (X_1b_2(t) + X_2b_1(t))dB_t + \int_{t_1}^{t_2} b_2(t)b_1(t)dt.$$

Demonstração: Suponhamos primeiramente que $a_i, b_i, i = 1, 2$ são processos estocásticos independentes do tempo. Então nesse caso e $X_i = a_it + b_iB_t$

$$\int_{t_1}^{t_2} X_1a_2(t)dt + \int_{t_1}^{t_2} X_1b_2(t)dB_t + \int_{t_1}^{t_2} X_2a_1(t)dt + \int_{t_1}^{t_2} X_2b_1(t)dB_t + \int_{t_1}^{t_2} b_2b_1(t)dt =$$

$$a_2 \int_{t_1}^{t_2} [a_1t + b_1B_t]dt + b_2 \int_{t_1}^{t_2} [a_1t + b_1B_t]dB_t +$$

$$+a_1 \int_{t_1}^{t_2} [a_2t + b_2B_t]dt + b_1 \int_{t_1}^{t_2} [a_2t + b_2B_t]dB_t + b_1b_2(t_2 - t_1) =$$

$$= a_1a_2(t_2^2 - t_1^2) + b_1b_2(t_2 - t_1) +$$

$$(a_2b_1 + a_1b_2) \int_{t_1}^{t_2} B_tdt + (b_2a_1 + a_2b_1) \int_{t_1}^{t_2} tdB_t + 2b_1b_2 \int_{t_1}^{t_2} B_tdB_t =$$

$$= a_1a_2(t_2^2 - t_1^2) + (a_2b_1 + a_1b_2)[\int_{t_1}^{t_2} B_tdt + \int_{t_1}^{t_2} tdB_t] + b_1b_2(B_{t_2}^2 - B_{t_1}^2) =$$

$$= a_1a_2(t_2^2 - t_1^2) + (a_2b_1 + a_1b_2)(t_2B_{t_2} - t_1B_{t_1}) + b_1b_2(B_{t_2}^2 - B_{t_1}^2) =$$

$$= (a_1t_2 + b_1B_{t_2})(a_2t_2 + b_2B_{t_2}) - (a_1t_1 + b_1B_{t_1})(a_2t_1 + b_2B_{t_1}).$$

Suponhamos agora que a_i, b_i são funções degrau com intervalos $I_1, I_2, \ldots, I_k$. A fórmula de Itô vale em cada um desses intervalos. Vale portanto para quaisquer $t_1, t_2 \in [0, T]$. Façamos agora o caso geral. Sejam a_{in}, b_{in}, sequências de funções degrau não-antecipativas tais que

$$\int_0^T |a_{in}(t) - a_i(t)|\, dt \overset{q.c.}{\to} 0 \quad \text{e} \quad \int_0^T |b_{in}(t) - b_i(t)|^2\, dt \overset{q.c.}{\to} 0.$$

Definamos $\zeta_{in}(t) = X_i(0) + \int_0^t a_{in}(s)ds + \int_0^t b_{in}(s)dB_s$. Como $\zeta_{in} - X_i$ é uma martingala separável temos pela desigualdade das martingalas que

$$\sup_{0 \leq t \leq T} |\zeta_{in}(t) - X_i(t)| \overset{P}{\to} 0.$$

Existe então uma subsequência $(\zeta_{in})_n$ que converge uniformemente para X_i, quase certamente. É claro então que

$$\int_0^t \zeta_{in}(t) b_{jn}(t) dB_t \overset{P}{\to} \int_0^t X_i(t) b_j(t) dB_t.$$

É claro também que

$$\int_0^t \zeta_{in}(t) a_{jn}(t) dB_t \overset{P}{\to} \int_0^t X_i(t) a_j(t) dB_t \text{ e } \int_0^t b_{1n}(s) b_{2n}(s) ds \to \int_0^t b_1(s) b_2(s) ds.$$

QED

Uma maneira mais fácil de se memorizar o teorema anterior é obtida com o conceito de diferencial. Se

$$\zeta(t_2) - \zeta(t_1) = \int_{t_1}^{t_2} a(s)ds + \int_{t_1}^{t_2} b(s)dB_s,\ \ 0 \leq t_1 \leq t_2 \leq T,$$

dizemos que a diferencial de ζ é $d\zeta(t) = a(t)dt + b(t)dB_t$. O teorema anterior pode ser reescrito assim:

$$d(\zeta_1\zeta_2)(t) = \zeta_1(t)d\zeta_2(t) + \zeta_2(t)d\zeta_1(t) + b_1(t)b_2(t)dt.$$

Note que a diferencial do produto é analoga à diferencial usual de um produto mais o termo adicional b_1b_2dt.

Passemos agora a demonstrar a fórmula de Itô.

Teorema 24 (A fórmula de Itô.) *Seja $(X_t)_t$ um processo estocástico tal que $dX_t = a(t)dt + b(t)dB_t$ sendo $a, b \in L^2_\Omega$ e seja $f : \mathbb{R} \times [0, \infty) \to \mathbb{R}$ com todas as derivadas até a segunda ordem contínuas. Então o processo estocástico $Y_t = f(X_t, t)$ é tal que, para $0 \leq u \leq v \leq T$, $Y_v - Y_u =$*

$$\int_u^v \left[\frac{\partial f}{\partial t}(X_t, t) + \frac{\partial f}{\partial x}(X_t, t)a(t) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(X_t, t)b^2(t)\right] dt + \int_u^v \frac{\partial f}{\partial x}(X_t, t)b(t)dB_t.$$

Ou seja, a diferencial estocástica do processo $(Y_t)_t$ *é*

$$dY_t = \left[\frac{\partial f}{\partial t}(X_t, t) + \frac{\partial f}{\partial x}(X_t, t)a(t) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(X_t, t)b^2(t)\right] dt + \frac{\partial f}{\partial x}(X_t, t)b(t)dB_t.$$

Demonstração: A demonstração será longa e a dividirei em várias etapas. Primeiramente demonstrarei por indução que

$$d(B_t)^m = mB_t^{m-1}dB_t + \frac{m(m-1)}{2}B_t^{m-2}dt,$$

para todo $m \in \mathbb{N}$. Facilmente verificamos que a relação acima vale para $m = 1, 2$. Suponhamos a relação válida para m e demonstremos que vale para $m + 1$. Calculando $d(B_t)^{m+1}$ temos,

$$d(B_t)^{m+1} = d(B_t^m B_t) = B_t^m dB_t + B_t d(B_t)^m + mB_t^{m-1}dt =$$

$$B_t^m dB_t + mB_t^m dB_t + \frac{m(m-1)}{2}B_t^{m-1}dt + mB_t^{m-1}dt =$$

$$= (m+1)B_t^m dB_t + \frac{m(m+1)}{2}B_t^{m-1}dt,$$

terminando a indução. Disto obtemos imediatamente, para toda função polinomial Q em uma variável,

$$d(Q(B_t)) = Q'(B_t)dB_t + \frac{Q''(B_t)}{2}dt.$$

Passemos agora a considerar funções da forma $G(x, t) = Q(x)g(t)$, Q um polinômio em x e g duas vezes diferenciável. Então

$$d(G(B_t, t)) = d(Q(B_t)g(t)) = g(t)d(Q(B_t)) + Q(B_t)dg(t) + Q'(B_t) \times 0 =$$

$$g(t)\left[Q'(B_t)dB_t + \frac{Q''(B_t)}{2}dt\right] + Q(B_t)g'(t)dt =$$

$$g(t)Q'(B_t)dB_t + (\frac{g(t)Q''(B_t)}{2} + Q(B_t)g'(t))dt.$$

Reescrevendo obtemos

$$G(B_{t_2}, t_2) - G(B_{t_1}, t_1) = \int_{t_1}^{t_2}\left[G_t + \frac{1}{2}G_{xx}\right]dt + \int_{t_1}^{t_2} G_x dB_t.$$

A fórmula acima é então válida para $G = \sum_{i=1}^m f_i(x)g_i(t)$, f_i polinômial e g_i contínuamente diferenciável. Pelo teorema de Stone-Weierstrass toda $f : [t_1, t_2] \times \mathbb{R} \to \mathbb{R}$, duas vezes contínuamente diferenciável é limite uniforme nos

compactos, de funções da forma $G = \sum_{i=1}^{m} f_i(x)g_i(t)$, com a convergência uniforme de todas as derivadas até a segunda ordem. Vale então

$$G_n(B_{t_2}, t_2) - G_n(B_{t_1}, t_1) = \int_{t_1}^{t_2} \left[\frac{\partial G_n}{\partial t}(B_t, t) + \frac{1}{2}\frac{\partial^2 G_n}{\partial x^2}(B_t, t) \right] dt + \int_{t_1}^{t_2} \frac{\partial G_n}{\partial x}(B_t, t) dB_t.$$

A convergência uniforme implica

$$P-\lim_{n\to\infty} \int_{t_1}^{t_2} \left[\frac{\partial G_n}{\partial t}(B_t, t) + \frac{1}{2}\frac{\partial^2 G_n}{\partial x^2}(B_t, t) \right] dt = \int_{t_1}^{t_2} \left[\frac{\partial f}{\partial t}(B_t, t) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(B_t, t) \right] dt.$$

(lembre-se de que $t \to B_t(\omega)$ é contínua quase-certamente e que portanto $\{B_t(\omega); t_1 \le t \le t_2\}$ é um conjunto compacto q.c.)
Temos também que

$$\lim_{n\to\infty} \int_{t_1}^{t_2} \left| \frac{\partial G_n}{\partial x}(B_t, t) - \frac{\partial f}{\partial x}(B_t, t) \right|^2 dt = 0 \text{ quase certamente.}$$

Vale então que $\int_{t_1}^{t_2} \frac{\partial G_n}{\partial x}(B_t, t) dB_t \xrightarrow{P} \int_{t_1}^{t_2} \frac{\partial f}{\partial x}(B_t, t) dB_t$. E portanto vale a fórmula de Itô para o processo estocástico B_t. Para extendermos o resultado a qualquer processo estocástico notemos primeiramente que a fórmula vale se $Y_t = f(\xi_1 + a_1 t + b_1 B_t, t)$ se ξ_1, a_1, b_1 são $\mathcal{A}_{t_1}$ mensuráveis. A demonstração segue os mesmos passos anteriores a partir da verificação de que, se $\tilde{\xi}(t) = \xi_1 + a_1 t + b_1 B_t$

$$d(\tilde{\xi}(t))^m = m\tilde{\xi}(t)^{m-1}[a_1 dt + b_1 dB_t] + \frac{m(m-1)}{2}(\tilde{\xi}(t))^{m-2} b_1^2 dt.$$

Esses detalhes estão ao alcance de qualquer leitor que entendeu a primeira parte desta demonstração. O próximo passo é notar que se $a(t), b(t)$ são funções degrau a mesma fórmula será valida. Isto porque a fórmula é valida em cada intervalo onde $a(t), b(t)$ são constantes no tempo. Sejam agora $a(t), b(t)$ funções não-antecipativas de L^2_Ω e $a_n, b_n \in L^2_\Omega$ sequência de funções degrau tais que

$$\lim_{n\to\infty} \int_0^T |a_n(t) - a(t)| dt = 0 \text{ e } \int_0^T |b_n(t) - b(t)|^2 dt \xrightarrow{P} 0.$$

Definamos $\xi_n(t) = \xi(0) + \int_0^t a_n(s) ds + \int_0^t b_n(s) dB_s$. Vale que $\sup_{0\le t\le T} |\xi_n(t) - \xi(t)| \xrightarrow{P} 0$ e portanto que converge q.c. numa subsequência k_n, que para simplificar a notação, consideraremos $k_n = n$. Vale então que

$$\int_0^T \left| \frac{\partial f}{\partial x}(\xi_n(t), t) b_n(t) - \frac{\partial f}{\partial x}(\xi(t), t) b(t) \right| \xrightarrow{P} 0$$

e portanto que

$$\int_{t_1}^{t_2} \frac{\partial f}{\partial x}(\xi_n(t), t) b_n(t) dB_t \xrightarrow{P} \int_{t_1}^{t_2} \frac{\partial f}{\partial x}(\xi(t), t) b(t) dB_t.$$

Finalmente da convergência uniforme obtemos que:

$$P - \lim_{n\to\infty} \int_{t_1}^{t_2} \left[\frac{\partial f}{\partial t}(\xi_n(t), t) + \frac{\partial f}{\partial x}(\xi_n(t), t) a_n(t) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(\xi_n(t), t) b_n^2(t) \right] dt =$$

$$\int_{t_1}^{t_2} \left[\frac{\partial f}{\partial t}(\xi(t), t) + \frac{\partial f}{\partial x}(\xi(t), t) a(t) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(\xi(t), t) b^2(t) \right] dt.$$

QED

7 A fórmula de Black-Scholes

O primeiro cálculo bem sucedido do preço de um ativo derivado foi feito por Black e Scholes. O conceito fundamental usado por eles foi o de arbitragem.

Suponhamos uma economia sem custos de transação, com uma ação com preço S_t e uma opção européia com vencimento no tempo T e preço de exercício K. Suponhamos ainda que exista um ativo sem risco com preço $\beta_t = e^{rt}, r \geq 0$. O preço da ação tem distribuição log-normal: $S_t = S_0 e^{\alpha t + \sigma B_t}$. Pela fórmula de Itô:

$$\frac{dS_t}{S_t} = (\alpha + \sigma^2/2)dt + \sigma\, dB_t.$$

O valor da opção no instante T é $Y_T = \max\{S_T - K, 0\}$. Queremos calcular Y_t, $0 \leq t < T$.

Definição 21 *Uma estratégia de compra e venda é um par $(a, b) \in M_\Omega^2 \times L_\Omega^2$. A estratégia é auto-financiável se para $0 \leq t \leq T$,*

$$a_t S_t + b_t \beta_t = a_0 S_0 + b_0 \beta_0 + \int_0^t a_s dS_s + \int_0^t b_s r e^{rs} ds.$$

Se (a, b) é auto-financiável e $a_T S_T + b_T \beta_T = Y_T$ então $a_0 S_0 + b_0 \beta_0$ é o preço da opção. É claro então que pelos mesmos motivos, $Y_t = a_t S_t + b_t \beta_t$ é o preço da opção no tempo t. Para calcular Y_t suporemos que $Y_t = f(S_t, t), f$ duas vezes diferenciável se $t < T$. Se a resposta obtida realmente for duas vezes diferenciável estaremos justificados. Pelo lema de Itô obtemos ($\mu = \alpha + \sigma^2/2$)

$$dY_t = \left(\mu \frac{\partial f}{\partial x}(S_t, t) S_t + \frac{\partial f}{\partial t}(S_t, t) + \frac{\sigma^2}{2}\frac{\partial^2 f}{\partial x^2}(S_t, t) S_t^2 \right) dt + \frac{\partial f}{\partial x}(S_t, t) \sigma S_t dB_t.$$

Queremos encontrar (a, b) auto-financiável tal que $a_t S_t + b_t \beta_t = Y_t$. Nesse caso teríamos

$$dY_t = a_t dS_t + b_t r e^{rt} dt = (a_t \mu S_t + b_t r e^{rt}) dt + a_t \sigma S_t dB_t.$$

Para isto ser possivel temos que $a_t \sigma S_t = \sigma \frac{\partial f}{\partial x}(S_t, t) S_t$ e portanto $a_t = \frac{\partial f}{\partial x}(S_t, t)$. Pela definição de (a, b) obtemos $b_t = e^{-rt}(Y_t - a_t S_t) = e^{-rt}(f(S_t, t) - \frac{\partial f}{\partial x}(S_t, t) S_t)$. Substituindo e igualando os coeficientes em dt nas equações acima vem:

$$-rf(S_t, t) + \frac{\partial f}{\partial t}(S_t, t) + rS_t \frac{\partial f}{\partial x}(S_t, t) + \frac{\sigma^2 S_t^2}{2} \frac{\partial^2 f}{\partial x^2}(S_t, t) = 0.$$

Ou seja é suficiente que f satisfaça a equação a derivadas parciais:

$$-rf(x, t) + \frac{\partial f}{\partial t}(x, t) + rx \frac{\partial f}{\partial x}(x, t) + \frac{\sigma^2 x^2}{2} \frac{\partial^2 f}{\partial x^2}(x, t) = 0,$$

com a condição de fronteira $f(x, T) = \max\{x - K, 0\}$. Em geral tenta-se primeiramente o método de variáveis separáveis. Neste problema entretanto a transformação $k(x, t) = y(\log(x/K) + Bt, Ct)$ permite escolhendo-se C e B obter a equação $y_t + y_{xx} = 0$. Esta equação se resolve fácilmente. Pelo momento nos contentaremos em dizer que

$$f(x, t) = \frac{x}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-s^2/2} ds - \frac{K e^{-r(T-t)}}{\sqrt{2\pi}} \int_{-\infty}^{z - \sigma\sqrt{T-t}} e^{-s^2/2} ds$$

sendo $z = \frac{\log(x/K)}{\sigma\sqrt{T-t}} + (r + \sigma^2/2)\sqrt{T-t}/\sigma$ é a solução que procuramos. O leitor deve verificar como exercício que $\lim_{t \to T} f(x, t) = \max\{x - K, 0\}$.

8 A fórmula de Feynman-Kac.

A fórmula de Feynman-Kac permite resolver equações diferenciais parciais do tipo obtido acima como o valor esperado de certo processo estocástico. Uma formulação muito útil com certeza.

Consideremos a equação

$$-r(x, t) f(x, t) + \frac{\partial f}{\partial t}(x, t) + r(x, t) x \frac{\partial f}{\partial x}(x, t) + \frac{\sigma^2(x, t)}{2} \frac{\partial^2 f}{\partial x^2}(x, t) = 0,$$

com condição de fronteira $f(x, T) = g(x)$ para g contínua. Esse tipo de equação surge quando no lugar de termos o preço final $Y_T = \max\{S_T - K, 0\}$ temos $Y_T = g(S_T)$, a taxa de juros r e a variância σ mais gerais do que na fórmula de Black-Scholes.

Seja então Z_s^{xt} o processo estocástico definido por $Z_s^{xt} = x$ se $s \leq t$ e

$$dZ_s^{xt} = r(Z_s^{xt}, s) Z_s^{xt} ds + \sigma(Z_s^{xt}, s) dB_s \;\; s > t.$$

Teorema 25 *Se r, σ, g tem módulo limitado superiormente por um polinômio em (x, t) o preço do ativo derivado cujo valor final é $Y_T = g(S_T)$ será $Y_t = f(S_t, t)$ onde f é dada por*

$$f(x, t) = E\left(e^{-\int_t^T r(Z_s^{xt}, s) ds} g(Z_T^{xt})\right)$$

satisfaz a $f(x,T)=g(x)$ *e resolve:*

$$-r(x,t)f(x,t)+\frac{\partial f}{\partial t}(x,t)+r(x,t)x\frac{\partial f}{\partial x}(x,t)+\frac{\sigma^2(x,t)}{2}\frac{\partial^2 f}{\partial x^2}(x,t)=0.$$

Por exemplo, para a fórmula de Black-Scholes, $Z_T^{xt}=xe^{(r-\frac{\sigma^2}{2})(T-t)+\sigma(B_T-B_t)}$.

Verifiquemos a fórmula de Feynman-Kac, supondo que exista uma solução $f(x,t)$ da equação acima. Definamos os processos $Z_s=\int_t^s r(X_\tau,\tau)d\tau$, $Y_s=f(x,t)$ para $s\leq t$ e $Y_s=f(X_s,s)e^{-Z_s}$ se $s\geq t$, onde $X_s=Z_s^{xt}$. Temos $dZ_s=r(X_s,s)ds$ e pela fórmula de Itô obtemos que $df(X_s,s)=$

$$\left[\frac{\partial f}{\partial t}(X_s,s)+\frac{\partial f}{\partial x}(X_s,s)r(X_s,s)X_s+\frac{\partial^2 f}{2\partial x^2}(X_s,s)\sigma^2(X_s,s)\right]ds+\frac{\partial f}{\partial x}(X_s,s)\sigma(Z_s^{xt},s)dB_s.$$

Portanto

$$dY_s=\left[\frac{\partial f}{\partial t}(X_s,s)+\frac{\partial f}{\partial x}(X_s,s)r(X_s,s)X_s+\frac{\partial^2 f}{2\partial x^2}(X_s,s)\sigma^2(X_s,s)\right]e^{-Z_s}ds-$$

$$-f(X_s,s)e^{-Z_s}r(X_s,s)ds+\frac{\partial f}{\partial x}(X_s,s)\sigma(Z_s^{xt},s)e^{-Z_s}dB_s+\frac{\partial f}{\partial x}(X_s,s)\times 0dt=$$

$$=\frac{\partial f}{\partial x}(X_s,s)\sigma(Z_s^{xt},s)e^{-Z_s}dB_s.$$

E portanto

$$Y_T-f(x,t)=\int_t^T\frac{\partial f}{\partial x}(X_s,s)e^{-Z_s}dB_s.$$

$$EY_T=f(x,t)+E\int_t^T\frac{\partial f}{\partial x}(X_s,s)e^{-Z_s}\sigma(X_s,s)dB_s=f(x,t).$$

Acima usamos que a integral estocástica é uma martingala. Finalmente temos

$$f(x,t)=E\left[f(X_T,T)e^{-\int_t^T r(X_\tau,\tau)d\tau}\right]=E\left[e^{-\int_t^T r(X_\tau,\tau)d\tau}g(X_T)\right].$$

Terminando a verificação.

9 Integração estocástica em várias dimensões.

Definição 22 *Um processo estocástico,* $B(t)=(B_1(t),\ldots,B_n(t))$, *é um movimento Browniano de dimensão* n *se* $B_j(t)$ *é um movimento Browninano (unidimensional) para todo* j, $1\leq j\leq n$ *e as* σ*-álgebras* $\sigma\{B_j(t);t\geq 0\}$, $1\leq j\leq n$ *são independentes.*

Se $b(t,\omega) = (b_{ij}(t,\omega))_{i\leq m, j\leq n}$ sendo que $b_{ij} \in L^2_\Omega$ para todo i,j a integral estocástica $\int_0^T b(t)dB(t)$ é definida por

$$\int_0^T b(t)dB(t) = \left(\sum_{j=1}^n \int_0^T b_{1j}(t)dB_j(t), \ldots, \sum_{j=1}^n \int_0^T b_{mj}(t)dB(t) \right).$$

É claro que a integral estocástica definida acima é linear em b. E é imediato que $E[\int_0^T b(t)dB(t)] = (0,\ldots,0) = 0$.

Proposição 11 *Se $f,g \in M^2_\Omega$ então:*

1. $E\left[\int_0^T f(t)dB_i(t) \cdot \int_0^T g(t)dB_i(t)\right] = E\left[\int_0^T f(t)g(t)dt\right]$.
2. $E\left[\int_0^T f(t)dB_i(t) \cdot \int_0^T g(t)dB_j(t)\right] = 0$ *se* $i \neq j$.

Demonstração: O segundo item decorre imediatamente pela independência. O primeiro item demonstra-se usando a identidade $4fg = (f+g)^2 - (f-g)^2$.

Corolário 2 $E\left|\left|\int_0^T b(t)dB(t)\right|\right|^2 = E\left[\int_0^T ||b(t)||^2 dt\right]$.

Definição 23 *Seja $(Y_t)_t$ um processo estocástico de dimensão n e suponhamos que*

$$Y(t) - Y(0) = \int_0^t a(u)du + \int_0^t b(u)dB(u),$$

sendo $a = (a_1,\ldots,a_m)$, $a_j \in L^1_\Omega$ e $b = (b_{ij})_{ij}$, $i \leq m, j \leq n$, $b_{ij} \in L^2_\Omega$. Dizemos que $(Y_t)_t$ tem diferencial estocástica

$$dY_t = a(t)dt + b(t)dB(t).$$

A fórmula de Itô multi-dimensional.

Teorema 26 *Seja $u(x,t)$ contínua juntamente com suas derivadas $\frac{\partial u}{\partial x_i}$, $\frac{\partial^2 u}{\partial x_i \partial x_j}$ e $\frac{\partial u}{\partial t}$.*[3] *Seja (Y_t) um processo estocástico de dimensão m com diferencial*

$$dY_t = a(t)dt + b(t)dB(t).$$

Então $u(Y_t,t)$ tem diferencial estocástica $du(Y_t,t) =$

$$\left[u_t(Y_t,t) + \sum_{i=1}^m u_{x_i}(Y_t,t)a_i(t) + \frac{1}{2}\sum_{i,j=1}^m u_{x_i x_j}(Y_t,t)(bb^T)_{ij}(t) \right] dt + \sum_{l=1}^n \sum_{i=1}^m u_{x_i}(Y_t,t) b_{il}(t) dB_l(t).$$

[3] Se alem disso $\frac{\partial^2 u}{\partial x_i \partial t}$ fosse contínua então u seria C^2.

A fórmula de Itô multi-dimensional não será demonstrada. A demonstração é analoga ao caso uni-dimensional com o auxílio do seguinte lema:

Lema 10 *Se* $dX_t^1 = a_1(t)dt + b_1(t)dB(t)$ *e* $dX_t^2 = a_2(t)dt + b_2(t)dB(t)$ *então*

$$d(X_t^1 X_t^2) = X_t^1 dX_t^2 + X_t^2 dX_t^1 + \sum_{j=1}^{n} b_{1j} b_{2j} dt.$$

A demonstração do lema acima decorre da bilinearidade de $d(X_t^1 X_t^2)$ e do seguinte lema:

Lema 11 *Se* $i \neq j$, *então* $d(B_i B_j) = B_i dB_j + B_j dB_i$.

Demonstração: Seja $W_t = \frac{B_i(t)+B_j(t)}{\sqrt{2}}$. Então W_t é um movimento Browniano normalizado. Logo $W_t^2 = t + 2\int_0^t W_t dW_t$. Usando que $W_t^2 = \frac{B_i^2+B_j^2}{2} + B_i B_j = t + \int_0^t B_i dB_i + \int_0^t B_j dB_j + B_i B_j$ e que

$$\int_0^t W_t dW_t = \int_0^t B_i dB_i + \int_0^t B_j dB_j$$

obtemos o resultado.

Existência e unicidade de equações diferenciais estocásticas.

Seja $b(x,t) = (b_1(x,t), \ldots, b_n(x,t))$ e $\sigma(x,t) = (\sigma_{ij}(x,t))_{i \leq n, j \leq n}$. Seja $(X_t)_t$ um processo estocástico tal que

$$\begin{cases} dX(t) = b(X(t),t)dt + \sigma(X(t),t)dB(t), \\ X(0) = \xi_0 \text{ quase certamente.} \end{cases} \tag{1}$$

Dizemos que $(X(t))$ satisfaz a equação diferencial estocástica (1) com condição inicial ξ_0.

Teorema 27 *Suponhamos* b, σ *como acima, definidas e mensuráveis em* $\mathbb{R}^n \times [0,T]$. *Suponhamos que para todo* x, y, t:

$$\begin{array}{l} |b(x,t) - b(y,t)| + |\sigma(x,t) - \sigma(y,t)| \leq K|x-y| \text{ e} \\ |b(x,t)| + |\sigma(x,t)| \leq K(1+|x|). \end{array}$$

Suponhamos também que ξ_0 *é um vetor aleatório independente de* $\sigma\{B(s); s \leq T\}$. *Então a equação diferencial estocástica (1) possui uma e única solução.*

Não demonstrarei o teorema acima. A sua demonstração é similar à demonstração do teorema de existência de equações diferenciais ordinárias. O exemplo $dX_t = X_t^2 dt, X_0 = 1$ mostra que a hipótese de Lipschitz em b é fundamental para obtermos solução em $[0, T]$. E o exemplo $dX_t = 3X^{\frac{2}{3}} dt$ mostra a importância da hipótese de Lipschitz para a unicidade da solução.

O conceito de solução definido acima é as vezes chamado de solução forte. A solução fraca corresponde a dados b, σ encontrarmos $(\tilde{X}_t, \tilde{B}_t)$ definidos em $(\Omega, \mathcal{A})$ tais que $d\tilde{X}_t = b(\tilde{X}_t, t)dt + \sigma(\tilde{X}_t, t)d\tilde{B}_t$, $\tilde{X}_0 = \xi_0$. Duas soluções fracas são consideradas únicas se tem a mesma distribuição. Nas condições do teorema de existência de solução forte de um sistema de equações diferenciais estocásticas há unicidade fraca também.

Processos de Markov

Definição 24 *Seja $p(s, x, t, A) \geq 0$ definida para $0 \leq s \leq t < \infty$, $x \in \mathbb{R}^n$, $A \in \mathcal{B}(\mathbb{R}^n)$. Dizemos que p é uma função de transição Markoviana se:*

1. *$x \to p(s, t, x, A)$ é Boreleana;*
2. *$A \to p(s, x, t, A)$ é uma medida de probabilidades;*
3. *p satisfaz a equação de Chapman-Kolmogorov:*

$$p(s, x, t, A) = \int_{\mathbb{R}^n} p(\lambda, y, t, A) d\nu_{s,x,\lambda}(y), \forall 0 \leq s < \lambda < t.$$

Sendo $\nu_{s,x,\lambda}(A) = p(s, x, \lambda, A)$.

Teorema 28 *Seja p uma função de transição Markoviana. E seja π uma medida de probabilidades no $\mathbb{R}^n$. Existe então um processo estocástico $(X_t)_{t \geq s}$, $X_t : \Omega \to \mathbb{R}^n$ tal que: $P(X_s \in A) = \pi(A), \forall A \in \mathcal{B}(\mathbb{R}^n)$ e*

$$P(X_t \in A | \sigma\{X_s; a \leq s \leq b\}) = p(b, X_b, t, A), a \leq b < t.$$

Aplicando o teorema anterior para $\pi = \delta_x$, $x \in \mathbb{R}^n$ obtemos uma medida de probabilidade $P_{s,x}$ num espaço Ω e $E_{s,x}$ a esperança tais que

$$P_{s,x}(X_s = x) = 1;$$

$$P_{s,x}(X_{t+h} \in A | \mathcal{F}_t^s) = p(t, X_t(\omega), t + h, A), \text{ q.c.}$$

Uma coleção $(\Omega, \mathcal{F}, \mathcal{F}_t^s, X_t, P_{s,x})$ satisfazendo as propriedades acima, sendo p uma função de transição Markoviana é chamado de processo de Markov. Vale

$$E_{s,x}[f(X_t)] = \int f(y) d\nu_{s,x,t}(y)$$

e

$$E_{s,x}\left[f(X_{t+h}) | \mathcal{F}_t^s\right] = \int f(y) d\nu_{t,X_t,t+h}(y) = E_{t,X_t}[f(X_{t+h})].$$

Um processo de Markov é homogêneo se $p(s, x, t, A) = p(0, x, t - s, A)$.

Propriedade de Markov para difusões

Definição 25 *Uma difusão de Itô, é um processo estocástico $(X_s)_s$ que satisfaz à equação diferencial estocástica:*

$$\begin{cases} dX_s = b(X_s, s)ds + \sigma(X_s, s)dB_s, s \geq t \\ X_t = x. \end{cases}$$

Supomos que b, σ satisfazem as condições de Lipschitz para a existência e unicidade da solução da equação diferencial estocástica. A solução desse sistema é denotada $X_s^{t,x}, s \geq t$.

Seja $\mathcal{F}_t = \sigma\{B_s, s \leq t\}$. Então:

$$P\left(X_{0,x}(t) \in A|\mathcal{F}_s\right) = P\left(X(t) \in A|X_s\right) = p(s, X_s, A)$$

Gerador infinitesimal de uma difusão.

Seja $(X_t)_t$ uma difusão de Itô, homogênea no tempo. O gerador infinitesimal, A, de $(X_t)_t$ é

$$(Af)(x) = \lim_{t \to 0^+} \frac{E_x[f(X_t)] - f(x)}{t}, x \in \mathbb{R}^n.$$

O domínio de A é o conjunto das f tal que o limite acima existe para quase todo x.

Lema 12 *Seja $Y_t = Y_t^x$ tal que $dY_s = u(s)ds + v(s)dB_s, s \geq t$ e $Y_t = x$, sendo B um movimento Browniano de dimensão m, $u = (u_1, \ldots, u_n)$, $v = (v_{ij})_{ij}$. Seja $\tau : \Omega \to [0, \infty]$ um tempo de parada com esperança $E_x[\tau] < \infty$. Seja $f \in C_0^2(\mathbb{R}^n)$, u, v limitadas no suporte de f. Então*

$$E_x[f(Y_\tau)] = f(x) + E_x\left[\int_0^\tau \left(\sum_{i=1}^n u_i(s,\omega)\frac{\partial f}{\partial x_i}(Y_s) + \frac{1}{2}\sum_{i,j=1}^n (vv^t)_{ij}\frac{\partial^2 f}{\partial x_i \partial x_j}(Y_s)\right) ds\right].$$

Demonstração: Seja $Z_t = f(Y_t)$. Então

$$dZ = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(Y_t)dY_i + \frac{1}{2}\sum_{i,j=1}^n (vv^t)_{ij}\frac{\partial^2 f}{\partial x_i \partial x_j}(Y_s)ds =$$

$$\left(\sum_{i=1}^n \frac{\partial f}{\partial x_i}(Y_t)u_i(s) + \frac{1}{2}\sum_{i,j=1}^n (vv^t)_{ij}\frac{\partial^2 f}{\partial x_i \partial x_j}(Y_s)\right) ds + \sum_{i=1}^n\sum_{j=1}^m \frac{\partial f}{\partial x_i}(Y_t)v_{ij}(t)dB_j(t).$$

Logo

$$f(Y_\tau) = f(x) + \int_0^\tau \left[\sum_{i=1}^n \frac{\partial f}{\partial x_i}u_i(s) + \frac{1}{2}\sum_{i,j=1}^n (vv^t)_{ij}\frac{\partial^2 f}{\partial x_i \partial x_j}\right] ds +$$

$$+\sum_{i=1}^{n}\sum_{j=1}^{m}\int_0^{\tau}\frac{\partial f}{\partial x_i}(Y_s)v_{ij}(s)dB_j(s).$$

Tomando o valor esperado terminaremos a demonstração, se

$$E_x\left[\int_0^{\tau}\frac{\partial f}{\partial x_i}(Y_s)v_{ij}(s)dB_j(s)\right]=0.$$

Isto decorre do lema a seguir fazendo-se $t\to\infty$.

Lema 13 *Se $(X_t)_t$ é uma martingala contínua à direita então para todo tempo de parada τ, $(X_{\min\{t,\tau\}})_t$ é uma martingala.*

Corolário 3 *Se $dX_t = b(X_t)dt + \sigma(X_t)dB_t$, $X(0) = x$ e f é contínua com suporte compacto, então*

$$Af(x)=\sum_{i=1}^{n}b_i(x)\frac{\partial f}{\partial x_i}(x)+\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(\sigma\sigma^t)_{ij}(x)\frac{\partial^2 f}{\partial x_i\partial x_j}(x).$$

Demonstração: Basta escolher $\tau = t$ e calcular o limite no lema anterior.

Corolário 4 (Fórmula de Dynkin) *Seja $f\in C_0^2(\mathbb{R}^n)$, τ um tempo de parada com $E_x[\tau]<\infty$. Então $E_x[f(X_\tau)] = f(x) + E_x[\int_0^\tau Af(X_s)dx]$.*

Exemplo 5 *Seja $B = (B_1,\ldots,B_n)$ Browniano começando em $a\in\mathbb{R}^n$, $n\geq 2$. Seja $R > |a| = \sqrt{a_1^2+\ldots+a_n^2}$. Qual o valor esperado do tempo de saída, τ_K de $(B_t)_t$ da bola $B_R(0)$? Dado $k\geq 1$ apliquemos a fórmula de Dynkin para $X_t = B_t$, $\tau = \min\{\tau_K, k\}$ e $f(x) = |x|^2$ se $|x|\leq R$, f de suporte compacto C^2. Temos*

$$E_a[f(B_\tau)] = f(a) + E_a[\int_0^\tau\frac{\Delta f(B_s)}{2}ds] = |a|^2 + nE_a[\tau]\leq R^2.$$

Logo $E_x[\min\{\tau_K,k\}]\leq\frac{R^2-|a|^2}{n}$. Fazendo $k\to\infty$ obtemos que $E_x[\tau_K]<\infty$. Portanto podemos usar τ_K na fórmula de Dynkin obtendo $E_a[\tau_K]=\frac{R^2-|a|^2}{n}$. Suponhamos agora que $|b|>R$. Qual a probabilidade de começando em b o Browniano atingir $B_R(0)$? Seja k o primeiro instante de saída do anel $A_k = \{x; R<|x|<2^kR\}$. Seja $f = f_{nk}$, C^2 com suporte compacto tal que

$$f(x)=\begin{cases}-\log(|x|) & se \quad n=2\\[1em] |x|^{2-n} & se \quad n>2.\end{cases}\qquad R\leq|x|\leq 2^kR$$

Pode-se verificar que $\Delta f = 0$ se $x\in A_k$. Pela fórmula de Dynkin $E_b[f(B_{\alpha_k})] = f(b)$. Se $p_k = P_b(|B_{\alpha_k}| = R)$, $q_k = 1-p_k$ então se $n\neq 2$, $p_kR^{2-n} + q_k(2^kR)^{2-n} = |b|^{2-n}$. Fazendo $k\to\infty$ vem $\lim_{k\to\infty}p_k = \left(\frac{|b|}{R}\right)^{2-n}<1$. Logo $P_b(\tau_K<\infty) = \lim_{k\to\infty}p_k = \left(\frac{|b|}{R}\right)^{2-n}$. Dizemos que o movimento é transitivo. Para $n=2$ temos $P(\tau_K<\infty)=1$ (movimento recorrente).

Exemplo 6 *O Browniano resolve $dX_t = dB_t$, logo $\sigma = I$ e portanto $Af = \frac{1}{2}\Delta f$.*

Exemplo 7 (Gráfico do movimento Browniano) *Seja $B = B_1$,*
$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ *solução de*

$$\begin{cases} dX_1 = dt & X_1(0) = t_0 \\ dX_2 = dB & X_2(0) = x_0 \end{cases}$$

ou $dX = bdt + \sigma dB$, $X(0) = \begin{pmatrix} t_0 \\ X_0 \end{pmatrix}$, $b = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\sigma = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Então $Af = \frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}$.

A equação retrospectiva de Kolmogorov.

Seja $(X_t)_t$ uma difusão de Itô no $\mathbb{R}^n$ com gerador infinitesimal A. Aplicando a fórmula de Dynkin para $f \in C_0^2(\mathbb{R}^n)$ e $\tau = t$ obtemos que $u(t,x) = E_x[f(X_t)]$ é diferenciável em t e $\frac{\partial u}{\partial t} = E_x[Af(X_t)]$. Podemos no entanto obter uma equação envolvendo somente u.

Teorema 29 (Equação retrospectiva de Kolmogorov) *Defina $u(t,x) = E_x[f(X_t)]$. Então $\frac{\partial u}{\partial t} = Au$.*

Demonstração: Seja $g(x) = u(t,x)$. Temos

$$\frac{E_x[g(X_r)] - g(x)}{r} = \frac{1}{r}E_x\left[E_{X_r}\left[f(X_t)\right] - E_x\left[f(X_t)\right]\right] =$$

$$\frac{1}{r}E_x\left[E_x\left[f(X_{t+r})|\mathcal{F}_r\right] - E_x\left[f(X_t)|\mathcal{F}_t\right]\right] =$$

$$\frac{u(t+r,x) - u(t,x)}{r} \to \frac{\partial u}{\partial t}.$$

Logo $Au = \lim_{r\to 0} \frac{E_x[g(X_r)]-g(x)}{r} = \frac{\partial u}{\partial t}$.

Definição 26 (Resolvente) *Para $\alpha > 0$ e $g \in C_b(\mathbb{R}^n)$ definimos o resolvente R_α por*

$$(R_\alpha g)(x) = E_x\left[\int_0^\infty e^{-\alpha t} g(X_t)dt\right].$$

Lema 14 *$R_\alpha g \in C_b(\mathbb{R}^n)$.*

Demonstração: De $(R_\alpha g)(x) = \int_0^\infty e^{-\alpha t}E_x[g(X_t)]dt$ então a continuidade de $E_x[g(X_t)]$ em x (propriedade de Feller) garante a continuidade de $R_\alpha g$. QED

Teorema 30 *Para todo $\alpha > 0$ temos $R_\alpha(\alpha - A)f = f$ se $f \in C_0^2(\mathbb{R}^n)$ e se $g \in C_b(\mathbb{R}^n)$ então $R_\alpha g$ está no domínio de A e $(\alpha - A)R_\alpha g = g$.*

Demonstração: Suponhamos que $f \in C_0^2(\mathbb{R}^n)$. Então $R_\alpha(\alpha - A)f(x) = (\alpha R_\alpha f - R_\alpha A f)(x) =$

$$\alpha \int_0^\infty e^{-\alpha t} E_x[f(X_t)]dt - \int_0^\infty e^{-\alpha t} E_x[Af(X_t)]dt =$$

$$-e^{-\alpha t} E_x[f(X_t)]|_0^\infty + \int_0^\infty e^{-\alpha t} \frac{d}{dt} E_x[f(X_t)]dt - \int_0^\infty e^{-\alpha t} E_x[Af(X_t)]dt = f(x).$$

Suponhamos agora que $g \in C_b(\mathbb{R}^n)$. Segue da propriedade forte de Markov

$$E_x[R_\alpha g(X_t)] = E_x\left[E_{X_t}\left[\int_0^\infty e^{-\alpha s} g(X_s)ds\right]\right] =$$

$$E_x\left[E_x\left[\int_0^\infty e^{-\alpha s} g(X_{t+s}ds|\mathcal{F}_t\right]\right] = E_x\left[\int_0^\infty e^{-\alpha s} g(X_{t+s}ds\right] =$$

$$\int_0^\infty e^{-\alpha s} E_x[g(X_{t+s})]ds.$$

Definindo $u(s, x) = E_x[g(X_s)]$ vem

$$\frac{1}{t}\left(E_x[R_\alpha g(X_t)] - R_\alpha g(x)\right) = \int_0^\infty e^{-\alpha s} \frac{u(t+s, x) - u(s, x)}{t} ds \to$$

$$\int_0^\infty e^{-\alpha s} \frac{\partial}{\partial s} u(s, x)ds.$$

Logo $R_\alpha g$ pertence ao domínio do gerador infinitesimal. Assim vem que

$$(\alpha - A)R_\alpha g(x) = \int_0^\infty \alpha e^{-\alpha s} u(s, x)ds - \int_0^\infty e^{-\alpha s} \frac{\partial}{\partial s} u(s, x)ds = g(x).$$

Teorema 31 (Fórmula de Feynman-Kac.) *Seja $f \in C_0^2(\mathbb{R}^n)$ e $v(t, x) = E_x\left[e^{-\int_0^t q(X_s)ds} f(X_t)\right]$, $q \in C_b(\mathbb{R}^n)$. Então*

$$\frac{\partial v}{\partial t} = Av - q(x)v(t, x).$$

Demonstração: Seja $Y_t = f(X_t)$, $Z_t = exp[-\int_0^t q(X_s)ds]$. Vale que $dZ_t = -Z_t q(X_t)dt$. E portanto $d(Z_t Y_t) = Y_t dZ_t + Z_t dY_t$. Logo $v(t, x) = E_x[Y_t Z_t]$ é diferenciável em t. Portanto

$$\frac{1}{r}\left(E_x[v(t, X_t)] - v(t, x)\right) = \frac{1}{r} E_x\left[E_{X_r}[Z_t f(X_t)] - E_x[Z_t f(X_t)]\right] =$$

$$\frac{1}{r}E_x\left[E_x[f(X_{t+r}\exp[-\int_0^t q(X_{s+r})ds]|\mathcal{A}_r]-E_x[Z_tf(X_t)]\right]=$$

$$\frac{1}{r}E_x\left[Z_{t+r}\exp[\int_0^r q(X_udu]f(X_{t+r})-Z_tf(X_t)\right]=$$

$$\frac{1}{r}E_x\left[f(X_{t+r})Z_{t+r}-f(X_t)Z_t\right]+\frac{1}{r}E_x\left[f(X_{t+r}Z_{t+r}\left(e^{\int_0^r q(X_u)du}-1\right)\right]\to$$

$$\frac{\partial v}{\partial t}+q(x)v(t,x) \text{ se } r\to 0.$$

E terminamos a demonstração.

10 A fórmula de Girsanov.

Comecemos com uma definição.

Definição 27 *Uma medida de probabilidade Q em $(\Omega,\mathcal{A})$ é equivalente à P se $P(A)=0$ se e somente se $Q(A)=0$*

Uma versão do teorema de Girsanov é a seguinte:

Teorema 32 *Seja (X_t) um processo estocástico da forma*

$$X_t=x+\int_0^t \mu_s ds+\int_0^t \sigma_s dB_s \;\; 0\le t\le T$$

sendo $\mu\in L^1_\Omega, \sigma\in L^2_\Omega$. Suponhamos que $(m_t)_t\in L^1_\Omega$ e que existe $\theta\in L^2_\Omega, Ee^{\frac{1}{2}\int_0^T\theta_s^2ds}<\infty$, tal que $\sigma_t\theta_t=\mu_t-m_t$ $0\le t\le T$. Então existe uma medida de probabilidade Q equivalente à P tal que $\hat{B}_t=B_t+\int_0^t\theta_s ds$ $0\le t\le T$ é um movimento Browniano padronizado em $(\Omega,\mathcal{A},Q)$. O processo (X_t) pode ser escrito na forma

$$X_t=x+\int_0^t m_s ds+\int_0^t \sigma_s d\hat{B}_s \;\; 0\le t\le T.$$

Além disso, para toda variável aleatória $Z, E_Q(|Z|)<\infty$ vale $E_Q(Z)=E_P(Z\xi_T)$ sendo

$$\xi_t=e^{(-\int_0^t\theta_s dB_s-\frac{1}{2}\int_0^t\theta_s^2ds)} \;\; 0\le t\le T.$$

Ou seja, o processo de deslocamento de (X_t) pode ser alterado, continuando (X_t) a ser uma integral estocástica.

Exemplo 8 *Suponhamos que $dX_t=\mu_tdt+\sigma_tdB_t$ e que existe $\phi\in L^2_\Omega$ tal que $\sigma_t\phi_t=\mu_t$ e que $E\exp[\int_0^T\phi_t^2dt]<\infty$ e $E\exp[-2\int_0^T\phi_tdB_t-\int_0^T\phi_t^2dt]<\infty$. Então existe uma medida de martingala equivalente. Com efeito: a medida equivalente é a derivada de Radon-Nykodin*

$$\xi_T=\exp\left(-\int_0^T\phi_s dB_s-\frac{1}{2}\int_0^T\phi_s^2ds\right)$$

dada pelo teorema de Girsanov com $m_t=0$.

Teorema 33 *Suponhamos que $f \in L^2_\Omega$ e existe $a > 0$ tal que $Ee^{af_t^2} \leq C < \infty$ $0 \leq t \leq T$. Então*

$$Ee^{\int_{t_1}^{t_2} f(s)dB_s - \frac{1}{2}\int_{t_1}^{t_2} f^2(s)ds} = 1 \quad se\ 0 \leq t_1 \leq t_2 \leq T.$$

Somente demonstrarei o teorema para f limitada. Seja $Y_t = \int_{t_1}^{t} f(s)dB(s) - \frac{1}{2}\int_{t_1}^{t} |f(s)|^2 ds$. Seja $U_t = e^{Y_t}$. Temos $dU_t = e^{Y_t} f(t)dB(t)$. Logo $EU_{t_2} - 1 = E\left[\int_{t_1}^{t_2} e^{Y_t} f(t)dB(t)\right] = 0$. QED

Em geral se $f \in L^2_\Omega$ vale somente a desigualdade ≤ 1 no lugar da igualdade.

Teorema 34 (Girsanov) *Seja (X_t) um processo de Itô e $\phi \in L^2_\Omega$. Definamos*

$$\zeta_s^t(\phi) = \int_s^t \phi(u)dB_u - \frac{1}{2}\int_s^t \phi(u)^2 du,$$

$$\tilde{B}(t) = B(t) - \int_0^t \phi(s)ds \ e\ \tilde{P}(A) = \int_A e^{\zeta_0^T(\phi)dP},\ A \in \mathcal{A}.$$

Se $\tilde{P}(\Omega) = 1$ então $\tilde{B}$ é um movimento Browniano em $(\Omega, \mathcal{A}, \tilde{P})$ adaptado à $\tilde{B}$ e

$$X_t - X_s = \int_s^t (b_u + \sigma_u \phi_u)du + \int_s^t \sigma_u d\tilde{B}_u.$$

Demonstração: Para demonstrar que $\tilde{B}_t$ é um movimento Browniano é suficiente segundo o teorema abaixo demonstrar que:

a. $\tilde{E}[\tilde{B}_i(t) - \tilde{B}_i(s)|\mathcal{A}_s] = 0$ quase certamente se $t > s$;

b. $\tilde{E}[(\tilde{B}_i(t) - \tilde{B}_i(s))(\tilde{B}_j(t) - \tilde{B}_j(s))|\mathcal{A}_s] = \delta_{ij}(t-s)$ quase certamente para $t > s$.

Suporemos ϕ limitada e somente verificaremos o caso (a). A demonstração completa é longa e será omitida(ver Friedman páginas 161-162.) Notemos inicialmente que $d\left(\int_s^t f(u(dB(u) \cdot \int_s^t g(u)dB(u)\right) =$

$$\left(g(t) \cdot \int_s^t f(u)dB(u)\right) dB(t) + \left(f(t) \cdot \int_s^t g(s)dB(s)\right) dB(t) + (f(t) \cdot g(t))dt.$$

Logo

$$\left(\int_s^t f(u)dB(u)\right)\left(\int_s^t g(u)dB(u)\right) = \int_s^t f(u) \cdot g(u)du +$$

$$\int_s^t \left(g(u)\int_s^u f(x)dB(x)\right) dB(u) + \int_s^t \left(f(u)\int_s^u g(x)dB(x)\right) dB(u).$$

Tambem precisamos de $\exp[\zeta_s^t] = 1 + \int_s^t \exp[\zeta_s^u \phi(u)dB(u)$. Portanto

$$\tilde{E}[(\tilde{B}_i(t) - \tilde{B}_j(s))|\mathcal{A}_s] =$$

$$E\left[(B_i(t)-B_i(s))\exp[\zeta_s^t]|\mathcal{A}_s\right]-E\left[\int_s^t \phi_i(u)du\exp[\zeta_s^t]|\mathcal{A}_s\right]=$$

$$E\left[(B_i(t)-B_i(s))\left(1+\int_s^t \exp[\zeta_s^u]\phi(u)dB_u\right)|\mathcal{A}_s\right]-E\left[\int_s^t \phi_i(u)du\exp[\zeta_s^t]|\mathcal{A}_s\right]=0,$$

pois $E\left[\int_s^t \phi_i(u)du\exp[\zeta_s^u]|\mathcal{A}_s\right]=\int_s^t E\left[\phi_i(u)\exp[\zeta_s^t]|\mathcal{A}_s\right]du$.

Na demonstração acima usamos a seguinte caracterização do movimento Browniano:

Teorema 35 *Seja $X_t=(X_1(t),\ldots,X_n(t))$ um processo estocástico contínuo e seja $(\mathcal{A}_t)_t$ uma família crescente de σ álgebras, X_t $\mathcal{A}_t$ mensurável e tal que*

a) $E[X_t-X_s|\mathcal{A}_s]=0$ *quase certamente, $s>t$;*

b) $E[(X_i(t)-X_i(s))(X_j(t)-X_j(s))|\mathcal{A}_s]=\delta_{ij}(t-s)$ *quase certamente.*

Então $(X_t)_t$ é um movimento Browniano normalizado.

Equações diferenciais parciais elípticas e parabólicas.

10.1 Princípio do máximo para equações elípticas.

Seja L um operador diferencial definido por

$$Lu=\frac{1}{2}\sum_{i,j=1}^n a_{ij}(x)\frac{\partial^2 u}{\partial x_i\partial x_j}+\sum_{i=1}^n b_i(x)\frac{\partial u}{\partial x_i}+c(x)u,$$

sendo $u(x,t)$ definida para $x\in G$, G aberto e conexo do $\mathbb{R}^n$. O operador L é elíptico em x^0 se $(a_{ij}(x^0)_{ij}$ é positiva definida.

Lema 15 *Seja $(a_{ij}(x))_{ij}$ positiva semidefinida e $c(x)\le 0$ para todo $x\in D$, $D\subset\mathbb{R}^n$ aberto e convexo. Se $u\in C^2(D)$ e alcança um máximo positivo em $x^0\in D$ então $Lu(x^0)\le 0$.*

Demonstração: Seja $A=(a_{ij}(x^0)_{ij}$, $B=-(u_{x_ix_j}(x))_{ij}$. Então $\sum_{ij}a_{ij}(x^0)u_{x_ix_j}(x^0)\le 0$, com $u_{x_i}(x^0)=0$. De $c(x^0)u(x^0)\le 0$ vem $Lu(x^0)\le 0$. QED

Teorema 36 (Princípio fraco do máximo) *Suponha $(a_{ij}(x))$ positiva semidefinida em D, D aberto convexo e limitado, $c(x)\le 0$. E $\frac{1}{2}a_{11}(x)\lambda^2+b_1(x)\lambda>0$ para algum $\lambda>0$ e para todo $x\in D$. Se $u\in C^2(D)\cap C^0(\bar{D})$ e $Lu\ge 0$ então* $\max_{x\in\bar{D}}u(x)>0$ *implica* $\sup_{x\in D}u(x)\le\max_{x\in\partial D}u(x)$.

Demonstração: Suponhamos para obter uma contradição que $\max_{x\in\partial D}<\max_{x\in D}u(x)$. Seja $x^0\in\mathbb{R}^n$ tal que $x_1^0>x_1$ para todo $x\in\bar{D}$. Seja $k(x)=e^{\lambda x_1^0}-e^{\lambda x}$. É claro que $k(x)>0$ e $Lk(x)<0$ em D. Se $\epsilon>0$ é suficientemente pequeno, $v=u-\epsilon k$ satisfaz a $v(\bar{x})>\max_{x\in\partial D}v(x)$. Logo, v alcança um máximo positivo para algum $x^*\in D$. Pelo lema anterior, $Lv(x^*)\le 0$ e logo $Lu(x^*)\le\epsilon Lk(x^*)<0$, contradição com $Lu\ge 0$.

Definição 28 *Se existe $\mu > 0$ tal que $\sum_{i,j=1}^n a_{ij}(x)y_iy_j \geq \mu|y|^2$ para todo $x \in D$, $y \in \mathbb{R}^n$, dizemos que L é uniformemente elíptico em D.*

Teorema 37 (Princípio forte do máximo.) *Seja L uniformemente elíptico com coeficientes limitados nas partes compactas de D, e suponhamos $c(x) \leq 0$ em D. Se $u \in C^2(D)$ e $Lu \geq 0$ e u não é constante então u não tem máximo positivo em D.*

Não demonstrarei este teorema.

Problema de Dirichlet.

O problema de Dirichlet é o de encontrar u que resolve

$$\begin{cases} Lu = f(x) & x \in D \\ u(x) = \phi(x) & x \in \partial D. \end{cases}$$

Definição 29 *Uma barreira em $y \in \partial D$ é uma função contínua não-negativa, $w_y : \bar{D} \to \mathbb{R}_+$, que se anula somente em y e tal que $Lw_y \leq -1$. Se existir uma bola fechada $K = B[x^0, r]$ tal que $K \cap \bar{D} = \{y\}$ então a função*

$$w_y(x) = k\left(|x^0 - y|^{-p} - |x - y|^{-p}\right),$$

é uma barreira em y para k e p sufientemente grandes.

Teorema 38 *Suponhamos L uniformemente elíptico, $c \leq 0$, a_{ij}, b_i, c, f Lipschitz em $\bar{D}$. Se para todo $y \in \partial D$ existir uma barreira e $\phi \in C(\partial D)$ então existe uma e única solução $u \in C^2(D) \cap C^0(\bar{D})$ do problema de Dirichlet.*

Não demonstraremos este teorema.

Definição 30 (Operador parabólico.) *O operador diferencial*

$$Mu = \frac{1}{2}\sum_{i,j=1}^n a_{ij}(x,t)\frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^n b_i(x,t)\frac{\partial u}{\partial x_i} + c(x,t)u - \frac{\partial u}{\partial t},$$

definido para $(x,t) \in \mathbb{R}^n \times \mathbb{R}$, é uniformemente parabólico em Q se existir $\mu > 0$ tal que $(a_{ij}(x,t))_{ij}$ seja uniformemente positiva definida.

O problema de Cauchy.

Seja

$$Lu = \frac{1}{2}\sum_{i,j=1}^n a_{ij}(x,t)\frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^n b_i(x,t)\frac{\partial u}{\partial x_i} + c(x,t)u,$$

um operador elíptico no $\mathbb{R}^n$ para todo $t \in [0,T]$. Consideremos a equação parabólica

$$\begin{cases} Mu = Lu(x,t) - \frac{\partial u}{\partial t}(x,t) = f(x,t), \\ u(x,0) = \phi(x) \end{cases}$$

Este problema chama-se problema de Cauchy.

Teorema 39 *Suponha $(a_{ij}(x,t))_{ij}$ positiva semidefinida e suponhamos que existe $C > 0$ tal que*

$$\begin{cases} \sum_{ij} |a_{ij}(x,t)| \le C(1+|x|^2) \\ \sum_i |b_i(x,t)| \le C(1+|x|) \\ c(x,t) \le C. \end{cases}$$

Então o problema de Cauchy tem no máximo uma solução tal que $|u(x,t)| \le N(1+|x|^q)$ para algum $N, q > 0$.

Definição 31 *Uma solução fundamental do operador parabólico $L - \frac{\partial}{\partial t}$ no $\mathbb{R}^n \times [0,T]$ é uma função $\Gamma(x,t,\xi,\tau)$, $t > \tau$ satisfazendo a seguinte condição para toda $f \in C_0(\mathbb{R}^n)$:*

$$\begin{array}{l} u(x,t) = \int_{\mathbb{R}^n} \Gamma(x,t,\xi,\tau) f(\xi) d\xi; \\ Lu - \frac{\partial u}{\partial t} = 0, \tau < t \le T; \\ \lim_{t\downarrow\tau} u(x,t) = f(x). \end{array}$$

A existência de solução fundamental depende de três condições:
C1) $(a_{ij}(x,t))_{ij}$ é uniformente elíptico;
C2) Os coeficientes de L são contínuos limitados e $t \to a_{ij}(x,t)$ é uniformemente contínuo.
C3) Os coeficientes de L são uniformemente contínuos à Hölder com expoente α;

Teorema 40 *Se vale $C1 - C3$, existe uma solução fundamental $\Gamma(x,t,\chi,\tau)$ de $L - \frac{\partial \chi}{\partial t}$ satisfazendo a desigualdade*

$$|D_x^m \Gamma(x,t,\chi,\tau)| \le C(t-\tau)^{-(n+|m|/2)} \exp[-C\frac{|x-\chi|^2}{t-\tau}],$$

para $|m| = 0,1$. E se $|m| \le 2$, $D_x^m \Gamma$ é contínua. Além disso $L\Gamma = \Gamma_t$. Finalmente para $f \in C_b(\mathbb{R}^n)$,

$$\int_{\mathbb{R}^n} \Gamma(x,t,\chi,\tau) f(x) dx \to f(\chi), \;\; se\; t \downarrow \tau.$$

Em condições gerais

$$u(x,t) = \int \Gamma(x,t,\chi,0)\phi(\chi)d\chi - \int_0^t \int \Gamma(x,t,\chi,\tau) f(\chi,\tau) d\chi d\tau,$$

resolve o problema de Cauchy.

O adjunto formal de M.

O adjunto formal M^* de $M = L - \frac{\partial}{\partial t}$ ó operador

$$\begin{aligned}
&M^*v = L^*v + \tfrac{\partial v}{\partial t}, \text{ sendo}\\
&L^*v = \tfrac{1}{2}\textstyle\sum_{i,j=1}^n a_{ij}(x,t)\frac{\partial^2 v}{\partial x_i \partial x_j} + \sum_{i=1}^n b_i^*(x,t)\frac{\partial v}{\partial x_i} + c^*(x,t)v\\
&b_i^* = -b_i + \tfrac{1}{2}\textstyle\sum_{j=1}^n \frac{\partial a_{ij}}{\partial x_j},\\
&c^* = c - \textstyle\sum_{i=1}^n \frac{\partial b}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^n \frac{\partial^2 a_{ij}}{\partial x_i \partial x_j}.
\end{aligned}$$

Notemos que $M^*v = \frac{1}{2}\sum_{i,j}\frac{\partial^2}{\partial x_i \partial x_j}(a_{ij}v) - \sum_i \frac{\partial}{\partial x_i}(b_i v) + cv + v_t$.

Definição 32 *Uma solução fundamental de* $L^* + \frac{\partial}{\partial t}$ *é uma função* $\Gamma^*(x,t,\xi,\tau)$, $t < \tau$ *satisfazendo a condição: para todo* $g \in C_0(\mathbb{R}^n)$,

$$v(x,t) = \int_{\mathbb{R}^n} \Gamma^*(x,t,\xi,\tau)g(\xi)d\xi$$

satisfaz à $L^*v + \frac{\partial v}{\partial t} = 0$, $0 \le t < \tau$ *e* $\lim_{t\uparrow\tau} v(x,t) = v(x,\tau)$.

Teorema 41 *Se* $C1 - C3$ *e*

$$a_{ij}, \frac{\partial}{\partial x_i}a_{ij}, \frac{\partial^2}{\partial x_i \partial x_j}a_{ij}, \frac{\partial b_i}{\partial x_i}, c$$

são limitadas, então existe uma solução fundamental $\Gamma^*(x,t,\chi,\tau)$ *de* $L^* + \partial_t$ *e* $\Gamma(x,t,\chi,\tau) = \Gamma^*(\chi,\tau,x,t)$ *se* $t > \tau$.

Representação estocástica de soluções de equações diferenciais parciais.

Suponha D aberto, convexo limitado e L operador uniformemente elíptico. Suponhamos $c \le 0$, a_{ij}, b_i e c Lipschitz. Suponhamos também que a fronteira de D é de classe C^2. Então o problema de Dirichlet tem solução única dados f e ϕ tais que: f Lipschitz em $\bar{D}$, ϕ contínua em $\bar{D}$. Queremos representar u em termos de soluções de equações diferenciais estocásticas. Seja $\sigma(x) = \sqrt{(a_{ij}(x))}$. σ é Lipschitz em $\bar{D}$. Seja $dX_t = \sigma(X_t)dB_t + b(X_t)dt$. Seja τ o tempo de saída de D. Vale que

$$u(x) = E_x\left[\phi(X_\tau)\exp\left[\int_0^\tau c(X_s)dx\right]\right] - E_x\left[\int_0^\tau f(X_t)\exp\left[\int_0^t c(X_s)ds\right]dt\right].$$

Consideremos agora o problema de valor inicial:

$$\begin{array}{ll}
Lu + \frac{\partial u}{\partial t} = f(x,t) & Q = B \times [0,T)\\
u(x,T) = \phi(x) & x \in B\\
u(x,t) = g(x,t) & S = \partial B \times [0,T).
\end{array}$$

Suponhamos $(a_{ij})_{ij}$ uniformemente Lipschitz, c, f uniformemente Hölder contínua, g contínua em $\bar{S}$, ϕ contínua em $\bar{B}$ e $g(x,T) = \phi(x)$, $x \in \partial B$

Teorema 42 *Se ∂B é C^2 então existe solução única do problema acima:*

$$u(x,t) = E_{x,t}\left[g(X_\tau,\tau)e^{\int_t^\tau c(X_s,s)ds}\xi_{\tau<T}\right] +$$

$$E_{x,t}\left[\phi(X_T)e^{\int_t^\tau c(X_s,s)ds}\xi_{\tau=T}\right] - E_{x,t}\left[\int_t^\tau f(X_s,s)e^{\int_t^s c(X_\lambda,\lambda)d\lambda}ds\right].$$

Sendo τ acima o primeiro instante $\lambda \in [t,T)$ tal que $X(\lambda)$ deixa B se isto ocorrer e $\tau = T$ caso o contrário.

Seja

$$L_0 = \frac{1}{2}\sum_{i,j} a_{ij}(x,t)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_i b_i(x,t)\frac{\partial}{\partial x_i}$$

e $\Gamma_0^*(x,s,y,t)$ a solução fundamental de $L_0 + \partial_s$, $s < t$. Considerando $c = 0, f = 0$ acima obtemos que $u(x,t) = E_{x,t}\left[\phi(X_T)\right]$. Tambem vale que $u(x,t) = \int_{\mathbb{R}^n} \Gamma_0^*(x,t,y,T)\phi(y)dy$. Logo $E_{x,t}[\phi(X_T)] = \int_{\mathbb{R}^n} \Gamma_0^*(x,t,y,T)\phi(y)dy$. Analogamente para $0 \leq s < t$

$$\int_{\mathbb{R}^n} \Gamma_0^*(x,s,y,t)\phi(y)dy = E_{s,x}\phi(X_t) = \int_{\mathbb{R}^n} \phi(y)P_{x,s}(X_t \in dy).$$

Logo $p(s,x,t,A) = P_{s,x}(X_t \in A) = P(X_{x,s}(t) \in A)$ tem densidade $\Gamma_0^*(x,s,y,t)$. Vale que Γ_0^* satisfaz

$$\frac{\partial}{\partial s}\Gamma_0^*(x,s,y,t) + \frac{1}{2}\sum_{i,j} a_{ij}\frac{\partial^2}{\partial x_i \partial x_j}\Gamma_0^*(x,s,y,t) + \sum_i b_i(x,s)\frac{\partial}{\partial x_i}\Gamma_0^*(x,s,y,t) = 0.$$

Equação chamada de parabólica restropectiva. Tambem vale que Γ_0^* satisfaz em (y,t) a

$$-\partial_t\Gamma_0^*(x,s,y,t) + \frac{1}{2}\sum_{i,j}\frac{\partial^2}{\partial x_i \partial x_j}\left[a_{ij}(y,t)\Gamma_0^*\right] - \sum_i \frac{\partial}{\partial y_i}\left[b_i(y,t)\Gamma_0^*\right] = 0.$$

A equação acima é chamada de equação parabólica prospectiva.

11 Exercícios.

1. a) Seja $X : \Omega \to \mathbb{R}^n$ um vetor aleatório. Suponha que $E\left[|X|^p\right] < \infty$, $0 < p < \infty$. Então $P(|X| \geq r) \leq \frac{1}{r^p}E[|X|^p]$.
   b) Suponha que existe $k > 0$ tal que $M := E[e^{k|X|}] < \infty$. Conclua então que $P(|X| \geq r) \leq Me^{-kr}$.

2. Seja $(B_t)_{t\geq 0}$ um movimento Browniano. Então $(B_{t+s} - B_s)_{t\geq 0}$ é um movimento Browniano.

3. Seja $(B_t)_t$ um movimento Browniano bidimensional. Calcule a probabilidade de que $|B_t| \leq 1$.

4. Seja $(B_t)_t$ um movimento Browniano de dimensão n. Suponha que $K \subset \mathbb{R}^n$ tem medida nula. Então o tempo médio que B_t passa em K é zero.

5. Sejam $f(t), g(t) \in L^2_\Omega$. Suponha que quase certamente em $A \subset \Omega$ $f(t,\omega) = g(t,\omega)$ para todo t. Conclua então que $\int_0^T f(t,\omega)dB(\omega) = \int_0^T g(t,\omega)dB(\omega)$ para quase todo $\omega \in A$.

6. Verifique se $(X_t)_t$ é uma martingala:
a)$X_t = B_t + 4t$;
b)$X_t = t^2 B_t - 2\int_0^t sB_s dB_s$;
c)$X_t = B_1(t)B_2(t)$;
d)$X_t = B_t^3 - 3tB_t$.

7. Obtenha a diferencial de X_t:
a) $X_t = 2 + t + e^{B_t}$;
b) $X_t = B_1^2(t) + B_2^2(t)$;
c) $X_t = (B_1(t) + B_2(t) + B_3(t), B_2^2(t) - B_1(t)B_3(t)$

8. (a) Seja B_t um movimento Browniano 2-dimensional com $B_0 = x \neq (0,0)$. Mostre que $X_t = \log(|B_t|^2)$ não é uma martingala.
   (b) Mostre que $dX_t = b(t)dB_t$. O item anterior não contradiz o fato de que toda integral estocástica é uma martingala? Justifique.

9. Para $c, \alpha_1, \ldots, \alpha_n \in \mathbb{R}$ e $B(t)$ movimento Browniano n-dimensional mostre que a diferencial estocástica de $X_t = e^{ct+\sum_{j=1}^n \alpha_j B_j(t)}$ é $dX_t = (c + \frac{1}{2}\sum_{j=1}^n \alpha_j^2)X_t dt + X_t \cdot (\sum_{j=1}^n \alpha_j dB_j)$.

10. Calcule a diferencial estocástica de $\log(S_t)$, e^{S_t} e $\operatorname{sen}(S_t)$ se $dS_t = a(t)dt + b(t)dB(t)$.

11. Seja B_t um movimento Browninano de dimensão n. Suponha que X_t seja tal que
$$dX_t = v(t,\omega)dB_t(\omega) \text{ onde } v \in \mathbb{R}^n.$$
De um exemplo no qual X_t^2 não é uma martingala. Mostre que em geral
$$M_t = X_t^2 - \int_0^t |v_s|^2 ds$$
é uma martingala.

12. Seja B_t Browniano de dimensão n e seja $f : \mathbb{R}^n \to \mathbb{R}$ C^2. Use a fórmula de Itô para demonstrar que
$$f(B_t) = f(B_0) + \int_0^t \nabla f(B_s)ds + \frac{1}{2}\int_0^t \triangle f(B_s)ds.$$

Acima $\triangle = \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2}$ é o Laplaciano.

13. Uma classe importante de martingalas é obtida como a seguir:
Seja $dZ_t = \alpha dt + \beta dB_t, Z_0 = 0$ sendo α, β constantes e B_t um Browniano normalizado. Defina

$$M_t = \exp\left(Z_t - (\alpha + \frac{1}{2}\beta^2)t\right) = \exp\left(-\frac{1}{2}\beta^2 t + \beta B_t\right).$$

Mostre que $dM_t = \beta M_t dB_t$. Em particular conclua que M_t é uma martingala. Formule condições em u e v para que a generalização a seguir seja válida: Seja $dZ_t = u(t,\omega)dt + v(t,\omega)dB_t, Z_0 = 0$ uma integral estocástica com valores reais. Defina

$$M_t = \exp\left(Z_t - \int_0^t \left[u(s,\omega) + \frac{1}{2}vv^T(s,\omega)\right] ds\right).$$

Então $(M_t)_t$ é uma martingala.

14. Suponha que $g : \mathbb{R} \to \mathbb{R}$ é C^1 em $\mathbb{R}$ e C^2 exceto em $z_1, \ldots, z_N$ com $|g''| \leq M$ se $x \not\in \{z_1, \ldots, z_N\}$. Seja B_t Browniano normalizado. Demonstre que

$$g(B_t) = g(0) + \int_0^t g'(B_s)dB_s + \frac{1}{2}\int_0^t g''(B_s)ds.$$

Sugestão: Tome $f_k \in C^2$ convergindo uniformemente para g com $f_k' \to g'$ uniformemente $|f_k''| \leq M$ e $f_k'' \to g''$ exceto em $\{z_1, \ldots, z_N\}$

15. O que acontece se tentarmos aplicar a fórmula de Itô a $g(B_t)$, B_t Browniano normalizado e $g(x) = |x|$? Modifiquemos g para ser C^2:

$$g_\epsilon(x) = \begin{cases} |x| & \text{se} \quad |x| \geq \epsilon \\ \frac{1}{2}(\epsilon + \frac{x^2}{\epsilon}) & \text{se} \quad |x| < \epsilon \end{cases}$$

Aplique o exercício anterior para mostrar que

$$g_\epsilon(B_t) = g_\epsilon(0) + \int_0^t g_\epsilon'(B_s)dB_s + \frac{1}{2\epsilon} \cdot \lambda(\{s \in [0,t]; |B_s| < \epsilon\})$$

sendo $\lambda(F)$ a medida de Lebesgue de F. Demosntre que

$$\int_0^t g_\epsilon'(B_s)\chi_{|B_s|<\epsilon}dB_s = \int_0^t \frac{B_s}{\epsilon}\chi_{|B_s|<\epsilon}dB_s \to 0$$

em $L^2(P)$ quando ϵ tende a zero. No limite $\epsilon \to 0$ obtenha

$$|B_t| = \int_0^t \text{sign}(B_s)dB_s + L_t(\omega),$$

onde

$$L_t = \lim_{\epsilon \to 0} \frac{1}{2\epsilon} \lambda(\{s \in [0,t]; |B_s| < \epsilon\}).$$

L_t é chamado de tempo local do movimento Browniano em 0 e a fórmula acima é a fórmula de Tanaka para o movimento Browniano.

16. Verifique que os processos a seguir resolvem as respectivas equações estocásticas para o Browniano unidimensional:
a) $X_t = e^{B_t}$ resolve $dX_t = \frac{1}{2}X_t dt + X_t dB_t$
b) $X_t = \frac{B_t}{t+1}$ resolve $dX_t = -\frac{1}{t+1}X_t dt + \frac{1}{t+1}dB_t$
c) $X_t = \operatorname{sen}(B_t)$ resolve $dX_t = -\frac{1}{2}X_t dt + \sqrt{1 - X_t^2}dB_t$ para $t < T(\omega) = \inf\{s > 0; B_s \notin [-\frac{\pi}{2}, \frac{\pi}{2}]\}$
d) $X_t = (t, e^t B_t)$ resolve

$$\begin{bmatrix} dX_1 \\ dX_2 \end{bmatrix} = \begin{bmatrix} 1 \\ X_2 \end{bmatrix} dt + \begin{bmatrix} 0 \\ e^{X_1} \end{bmatrix} dB_t.$$

e) $X_t = (\cosh(B_t), \sinh(B_t))$ resolve

$$\begin{bmatrix} dX_1 \\ dX_2 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} dt + \begin{bmatrix} X_2 \\ X_1 \end{bmatrix} dB_t.$$

17. Um candidato natural para o movimento Browniano na elipse é o processo $X_t = (a \cos B_t, b \operatorname{sen} B_t)$. Mostre que X_t é solução da equação

$$dX_t = -\frac{1}{2}X_t dt + MX_t dB_t, \text{ sendo } M = \begin{bmatrix} 0 & -\frac{a}{b} \\ \frac{b}{a} & 0 \end{bmatrix}.$$

18. Seja B_t um movimento Browniano n-dimensional, $\alpha_1, \ldots, \alpha_n$ constantes. Resolva a equação diferencial estocástica

$$dX_t = rX_t dt + X_t \cdot (\sum_{k=1}^{n} \alpha_k dB_k(t)); X_0 > 0.$$

19. Mostre que existe uma única solução da equação uni-dimensional:

$$dX_t = \log(1 + X_t^2)dt + \chi_{\{X_t > 0\}} X_t dB_t, X_0 = a \in \mathbb{R}.$$

20. A ponte Browniana: Para $a, b \in \mathbb{R}$ considere a equação

$$dY_t = \frac{b - Y_t}{1 - t}dt + dB_t; 0 \leq t < 1, X_0 = a.$$

Verifique que

$$Y_t = a(1-t) + bt + (1-t)\int_0^t \frac{dB_s}{1-s}$$

resolve a equação e mostre que $\lim_{t \to 1} Y_t = b$.

21. Suponha que o preço de uma ação, S_t, siga o processo estocástico definido por
$$dS_t = \mu(S_t, t)dt + \sigma(S_t, t)dB_t, S(0) = x.$$
A ação paga dividendos à taxa $\delta(S_t, t)$ sendo portanto $D_t = \int_0^t \delta(S_u, u)du$ o processo de dividendos acumulativo. O processo de ganho é definido por $G_t = S_t + D_t$ e uma estratégia θ de compra e venda gera um processo de ganho $\int \theta dG$. Obtenha a equação a diferenças parciais deste modelo e forneça condições de regularidade suficientes para a solução de Feynman-Kac para o cálculo do ativo derivado que paga $g(S_T)$ no tempo T.

22. Suponha que $(S_t)_t$ é o preço da ação definido por
$$dS_t = \mu(S_t, t)dt + \sigma(S_t, t)dB_t; S(0) = x.$$
Há um ativo sem risco de preço $\beta_t = \exp\left[\int_0^t r(S_u, u)du\right]$ e um ativo derivado que paga $g(S_T)$ no tempo T. O processo cumulativo de dividendos é $(H_t)_t$, $H_t = \int_0^t h(S_u, u)du$. Uma estratégia (a, b) financia o ativo derivado se
$$a_tS_t+b_t\beta_t = a_0S_0+b_0\beta_0+\int_0^t a_u dS_u+\int_0^t b_u dB_u-H_t, t \leq T; a_TS_T+b_T\beta_T = g(S_T).$$
Obtenha a equação a diferenças parciais desse problema e aplique a fórmula de Feynman-Kac.

23. Suponha que a taxa instântanea de juros $(r_t)_t$ é limitada. O preço da ação é definido por $dS_t = \mu(S_t)dt + \sigma(S_t)dB_t$, μ e σ Lipschitz. Seja $D_t = \int_0^t \delta_t S_u du$ o processo acumulativo de dividendos, δ limitada.

    A. Suponha que há mercados para opções Européias de compra e de venda para a ação com processo de preços acima, com preço de exercício K e data de expiração T. Sejam C_0 e V_0 os respectivos preços das opções de compra e venda. Obtenha uma expressão explicita para P_0 em termos de C_0 na ausência de arbitragem e custos de transação.

    B. Suponha $r_t = 0,1$; $\delta_t = 0,08$; $T = 0,25$, $K = 50$;$S_0 = 45$. Se $C_0 = 3,75$ quanto é P_0?

    C. Suponha agora que o processo de dividendos (D_t) seja definido por $D_t = \int_0^t \delta_s \log(S_s)S_s ds$. Suponha que $\sigma(x) = \epsilon x$. Faça a parte (A) de novo. E calcule C_0 se $K = 35$; $S_0 = 40$; $\epsilon = 0,2$; $r_t = 0,1$ e $\delta_t = 0,08$. Justifique sua resposta.

24. Encontre o gerador infinitesimal das difusões a seguir:

    a. $dX_t = -cX_tdt + \beta dB_t$ sendo o Browniano uni-dimensional, $c > 0, \beta$ constantes. Este processo é chamado de processo de Ornstein-Uhlenbeck)

b. $dX_t = rX_t dt + dX_t dB_t$ (movimento Browniano geométrico)

c.

$$dY_t = \begin{bmatrix} dt \\ dX_t \end{bmatrix}$$

sendo $(X_t)_t$ como no item [a.]

d.

$$\begin{bmatrix} dX_1 \\ dX_2 \end{bmatrix} = \begin{bmatrix} 1 \\ dt \end{bmatrix} dt + \begin{bmatrix} 0 \\ e^{X_1} \end{bmatrix} dB_t$$

sendo o Browniano uni-dimensional.

e.

$$\begin{bmatrix} dX_1 \\ dX_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} dt + \begin{bmatrix} 1 & 0 \\ 0 & X_1 \end{bmatrix} \begin{bmatrix} dB_1 \\ dB_2 \end{bmatrix}$$

f. $X_t = (X_1, \ldots, X_n)$, sendo

$$dX_k(t) = rX_k dt + X_k \cdot \sum_{j=1}^{n} \alpha_{kj} dB_j; 1 \leq k \leq n$$

25. Encontre difusões de Itô com geradores como a seguir:

a. $Af(x) = f'(x) + f''(x)$;

b. $Af(t,x) = \frac{\partial f}{\partial t} + cx\frac{\partial f}{\partial x} + \frac{1}{2}\alpha^2 x^2 \frac{\partial^2 f}{\partial x^2}$ sendo c, α constantes.

c.

$$Af(x,y) = 2y\frac{\partial f}{\partial x} + \log(1+x^2+y^2)\frac{\partial f}{\partial y} + \frac{1}{2}(1+x^2)\frac{\partial^2 f}{\partial x^2} + x\frac{\partial^2 f}{\partial x \partial y} + \frac{1}{2}\frac{\partial^2 f}{\partial y^2}.$$

26. Seja B_t^x o movimento Browniano na reta começando em $x > 0$. Seja $\tau = \inf\{t > 0; B_t^x = 0\}$.

a) Demonstre que $E_x[\tau] = \infty$.

b) Demonstre que $\tau < \infty$ quase certamente.

27. Seja X_t movimento Browniano geométrico:

$$dX_t = rX_t dt + \alpha X_t dB_t, X_0 = x > 0$$

onde o Browniano é na reta, r, α são constantes.

a. Encontre o gerador A de X_t e compute $Af(x)$ para $f(x) = x^\gamma; x > 0$.

b. Se $r < \frac{1}{2}\alpha^2$ então $X \to 0$ se $t \to \infty$, quase certamente. Mas qual a probabilidade p de que X_t começando de $x < R$, tome o valor R? Use a fórmula de Dynkin com $f(x) = x^{\gamma_1}, \gamma_1 = 1 - \frac{2r}{\alpha^2}$ para demonstrar que $p = \left(\frac{x}{R}\right)^{\gamma_1}$.

c. Se $r > \frac{1}{2}\alpha^2$ então $X_t \to \infty$ quase certamente. Defina $\tau = \inf\{t > 0; X_t \geq R\}$. Use a fórmula de Dynkin com $f(x) = \log(x), x > 0$ para demonstrar que

$$E_x[\tau] = \frac{\log \frac{R}{x}}{r - \frac{1}{2}\alpha^2}.$$

Sugestão: Considere primeiro os tempos de saída de (ρ, R), $\rho > 0$ e faça então $\rho \to 0$. É necessário encontrar estimativas para $(1 - p(\rho)) \log \rho$, sendo

$$p(\rho) = P_x\left[X_t \text{ alcança o valor } R \text{ antes de } \rho\right],$$

que podem ser obtidas dos cálculos em (a) e (b).

28. **a)** Escreva em termos do movimento Browniano uma solução g do problema de Cauchy

$$\begin{cases} \frac{\partial g(t,x)}{\partial t} - \frac{1}{2}\Delta_x g(t,x) = 0 & \text{para } t > 0, x \in \mathbb{R}^n \\ g(0,x) = \phi(x) & \end{cases}$$

sendo $\phi \in C_b(\mathbb{R}^n)$ e $\alpha > 0$.

b) Seja $\psi \in C_b(\mathbb{R}^n)$ e $\alpha > 0$. Encontre uma solução, u, da equação

$$\left(\alpha - \frac{1}{2}\Delta\right) u = \psi \text{ em } \mathbb{R}^n.$$

Demonstre que a solução é única.

29. Seja $B_t = (B_1(t), \ldots, B_n(t))$ um movimento Browniano e seja F boreleano do $\mathbb{R}^n$. Demonstre que a duração total do tempo t que B_t passa em F é zero se e somente se F tem medida de Lebesgue nula.
Sugestão: Considere o resolvente R_α e faça $\alpha \to 0$.

Referências.

1. Bartle, R. G., The elements of integration, 1966, John Wiley & Sons.
2. D.Duffie,"Dynamic asset pricing theory", 1992, Princeton University Press
3. A. Friedman, "Stochastic differential equations and applications", vol.I, 1975, Academic Press, Nova Iorque

4. Bernt Øksendal, Stochastic Differential Equations, 3ª edição, Springer-Verlag

5. Philip Protter, Stochastic Integration and Differential Equations, A new approach, *Applications of Mathematics 21*, Springer-Verlag

12 Tabela de símbolos

Símbolo	Significado usual	Página da primeira ocorência
$\mathcal{P}(\Omega)$	partes de Ω	4
$A\Delta B$	diferença simétrica de A e B	4
Δ	laplaciano	4
$A_n \uparrow A$	A_n cresce para A	4
Ω	conjunto/espaço amostral	6
$\mathcal{A}$	σ-álgebra	6
$\cap_{t\in T}\mathcal{A}_t$	intersecção de σ-álgebras	6
$\sigma(\mathbf{U})$	σ-álgebra gerada por $\mathbf{U}$	6-Prop. 1
$\mathcal{B}_n$	σ-álgebra dos boreleanos do $\mathbb{R}^n$	6-Def. 4
$f^{-1}(B)$	imagem inversa do conjunto B	6-Def. 5
$\mathcal{B}$	boreleanos de $\mathbb{R}$	6-Def. 5
$\liminf, \limsup$	limite inferior, limite superior	7-Prop. 3
$\tilde{\mathcal{A}}$	completamento da σ-álgebra $\mathcal{A}$	8
$\int f d\mu$	integral de Lebesgue	9
$E[X]$	esperança de X	14
A_n i.v.	A_n infinitas vezes	15
$\{X_t\}_{t\geq 0}$	processo estocásitco	15-Def. 16
B_t	movimento Browniano	16
$E(X\|\mathcal{B})$	esperança condicional	16
τ	tempo de parada	17-Def. 17
$\int_0^T f(t)dB_t(\omega)$	integral estocástica	19
L^2_Ω		20
M^2_Ω		20
dX_t	diferencial estocástica	27
Af	gerador infinitesimal	36
R_α	resolvente	38
$C_b(\mathbb{R}^n)$	funções contínuas limitadas	38
Lu	operador diferencial elíptico	42
$\Gamma(x,t,\xi,\tau)$	solução fundamental	44-Def. 31