Universidade Federal do Rio Grande do Sul

Instituto de Matemática

Programa de Pós-Graduação em Matemática

Movimento Browniano, Integral de Itô e Introdução às Equações Diferenciais Estocásticas

Ricardo Misturini

Dissertação de Mestrado

Porto Alegre, 8 de Julho de 2010.

Dissertação submetida por Ricardo Misturini[1] como requisito parcial para a obtenção do grau de Mestre em Matemática pelo Programa de Pós-Graduação em Matemática do Instituto de Matemática da Universidade Federal do Rio Grande do Sul.

Professor Orientador:
Dr. Rafael Rigão Souza (UFRGS)

Banca Examinadora:
Dr. Artur Oscar Lopes (UFRGS)
Dr. Elismar da Rosa Oliveira (UFRGS)
Dr. Rogério Ricardo Steffenon (UNISINOS)

Data da Apresentação: 8 de Julho de 2010.

[1]Bolsista da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Capes

Agradecimentos

Primeiramente, agradeço à minha família (Mãe, Pai, Dé e Tiago) e à Cari. Não apenas por me incentivarem nos estudos. Por tudo. Saber que eu sempre tenho com quem contar torna tudo mais fácil. Dedico a vocês esta conquista.

Agradeço aos colegas, em especial ao Matheus, um baita parceiro, e a meu primo Rangel que acompanhou de perto o andamento deste trabalho, sempre disposto a dar uma opinião quando solicitei.

Agradeço também aos professores. Ao Rogério, meu primeiro orientador, que continua sempre interessado e prestativo. Ao Rafael, orientador super acessível e comprometido com minha formação. Também ao Luiz Fernando Carvalho da Rocha e ao Artur Lopes, por toda ajuda e incentivo.

A todos, muito obrigado.

Resumo

Este texto apresenta alguns dos elementos básicos envolvidos em um estudo introdutório das equações diferencias estocásticas. Tais equações modelam problemas a tempo contínuo em que as grandezas de interesse estão sujeitas a certos tipos de perturbações aleatórias. Em nosso estudo, a aleatoriedade nessas equações será representada por um termo que envolve o processo estocástico conhecido como Movimento Browniano. Para um tratamento matematicamente rigoroso dessas equações, faremos uso da Integral Estocástica de Itô. A construção dessa integral é um dos principais objetivos do texto. Depois de desenvolver os conceitos necessários, apresentaremos alguns exemplos e provaremos existência e unicidade de solução para equações diferenciais estocásticas satisfazendo certas hipóteses.

Palavras-chave: Movimento Browniano. Martingales. Integral Estocástica. Fórmula de Itô. Equações Diferenciais Estocásticas.

Abstract

This text presents some of the basic elements involved in an introductory study of stochastic differential equations. Such equations describe certain kinds of random perturbations on continuous time models. In our study, the randomness in these equations will be represented by a term involving the stochastic process known as Brownian Motion. For a mathematically rigorous treatment of these equations, we use the Itô Stochastic Integral. The construction of this integral is one of the main goals of the text. After developing the necessary concepts, we present some examples and prove existence and uniqueness of solution of stochastic differential equations satisfying some hypothesis.

Key-words: Brownian Motion. Martingales. Stochastic Integral. Itô's Formula. Stochastic Differential Equations.

Sumário

Capítulo 1

Introdução

Consideremos uma equação diferencial ordinária da seguinte forma

$$\begin{cases} x'(t) &= f(x(t),t) \quad t \in (0,T) \\ x(0) &= x_0 \end{cases}, \tag{1.1}$$

onde $f : \mathbb{R} \times [0,T] \to \mathbb{R}$ é uma função com um certo grau de regularidade e x_0 é uma constante fixada. Neste caso, a solução $x(\cdot)$ procurada é uma função $x : [0,T] \to \mathbb{R}$ satisfazendo as igualdades acima. É também comum escrever a equação (1.1) em sua forma diferencial:

$$\begin{cases} dx(t) &= f(x(t),t)dt \\ x(0) &= x_0 \end{cases}.$$

Ocorre que, em diversas situações, podemos estar interessados em um modelo que descreva também a ação de algum efeito aleatório que influencia o sistema, tal como algum tipo de perturbação ou ruído. Como uma possível descrição para algumas situações desse tipo, somos levados às chamadas *equações diferenciais estocásticas:*

$$\begin{cases} dX_t &= f(X_t,t)dt + g(X_t,t)dB_t \\ X_0 &= x_0 \end{cases}, \tag{1.2}$$

onde f, $g : \mathbb{R} \times [0,T] \to \mathbb{R}$ são funções e $(B_t)_{t\in[0,T]}$ é um processo estocástico conhecido como *movimento browniano*. Neste caso, a solução procurada é um processo estocástico $(X_t)_{t\in[0,T]}$.

Neste texto será dada uma interpretação matemática para a equação diferencial (1.2), a qual, na verdade, terá o sentido de uma equação integral

$$X_t = x_0 + \int_0^t f(X_s,s)ds + \int_0^t g(X_s,s)dB_s. \tag{1.3}$$

Para tanto, de imediato, somos conduzidos aos seguintes problemas:

1. Definir o movimento browniano $(B_t)_{t\in[0,T]}$;

2. Dar um sentido para a *integral estocástica* $\displaystyle\int_0^t g(X_s, s)dB_s$ que aparece em (1.3).

O problema 1 será abordado no Capítulo 2. Definiremos movimento browniano como um processo estocástico com trajetórias contínuas cujos incrementos são independentes e normalmente distribuídos, com média zero e variância igual ao correspondente incremento temporal.

Ainda no Capítulo 2, apresentaremos algumas propriedades dos processos estocásticos chamados de *martingales*, os quais incluem o movimento browniano. Para isso, precisaremos tratar, entre outras coisas, dos importantes conceitos de *filtração* e de *esperança condicional em relação a uma σ-álgebra*, os quais, portanto, não são supostos pré-requisitos para a leitura deste texto.

O problema 2 é assunto para o Capítulo 3, no qual definiremos a *Integral Estocástica de Itô*, para uma classe especial de processos que iremos definir. O resultado mais importante do capítulo é a chamada *Fórmula de Itô*, que é a correspondente regra da cadeia do cálculo estocástico.

No Capítulo 4, de posse das ferramentas construídas nos capítulos precedentes, estaremos aptos a apresentar as equações diferenciais estocásticas, em um contexto um pouco mais amplo do que em (1.2). Com uso da Fórmula de Itô, serão resolvidos alguns exemplos simples. Concluiremos o texto provando um teorema que, sob certas hipóteses, assegura existência e unicidade de solução para equações diferenciais estocásticas.

Capítulo 2

Movimento Browniano

2.1 Motivação e Definição

O movimento browniano, como fenômeno físico, foi proposto em 1827 pelo botânico inglês Robert Brown ao observar o movimento irregular de partículas microscópicas de pólen suspensas na água. Em 1905, A. Einstein apresentou uma descrição matemática do movimento como consequência de leis da física. Posteriormente, a teoria foi matematicamente formalizada por N. Wiener na década de 1920. Por isso, na literatura, o movimento browniano é também chamado de *Processo de Wiener*.

Para uma ideia intuitiva do movimento browniano, imagine um passeio aleatório simétrico no espaço euclidiano dando saltos infinitesimais com frequência infinita. Com isso em mente, como motivação heurística para a definição a ser dada a seguir, construímos um reticulado bidimensional conectando os vértices $\{(m\Delta x, n\Delta t) | m \in \mathbb{Z}, n \in \mathbb{N}\}$. Considere uma partícula que no tempo $t = 0$ se encontra na posição $x = 0$ e que, a cada tempo $n\Delta t$ move-se uma distância Δx para a direita, com probabilidade $1/2$, ou para a esquerda, com probabilidade $1/2$ (veja Figura 2.1, adiante). Chamemos de $p(m, n)$ a probabilidade de que a partícula esteja na posição $m\Delta x$ no instante $n\Delta t$. Temos, então, que

$$p(m, 0) = \begin{cases} 0 & m \neq 0 \\ 1 & m = 0 \end{cases}$$

e

$$p(m, n+1) = \frac{1}{2} p(m-1, n) + \frac{1}{2} p(m+1, n),$$

donde segue que

$$p(m,n+1)-p(m,n)=\frac{1}{2}(p(m+1,n)-2p(m,n)+p(m-1,n)).$$

Além disso, estaremos assumindo também que $(\Delta x)^2=\Delta t$, ou seja, que o passeio aleatório está sendo espacialmente reescalado pela raiz quadrada da escala do tempo. Então

$$\frac{p(m,n+1)-p(m,n)}{\Delta t}=\frac{1}{2}\left(\frac{p(m+1,n)-2p(m,n)+p(m-1,n)}{(\Delta x)^2}\right). \tag{2.1}$$

Agora fazemos $\Delta t\to 0$, $\Delta x\to 0$, $m\Delta x\to x$, $n\Delta t\to t$ mantendo a relação $(\Delta x)^2=\Delta t$. Então, de uma maneira intuitiva, $p(m,n)\to f(x,t)$, onde $f(x,t)$ é interpretada como a densidade de probabilidade da partícula estar na posição x no tempo t. Formalmente, no limite, a equação de diferenças (2.1) torna-se

$$f_t=\frac{1}{2}f_{xx}. \tag{2.2}$$

De fato, o segundo termo em (2.1) pode ser expresso como

$$\frac{1}{2}\left(\frac{\dfrac{p(m+1,n)-p(m,n)}{\Delta x}-\dfrac{p(m,n)-p(m-1,n)}{\Delta x}}{\Delta x}\right)$$

$$\cong\frac{1}{2}\frac{\dfrac{\partial p}{\partial x}(m,n)-\dfrac{\partial p}{\partial x}(m-1,n)}{\Delta x}\to\frac{1}{2}f_{xx}(x,t).$$

A expressão em (2.2) é a *equação de difusão*, também chamada de *equação do calor*. Podemos verificar que

$$f(x,t)=\frac{1}{\sqrt{2\pi t}}e^{-\frac{x^2}{2t}} \tag{2.3}$$

é uma, e de fato a única, solução de (2.2) com condição inicial $f(x,0)=\delta_0$. Note que (2.3) é a densidade de uma variável aleatória com distribuição $N(0,t)$.

Agora, usando o Teorema Central do Limite tornaremos mais rigorosa a passagem ao limite feita acima. Novamente, suponha que a partícula move-se para esquerda ou direita uma quantidade Δx com probabilidade $1/2$ para

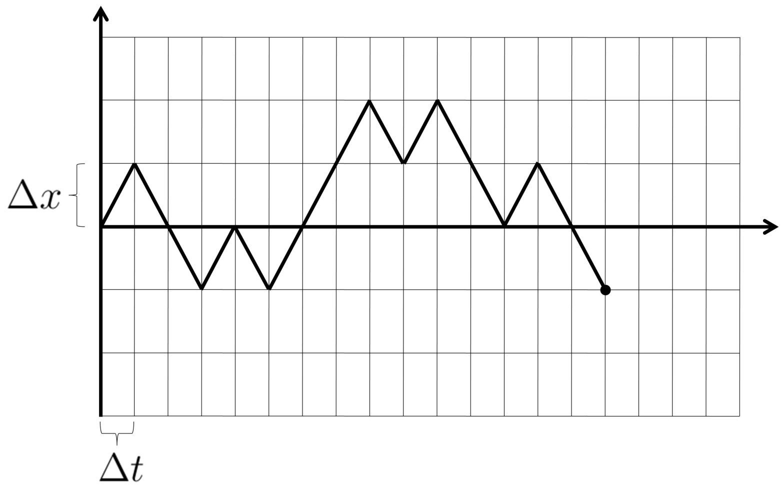

Figura 2.1: Passeio aleatório

cada direção. Denotemos por X_t a posição da partícula no instante $t = n\Delta t$. Seja Y_k, $k \in \mathbb{N}$, uma sequência de variáveis aleatórias independentes com

$$\left\{ \begin{array}{rcl} P(Y_k = 1) & = & 1/2 \\ P(Y_k = -1) & = & 1/2 \end{array} \right. \quad \text{para} \quad k = 1, 2, \dots \ .$$

Então $\mathbb{E}(Y_k) = 0$ e $Var(Y_k) = 1$. Definamos

$$S_n := \sum_{k=1}^{n} Y_k.$$

Podemos modelar X_t através da expressão

$$X_t = S_n \Delta x.$$

Como antes, impomos que $(\Delta x)^2 = \Delta t$. Então

$$X_t = S_n \Delta x = \left(\frac{S_n}{\sqrt{n}} \right) \sqrt{n} \Delta x = \left(\frac{S_n}{\sqrt{n}} \right) \sqrt{n \Delta t} = \left(\frac{S_n}{\sqrt{n}} \right) \sqrt{t}.$$

Portanto, pelo Teorema Central do Limite, temos

$$\begin{aligned}
\lim_{\substack{n \to \infty \\ t = n\Delta t}} \mathbb{P}(a < X_t < b) &= \lim_{n\to\infty} \mathbb{P}\left(\frac{a}{\sqrt{t}} \leq \frac{S_n}{\sqrt{n}} \leq \frac{b}{\sqrt{t}}\right) \\
&\overset{\text{TCL}}{=} \frac{1}{\sqrt{2\pi}} \int_{\frac{a}{\sqrt{t}}}^{\frac{b}{\sqrt{t}}} e^{-\frac{x^2}{2}} dx \\
&= \frac{1}{\sqrt{2\pi t}} \int_a^b e^{-\frac{u^2}{2t}} du.
\end{aligned}$$

Onde usamos a substituição $u = x\sqrt{t}$ na última igualdade. Assim, novamente, obtemos a distribuição $N(0,t)$.

Passemos agora à definição do Movimento Browniano. Inicialmente, consideremos o caso unidimensional.

Definição 2.1.1. *Um Movimento Browniano $(B_t)_{t\geq 0}$ em $\mathbb{R}$ é um processo estocástico $B(\cdot,\cdot) : \Omega \times [0,+\infty) \to \mathbb{R}$, tal que*

1. *Para quase todo $\omega \in \Omega$ fixado, $B(\omega,\cdot) : [0,+\infty) \to \mathbb{R}$ é uma função contínua e $B_0(\omega) = B(\omega,0) = 0$;*
2. *Os incrementos $B_t - B_s$, $0 \leq s < t$, são variáveis aleatórias normais de média zero e variância $t - s$, ou seja, $B_t - B_s \sim N(0,t-s)$;*
3. *Incrementos em intervalos de tempo disjuntos são independentes. Ou seja, as variáveis aleatórias $B_{t_4} - B_{t_3}$ e $B_{t_2} - B_{t_1}$ são independentes sempre que $0 \leq t_1 < t_2 \leq t_3 < t_4$, o mesmo valendo para n intervalos disjuntos, onde n é arbitrário.*

Obs. Usamos como notação $B_t(\omega) = B(\omega,t)$.

O item *3* estabelece que, para $s < t$, $B_t - B_s$ é independente do que aconteceu com o processo no intervalo $[0,s]$. Ou seja, se sabemos que $B_s = x_0$, nenhuma outra informação sobre os valores de B_τ para $\tau < s$ tem efeito sobre o nosso conhecimento da distribuição de probabilidade de $B_t - B_s$. Formalmente, isto significa que, se $t_0 < t_1 < \ldots < t_n < t$ então:

$$\mathbb{P}\left(B_t \leq x | B_{t_0} = x_0, B_{t_1} = x_1, \ldots, B_{t_n} = x_n\right) = \mathbb{P}(B_t \leq x | B_{t_n} = x_n). \quad (2.4)$$

Processos com esta característica são chamados de *Processos de Markov*.

Notemos também que, pelo item *2*, se $s < t$ temos

$$\begin{aligned}\mathbb{P}(B_t \leq x | B_s = x_0) &= \mathbb{P}(B_t - B_s \leq x - x_0)\\ &= \frac{1}{\sqrt{2\pi(t-s)}} \int_{-\infty}^{x-x_0} e^{-\frac{y^2}{2(t-s)}} dy\\ &= \frac{1}{\sqrt{2\pi(t-s)}} \int_{-\infty}^{x} e^{-\frac{(y-x_0)^2}{2(t-s)}} dy.\end{aligned}$$

Ou seja, dado que $B_s = x_0$ então $B_t \sim N(x_0, t-s)$.

As condições que definem $(B_t)_{t\geq 0}$ são suficientes para determinar todas as probabilidades associadas ao processo. Como sempre podemos associar cada $\omega \in \Omega$ com a função $t \mapsto B_t(\omega)$, isto significa que a lei do movimento browniano determina, de maneira única, uma medida no chamado *Espaço de Wiener*, $(W, \mathcal{F})$, onde

$$W = \{w : [0, \infty) \to \mathbb{R} \text{ contínua, com } w(0) = 0\}$$

e $\mathcal{F}$ é a σ-álgebra gerada pelos cilindros de dimensão finita. Entretanto, não é óbvio que um processo com estas características, de fato, existe. Este é o assunto da próxima seção.

Como consequência simples da definição, notemos uma propriedade importante do movimento browniano.

Proposição 2.1.2. *Se $(B_t)_{t\geq 0}$ é um movimento browniano, então para quaisquer tempos s e t temos $\mathbb{E}[B_t B_s] = \min\{s, t\}$.*

Demonstração. Suponhamos $0 \leq s \leq t$. Então

$$\begin{aligned}\mathbb{E}[B_s B_t] &= \mathbb{E}[B_s(B_s + B_t - B_s)]\\ &= \mathbb{E}[B_s^2] + \mathbb{E}[B_s(B_t - B_s)]\\ &= s + \mathbb{E}[B_s]E[B_t - B_s]\\ &= s + 0 = s = \min\{s, t\}.\end{aligned}$$

Na terceira igualdade usamos $Var[B_s] = \mathbb{E}[B_s^2] = s$, pois $B_s \sim N(0, s)$, e a independência entre B_s e $B_t - B_s$. □

2.2 Existência do Movimento Browniano

Existem várias maneiras de se construir o movimento browniano. Faremos em detalhes, na Seção 2.2.1, uma construção explícita que expressa o processo como uma série quase certamente uniformemente convergente de funções contínuas com coeficientes aleatórios. Na Seção 2.2.2, mostraremos, de maneira sucinta, como a existência do processo pode ser obtida através do *Teorema de Extensão de Kolmogorov* somado a um critério que garante a continuidade das trajetórias.

2.2.1 Uma construção explícita

Inicialmente, construiremos um movimento browniano $(B_t)_{t\in[0,1]}$ definido apenas para $t \in [0,1]$. O processo indexado por $[0,+\infty)$ será, posteriormente, construído a partir deste. Nosso objetivo é definir $(B_t)_{t\in[0,1]}$ como

$$B_t := \sum_{k=0}^{\infty} A_k \int_0^t h_k(s)ds$$

onde A_k é uma sequência de variáveis aleatórias independentes com distribuição $N(0,1)$ e as h_k são uma base ortonormal do espaço $L^2([0,1],\mathbb{R})$. O primeiro passo, é definir as funções h_k.

Definição 2.2.1. *As funções de Haar $h_k : [0,1] \to \mathbb{R}$, $k = 0,1,2,\ldots$, são definidas por:*

$$h_0(t) = 1 \qquad 0 \le t \le 1; \tag{2.5}$$

$$h_1(t) = \begin{cases} 1, & 0 \le t \le \frac{1}{2} \\ -1, & \frac{1}{2} \le t \le 1 \end{cases}; \tag{2.6}$$

e, para $2^n \le k < 2^{n+1}$, $n = 1,2,\ldots$,

$$h_k(t) = \begin{cases} 2^{n/2}, & \frac{k-2^n}{2^n} \le t \le \frac{k-2^n+1/2}{2^n} \\ -2^{n/2}, & \frac{k-2^n+1/2}{2^n} < t \le \frac{k-2^n+1}{2^n} \\ 0, & \text{demais casos} \end{cases}. \tag{2.7}$$

Lema 2.2.2. *As funções de Haar $\{h_k\}_{k=0}^{\infty}$ formam um base ortonormal do espaço de Hilbert $L^2([0,1],\mathbb{R})$.*

Demonstração. Fixado k, seja n tal que $2^n \le k < 2^{n+1}$. Temos que

$$\int_0^1 h_k^2(t)dt = \int_{\frac{k-2^n}{2^n}}^{\frac{k-2^n+1/2}{2^n}} 2^n dt + \int_{\frac{k-2^n+1/2}{2^n}}^{\frac{k-2^n+1}{2^n}} 2^n dt = 2^n\left(\frac{1}{2^{n+1}} + \frac{1}{2^{n+1}}\right) = 1.$$

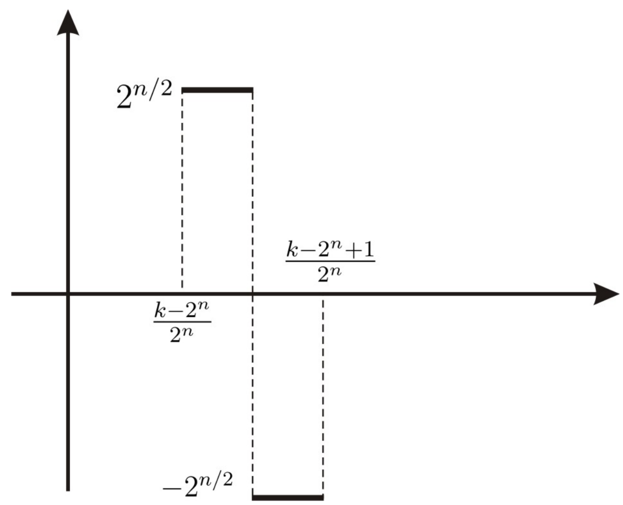

Figura 2.2: Gráfico da k-ésima função de Haar, $2^n \leq k < 2^{n+1}$

Seja agora $j > k$. Se $j < 2^{n+1}$ os suportes de h_k e h_j são disjuntos, portanto $h_k h_j \equiv 0$. Se, por outro lado, $2^{n+1} \leq j$ então a função h_k é constante no suporte de h_j de modo que

$$\int_0^1 h_k(t)h_j(t)dt = \pm 2^{n/2} \int_0^1 h_j(t)dt = 0.$$

Para completar a demonstração, dada uma função $f \in L^2$ tal que o produto interno $\int_0^1 f(t)h_k(t) = 0$, para todo k, devemos mostrar que $f = 0$ no sentido de L^2. De fato, isto significa que o complemento ortogonal das funções de Haar é o espaço nulo e, consequentemente, este conjunto forma um sistema ortonormal completo no espaço de Hilbert $L^2([0,1],\mathbb{R})$. Para mais detalhes referentes a conjuntos ortonormais completos em espaços de Hilbert veja, por exemplo, o Capítulo 10 de Bachman [1]. Vejamos:

$$k = 0 \Rightarrow \int_0^1 f(t)dt = 0; \tag{2.8}$$

$$k = 1 \Rightarrow \int_0^{\frac{1}{2}} f(t)dt - \int_{\frac{1}{2}}^1 f(t)dt = 0 \Rightarrow \int_0^{\frac{1}{2}} f(t)dt = \int_{\frac{1}{2}}^1 f(t)dt. \tag{2.9}$$

Logo, por (2.8) e (2.9), temos que

$$0 = \int_0^1 f(t)dt = \int_0^{\frac{1}{2}} f(t)dt + \int_{\frac{1}{2}}^1 f(t)dt = 2\int_0^{\frac{1}{2}} f(t)dt.$$

Portanto, $\int_0^1 f(t)dt = \int_0^{\frac{1}{2}} f(t)dt = \int_{\frac{1}{2}}^1 f(t)dt = 0$. Indutivamente, mostra-se que

$$\int_{\frac{k}{2^{n+1}}}^{\frac{k+1}{2^{n+1}}} f(t)dt = 0$$

para todo $0 \leq k < 2^{n+1}$ e para todo n. Logo, $\int_r^s f(t)dt = 0$ para todos os racionais $0 \leq r \leq s \leq 1$ diádicos, isto é, da forma $k/2^n$. Como os intervalos da forma $[r, s]$, com r e s diádicos, geram a σ-álgebra de Borel do intervalo $[0, 1]$, concluímos que

$$\int_B f(t)dt = 0 \tag{2.10}$$

para todo boreliano B de $[0, 1]$. Considere os conjuntos da forma

$$A_n = \{t \in [0, 1] : f(t) > 1/n\}.$$

Note que $\bigcup_{n=1}^{\infty} A_n = A \equiv \{t \in [0, 1] : f(t) > 0\}$. Denotemos por $Leb(A)$ a medida de Lebesgue do conjunto A. Para cada n, temos, por (2.10), que

$$0 = \int_{A_n} f(t)dt \geq \frac{1}{n} Leb(A_n) \geq 0.$$

Logo, $Leb(A_n) = 0$ para todo n e, consequentemente, $Leb(A) = 0$. Analogamente, se mostra que $Leb(\{t \in [0, 1] : f(t) < 0\}) = 0$. Portanto, f é nula q.t.p., conforme desejávamos. □

As *funções de Schauder* são as integrais das funções de Haar.

Definição 2.2.3. *Para $k = 0, 1, \ldots$, a k-ésima função de Schauder é definida por:*

$$s_k(t) = \int_0^t h_k(y)dy, \tag{2.11}$$

onde as h_k são as funções de Haar dadas em (2.5), (2.6) *e* (2.7).

Na Figura 2.3 é mostrado o gráfico da k-ésima função de Schauder. Podemos ver que

$$\max_{0 \leq t \leq 1} |s_k(t)| = 2^{-\frac{n}{2}-1}, \text{ se } 2^n \leq k < 2^{n+1}. \tag{2.12}$$

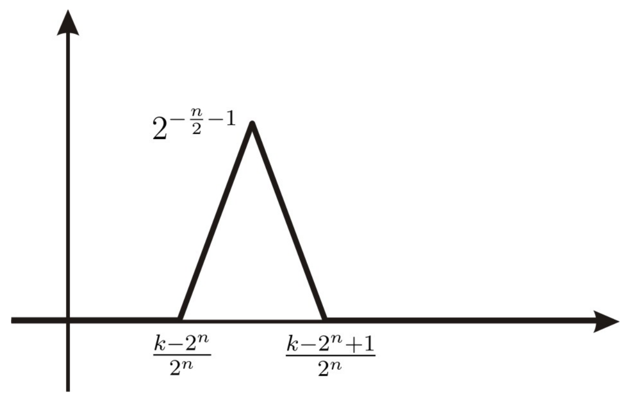

Figura 2.3: Gráfico da k-ésima função de Schauder, $2^n \leq k < 2^{n+1}$

Seja $\{A_k\}_{k=0}^{\infty}$ uma sequência de variáveis aleatórias independentes com distribuição $N(0,1)$ definidas em algum espaço de probabilidade. Definiremos o processo $(B_t)_{t\in[0,1]}$ por

$$B_t := \sum_{k=0}^{\infty} A_k s_k(t). \tag{2.13}$$

Primeiramente, para que isto faça sentido, precisamos analisar a convergência desta série.

Lema 2.2.4. *Sejam $\{a_k\}_{k=0}^{\infty}$ uma sequência de números reais e $\{s_k\}_{k=0}^{\infty}$ as funções de Haar definidas em* (2.11)*. Defina*

$$y_j = \max_{2^j \leq k < 2^{j+1}} |a_k| \tag{2.14}$$

e suponha que

$$\sum_{j=0}^{\infty} y_j 2^{-\frac{j}{2}-1} < \infty. \tag{2.15}$$

Então a série

$$x(t) = \sum_{k=0}^{\infty} a_k s_k(t), \tag{2.16}$$

definida em $[0,1]$, converge uniformemente para uma função contínua de t.

Demonstração. É suficiente mostrar que

$$\lim_{m,n\to\infty} \left| \sum_{k=m}^{m+n} a_k \max_{0\leq t\leq 1} s_k(t) \right| = 0. \tag{2.17}$$

Notando que para $2^j \leq k < 2^{j+1}$ as funções s_k têm suportes disjuntos e usando (2.12) e (2.14) obtemos

$$\left| \sum_{k=2^j}^{2^{j+1}-1} a_k \max_{0 \leq t \leq 1} s_k(t) \right| \leq y_j 2^{-\frac{j}{2}-1}. \tag{2.18}$$

Então, agrupando os termos em (2.17) de acordo com (2.18) e usando a desigualdade triangular, temos que, para $m > 2^N$, a soma de Cauchy em (2.17) é menor que $\displaystyle\sum_{j=N}^{\infty} y_j 2^{-\frac{j}{2}-1}$, a qual converge para zero, quando $N \to \infty$, de acordo com (2.15). Então, a série em (2.16) converge uniformemente e, como cada parcela é contínua, concluímos que $x(\cdot)$ é uma função contínua de t. □

Lema 2.2.5. *Seja* $\{A_k\}_{k=0}^{\infty}$ *uma sequência de variáveis aleatórias independentes com distribuição* $N(0,1)$*. Defina*

$$Y_j = \max_{2^j \leq k < 2^{j+1}} |A_k|. \tag{2.19}$$

Então

$$\mathbb{P}\left(\sum_{j=0}^{\infty} Y_j 2^{-\frac{j}{2}-1} < \infty \right) = 1.$$

Demonstração.

$$\begin{aligned}
\mathbb{P}(|A_k| > x) &= \frac{2}{\sqrt{2\pi}} \int_x^{\infty} e^{-\frac{u^2}{2}} du \\
&= \frac{2}{\sqrt{2\pi}} \int_x^{\infty} e^{-\frac{u^2}{4}} e^{-\frac{u^2}{4}} du \\
&\leq \frac{2}{\sqrt{2\pi}} e^{-\frac{x^2}{4}} \int_x^{\infty} e^{-\frac{u^2}{4}} du \leq K e^{-\frac{x^2}{4}}
\end{aligned}$$

onde K é uma constante. Em particular, para $x = 4\sqrt{\log k}$, temos que

$$\mathbb{P}\left(|A_k| > 4\sqrt{\log k} \right) \leq K e^{-4 \log k} = K \frac{1}{k^4}.$$

Como $\sum \frac{1}{k^4}$ converge, temos, pelo Lema de Borel-Cantelli (veja, por exemplo, James [7], página 200), que

$$\mathbb{P}\left(|A_k| > 4\sqrt{\log k} \text{ para infinitos valores de } k \right) = 0.$$

Ou seja, com probabilidade um, apenas um número finito dos A_k excede $4\sqrt{\log k}$ e, consequentemente, apenas um número finito dos $Y_j = \max_{2^j \leq k < 2^{j+1}} |A_k|$ excede $y_j := 4\sqrt{\log(2^{j+1})} = 4\sqrt{j+1}\sqrt{\log 2}$. Mas, aplicando o teste da razão, vemos que $\sum y_j 2^{-\frac{j}{2}-1} < \infty$. Isto estabelece, com probabilidade um, a convergência da série $\sum Y_j 2^{-\frac{j}{2}-1}$. $\square$

Juntos, os Lemas 2.2.4 e 2.2.5 garantem que a expressão (2.13) define um processo $(B_t)_{t\in[0,1]}$ com trajetórias contínuas, quase certamente. Além disso, é imediato que $B_0 = 0$. Para provar que $(B_t)_{t\in[0,1]}$ é um movimento browniano usaremos ainda o seguinte resultado:

Lema 2.2.6. *Sejam s_k, $k = 0, 1, \ldots$, as funções de Schauder definidas em* (2.11). *Então:*

$$\sum_{k=0}^{\infty} s_k(s) s_k(t) = \min\{s, t\}.$$

Demonstração. Para $0 \leq t \leq 1$ definimos

$$\phi_t(x) = \begin{cases} 1 & 0 \leq x \leq t \\ 0 & t < x \leq 1 \end{cases}.$$

Suponhamos $s \leq t$. Então, temos que

$$\int_0^1 \phi_s(x)\phi_t(x)dx = s.$$

Mas esta expressão é o produto interno entre ϕ_s e ϕ_t no espaço $L^2[0,1]$. Portanto, pelo Lema 2.2.2, podemos escrever

$$s = \int_0^1 \phi_s(x)\phi_t(x)dx = \sum_{k=0}^{\infty} a_k b_k,$$

onde os coeficientes a_k e b_k são as coordenadas das funções ϕ_s e ϕ_t na base formada pelas funções de Haar, ou seja

$$a_k = \int_0^1 \phi_s(x)h_k(x)dx = \int_0^s h_k(x)dx = s_k(s);$$

$$b_k = \int_0^1 \phi_t(x)h_k(x)dx = \int_0^t h_k(x)dx = s_k(t).$$

$\square$

Teorema 2.2.7. *A expressão* $B_t := \sum_{k=0}^{\infty} A_k s_k(t)$ *define um movimento browniano* $(B_t)_{t\in[0,1]}$.

Demonstração. Já vimos, através dos Lemas 2.2.4 e 2.2.5, que as trajetórias de $(B_t)_{t\in[0,1]}$ são contínuas quase certamente e iniciadas em zero. Precisamos mostrar, ainda, que, para $0 = t_0 < t_1 < \ldots < t_n \leq 1$ os incrementos $\{B_{t_j} - B_{t_{j-1}}\}_{j=1}^n$ são independentes, normalmente distribuídos com média zero e variância $t_j - t_{j-1}$. Para isso, consideraremos os vetores somas parciais

$$\begin{aligned}&\left(B_{t_1}^m - B_{t_0}^m, \ldots, B_{t_n}^m - B_{t_{n-1}}^m\right)\\ &= \left(\sum_{k=0}^{m} A_k s_k(t_1) - \sum_{k=0}^{m} A_k s_k(t_0), \ldots, \sum_{k=0}^{m} A_k s_k(t_n) - \sum_{k=0}^{m} A_k s_k(t_{n-1})\right)\end{aligned}$$

e mostraremos que as função características destes vetores, a saber

$$\varphi_m(\lambda_1, \ldots, \lambda_n) = \mathbb{E}\left[e^{i\sum_{j=1}^n \lambda_j\left(\sum_{k=0}^m A_k s_k(t_j) - \sum_{k=0}^m A_k s_k(t_{j-1})\right)}\right], \tag{2.20}$$

convergem pontualmente para

$$\varphi(\lambda_1, \ldots, \lambda_n) = \prod_{j=1}^{n} e^{-\frac{1}{2}\lambda_j^2(t_j - t_{j-1})} \tag{2.21}$$

que é a função característica de um vetor aleatório onde as componentes são independentes com distribuição $N(0, t_j - t_{j-1})$. Por simplicidade, consideremos o caso $n = 2$, o caso geral segue de maneira similar. Fazendo $n = 2$ em (2.20), lembrando que $s_k(t_0) = s_k(0) = 0$, obtemos

$$\begin{aligned}
\varphi_m(\lambda_1, \lambda_2) &= \mathbb{E}\left[e^{i\left[\lambda_1 \sum_{k=0}^m A_k s_k(t_1) + \lambda_2\left(\sum_{k=0}^m A_k s_k(t_2) - \sum_{k=0}^m A_k s_k(t_1)\right)\right]}\right]\\
&= \mathbb{E}\left[e^{\sum_{k=0}^m iA_k[(\lambda_1-\lambda_2)s_k(t_1) + \lambda_2 s_k(t_2)]}\right]\\
\substack{\text{por}\\ \text{independência}} \quad &= \prod_{k=0}^{m} \mathbb{E}\left[e^{iA_k[(\lambda_1-\lambda_2)s_k(t_1)+\lambda_2 s_k(t_2)]}\right]\\
\substack{\text{como}\\ A_k \sim N(0,1)} \quad &= \prod_{k=0}^{m} e^{-\frac{1}{2}[(\lambda_1-\lambda_2)s_k(t_1)+\lambda_2 s_k(t_2)]^2}\\
&= e^{-\frac{1}{2}\sum_{k=0}^m[(\lambda_1-\lambda_2)s_k(t_1)+\lambda_2 s_k(t_2)]^2}\\
&= e^{-\frac{1}{2}\left[(\lambda_1-\lambda_2)^2\sum_{k=0}^m s_k^2(t_1) + 2(\lambda_1-\lambda_2)\lambda_2\sum_{k=0}^m s_k(t_1)s_k(t_2) + \lambda_2^2\sum_{k=0}^m s_k^2(t_2)\right]}.
\end{aligned}$$

Fazendo $m \to \infty$ e usando o Lema 2.2.6, obtemos:

$$\begin{aligned}
\lim_{m\to\infty} \varphi_m(\lambda_1, \lambda_2) &= e^{-\frac{1}{2}\left[(\lambda_1-\lambda_2)^2 t_1 + 2(\lambda_1-\lambda_2)\lambda_2 t_1 + \lambda_2^2 t_2\right]} \\
&= e^{-\frac{1}{2}\lambda_1^2 t_1} e^{-\frac{1}{2}\lambda_2^2 (t_2 - t_1)}. \qquad (2.22)
\end{aligned}$$

Note que (2.22) é igual ao segundo termo de (2.21) com $n = 2$, conforme desejávamos. □

Tendo mostrado a existência do movimento browniano com tempo variando no intervalo $[0, 1]$, o próximo passo é estender o resultado para o intervalo $[0, \infty)$. Basicamente, a ideia consiste em "emendar" uma sequência de movimentos brownianos independentes definidos em $[0, 1]$. Mais formalmente, pelo resultado já provado, podemos construir uma sequência de movimentos brownianos independentes $\{(B_t^{(n)})_{t\in[0,1]}\}_{n=0}^{\infty}$. De fato, basta tomar uma sequência de variáveis aleatórias $N(0, 1)$ independentes, reindexá-la de modo a obter uma família enumerável de sequências e, a partir de cada sequência definir um movimento browniano usando o procedimento descrito anteriormente. Em seguida, definimos o processo $(B_t)_{t\in[0,\infty)}$ da seguinte maneira:

$$\begin{aligned}
B_t &= B_t^{(0)} && \text{se } 0 \le t \le 1 \\
&= B_1^{(0)} + B_{t-1}^{(1)} && \text{se } 1 < t \le 2 \\
&= B_1^{(0)} + B_1^{(1)} + B_{t-2}^{(2)} && \text{se } 2 < t \le 3 \\
&\ \vdots \\
&= B_1^{(0)} + \ldots + B_1^{(n-1)} + B_{t-n}^{(n)} && \text{se } n < t \le n+1.
\end{aligned}$$

Verifiquemos que $(B_t)_{t\ge 0}$ é, de fato, um movimento browniano. É imediato que $B_0 = 0$. A continuidade quase certa das trajetórias também é clara, pois a interseção enumerável de conjuntos com probabilidade um é também um conjunto com probabilidade um e, além disso, a continuidade nos naturais, onde poderia ocorrer problemas, é evidente a partir da construção, visto que $B_0^{(k)} = 0$. Fixemos tempos $0 \le t_1 < t_2 \le t_3 < t_4$. Precisamos mostrar que $B_{t_2} - B_{t_1}$ e $B_{t_4} - B_{t_3}$ são independentes e têm distribuição $N(0, t_2 - t_1)$ e $N(0, t_4 - t_3)$, respectivamente. Se tivermos $m < t_2 \le m + 1$ e $n < t_3 \le n + 1$ com $m < n$, então nota-se, claramente, que o incremento $B_{t_2} - B_{t_1}$ depende apenas de movimentos brownianos $(B_t^{(k)})_{t\in[0,1]}$ para $k < m$ e, por outro lado, o incremento $B_{t_4} - B_{t_3}$ só depende de movimentos brownianos

$(B_t^{(k)})_{t\in[0,1]}$ para $k \geq m$, os quais, por construção, são independentes de $(B_t^{(k)})_{t\in[0,1]}$ para $k < m$. Logo, vale a independência entre os incrementos neste caso. Se, entretanto, $m = n$ basta usar, também, a independência entre incrementos em intervalos de tempos disjuntos do processo $(B_t^{(m)})_{t\in[0,1]}$ e chegamos à mesma conclusão. Claramente, o mesmo argumento funciona para um número qualquer de incrementos.

Agora mostremos que, se $s < t$, então o incremento $B_t - B_s$ tem distribuição $N(0, t-s)$. Suponhamos $m < s \leq m+1$ e $n < t \leq n+1$. Se $m = n$, o resultado é claro, pois $B_t - B_s = B_{s-m}^{(m)} - B_{t-m}^{(m)}$ que tem distribuição $N(0, (t-m) - (s-m)) = N(0, t-s)$. Se, por outro lado, $m < n$, então

$$\begin{array}{l} B_t = B_1^{(0)} + \ldots + B_1^{(m-1)} + B_1^{(m)} + \ldots + B_1^{(n-1)} + B_{t-n}^{(n)} \\ B_s = B_1^{(0)} + \ldots + B_1^{(m-1)} + B_{s-m}^{(m)} \end{array}$$

logo

$$B_t - B_s = (B_1^{(m)} - B_{s-m}^{(m)}) + B_1^{(m+1)} + \ldots + B_1^{(n-1)} + B_{t-n}^{(n)}$$

é soma de variáveis aleatórias normais independentes de médias zero. Além disso, usando a independência, temos que

$$\begin{aligned} V(B_t - B_s) &= V(B_1^{(m)} - B_{s-m}^{(m)}) + \underbrace{V(B_1^{(m+1)}) + \ldots + V(B_1^{(n-1)})}_{n-m-1 \text{ termos}} + V(B_{t-n}^{(n)}) \\ &= 1 - (s-m) + \underbrace{1 + \ldots + 1}_{n-m-1 \text{ termos}} + (t-n) = t - s. \end{aligned}$$

Logo, $(B_t - B_s) \sim N(0, t-s)$ e, consequentemente $(B_t)_{t\geq 0}$ é um movimento browniano.

2.2.2 Construção alternativa

Usando resultados gerais da teoria dos processos estocásticos, podemos obter a existência do movimento browniano sem recorrer a uma construção explícita. Não entraremos em muitos detalhes, mas a ideia é notar que, a menos da condição da continuidade das trajetórias, o movimento browniano é definido a partir de suas *distribuições finito-dimensionais*, ou seja, da distribuição de probabilidade dos vetores $(B_{t_1}, \ldots, B_{t_n})$, onde $t_1, \ldots, t_n$ é qualquer sequência finita de tempos fixada. Ocorre que, sob certas condições naturais de consistência, o clássico *Teorema de Extensão de Kolmogorov*[1] garante

[1]Veja, por exemplo, Billingsley [3], páginas 510-517.

a existência de um processo estocástico com uma família de distribuições finito-dimensionais dada. Assim, a existência do movimento browniano fica parcialmente estabelecida, faltando-se apenas garantir a continuidade das trajetórias.

Há algumas sutilezas nesta questão. É natural que o Teorema de Extensão de Kolmogorov nada afirme sobre a continuidade das trajetórias do processo, uma vez que a continuidade de uma função não é uma característica que possa ser determinada pela especificação de apenas uma quantidade enumerável de coordenadas. Ou, dito de outra forma, as distribuições finito-dimensionais sozinhas não são capazes de determinar se um processo possui ou não trajetórias contínuas. Por exemplo, se $(X_t)_{t\geq 0}$ tem trajetórias contínuas quase certamente, defina:

$$Y_t = \begin{cases} X_t & \text{se } t \neq T \\ 0 & \text{se } t = T \end{cases},$$

onde T é uma variável aleatória com distribuição contínua em $[0,\infty)$. Então, as trajetórias típicas de $(Y_t)_{t\geq 0}$ não serão contínuas, embora $(X_t)_{t\geq 0}$ e $(Y_t)_{t\geq 0}$ tenham as mesmas distribuições finito-dimensionais, com é fácil ver.

O que faremos é garantir a continuidade a menos de uma *modificação*, através de um resultado conhecido como *Critério de Kolmogorov*, o qual apenas enunciaremos[2].

Definição 2.2.8. *Um processo estocástico $(Y_t)_{t\geq 0}$ é dito uma modificação de $(X_t)_{t\geq 0}$ se, para todo $t \geq 0$, tivermos $\mathbb{P}(X_t = Y_t) = 1$.*

Naturalmente, se um processo é modificação de outro então ambos têm o mesmo sistema de distribuições finito-dimensionais.

Teorema 2.2.9 (Critério de Kolmogorov)**.** *Seja $(X_t)_{t\geq 0}$ um processo estocástico. Suponhamos que existam constantes positivas α, β e C tais que*

$$\mathbb{E}\left[|X_t - X_s|^\alpha\right] \leq C(t-s)^{1+\beta}$$

para quaisquer tempos $s < t$. Então existe um modificação $(\tilde{X}_t)_{t\geq 0}$ de $(X_t)_{t\geq 0}$ com trajetórias contínuas quase certamente.

Para verificar o critério no caso do movimento browniano, lembrando que $B_t - B_s$ é $N(0, t-s)$, usaremos as derivadas no zero da função geradora de momentos

$$M(\lambda) = \mathbb{E}\left(e^{\lambda(B_t - B_s)}\right) = e^{\frac{t-s}{2}\lambda^2}.$$

[2]Para demonstração veja, por exemplo, Karatzas & Shreve [8], páginas 53-55.

Fazendo as contas, encontramos

$$\mathbb{E}\left[|B_t - B_s|^4\right] = \frac{d^4}{d\lambda^4} M(\lambda)|_{\lambda=0} = 3(t-s)^2.$$

Portanto, o critério é satisfeito com $\alpha = 4$, $\beta = 1$, e $C = 3$.

Assim, fica estabelecida, de uma maneira alternativa, a existência do movimento browniano com trajetórias contínuas quase certamente.

2.3 Movimento Browniano n-dimensional

A partir do movimento browniano na reta, é fácil estender a definição para o caso n-dimensional.

Definição 2.3.1. *Dizemos que um processo* $(B_t)_{t\geq 0} = \left(B_t^{(1)}, \ldots, B_t^{(n)}\right)_{t\geq 0}$ *tomando valores em* $\mathbb{R}^n$ *é um movimento browniano* n*-dimensional se*

1. *Para cada* $k = 1, \ldots, n$, *a componente* $(B_t^{(k)})_{t\geq 0}$ *é um movimento browniano unidimensional.*

2. *Os processos* $(B_t^{(k)})_{t\geq 0}$, $k = 1, \ldots, n$, *são independentes.*

O item *2* significa que, para quaisquer sequências finitas de tempos

$$\begin{array}{llll} t_{11}, & t_{12}, & \ldots, & t_{1,k_1} \\ t_{21}, & t_{22}, & \ldots, & t_{2,k_2} \\ \vdots & & & \\ t_{n,1}, & t_{n,2} & \ldots, & t_{n,k_n}, \end{array}$$

os n vetores

$$X_1 = \left(B_{t_{11}}^{(1)}, B_{t_{12}}^{(1)}, \ldots, B_{t_{1,k_1}}^{(1)}\right)$$

$$X_2 = \left(B_{t_{21}}^{(2)}, B_{t_{22}}^{(2)}, \ldots, B_{t_{2,k_2}}^{(2)}\right)$$

$$\vdots$$

$$X_n = \left(B_{t_{n,1}}^{(n)}, B_{t_{n,2}}^{(n)}, \ldots, B_{t_{n,k_n}}^{(n)}\right)$$

são independentes.

É possível, também, uma caracterização em termos da distribuição normal multivariada.

Definição 2.3.2. *Dizemos que um vetor aleatório $X = (X_1, \ldots, X_n)$ tem distribuição normal n-variada com média $m \in \mathbb{R}^n$ e matriz de covariâncias C, abreviadamente $X \sim N(m, C)$, se sua função de densidade de probabilidade conjunta é dada por*

$$f(x_1, \ldots, x_n) = \frac{1}{(2\pi)^{n/2}\sqrt{\det C}} \exp\left\{-\frac{\langle C^{-1}(x-m), (x-m)\rangle}{2}\right\},$$

onde x representa o vetor coluna de entradas $x_1, \ldots, x_n$ e $\langle .,. \rangle$ denota o produto interno usual do R^n.

A proposição a seguir permite uma definição equivalente de movimento browniano n-dimensional, igual ao caso unidimensional, apenas trocando, no item *2* da Definição 2.1.1, $N(0, t-s)$ por $N(0, (t-s)I_n)$.

Proposição 2.3.3. *Se $(B_t)_{t \geq 0}$ é um movimento browniano n-dimensional então $B_t \sim N(0, tI_n)$, para cada tempo $t > 0$. Portanto*

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

para cada boreliano $A \subseteq \mathbb{R}^n$.

Demonstração. Para cada t, as coordenadas B_t^k do processo são variáveis aleatórias independentes com distribuição $N(0, t)$. Denotemos por f a função densidade conjunta do vetor aleatório B_t e por $f_1, \ldots, f_n$ as densidades das coordenadas. Então, por independência, temos

$$\begin{aligned} f(x_1, \ldots, x_n) &= f_1(x_1) \ldots f_n(x_n) \\ &= \frac{1}{(2\pi t)^{1/2}} e^{-\frac{x_1^2}{2t}} \ldots \frac{1}{(2\pi t)^{1/2}} e^{-\frac{x_n^2}{2t}} \\ &= \frac{1}{(2\pi t)^{n/2}} e^{-\frac{|x|^2}{2t}}. \end{aligned}$$

□

2.4 Não-diferenciabilidade das trajetórias

Por definição, as trajetórias do movimento browniano são contínuas, quase certamente. Nesta seção mostraremos que, com probabilidade um, as trajetórias não são diferenciáveis em nenhum ponto. Este resultado pode ser intuído através da propriedade markoviana (2.4). De fato, se $B_s = x_0$, o comportamento futuro de $(B_t)_{t>s}$ depende apenas deste fato e não leva em conta de que maneira o processo $(B_t)_{t\in[0,s]}$ se aproximou do ponto x_0 quando t cresceu a s. Assim, não temos por que esperar que o processo saia de x_0 de modo que haja uma reta tangente neste ponto.

Teorema 2.4.1. *Com probabilidade um, as trajetórias do movimento browniano não são diferenciáveis em nenhum ponto.*

A demonstração é bastante técnica e pode ser omitida em uma primeira leitura, sem perda de continuidade.

Demonstração. Basta considerar o caso unidimensional e, sem perda de generalidade, tempos $0 \le t \le 1$. Suponhamos que $(B_t)_{t\in[0,1]}$ seja diferenciável em algum ponto $0 \le t_0 < 1$. Então, em algum intervalo contendo t_0, devemos ter

$$|B_t - B_{t_0}| \le K\,|t - t_0|$$

para alguma constante K. Logo, para pontos t_1 e t_2 suficientemente próximos de t_0, temos, pela desigualdade triangular, que

$$\begin{aligned} |B_{t_2} - B_{t_1}| &\le |B_{t_2} - B_{t_0}| + |B_{t_0} - B_{t_1}| \\ &\le K\left(|t_2 - t_0| + |t_1 - t_0|\right). \end{aligned} \tag{2.23}$$

Para $i = 0, 1$ e 2 e $n = 1, 2, 3, \ldots$ definamos

$$t_2(n,i) = \frac{\lfloor nt_0 \rfloor + i + 1}{n} \quad \text{e} \quad t_1(n,i) = \frac{\lfloor nt_0 \rfloor + i}{n}.$$

Então, para $i = 0, 1$ e 2 temos que $\lim_{n\to\infty} t_2(n,i) = \lim_{n\to\infty} t_1(n,i) = t_0$. Portanto, para n suficientemente grande, de acordo com (2.23) temos que, para $i = 0, 1$ e 2

$$\begin{aligned} \left|B_{\frac{\lfloor nt_0 \rfloor + i + 1}{n}} - B_{\frac{\lfloor nt_0 \rfloor + i}{n}}\right| &\le K\left(\left|\frac{\lfloor nt_0 \rfloor + i + 1}{n} - t_0\right| + \left|\frac{\lfloor nt_0 \rfloor + i}{n} - t_0\right|\right) \\ &\le K\left(\frac{i+1}{n} + \frac{i}{n}\right) \le \frac{5K}{n}. \end{aligned}$$

Portanto, o evento "$(B_t)_{t\in[0,1]}$ é diferenciável em algum $t_0 \in [0,1)$" está contido no evento

$$A \equiv \bigcup_{k=1}^{\infty} \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} \bigcup_{j=0}^{n-1} \bigcap_{i=0}^{2} \left(\left|B_{\frac{j+i+1}{n}} - B_{\frac{j+i}{n}}\right| \le \frac{5k}{n} \right).$$

Mostraremos que A é união enumerável de eventos de probabilidade zero, donde concluiremos que $\mathbb{P}(A) = 0$. Para cada k definamos

$$A_n^k \equiv \bigcup_{j=0}^{n-1} \bigcap_{i=0}^{2} \left(\left| B_{\frac{j+i+1}{n}} - B_{\frac{j+i}{n}} \right| \leq \frac{5k}{n} \right).$$

Então, A pode ser expresso como[3] $A = \bigcup_{k=1}^{\infty} \left(\liminf_n A_n^k \right)$. Portanto, basta mostrar que, para cada k, $\mathbb{P}\left(\liminf_n A_n^k\right) = 0$. Mas, em geral, vale que[4]

$$\mathbb{P}\left(\liminf_n A_n^k\right) \leq \liminf_{n\to\infty} \mathbb{P}(A_n^k)$$

e, portanto

$$\begin{aligned}
\mathbb{P}\left(\liminf_n A_n^k\right) &\leq \liminf_{n\to\infty} \mathbb{P}\left[\bigcup_{j=0}^{n-1} \bigcap_{i=0}^{2} \left(\left| B_{\frac{j+i+1}{n}} - B_{\frac{j+i}{n}} \right| \leq \frac{5k}{n} \right)\right] \\
&\leq \liminf_{n\to\infty} \sum_{j=0}^{n-1} \mathbb{P}\left[\bigcap_{i=0}^{2} \left(\left| B_{\frac{j+i+1}{n}} - B_{\frac{j+i}{n}} \right| \leq \frac{5k}{n} \right)\right] \\
\text{por independência incrementos disjuntos} \quad &= \liminf_{n\to\infty} n \left(\mathbb{P}\left[\left| B_{\frac{1}{n}} \right| \leq \frac{5k}{n} \right] \right)^3 \\
\sim N(0,1/n) \quad &= \liminf_{n\to\infty} n \left(\frac{2}{\sqrt{2\pi/n}} \int_0^{\frac{5k}{n}} e^{-\frac{nx^2}{2}} dx \right)^3 \\
&\leq \liminf_{n\to\infty} n \left(\frac{2}{\sqrt{2\pi/n}} \int_0^{\frac{5k}{n}} 1 dx \right)^3 \\
&= \lim_{n\to\infty} \left(\frac{10k}{\sqrt{2\pi}} \right)^3 n^{-1/2} = 0.
\end{aligned}$$

Pois as variáveis aleatórias $\left(B_{\frac{j+i+1}{n}} - B_{\frac{j+i}{n}}\right)$, $i = 0, 1$ e 2, são independentes e têm distribuição $N(0, \frac{1}{n})$. Isto conclui a demonstração. $\square$

2.5 Martingales

Outra importante propriedade do movimento browniano é que ele faz parte de uma classe de processos estocásticos chamados de *Martingales*. Neste seção,

[3] Veja Bartle [2], página 15

[4] Veja Bartle [2], Exer. 3I. página 24

explicaremos o que é um martingale e apresentaremos algumas propriedades básicas que serão úteis nos capítulos seguintes. Para isso, antes de tudo, precisaremos definir alguns termos e fixar notações correspondentes.

Nas definições a seguir, o tempo pode ser tanto discreto como contínuo.

Definição 2.5.1. *Dada um variável aleatória X definida em $(\Omega, \mathcal{F})$ denotemos por $\sigma(X)$ a σ-álgebra gerada por X. Ou seja, a menor σ-álgebra em relação a qual X é mensurável. Da mesma forma, dado um processo estocástico $(X_t)_{t\geq 0}$, $\sigma(X_s, 0 \leq s \leq t)$ denota a σ-álgebra gerada pelas variáveis aleatórias X_s, para $0 \leq s \leq t$.*

Definição 2.5.2. *Uma família de sub-σ-álgebras $(\mathcal{F}_t)_{t\geq 0}$ é chamada de filtração no espaço mensurável $(\Omega, \mathcal{F})$ se*

$$0 \leq s < t \Rightarrow \mathcal{F}_s \subseteq \mathcal{F}_t$$

isto é, $(\mathcal{F}_t)_{t\geq 0}$ é crescente.

Definição 2.5.3. *Um processo estocástico $(X_t)_{t\geq 0}$ é dito adaptado a uma filtração $(\mathcal{F}_t)_{t\geq 0}$ se, para cada tempo $t \geq 0$, a variável aleatória X_t é mensurável em relação à σ-álgebra $\mathcal{F}_t$.*

Todo processo é adaptado à sua filtração natural $\mathcal{F}_t = \sigma(X_s, 0 \leq s \leq t)$, também chamada de *história* do processo até o tempo t.

É intuitivo interpretar uma σ-álgebra como sendo a quantidade de informação disponível sobre as variáveis aleatórias mensuráveis em relação a ela. De modo que, quanto maior for a σ-álgebra, mais informação há disponível. Nesta linha de pensamento, dizer que um processo estocástico$(X_t)_{t\geq 0}$ é adaptado à filtração $(\mathcal{F}_t)_{t\geq 0}$ significa que, para cada tempo $t \geq 0$, a variável aleatória X_t depende apenas da informação disponível na σ-álgebra $\mathcal{F}_t$.

A definição mais geral de martingale, a ser dada a seguir, faz uso de uma importante noção de esperança condicional: a esperança condicional em relação a uma σ-álgebra.

Definição 2.5.4. *Seja X uma variável aleatória integrável em $(\Omega, \mathcal{F}, \mathbb{P})$ e seja $\mathcal{G}$ uma sub-σ-álgebra de $\mathcal{F}$. A esperança condicional $\mathbb{E}[X|\mathcal{G}]$ é uma variável aleatória $\mathcal{G}$-mensurável tal que*

$$\int_G X d\mathbb{P} = \int_G \mathbb{E}[X|\mathcal{G}] d\mathbb{P} \tag{2.24}$$

para todo conjunto $G \in \mathcal{G}$.

Interpretamos $\mathbb{E}[X|\mathcal{G}]$ como uma estimativa da variável aleatória X com base na informação dada pela σ-álgebra $\mathcal{G}$. A expressão (2.24) significa que nos conjuntos de $\mathcal{G}$ as médias de X e de $\mathbb{E}[X|\mathcal{G}]$ são as mesmas.

Exemplo 2.5.5 (Espaço de probabilidade finito)**.** Seja $\Omega = \{1, \ldots, n\}$, $\mathcal{F}$ a σ-álgebra das partes de Ω e $\mathbb{P} = \sum_{i=1}^{n} p_i \delta_i$ em que $p_i \geq 0$, $\sum_{i=1}^{n} p_i = 1$ e δ_i é a medida que concede massa total ao elemento i (delta de Dirac em i). Seja $A_1, \ldots, A_m$ uma partição de Ω, isto é, conjuntos disjuntos cuja união é Ω. Consideremos $\mathcal{G}$ a σ-álgebra gerada pelos conjuntos A_j, $j = 1, \ldots, m$. As funções mensuráveis em relação a $\mathcal{G}$ são constantes nos conjuntos A_j's, pois $\mathcal{G}$ não possui subconjuntos próprios não vazios dos A_j's. Logo, dada uma variável aleatória X em $(\Omega, \mathcal{F}, \mathbb{P})$, definida por $X(i) = x_i$, $i = 1, \ldots, n$, a esperança condicional $\mathbb{E}[X|\mathcal{G}]$ é uma função constante nos conjuntos A_j's tal que, se $i \in A_j$ então:

$$\mathbb{E}[X|\mathcal{G}](i) = \frac{1}{\mathbb{P}(A_j)} \sum_{k \in A_j} p_k x_k = \frac{1}{\mathbb{P}(A_j)} \int_{A_j} X d\mathbb{P}$$

caso $\mathbb{P}(A_j) \neq 0$. Se $\mathbb{P}(A_j) = 0$, podemos impor qualquer valor a $\mathbb{E}[X|\mathcal{G}](i)$. $\square$

Teorema 2.5.6 (Existência e unicidade da esperança condicional)**.** *Dada uma variável aleatória X integrável, definida em $(\Omega, \mathcal{F}, \mathbb{P})$, e uma sub-$\sigma$-álgebra $\mathcal{G} \subseteq \mathcal{F}$ existe, e é única $\mathbb{P}$-quase certamente, a esperança condicional $\mathbb{E}[X|\mathcal{G}]$ apresentada na Definição 2.5.4.*

A demonstração do Teorema 2.5.6, a ser dada a seguir, faz uso do *Teorema de Radon-Nikodyn* que é assunto clássico nos cursos de Teoria da Medida. Aqui apenas o enunciaremos. Para a demonstração, veja, por exemplo Bartle [2], p.85.

Definição 2.5.7. *Uma medida λ em $(\Omega, \mathcal{A})$ é dita absolutamente contínua em relação a uma medida μ em $(\Omega, \mathcal{A})$, denota-se por $\lambda \ll \mu$, se, para $A \in \mathcal{A}$, $\mu(A) = 0$ implicar $\lambda(A) = 0$.*

Teorema 2.5.8 (Radon-Nikodyn)**.** *Sejam λ e μ medidas σ-finitas definidas em $(\Omega, \mathcal{A})$. Suponha que $\lambda \ll \mu$. Então existe uma função $f : \Omega \to \mathbb{R}$, $\mathcal{A}$-mensurável, não negativa, tal que*

$$\lambda(A) = \int_A f d\mu$$

para todo $A \in \mathcal{A}$.

Demonstração do Teorema 2.5.6. Inicialmente, consideremos o caso em que a variável aleatória X é não negativa. Defina uma medida λ em $(\Omega, \mathcal{G})$ por

$$\lambda(A) = \int_A X d\mathbb{P}.$$

Temos que λ é medida finita pois X é integrável. Notemos também que $\lambda \ll \mathbb{P}|_{(\Omega,\mathcal{G})}$ e, portanto, pelo Teorema de Radon-Nikodyn, existe uma função f, $\mathcal{G}$-mensurável, tal que

$$\lambda(A) = \int_A f d\mathbb{P}$$

para todo $A \in \mathcal{G}$. Assim, $\mathbb{E}[X|\mathcal{G}] \equiv f$ satisfaz as propriedades requeridas. Para uma variável aleatória X qualquer, basta decompor $X = X^+ - X^-$ em suas partes positiva e negativa, e tomar $\mathbb{E}[X|\mathcal{G}] \equiv \mathbb{E}[X^+|\mathcal{G}] - \mathbb{E}[X^-|\mathcal{G}]$. Quanto à unicidade, suponhamos que exista outra variável aleatória Y, $\mathcal{G}$-mensurável satisfazendo a propriedade (2.24). Então

$$\int_G \mathbb{E}[X|\mathcal{G}] - Y d\mathbb{P} = 0$$

para todo $G \in \mathcal{G}$. Em particular, para todo n, tomando

$$G_n = \{\omega \in \Omega : \mathbb{E}[X|\mathcal{G}] - Y > 1/n\} \in \mathcal{G}$$

temos que

$$0 = \int_{G_n} \mathbb{E}[X|\mathcal{G}] - Y d\mathbb{P} \geq \int_{G_n} \frac{1}{n} d\mathbb{P} = \frac{1}{n}\mathbb{P}(G_n).$$

Logo, $\mathbb{P}(G_n) = 0$ e, consequentemente, $\mathbb{P}(E[X|\mathcal{G}] - Y > 0) = 0$. Da mesma forma, se mostra que $\mathbb{P}(E[X|\mathcal{G}] - Y < 0) = 0$. Logo, $\mathbb{E}[X|\mathcal{G}] = Y$ com probabilidade um. $\square$

Teorema 2.5.9 (Algumas propriedades da esperança condicional)**.** *Sejam X e Y variáveis aleatórias integráveis em $(\Omega, \mathcal{F}, \mathbb{P})$, e $\mathcal{G}$ uma sub-σ-álgebra de $\mathcal{F}$.*

1. *(invariância da média)* $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[X]$.

2. *(propriedade de projeção) Se X é $\mathcal{G}$-mensurável então* $\mathbb{E}[X|\mathcal{G}] = X$.

3. *(linearidade) Se a e b são constantes então*

$$\mathbb{E}[aX + bY|\mathcal{G}] = a\mathbb{E}[X|\mathcal{G}] + b\mathbb{E}[Y|\mathcal{G}].$$

4. *(monotocidade) Se $X \leq Y$ quase certamente, então $\mathbb{E}[X|\mathcal{G}] \leq \mathbb{E}[Y|\mathcal{G}]$ quase certamente.*

5. *(convergência dominada) Se $\lim_{n\to\infty} X_n = X$ com probabilidade um, $|X_n| \leq Y$, e Y é integrável, então $\lim_{n\to\infty} \mathbb{E}[X_n|\mathcal{G}] = \mathbb{E}[X|\mathcal{G}]$.*

6. *(convergência monótona) Se $X_n \geq 0$ e $X_n \uparrow X$ com $\mathbb{E}[X] < \infty$, então $\mathbb{E}[X_n|\mathcal{G}] \uparrow \mathbb{E}[X|\mathcal{G}]$.*

7. *($\mathcal{G}$-mensuráveis agem como constantes) Se X é $\mathcal{G}$-mensurável e XY é integrável, então $\mathbb{E}[XY|\mathcal{G}] = X\mathbb{E}[Y|\mathcal{G}]$.*

8. *(independência) Se X é independente de $\mathcal{G}$, então $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$.*

9. *(filtração) Se $\mathcal{G}_1 \subseteq \mathcal{G}_2$ então $\mathbb{E}[\mathbb{E}[X|\mathcal{G}_1]]|\mathcal{G}_2] = \mathbb{E}[\mathbb{E}[X|\mathcal{G}_2]]|\mathcal{G}_1] = \mathbb{E}[X|\mathcal{G}_1]$.*

Demonstração.

1. Basta fazer $G = \Omega$ em (2.24).

2. Evidentemente, definindo $\mathbb{E}[X|\mathcal{G}] \equiv X$ temos que $\mathbb{E}[X|\mathcal{G}]$ é $\mathcal{G}$-mensurável e satisfaz (2.24). Logo, o resultado segue da unicidade.

3. $a\mathbb{E}[X|\mathcal{G}] + b\mathbb{E}[Y|\mathcal{G}]$ é $\mathcal{G}$-mensurável e satisfaz (2.24) pela linearidade da integral. O resultado segue da unicidade.

4. $$\int_A \mathbb{E}[X|\mathcal{G}]d\mathbb{P} = \int_A X d\mathbb{P} \leq \int_A Y d\mathbb{P} = \int_A \mathbb{E}[Y|\mathcal{G}]d\mathbb{P}$$
para todo $A \in \mathcal{G}$. Agora, tomando $A_n = \{\mathbb{E}[X|\mathcal{G}] - \mathbb{E}[Y|\mathcal{G}] > 1/n\}$, vemos que $\mathbb{P}(A_n) = 0$ para todo n e, consequentemente,
$$\mathbb{P}(\mathbb{E}[X|\mathcal{G}] > \mathbb{E}[Y|\mathcal{G}]) = 0.$$

5. Seja $Z_n = \sup_{k\geq n} |X_k - X|$. Então $Z_n \downarrow 0$ com probabilidade um e, pelas propriedades *3* e *4*, $|\mathbb{E}[X_n|\mathcal{G}] - \mathbb{E}[X|\mathcal{G}]| \leq \mathbb{E}[Z_n|\mathcal{G}]$. Portanto, é suficiente mostrar que $\mathbb{E}[Z_n|\mathcal{G}] \downarrow 0$ com probabilidade um. Como $Z_n \geq 0$ é não crescente, pela propriedade *4*, $\mathbb{E}[Z_n|\mathcal{G}]$ também o é e, portanto, tem um limite Z. Note que $0 \leq Z \leq \mathbb{E}[Z_n|\mathcal{G}]$ para todo n. Queremos mostrar que $Z = 0$ com probabilidade um. Sendo Z não negativa, é suficiente mostrar que $\mathbb{E}[Z] = 0$. Mas $0 \leq Z_n \leq 2Y$, então,

por Convergência Dominada usual (veja Bartle [2], página 44), temos que

$$\mathbb{E}[Z] = \int Z d\mathbb{P} \leq \int \mathbb{E}[Z_n|\mathcal{G}] d\mathbb{P} = \mathbb{E}[Z_n] \to 0.$$

6. Seja $Y_n = X - X_n$. Por linearidade, é suficiente mostrar que $Z_n \equiv \mathbb{E}[Y_n|\mathcal{G}] \downarrow 0$. Como $Y_n \downarrow 0$, temos, pela propriedade *4*, que também Z_n é não crescente, logo tem um limite Z, $0 \leq Z \leq Z_n$ para todo n. Notemos que Y_n é dominada por X que é integrável portanto, pela propriedade *1* e novamente pelo Teorema da Convergência Dominada (usual),

$$\mathbb{E}[Z] \leq \mathbb{E}[Z_n] = \mathbb{E}[Y_n] \to 0.$$

Logo, $Z = 0$ quase certamente, como queríamos.

7. Como $X\mathbb{E}[Y|\mathcal{G}]$ é $\mathcal{G}$-mensurável, é suficiente mostrar que

$$\int_A X\mathbb{E}[Y|\mathcal{G}] = \int_A XY d\mathbb{P} \qquad \text{para todo } A \in \mathcal{G}. \tag{2.25}$$

Inicialmente, suponhamos que $X = I_B$ com $B \in \mathcal{G}$. Neste caso, se $A \in \mathcal{G}$, então, como $A \bigcap B \in \mathcal{G}$,

$$\int_A I_B \mathbb{E}[Y|\mathcal{G}] d\mathbb{P} = \int_{A \bigcap B} \mathbb{E}[Y|\mathcal{G}] d\mathbb{P} = \int_{A \bigcap B} Y d\mathbb{P} = \int_A I_B Y d\mathbb{P}.$$

Logo, vale (2.25). Isto se estende para X função simples $\mathcal{G}$-mensurável, $X = \sum_{i=1}^n a_i I_{B_i}$, por linearidade. Para X $\mathcal{G}$-mensurável geral, tomemos uma sequência de funções simples $\mathcal{G}$-mensuráveis tais que $|X_n| \leq X$ e $\lim_n X_n = X$. Como $|X_n Y| \leq |XY|$ e $|XY|$ é integrável, pela propriedade *5* temos que $\lim_n \mathbb{E}[X_n Y|\mathcal{G}] = \mathbb{E}[XY|\mathcal{G}]$. Mas, por outro lado, $\mathbb{E}[X_n Y|\mathcal{G}] = X_n \mathbb{E}[Y|\mathcal{G}]$, pois as X_n's são funções simples, e $\lim_n X_n \mathbb{E}[Y|\mathcal{G}] = X\mathbb{E}[Y|\mathcal{G}]$. Portanto, vale a propriedade *7*.

8. X ser independente de $\mathcal{G}$ significa que a variável aleatória X é independente da informação dada por $\mathcal{G}$, ou ainda, que X é independente das variáveis aleatórias $\mathcal{G}$-mensuráveis. Como $\mathbb{E}[X]$ é uma constante, e portanto $\mathcal{G}$-mensurável, é suficiente mostrar que, para todo $A \in \mathcal{G}$, vale $\int_A X d\mathbb{P} = \int_A \mathbb{E}[X] d\mathbb{P}$. Vejamos:

$$\int_A X d\mathbb{P} = \int I_A X d\mathbb{P} = \mathbb{E}[I_A X] = \mathbb{E}[I_A]\mathbb{E}[X] = \mathbb{P}(A)\mathbb{E}[X] = \int_A \mathbb{E}[X] d\mathbb{P},$$

onde usamos a independência entre X e a variável aleatória $\mathcal{G}$-mensurável I_A na terceira igualdade.

9. Notemos que $\mathbb{E}[X|\mathcal{G}_1]$ é $\mathcal{G}_2$-mensurável, pois $\mathcal{G}_1 \subseteq \mathcal{G}_2$. Logo, pela propriedade *2*, $\mathbb{E}[\mathbb{E}[X|\mathcal{G}_1]]|\mathcal{G}_2] = \mathbb{E}[X|\mathcal{G}_1]$. Para provar que também $\mathbb{E}[\mathbb{E}[X|\mathcal{G}_2]]|\mathcal{G}_1] = \mathbb{E}[X|\mathcal{G}_1]$, notemos que $\mathbb{E}[X|\mathcal{G}_1]$ é $\mathcal{G}_1$-mensurável e se $A \in \mathcal{G}_1 \subseteq \mathcal{G}_2$ então

$$\int_A \mathbb{E}[X|\mathcal{G}_1] d\mathbb{P} = \int_A X d\mathbb{P} = \int_A \mathbb{E}[X|\mathcal{G}_2] d\mathbb{P}.$$

□

Para a esperança condicional também vale a desigualdade de Jensen, conforme o próximo teorema.

Teorema 2.5.10 (Desigualdade de Jensen)**.** *Se* $\varphi : \mathbb{R} \to \mathbb{R}$ *é convexa e* $\mathbb{E}[|X|]$, $\mathbb{E}[|\varphi(X)|] < \infty$ *então, com probabilidade um,*

$$\varphi(\mathbb{E}[X|\mathcal{F}]) \leq \mathbb{E}[\varphi(X)|\mathcal{F}].$$

Demonstração. Se φ é linear o resultado já foi provado, valendo a igualdade. Suponhamos, então que φ não é linear. Neste caso, se definirmos

$$S = \{(a,b) : a, b \in \mathbb{Q},\ ax + b \leq \varphi(x) \ \text{ para todo } x \in \mathbb{R}\},$$

então, pela convexidade, pode-se provar que

$$\varphi(x) = \sup_{(a,b)\in S} (ax + b).$$

Se para algum par (a,b) tivermos $\varphi(x) \geq ax + b$ então, pelos itens *4*, *3* e *2* do Teorema 2.5.9, existe um conjunto $A_{(a,b)}$, de probabilidade um, tal que

$$\mathbb{E}[\varphi(X)|\mathcal{F}](\omega) \geq a\mathbb{E}[X|\mathcal{F}](\omega) + b$$

para todo $\omega \in A_{(a,b)}$. Seja A a interseção (enumerável) dos conjuntos $A_{(a,b)}$ sobre todos os pares $(a,b) \in S$. Então, $\mathbb{P}(A) = 1$, e para todo $\omega \in A$ temos

$$\mathbb{E}[\varphi(X)|\mathcal{F}](\omega) \geq a\mathbb{E}[X|\mathcal{F}](\omega) + b$$

para todo par $(a,b) \in S$. Tomando o supremo sobre todos $(a,b) \in S$ concluímos que, com probabilidade um,

$$\mathbb{E}[\varphi(X)|\mathcal{F}] \geq \varphi(\mathbb{E}[X|\mathcal{F}]).$$

□

O próximo teorema fornece uma interessante interpretação geométrica da esperança condicional como uma projeção. Dado um espaço $(\Omega, \mathcal{F}, \mathbb{P})$ e uma sub-σ-álgebra $\mathcal{G} \subseteq \mathcal{F}$, o conjunto

$$L^2(\Omega, \mathcal{F}, \mathbb{P}) \equiv \{Y |\ Y \text{ é } \mathcal{F}\text{-mensurável e } \mathbb{E}[Y^2] < \infty\}$$

é um espaço de Hilbert do qual $L^2(\Omega, \mathcal{G}, \mathbb{P})$ é subespaço fechado. O produto interno considerado é $\langle X, Y \rangle \equiv \mathbb{E}[XY]$. Se $\mathbb{E}[X^2] < \infty$, então o teorema seguinte mostra que $\mathbb{E}[X|\mathcal{G}]$ é a projeção de X sobre o subespaço $L^2(\Omega, \mathcal{G}, \mathbb{P})$ isto é, o ponto no subespaço $L^2(\Omega, \mathcal{G}, \mathbb{P})$ mais próximo de X, na norma dada pelo produto interno.

Teorema 2.5.11. *Suponha que $\mathbb{E}[X^2] < \infty$. Então a esperança condicional $\mathbb{E}[X|\mathcal{G}]$ é a variável aleatória Y, $\mathcal{G}$-mensurável, que minimiza o erro quadrático médio $\mathbb{E}[(X - Y)^2]$.*

Demonstração. Comecemos observando que se $Z \in L^2(\mathcal{G})$ então, pela desigualdade de Cauchy-Schwarz,

$$\mathbb{E}[|ZX|] \leq \sqrt{\mathbb{E}[Z^2]\mathbb{E}[X^2]} < \infty.$$

Logo, pelo Teorema 2.5.9 (item *7*),

$$Z\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[ZX|\mathcal{G}]. \tag{2.26}$$

Tomando a esperança, obtemos

$$\mathbb{E}(Z\mathbb{E}[X|\mathcal{G}]) = \mathbb{E}(\mathbb{E}[ZX|\mathcal{G}]) = \mathbb{E}[ZX]. \tag{2.27}$$

Por linearidade, podemos reescrever (2.27) como

$$\mathbb{E}(Z(X - \mathbb{E}[X|\mathcal{G}])) = 0 \quad \text{para } Z \in L^2(\mathcal{G}). \tag{2.28}$$

Seja $Y \in L^2(\mathcal{G})$. Faça $Z = Y - \mathbb{E}[X|\mathcal{G}] \in L^2(\mathcal{G})$. Então

$$\begin{aligned} \mathbb{E}[(X - Y)^2] &= \mathbb{E}[(X - \mathbb{E}[X|\mathcal{G}] - Z)^2] \\ &= \mathbb{E}[(X - \mathbb{E}[X|\mathcal{G}])^2] - 2\mathbb{E}[Z(X - \mathbb{E}[X|\mathcal{G}])] + \mathbb{E}[Z^2] \\ &= \mathbb{E}[(X - \mathbb{E}[X|\mathcal{G}])^2] + \mathbb{E}[Z^2]. \end{aligned} \tag{2.29}$$

Usamos a equação (2.28) na última igualdade. A equação (2.29) mostra que o erro quadrático médio $\mathbb{E}[(X - Y)^2]$ é mínimo quando $Z = 0$, ou seja, quando $Y = \mathbb{E}[X|\mathcal{G}]$. $\square$

A noção de esperança condicional em relação a uma variável aleatória torna-se um caso particular da esperança condicional em relação a uma σ-álgebra definindo-se $\mathbb{E}[X|Y] \equiv \mathbb{E}[X|\sigma(Y)]$. De fato, na esperança condicional $\mathbb{E}[X|Y]$, a característica importante da variável Y não é o valor que ela assume em cada $\omega \in \Omega$, mas sim a informação que a ocorrência de cada valor de Y pode fornecer a respeito da ocorrência dos eventos $\omega \in \Omega$. Esta informação é captada pela σ-álgebra gerada $\sigma(Y)$.

Definição 2.5.12. *Seja $(\Omega, \mathcal{F}, \mathbb{P})$ um espaço de probabilidade e $(\mathcal{F}_t)_{t \in T}$ uma filtração em σ-álgebras. O tempo T pode ser contínuo ou discreto. Um processo estocástico $(X_t)_{t \in T}$ a valores reais definido em $(\Omega, \mathcal{F}, \mathbb{P})$ é dito um martingale em relação à filtração $(\mathcal{F}_t)_{t \in T}$ (e à medida $\mathbb{P}$) se*

1. $\mathbb{E}[|X_t|] < \infty$ *para todo* $t \in T$*;*
2. $(X_t)_{t \in T}$ *é adaptado à filtração* $(\mathcal{F}_t)_{t \in T}$*;*
3. $\mathbb{E}[X_t|\mathcal{F}_s] = X_s$ *q.c. para todo* $t \geq s$*.*

Quando a filtração não for mencionada, subentende-se que se trata da filtração natural $\mathcal{F}_t = \sigma(X_s, 0 \leq s \leq t)$ gerada pelo próprio processo. A condição *3* expressa a ideia de que, tendo toda informação sobre o processo até o tempo s, a melhor estimativa para o estado futuro do processo no tempo $t \geq s$ é justamente X_s, o último valor observado. Esta característica aparece em vários contextos. Para um exemplo intuitivo, considere o processo estocástico discreto S_n que representa a fortuna de um jogador na n-ésima rodada de um jogo. Neste caso, a propriedade de martingale significa que o jogo é justo, no seguinte sentido: a fortuna do jogador numa rodada futura S_n, $n > m$, é, em média, a fortuna presente S_m. Além disso, este fato independe das informações sobre a evolução passada da fortuna (S_k para $k < n$). O próximo exemplo pode servir como modelo para uma situação deste tipo.

Exemplo 2.5.13. Seja $X_1, X_2, \ldots$ uma sequência de variáveis aleatórias independentes com $\mathbb{E}[|X_n|] < \infty$ e $\mathbb{E}[X_n] = 0$ para todo n. Consideremos a filtração $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$. Definimos o processo S_n como a soma $S_n = \sum_{k=1}^{n} X_k$. Então S_n é um martingale em relação à σ-álgebra $\mathcal{F}_n$. Vejamos:

1. Para todo n, $\mathbb{E}[|S_n|] \leq \mathbb{E}[|X_1|] + \ldots + \mathbb{E}[|X_n|] < \infty$;

2. Para todo n, S_n é $\mathcal{F}_n$-mensurável, pois é soma das variáveis X_k, $k = 1, \ldots, n$, que são $\mathcal{F}_n$-mensuráveis por definição de $\mathcal{F}_n$;

3. Se $n \geq m$ então

$$\begin{aligned}
\mathbb{E}[S_n|\mathcal{F}_m] &= \mathbb{E}[X_n + X_{n-1} + \ldots + X_{m+1} + S_m|\mathcal{F}_m] \\
&= \mathbb{E}[X_n + X_{n-1} + \ldots + X_{m+1}|\mathcal{F}_m] + \mathbb{E}[S_m|\mathcal{F}_m] \\
&= \mathbb{E}[X_n + X_{n-1} + \ldots + X_{m+1}] + S_m \\
&= 0 + S_m = S_m.
\end{aligned}$$

A segunda igualdade decorre da linearidade da esperança condicional. Na terceira igualdade usamos que $X_n + \ldots + X_{m+1}$ é independente da σ-álgebra $\mathcal{F}_m = \sigma(X_1, \ldots, X_m)$, que S_m é $\mathcal{F}_m$-mensurável e o Teorema 2.5.9 (itens *2* e *8*, respectivamente).

Note que, se as variáveis aleatórias X_k forem i.i.d $N(0,1)$ então S_n é uma aproximação discreta do movimento browniano. □

Exemplo 2.5.14 (Movimento Browniano)**.** Denotemos por $(\mathcal{F}_t)_{t\geq 0}$ a filtração natural do movimento browniano $(B_t)_{t\geq 0}$. Verifiquemos que $(B_t)_{t\geq 0}$ é um martingale:

1. Para cada $t \geq 0$, $\mathbb{E}[|B_t|] < \infty$, pois $B_t \sim N(0,t)$;

2. $(B_t)_{t\geq 0}$ é adaptado à $(\mathcal{F}_t)_{t\geq 0}$ (filtração natural);

3. Se $t \geq s$ então

$$\begin{aligned}
\mathbb{E}[B_t|\mathcal{F}_s] &= \mathbb{E}[(B_t - B_s) + B_s|\mathcal{F}_s] \\
&= \mathbb{E}[(B_t - B_s)|\mathcal{F}_s] + \mathbb{E}[B_s|\mathcal{F}_s] \\
&= \mathbb{E}[(B_t - B_s)] + B_s = 0 + B_s = B_s,
\end{aligned}$$

pois $(B_t - B_s)$ é independente da σ-álgebra $\mathcal{F}_s$ e B_s é $\mathcal{F}_s$-mensurável. □

Definição 2.5.15. *Se na definição 2.5.12 a condição 3 for substituída por*

3'. $\mathbb{E}[X_t|\mathcal{F}_s] \geq X_s$ *q.c. para todo* $t \geq s$

então dizemos que o processo $(X_t)_{t\in T}$ *é um submartingale. Se, por outro lado, tivermos*

3". $\mathbb{E}[X_t|\mathcal{F}_s] \leq X_s$ q.c. para todo $t \geq s$

então dizemos que o processo $(X_t)_{t\in T}$ é um supermartingale.

Usando a mesma analogia anteriormente apresentada, um submartingale pode ser pensado como um processo que descreve a fortuna de um jogador em um jogo favorável ao jogador (a esperança da fortuna futura é maior ou igual que a fortuna presente) enquanto que, um supermartingale descreveria um jogo desfavorável. Convém notar que um martingale é um submartingale e um supermartingale ao mesmo tempo.

Proposição 2.5.16. *Se o processo $(X_t)_{t\geq 0}$ é um martingale em relação à filtração $(\mathcal{F}_t)_{t\geq 0}$ e $\varphi : \mathbb{R} \to \mathbb{R}$ é convexa, com $\mathbb{E}[|\varphi(X_t)|] < \infty$ para todo $t \geq 0$, então $(\varphi(X_t))_{t\geq 0}$ é um submartingale em relação à $(\mathcal{F}_t)_{t\geq 0}$.*

Demonstração.

1. Por hipótese, $\mathbb{E}[|\varphi(X_t)|] < \infty$ para todo $t \geq 0$.

2. Toda função convexa é contínua e portanto borel-mensurável. Logo, $\varphi(X_t)$ é $\mathcal{F}_t$-mensurável pra todo t.

3. Pela desigualdade de Jensen, se $t \geq s$

$$\mathbb{E}[\varphi(X_t)|\mathcal{F}_s] \geq \varphi(\mathbb{E}[X_t|\mathcal{F}_s]) = \varphi(X_s).$$

□

Agora apresentaremos algumas importantes desigualdades. Tratemos, primeiramente, do caso em que o tempo é discreto.

Teorema 2.5.17 (Desigualdade L^p de Doob - tempo discreto)**.**

1. *Se $\{X_n\}_{n=1}^{\infty}$ é um submartingale, então para todo $n = 1, 2, \ldots$ e $\lambda > 0$*

$$\mathbb{P}\left(\max_{1\leq k\leq n} X_k \geq \lambda\right) \leq \frac{1}{\lambda}\mathbb{E}\left[X_n^+\right], \tag{2.30}$$

onde $X_n^+ = \max\{X_n, 0\}$.

2. *Se $\{X_n\}_{n=1}^{\infty}$ é um martingale e $p > 1$ então*

$$\mathbb{E}\left(\max_{1\leq k\leq n} |X_k|^p\right) \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}\left(|X_n|^p\right). \tag{2.31}$$

Corolário 2.5.18. *Note que, pela desigualdade* (2.30)*, se* $\{X_n\}_{n=1}^{\infty}$ *é um martingale então, para todo* $p \geq 1$ *e* $\lambda > 0$

$$\mathbb{P}\left(\max_{1\leq k\leq n} |X_k| \geq \lambda\right) \leq \frac{1}{\lambda^p}\mathbb{E}[|X_n|^p] \tag{2.32}$$

pois, pela Proposição 2.5.16, $\{|X_n|^p\}_{n=0}^{\infty}$ *é um submartingale, já que* $\varphi(x) = |x|^p$ *é convexa.*

Demonstração.
1. Suponhamos que $\{X_n\}_{n=1}^{\infty}$ seja um submartingale. Seja

$$A = \left[\max_{1\leq k\leq n} X_k \geq \lambda\right].$$

Vamos decompor A de acordo com a primeira vez em que $X_k \geq \lambda$. Para isso, definimos

$$\begin{aligned} A_1 &= [X_1 \geq \lambda] \\ A_2 &= [X_1 < \lambda, X_2 \geq \lambda] \\ A_k &= [X_1 < \lambda, X_2 < \lambda, \ldots, X_k \geq \lambda]\,, \text{ para } 2 < k \leq n \end{aligned}$$

Então os conjuntos A_k são 2 a 2 disjuntos e $A = \bigcup_{k=1}^{n} A_k$. Logo, $I_A = \sum_{k=1}^{n} I_{A_k}$.

Portanto

$$\begin{aligned} \mathbb{E}(X_n^+) \geq \mathbb{E}(X_n^+ I_A) &= \sum_{k=1}^{n} \mathbb{E}(X_n^+ I_{A_k}) \\ \text{(pela invariância da média)} &= \sum_{k=1}^{n} \mathbb{E}\left(\mathbb{E}[X_n^+ I_{A_k}|\mathcal{F}_k]\right) \\ (I_{A_k} \text{ é } \mathcal{F}_k\text{-mensurável}) &= \sum_{k=1}^{n} \mathbb{E}\left(I_{A_k}\mathbb{E}[X_n^+|\mathcal{F}_k]\right) \\ \text{(monotocidade)} &\geq \sum_{k=1}^{n} \mathbb{E}\left(I_{A_k}\mathbb{E}[X_n|\mathcal{F}_k]\right) \\ \text{(submartingale e monotocidade)} &\geq \sum_{k=1}^{n} \mathbb{E}\left(I_{A_k} X_k\right) \\ &\geq \sum_{k=1}^{n} \mathbb{E}\left(I_{A_k}\lambda\right) = \lambda \sum_{k=1}^{n} \mathbb{P}(A_k) = \lambda\mathbb{P}(A). \end{aligned}$$

Logo, vale (2.30).
2. Note que a sequência de desigualdades anteriores demonstra algo mais forte do que (2.30). De fato, mostramos que

$$\lambda \mathbb{P}(A) \leq \mathbb{E}(X_n^+ I_A),$$

ou seja,

$$\lambda \mathbb{P}\left(\max_{1\leq k\leq n} X_k \geq \lambda\right) \leq \int_\Omega X_n^+ I_{[\max_{1\leq k\leq n} X_k \geq \lambda]} d\mathbb{P}. \tag{2.33}$$

Agora, suponhamos que $\{X_n\}_{n=1}^\infty$ seja um martingale e consideremos o processo $\{|X_n|\}_{n=1}^\infty$. Pelo Teorema 2.5.16, temos que $\{|X_n|\}_{n=1}^\infty$ é um submartingale. Aplicando a desigualdade (2.33) a este processo, obtemos

$$\lambda \mathbb{P}\left(X^* \geq \lambda\right) \leq \int_\Omega |X_n| I_{[X^* \geq \lambda]} d\mathbb{P}, \tag{2.34}$$

onde $X^* = \max_{1\leq k\leq n} |X_k|$. Seja $p > 1$. Então

$$\begin{aligned}
\mathbb{E}[(X^*)^p] &= \mathbb{E}\left[\int_0^{X^*} p\lambda^{p-1} d\lambda\right] = \mathbb{E}\left[\int_0^\infty p\lambda^{p-1} I_{[X^*\geq\lambda]} d\lambda\right] \\
&= \int_\Omega \int_0^\infty p\lambda^{p-1} I_{[X^*\geq\lambda]} d\lambda d\mathbb{P} = \int_0^\infty \int_\Omega p\lambda^{p-1} I_{[X^*\geq\lambda]} d\mathbb{P} d\lambda \\
&= p\int_0^\infty \lambda^{p-1}\mathbb{P}[X^* \geq \lambda] d\lambda, \qquad (2.35)
\end{aligned}$$

onde a mudança na ordem de integração é justificada pelo Teorema de Tonelli[5]. Usando a desigualdade (2.34) em (2.35) e mudando a ordem de integração novamente, obtemos

$$\begin{aligned}
\mathbb{E}[(X^*)^p] &\leq p\int_0^\infty \lambda^{p-2} \int_\Omega |X_n| I_{[X^*\geq\lambda]} d\mathbb{P} d\lambda = p\int_\Omega |X_n| \int_0^\infty \lambda^{p-2} I_{[X^*\geq\lambda]} d\lambda d\mathbb{P} \\
&= p\int_\Omega |X_n| \int_0^{X^*} \lambda^{p-2} d\lambda d\mathbb{P} = \frac{p}{p-1}\int_\Omega |X_n|(X^*)^{p-1} d\mathbb{P}. \qquad (2.36)
\end{aligned}$$

Aplicando a Desigualdade de Hölder, obtemos

$$\mathbb{E}[(X^*)^p] \leq \frac{p}{p-1}\left(\int_\Omega |X_n|^p d\mathbb{P}\right)^{\frac{1}{p}} \left(\int_\Omega (X^*)^{(p-1)q} d\mathbb{P}\right)^{\frac{1}{q}}, \tag{2.37}$$

[5] Veja, por exemplo, Bartle [2], página 118.

onde q é tal que $\frac{1}{p}+\frac{1}{q}=1$. Mas então, $(p-1)q=p$ e $\frac{1}{q}=1-\frac{1}{q}$. Logo, (2.37) se torna

$$\mathbb{E}[(X^*)^p] \leq \frac{p}{p-1}\left(\int_\Omega |X_n|^p d\mathbb{P}\right)^{\frac{1}{p}}\left(\int_\Omega (X^*)^p d\mathbb{P}\right)^{1-\frac{1}{p}}$$

ou, em outros termos,

$$\mathbb{E}[(X^*)^p] \leq \frac{p}{p-1}\left(\mathbb{E}\left(|X_n|^p\right)\right)^{\frac{1}{p}}\left(\mathbb{E}[(X^*)^p]\right)^{1-\frac{1}{p}}.$$

Ou ainda

$$\mathbb{E}[(X^*)^p] \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}\left(|X_n|^p\right),$$

conforme desejávamos. □

É interessante notar que o Corolário 2.5.18 generaliza a Desigualdade de Kolmogorov. Sejam $X_1,\ldots,X_n$ variáveis aleatórias independentes tais que $\mathbb{E}[X_k]=0$ e $Var[X_k]<\infty$, $k=1,\ldots,n$. Então, vimos no Exemplo 2.5.13 que S_n é um martingale e, portanto, aplicando a desigualdade (2.32) com $p=2$ obtemos, para todo $\lambda>0$,

$$\mathbb{P}\left(\max_{1\leq k\leq n}|S_k|\geq\lambda\right)\leq\frac{1}{\lambda^2}\mathbb{E}[S_n^2]$$

que é, exatamente, o que diz a desigualdade de Kolmogorov[6].

Teorema 2.5.19 (Desigualdade L^p de Doob - tempo contínuo)**.** *Seja $(X_t)_{t\geq 0}$ um processo estocástico a tempo contínuo com trajetórias contínuas quase certamente.*

1. *Se $(X_t)_{t\geq 0}$ é um submartingale, então para todo $t\geq 0$ e $\lambda>0$*

$$\mathbb{P}\left(\sup_{0\leq s\leq t}X_s\geq\lambda\right)\leq\frac{1}{\lambda}\mathbb{E}\left[X_t^+\right], \tag{2.38}$$

onde $X_t^+=\max\{X_t,0\}$.

2. *Se $(X_t)_{t\geq 0}$ é um martingale e $p>1$ então*

$$\mathbb{E}\left(\sup_{0\leq s\leq t}|X_s|^p\right)\leq\left(\frac{p}{p-1}\right)^p\mathbb{E}\left(|X_t|^p\right). \tag{2.39}$$

[6]Veja, por exemplo, James [7], página 205.

Corolário 2.5.20. *Da parte 1, obtemos que, se $(X_t)_{t\geq 0}$ é um martingale com trajetórias contínuas q.c. então para todo $p \geq 1$ e $\lambda > 0$*

$$\mathbb{P}\left(\sup_{0\leq s\leq t} |X_s| \geq \lambda\right) \leq \frac{1}{\lambda^p}\mathbb{E}[|X_t|^p], \tag{2.40}$$

já que $(|X_t|^p)_{t\geq 0}$ é um submartingale.

Demonstração.
1. A prova consiste em aproximar o processo $(X_t)_{t\geq 0}$ por processos a tempo discreto e usar o teorema anterior. Em mais detalhes: seja $(X_t)_{t\geq 0}$ um submartingale. Fixemos $\lambda > 0$ e $t > 0$ e particionemos o intervalo $[0, t]$, ou seja, escolhemos pontos $0 = t_0 < t_1 < \ldots < t_n = t$. Notemos que o processo $\{Y_k\}_{k=0}^n$ definido por $Y_k = X_{t_k}$, para $k = 0, 1, \ldots, n$, é um submartingale discreto. De fato,

$$\begin{aligned}
\mathbb{E}[Y_k|\sigma(Y_0,\ldots,Y_{k-1})] &= \mathbb{E}[X_{t_k}|\sigma(X_{t_0},\ldots,X_{t_{k-1}})] \\
&= \mathbb{E}\big[\mathbb{E}[X_{t_k}|\sigma(X_s, 0\leq s\leq t_{k-1})]|\sigma(X_{t_0},\ldots,X_{t_{k-1}})\big] \\
&\geq \mathbb{E}[X_{t_{k-1}}|\sigma(X_{t_0},\ldots,X_{t_{k-1}})] = X_{t_{k-1}} = Y_{k-1}.
\end{aligned}$$

Na segunda igualdade usamos que $\sigma(X_{t_0},\ldots,X_{t_{n-1}}) \subseteq \sigma(X_s, 0 \leq s \leq t_{n-1})$ e o item *9* do Teorema 2.5.9. A desigualdade é justificada pela propriedade de submartingale do processo original e pela monotocidade da esperança condicional. Para a igualdade seguinte, usamos que $X_{t_{k-1}}$ é mensurável em relação à $\sigma(X_{t_0},\ldots,X_{t_{k-1}})$ e o item *2* do Teorema 2.5.9.

Portanto, para todo $\mu > 0$, pelo Teorema 2.5.17, vale

$$\mathbb{P}\left(\max_{s\in\{t_0,\ldots,t_n\}} X_s \geq \mu\right) \leq \frac{1}{\mu}\mathbb{E}\left[X_t^+\right]. \tag{2.41}$$

Agora, consideremos uma sequência crescente $\{F_n\}_{n=1}^\infty$ de conjuntos finitos cuja união é $F = ([0,t]\cap\mathbb{Q})\cup\{t\}$. Então, para todo $\lambda > 0$ e todo $\varepsilon > 0$, temos que

$$\begin{aligned}
\mathbb{P}\left(\sup_{s\in F} X_s \geq \lambda\right) &\leq \mathbb{P}\left(\bigcup_{n=1}^\infty [\max_{s\in F_n} X_s \geq \lambda-\varepsilon]\right) \\
&= \lim_{n\to\infty}\mathbb{P}\left(\max_{s\in F_n} X_s \geq \lambda-\varepsilon\right) \\
&\leq \frac{1}{\lambda-\varepsilon}\mathbb{E}\left[X_t^+\right].
\end{aligned} \tag{2.42}$$

Onde, na última desigualdade, usamos (2.41) com $\mu = \lambda - \varepsilon$. Fazendo $\varepsilon \to 0$ em (2.42) obtemos

$$\mathbb{P}\left(\sup_{s\in F} X_s \geq \lambda\right) \leq \frac{1}{\lambda}\mathbb{E}\left[X_t^+\right]. \tag{2.43}$$

Mas, pela continuidade quase certa das trajetórias, temos que $\sup_{0\leq s\leq t} X_s = \sup_{s\in F} X_s$ com probabilidade um. Usando este fato em (2.43), obtemos

$$\mathbb{P}\left(\sup_{0\leq s\leq t} X_s \geq \lambda\right) \leq \frac{1}{\lambda}\mathbb{E}\left[X_t^+\right].$$

De maneira similar se mostra (2.39). □

Concluímos este capítulo com uma importante aplicação destas desigualdades ao movimento browniano, conhecida como *desigualdade exponencial.*

Proposição 2.5.21. *Seja $(B_t)_{t\geq 0}$ um movimento browniano. Considere o processo $(S_t)_{t\geq 0} = (\sup_{0\leq s\leq t} B_s)_{t\geq 0}$ que indica o máximo da trajetória do movimento browniano até o tempo t. Então, para todo $c \geq 0$ e todo $t \geq 0$,*

$$\mathbb{P}\left(S_t \geq ct\right) \leq e^{-\frac{c^2}{2}t}. \tag{2.44}$$

Demonstração. Primeiramente notemos que, para todo $\alpha > 0$, o processo $(X_t^{(\alpha)})_{t\geq 0} = (e^{\alpha B_t - \frac{\alpha^2}{2}t})_{t\geq 0}$ é um martingale em relação à filtração natural de $(B_t)_{t\geq 0}$, com trajetórias contínuas quase certamente. Vejamos:

$$\begin{aligned}
\mathbb{E}[e^{\alpha B_t - \frac{\alpha^2}{2}t}|\mathcal{F}_s] &= e^{-\frac{\alpha^2}{2}t}\mathbb{E}[e^{\alpha B_t}|\mathcal{F}_s] = e^{-\frac{\alpha^2}{2}t}\mathbb{E}[e^{\alpha(B_t-B_s)}e^{\alpha B_s}|\mathcal{F}_s] \\
&= e^{-\frac{\alpha^2}{2}t}e^{\alpha B_s}\mathbb{E}[e^{\alpha(B_t-B_s)}] = e^{-\frac{\alpha 2}{2}t}e^{\alpha B_s}e^{\frac{(t-s)\alpha^2}{2}} = e^{\alpha B_s - \frac{\alpha^2}{2}s}.
\end{aligned}$$

Na terceira igualdade usamos que $e^{\alpha B_s}$ é $\mathcal{F}_s$-mensurável e que $e^{\alpha(B_t-B_s)}$ é independente de $\mathcal{F}_s$. Em seguida, usamos a função geradora de momentos de $(B_t - B_s) \sim N(0, t-s)$. Como $e^{\alpha S_t - \frac{\alpha^2}{2}t} \leq \sup_{0\leq s\leq t} X_t^{(\alpha)}$, obtemos

$$\begin{aligned}
\mathbb{P}(S_t \geq ct) &= \mathbb{P}\left(e^{\alpha S_t - \frac{\alpha^2}{2}t} \geq e^{\alpha ct - \frac{\alpha^2}{2}t}\right) \\
&\leq \mathbb{P}\left(\sup_{0\leq s\leq t} X_t^{(\alpha)} \geq e^{\alpha ct - \frac{\alpha^2}{2}t}\right) \\
&\leq e^{-\alpha ct + \frac{\alpha^2}{2}t}\mathbb{E}[X_t^{(\alpha)}], \qquad (2.45)
\end{aligned}$$

onde a última desigualdade é consequência da desigualdade (2.38) aplicada ao martingale $(X_t^{(\alpha)})_{t\geq 0}$. Mas $\mathbb{E}[X_t^{(\alpha)}] = \mathbb{E}[X_0^{(\alpha)}] = 1$. Para o α particular que minimiza o lado direito de (2.45), a saber $\alpha = c$, obtemos (2.44). □

Capítulo 3

Integral Estocástica de Itô

Neste capítulo daremos um sentido à expressão

$$\int_{\alpha}^{\beta} X_t dB_t \tag{3.1}$$

onde $(B_t)_{t\geq 0}$ é um movimento browniano unidimensional (inicialmente) e $(X_t)_{t\in[\alpha,\beta]}$ é um processo estocástico satisfazendo algumas condições a serem dadas a seguir. Expressões deste tipo aparecem no contexto das equações diferenciais estocásticas, assunto do próximo capítulo.

3.1 Integral de Itô para processes elementares

Para definir a integral em relação ao movimento browniano (3.1) é natural começarmos com processos $(X_t)_{t\in[\alpha,\beta]}$ mais simples e então estender a definição para processos mais gerais por procedimentos de aproximação.

Definição 3.1.1. *Um processo estocástico $\phi : [\alpha, \beta] \times \Omega \to \mathbb{R}$ é dito simples se for constante por partes na variável t, isto é, se existir uma partição*

$$\alpha = t_0 < t_1 < \ldots < t_n = \beta$$

do intervalo $[\alpha, \beta]$ tal que

$$\phi(t) = \phi(t_i) \quad se \quad t_i \leq t < t_{i+1}, \quad 0 \leq i \leq n-1.$$

Notação. Quando escrevemos $\phi(t)$ estamos nos referindo à variável aleatória, definida em Ω, que leva ω em $\phi(t,\omega)$. Evitamos a notação ϕ_t, a qual vínhamos usando até aqui, porque adiante trataremos de sequências de processos simples e, nesse caso, o sub-índice terá outro sentido.

Obs. Note na Definição 3.1.1 que a partição é fixada à priori e em cada subintervalo o processo é igual a uma variável aleatória. Um Processo de Poisson, por exemplo, *não* se enquadra nesta definição, embora suas trajetórias sejam constantes por partes.

Para um processo simples ϕ definido em $[\alpha,\beta]$, tendo a forma

$$\phi(t)=\sum_{i=0}^{n-1}\phi(t_i)I_{[t_i,t_{i+1})}(t), \tag{3.2}$$

onde $I_{[t_i,t_{i+1})}(t)$ é a função indicadora do intervalo $[t_i,t_{i+1})$, é natural definirmos

$$\int_\alpha^\beta \phi(t)dB_t=\sum_{i=0}^{n-1}\phi(t_i)(B_{t_{i+1}}-B_{t_i}). \tag{3.3}$$

O próximo exemplo indica uma dificuldade que surge quando tentamos estender esta definição para processos mais gerais procedendo por aproximação conforme a integral de Riemann-Stieltjes. Esta dificuldade está relacionada com o fato das trajetórias do movimento browniano terem variação ilimitada (Seção 2.4).

Exemplo 3.1.2. Dado um movimento browniano $(B_t)_{t\geq 0}$ e um intervalo $[0,\beta]$ consideremos, para cada $n=1,2,\ldots$, os seguintes processos simples:

$$\phi_n(t)=\sum_{i=0}^{m-1}B_{\frac{i}{2^n}}I_{[\frac{i}{2^n},\frac{i+1}{2^n})}(t)+B_{\frac{m}{2^n}}I_{[\frac{m}{2^n},\beta]}(t)\ \text{ se }\ \frac{m}{2^n}\leq\beta<\frac{m+1}{2^n};$$

$$\psi_n(t)=\sum_{i=0}^{m-1}B_{\frac{i+1}{2^n}}I_{[\frac{i}{2^n},\frac{i+1}{2^n})}(t)+B_\beta I_{[\frac{m}{2^n},\beta]}(t)\ \text{ se }\ \frac{m}{2^n}\leq\beta<\frac{m+1}{2^n}.$$

Então, de acordo com (3.3), temos

$$
\begin{aligned}
\mathbb{E}\left(\int_0^\beta \phi_n(t)dB_t\right) &= \sum_{i=0}^{m-1} \mathbb{E}\left[B_{\frac{i}{2^n}}\left(B_{\frac{i+1}{2^n}} - B_{\frac{i}{2^n}}\right)\right] + \mathbb{E}\left[B_{\frac{m}{2^n}}\left(B_\beta - B_{\frac{m}{2^n}}\right)\right] \\
&= \sum_{i=0}^{m-1} \mathbb{E}\left(B_{\frac{i}{2^n}}\right)\mathbb{E}\left(B_{\frac{i+1}{2^n}} - B_{\frac{i}{2^n}}\right) + \mathbb{E}\left(B_{\frac{m}{2^n}}\right)\mathbb{E}\left(B_\beta - B_{\frac{m}{2^n}}\right) \\
&= 0
\end{aligned}
$$

pela independência dos incrementos do movimento browniano, os quais têm média zero. Por outro lado

$$
\begin{aligned}
\mathbb{E}\left(\int_0^\beta \psi_n(t)dB_t\right) &= \sum_{i=0}^{m-1} \mathbb{E}\left[B_{\frac{i+1}{2^n}}\left(B_{\frac{i+1}{2^n}} - B_{\frac{i}{2^n}}\right)\right] + \mathbb{E}\left[B_\beta\left(B_\beta - B_{\frac{m}{2^n}}\right)\right] \\
&= \sum_{i=0}^{m-1} \left[\mathbb{E}\left(B^2_{\frac{i+1}{2^n}}\right) - \mathbb{E}\left(B_{\frac{i+1}{2^n}} B_{\frac{i}{2^n}}\right)\right] + \mathbb{E}\left(B^2_\beta\right) - \mathbb{E}\left(B_\beta B_{\frac{m}{2^n}}\right) \\
&= \sum_{i=0}^{m-1} \left[\frac{i+1}{2^n} - \frac{i}{2^n}\right] + \beta - \frac{m}{2^n} = \beta,
\end{aligned}
$$

onde usamos a Proposição 2.1.2 e notamos que o somatório é telescópico.

Percebe-se neste exemplo que, embora os processos simples ϕ_n e ψ_n aparentem ambos serem boas aproximações para o processo $(B_t)_{t\in[0,\beta]}$, suas integrais de acordo com (3.3) não são próximas entre si, não importando quão grande n seja escolhido.

É razoável, conforme fizemos no exemplo acima, aproximar um dado processo $(X_t)_{t\in[\alpha,\beta]}$ considerando-se processos simples $\sum_{j=0}^{m_n-1} X_{\xi_j^n} I_{[t_j^n, t_{j+1}^n)}(t)$ onde $P_n = \{\alpha = t_0^n < \ldots < t_{m_n}^n = \beta\}$ é uma partição de $[\alpha, \beta]$ e ξ_j^n é um ponto do intervalo $[t_j^n, t_{j+1}^n]$. A partir disso, esperaríamos definir $\int_\alpha^\beta X_t dB_t$ como o limite (num certo sentido) das integrais destes processos simples quando a norma da partição $|P_n| \equiv \max_{0<j<m_{n-1}}(t_{j+1}^n - t_j^n)$ tende a zero. Entretanto, o exemplo acima mostra que este limite vai depender dos pontos ξ_j^n's escolhidos. A integral estocástica de Itô corresponderá, como veremos a seguir, à escolha $\xi_j^n = t_j^n$, o extremo esquerdo do intervalo $[t_j^n, t_{j+1}^n]$. Uma outra noção de integração estocástica, a *Integral de Stratonovich*, que não será abordada neste texto, corresponde à escolha $\xi_j^n = \frac{t_j^n + t_{j+1}^n}{2}$. Em diversas situações é mais apropriado o uso da integral de Stratonovich. Entretanto, o fato de haver

uma conexão explícita entre as duas integrais torna suficiente, para a maioria dos propósitos, o tratamento de apenas um dos dois tipos. Para uma comparação entre as integrais de Itô e de Stratonovich veja, por exemplo, Øksendal [11], páginas 32-34.

Agora discutiremos para qual classe de processos estocásticos $(X_t)_{t\in[\alpha,\beta]}$ iremos definir a integral estocástica de Itô (3.1). Como veremos, até mesmo os processos simples, da forma (3.2), estarão sujeitos a restrições. Intuitivamente, o que faremos é, grosso modo, exigir, para cada t, que X_t dependa apenas do comportamento do movimento browniano até o tempo t (na verdade X_t poderá depender também de outras informações, mas não do comportamento futuro do movimento browniano). Esta ideia intuitiva é formalizada em termos de σ-álgebras.

Dado o movimento browniano $(B_t)_{t\geq 0}$ definido em um espaço de probabilidade abstrato $(\Omega, \mathcal{F}, \mathbb{P})$, consideremos uma filtração $(\mathcal{F}_t)_{t\geq 0}$ do espaço $(\Omega, \mathcal{F}, \mathbb{P})$ tal que

(i) $\sigma(B_s, 0 \leq s \leq t) \subset \mathcal{F}_t$ para todo $t \geq 0$;

(ii) $\sigma(B_{t+\lambda} - B_t, \lambda \geq 0)$ é independente de $\mathcal{F}_t$ para todo $t \geq 0$.

Obs. Podemos tomar, em particular, $\mathcal{F}_t = \sigma(B_s, 0 \leq s \leq t)$ a filtração natural do movimento browniano. Entretanto, a construção que faremos só exige que valham as condições (i) e (ii), sendo assim mais abrangente. Notemos também que, valendo as condições acima, $(B_t)_{t\geq 0}$ é um martingale em relação à $(\mathcal{F}_t)_{t\geq 0}$, mesmo que esta não seja a filtração natural (a demonstração é a mesma).

Definição 3.1.3. *Sendo dados um intervalo* $[\alpha, \beta] \subset [0, \infty)$, *um espaço de probabilidade* $(\Omega, \mathcal{F}, \mathbb{P})$, *onde é definido o movimento browniano* $(B_t)_{t\geq 0}$ *e uma filtração* $(\mathcal{F}_t)_{t\geq 0}$ *satisfazendo* (i) *e* (ii), *definimos* $\mathbb{M}^2 = \mathbb{M}^2(\alpha, \beta)$ *como a classe dos processos estocásticos* $(X_t)_{t\in[\alpha,\beta]}$ *que satisfazem:*

1. $(X_t)_{t\in[\alpha,\beta]}$ *é um processo mensurável, isto é, a função* $(t, \omega) \mapsto X_t(\omega)$ *é* $\mathcal{B} \times \mathcal{F}$*-mensurável, onde* $\mathcal{B}$ *denota a* σ*-álgebra de Borel do intervalo* $[\alpha, \beta]$;

2. $(X_t)_{t\in[\alpha,\beta]}$ *é adaptado à* $(\mathcal{F}_t)_{t\geq 0}$, *ou seja, para cada* $t \in [\alpha, \beta]$, *a variável aleatória* X_t *é* $\mathcal{F}_t$*-mensurável;*

3. $\mathbb{E}\left(\int_\alpha^\beta |X_t|^2 dt\right) < \infty$.

Definição 3.1.4. *Chamemos de elementares os processos simples em* $\mathbb{M}^2$.

Note que, no Exemplo 3.1.2, os ϕ_n's são processos elementares, o que não ocorre com os ψ_n's.

Definiremos a integral de Itô para todo processo $(X_t)_{t\in[\alpha,\beta]} \in \mathbb{M}^2$. Primeiramente consideremos os processos elementares.

Definição 3.1.5. *Seja* ϕ *um processo elementar,* $\phi(t) = \sum_{i=0}^{n-1} \phi(t_i) I_{[t_i,t_{i+1})}(t)$, *onde* $\alpha = t_0 < t_1 < \ldots < t_n = \beta$ *é uma partição do intervalo* $[\alpha, \beta]$. *Definimos a integral de Itô de* ϕ *no intervalo* $[\alpha, \beta]$ *pondo*

$$\int_\alpha^\beta \phi(t) dB_t = \sum_{i=0}^{n-1} \phi(t_i)(B_{t_{i+1}} - B_{t_i}). \tag{3.4}$$

Obs. Salientamos que a integral (3.4) é uma variável aleatória definida em $(\Omega, \mathcal{F}, \mathbb{P})$.

Lema 3.1.6 (Propriedades da integral de Itô para processos elementares)**.** *Sejam* ϕ *e* ψ *processos elementares,* a *e* b *números reais, então:*

$$\int_\alpha^\beta a\phi(t) + b\psi(t) dB_t = a \int_\alpha^\beta \phi(t) dB_t + b \int_\alpha^\beta \psi(t) dB_t; \tag{3.5}$$

$$\mathbb{E}\left(\int_\alpha^\beta \phi(t) dB_t\right) = 0; \tag{3.6}$$

$$\mathbb{E}\left[\left(\int_\alpha^\beta \phi(t) dB_t\right)^2\right] = \mathbb{E}\left(\int_\alpha^\beta \phi^2(t) dt\right). \tag{3.7}$$

Demonstração.

1. A verificação de (3.5) é simples e será omitida.

2. Por linearidade

$$\mathbb{E}\left(\int_\alpha^\beta \phi^2(t) dt\right) = \sum_{i=0}^{n-1} \mathbb{E}[\phi^2(t_i)](t_{i+1} - t_i). \tag{3.8}$$

   Além disso, este número é finito pela hipótese *3* da Definição 3.1.3. Portanto, para todo i, $\mathbb{E}\left[\phi^2(t_i)\right] < \infty$ e, pela desigualdade de Cauchy-Schwarz, também $\mathbb{E}[|\phi(t_i)|] < \infty$. Novamente pela linearidade da esperança, obtemos

$$\mathbb{E}\left(\int_\alpha^\beta \phi(t) dB_t\right) = \sum_{i=0}^{n-1} \mathbb{E}\left[\phi(t_i)(B_{t_{i+1}} - B_{t_i})\right].$$

Como $\phi \in \mathbb{M}^2$, temos que $\phi(t_i)$ é $\mathcal{F}_{t_i}$-mensurável. Por outro lado, $(B_{t_{i+1}} - B_{t_i})$ é independente de $\mathcal{F}_{t_i}$ e, assim, $\phi(t_i)$ e $(B_{t_{i+1}} - B_{t_i})$ são independentes. Portanto, para $i = 0, \ldots, n-1$ temos

$$\mathbb{E}\left[\phi(t_i)(B_{t_{i+1}} - B_{t_i})\right] = \underbrace{\mathbb{E}\left[\phi(t_i)\right]}_{<\infty} \underbrace{\mathbb{E}\left[B_{t_{i+1}} - B_{t_i}\right]}_{=0} = 0.$$

Somando em i, (3.6) segue.

3. Como, para cada i, $\phi^2(t_i)$ e $\left(B_{t_{i+1}} - B_{t_i}\right)^2$ são independentes e têm esperança finita, também o produto $\phi^2(t_i)\left(B_{t_{i+1}} - B_{t_i}\right)^2$ tem esperança finita. Por Cauchy-Schwarz (aplicado duas vezes) segue então que, para quaisquer i e j,

$$\mathbb{E}\left(\left|\phi(t_i)\phi(t_j)(B_{t_{i+1}} - B_{t_i})\right|\right) < \infty.$$

Se $i < j$, então $(B_{t_{j+1}} - B_{t_j})$ e $\phi(t_i)\phi(t_j)(B_{t_{i+1}} - B_{t_i})$ são independentes, donde

$$\begin{aligned} &\mathbb{E}\left[\phi(t_i)\phi(t_j)(B_{t_{i+1}} - B_{t_i})(B_{t_{j+1}} - B_{t_j})\right] \\ &= \underbrace{\mathbb{E}\left[\phi(t_i)\phi(t_j)(B_{t_{i+1}} - B_{t_i})\right]}_{<\infty} \underbrace{\mathbb{E}(B_{t_{j+1}} - B_{t_j})}_{=0} = 0. \end{aligned}$$

Consequentemente, os termos cruzados se anulam e

$$\begin{aligned} \mathbb{E}\left[\left(\int_\alpha^\beta \phi(t)dB_t\right)^2\right] &= \sum_{i=0}^{n-1} \mathbb{E}\left[\phi^2(t_i)(B_{t_{i+1}} - B_{t_i})^2\right] \\ &= \sum_{i=0}^{n-1} \mathbb{E}\left[\phi^2(t_i)\right]\mathbb{E}\left[(B_{t_{i+1}} - B_{t_i})^2\right] \\ &= \sum_{i=0}^{n-1} \mathbb{E}\left[\phi^2(t_i)\right](t_{i+1} - t_i) \\ &= \mathbb{E}\left(\int_\alpha^\beta \phi^2(t)dt\right) \end{aligned}$$

por (3.8). Logo, vale (3.7).

□

A igualdade (3.7) é chamada de *Isometria de Itô*. Esta denominação é justificada pela seguinte observação. Pelo fato de ser mensurável, o processo estocástico ϕ também pode ser pensado como uma variável aleatória definida no espaço $([\alpha,\beta]\times\Omega, \mathcal{B}\times\mathcal{F}, Leb\times\mathbb{P})$ e, mais ainda, pelo item *3* da Definição 3.1.3, temos que $\phi\in L^2([\alpha,\beta]\times\Omega, \mathcal{B}\times\mathcal{F}, Leb\times\mathbb{P})$. Sendo assim, a igualdade (3.7) afirma que as variáveis aleatórias

$$Y \equiv \int_\alpha^\beta \phi(t)dB_t \in L^2(\Omega,\mathcal{F},\mathbb{P})$$

e

$$\phi\in L^2([\alpha,\beta]\times\Omega, \mathcal{B}\times\mathcal{F}, Leb\times\mathbb{P})$$

têm as mesmas normas nos respectivos espaços. De fato,

$$\|Y\|^2_{L^2(\Omega)} = \int_\Omega Y^2 d\mathbb{P} = \int_\Omega \left(\int_\alpha^\beta \phi(t)dB_t\right)^2 d\mathbb{P} = \mathbb{E}\left[\left(\int_\alpha^\beta \phi(t)dB_t\right)^2\right]$$

e

$$\|\phi\|^2_{L^2([\alpha,\beta]\times\Omega)} = \int_{[\alpha,\beta]\times\Omega} \phi^2 d(Leb\times\mathbb{P}) = \int_\Omega \left(\int_\alpha^\beta \phi^2(t)dt\right)d\mathbb{P} = \mathbb{E}\left(\int_\alpha^\beta \phi^2(t)dt\right),$$

a igualdade central sendo justificada pelo Teorema de Fubini.

3.2 Aproximação por processos elementares

O lema seguinte mostra que os processos elementares são densos em $\mathbb{M}^2$, considerando-se a norma referida acima. Este fato, juntamente com a isometria (3.7), nos permitirá estender a definição da integral de Itô para toda a classe $\mathbb{M}^2$.

Lema 3.2.1. *Seja $(X_t)_{t\in[\alpha,\beta]}$ um processo estocástico em $\mathbb{M}^2(\alpha,\beta)$. Então existe uma sequência $\{\phi_n\}_{n=1}^\infty$ de processos elementares tais que*

$$\lim_{n\to\infty} \mathbb{E}\left(\int_\alpha^\beta |X_t-\phi_n(t)|^2 dt\right) = 0. \tag{3.9}$$

A demonstração é feita em 3 passos.

Passo 1. *Seja $g \in \mathbb{M}^2$ limitada e tal que $g(\cdot, \omega)$ é contínua para todo ω. Então existe uma sequência de funções elementares $\{\phi_n\}_{n=1}^{\infty}$ tal que*

$$\lim_{n\to\infty} \mathbb{E}\left(\int_\alpha^\beta |g-\phi_n|^2 dt\right) = 0.$$

Demonstração. Seja M tal que $|g(t,\omega)| < M$ para quaisquer t e ω. Para cada n tomemos uma partição $P_n = \{\alpha = t_0^n < \ldots < t_{m_n}^n = \beta\}$ de tal modo que $|P_n| \to 0$ quando $n \to \infty$. Em seguida, para cada n definimos ϕ_n pondo

$$\phi_n(t,\omega) = \sum_{j=0}^{m_n-1} g(t_j^n, \omega) I_{[t_j^n, t_{j+1}^n]}(t).$$

Então ϕ_n é elementar, pois $g \in \mathbb{M}^2$. Além disso, temos que

$$\lim_{n\to\infty} \int_\alpha^\beta |g(t,\omega) - \phi_n(t,\omega)|^2 dt = 0 \text{ para cada } \omega, \tag{3.10}$$

pois $g(\cdot,\omega)$ é contínua para cada $\omega \in \Omega$. Notemos também que, para quaisquer t, ω e n, temos $|g(t,\omega) - \phi_n(t,\omega)| < 2M$ e consequentemente

$$\int_\alpha^\beta |g(t,\omega) - \phi_n(t,\omega)|^2 dt < 4M^2(\beta - \alpha). \tag{3.11}$$

Portanto, por (3.10), (3.11) e pelo Teorema da Convergência Dominada,

$$\lim_{n\to\infty} \mathbb{E}\left(\int_\alpha^\beta |g-\phi_n|^2 dt\right) = 0.$$

□

Passo 2. *Seja $h \in \mathbb{M}^2$ limitada. Então existe uma sequência $\{g_n\}_{n=1}^{\infty}$ de funções limitadas em $\mathbb{M}^2$ tal que $g_n(\cdot,\omega)$ é contínua para todos ω e n, e*

$$\lim_{n\to\infty} \mathbb{E}\left(\int_\alpha^\beta |h-g_n|^2 dt\right) = 0.$$

Demonstração. Seja M tal que $|h(t,\omega)| \leq M$ para quaisquer t e ω. Para cada n, consideremos uma função contínua $\varphi_n : \mathbb{R} \to \mathbb{R}$ satisfazendo às seguintes condições:

(i) $\varphi_n(x) = 0$ para $x < -\frac{1}{n}$ e $x \geq 0$;

(ii) $\displaystyle\int_{-\infty}^{\infty} \varphi_n(x)dx = 1$.

A sequência de funções $\varphi_n : \mathbb{R} \to \mathbb{R}$ é chamada uma sucessão de núcleos de Dirac. A partir disso, para cada n, definimos

$$g_n(t,\omega) = \int_0^t \varphi_n(s-t)h(s,\omega)ds. \tag{3.12}$$

Então $g_n(\cdot,\omega)$ é contínua para quaisquer ω e n, pois o integrando acima é limitado. Temos também que $|g_n(t,\omega)| \leq M$, de fato

$$\begin{aligned} \left|\int_0^t \varphi_n(s-t)h(s,\omega)ds\right| &\leq \int_0^t \varphi_n(s-t)|h(s,\omega)|ds \\ &\leq \int_0^t \varphi_n(s-t)Mds \\ &\leq M\int_{-\infty}^{\infty} \varphi_n(s-t)ds = M. \end{aligned}$$

Pelas propriedades de $h \in \mathbb{M}^2$ e pelo fato de que a integral que define $g_n(t,\cdot)$ não envolve valores de $h(s,\cdot)$ para $s > t$ temos que também $g_n \in \mathbb{M}^2$. Além disso, para cada ω, pode-se provar que

$$\lim_{n\to\infty} \int_\alpha^\beta |h(t,\omega) - g_n(t,\omega)|^2 dt = 0, \tag{3.13}$$

pois as φ_n são aproximações da distribuição delta de Dirac. Notemos também que para quaisquer t, ω e n temos $|h(t,\omega) - g_n(t,\omega)| \leq 2M$ de modo que

$$\int_\alpha^\beta |h(t,\omega) - g_n(t,\omega)|^2 dt < 4M^2(\beta - \alpha). \tag{3.14}$$

Por (3.13), (3.14) e pelo Teorema da Convergência Dominada, concluímos que

$$\lim_{n\to\infty} \mathbb{E}\left(\int_\alpha^\beta |h - g_n|^2 dt\right) = 0.$$

□

Passo 3. *Seja* $(X_t)_{t\in[\alpha,\beta]} \in \mathbb{M}^2$*. Então existe uma sequência* $\{h_n\}_{n=1}^{\infty}$ *de funções limitadas em* $\mathbb{M}^2$ *tal que*

$$\lim_{n\to\infty} \mathbb{E}\left(\int_\alpha^\beta |X_t - h_n(t)|^2 dt\right) = 0.$$

Demonstração. Definamos

$$h_n(t,\omega) = \begin{cases} -n, & \text{se} \quad X_t(\omega) < -n \\ X_t(\omega), & \text{se} \quad -n \leq X_t(\omega) \leq n \\ n, & \text{se} \quad X_t(\omega) > n \end{cases}.$$

Então, para cada n, claramente, h_n é limitada e está em $\mathbb{M}^2$. Para cada par $(t,\omega) \in [\alpha,\beta] \times \Omega$ temos que

$$\lim_{n\to\infty} h_n(t,\omega) = X_t(\omega)$$

e

$$|h_n(t,\omega)| \leq |X_t(\omega)|, \quad \text{para todo } n.$$

Portanto, o resultado segue do Teorema da Convergência Dominada, no espaço $L^2([\alpha,\beta]\times\Omega)$. □

É claro que, juntos, os passos 1, 2 e 3 demonstram o Lema 3.2.1.

3.3 Integral de Itô para processos mais gerais

Agora, estamos aptos a definir a Integral de Itô para toda a classe $\mathbb{M}^2$.

Definição 3.3.1. *Seja* $(X_t)_{t\in[\alpha,\beta]}$ *um processo estocástico em* $\mathbb{M}^2(\alpha,\beta)$*. Fixemos uma sequência* $\{\phi_n\}_{n=1}^{\infty}$ *de processos elementares tal que*

$$\lim_{n\to\infty} \mathbb{E}\left(\int_\alpha^\beta |X_t - \phi_n(t)|^2 dt\right) = 0$$

e definimos

$$\int_\alpha^\beta X_t dB_t \equiv \lim_{n\to\infty} \int_\alpha^\beta \phi_n(t) dB_t,$$

como um limite em $L^2(\Omega)$*.*

O limite acima existe, pois pela Isometria de Itô para processos elementares (3.7) nota-se que

$$\mathbb{E}\left[\left(\int_\alpha^\beta \phi_n(t)dB_t - \int_\alpha^\beta \phi_m(t)dB_t\right)^2\right] = \mathbb{E}\left(\int_\alpha^\beta |\phi_n(t)-\phi_m(t)|^2 dt\right) \to 0$$

quando $n, m \to \infty$. Portanto, $\int_\alpha^\beta \phi_n(t)dB_t$ é uma sequência de Cauchy no espaço completo $L^2(\Omega)$. Além disso, o limite não depende da particular sequência ϕ_n escolhida. De fato, suponhamos que $\{\psi_n\}_{n=1}^\infty$ é outra sequência também satisfazendo $\lim_{n\to\infty} \mathbb{E}\left(\int_\alpha^\beta |X_t - \psi_n(t)|^2 dt\right) = 0$. Definimos uma sequência $\{\Phi_n\}_{n=1}^\infty$ pondo $\Phi_{2n} = \phi_n$ e $\Phi_{2n-1} = \psi_n$. Então, teremos que também $\lim_{n\to\infty} \mathbb{E}\left(\int_\alpha^\beta |X_t - \Phi_n(t)|^2 dt\right) = 0$ e, como vimos acima, isto implica que $\lim_{n\to\infty} \int_\alpha^\beta \Phi_n(t)dB_t$ existe em $L^2(\Omega)$. Consequentemente, as subsequências $\{\int_\alpha^\beta \phi_n(t)dB_t\}_{n=1}^\infty$ e $\{\int_\alpha^\beta \psi_n(t)dB_t\}_{n=1}^\infty$ devem convergir para o mesmo limite. Em termos práticos, uma sequência padrão $\{\phi_n\}_{n=1}^\infty$ costuma ser tomada discretizando-se o processo $(X_t)_{t\in[\alpha,\beta]}$ em partições do intervalo $[\alpha, \beta]$ que têm norma tendendo a zero com avaliação no extremo esquerdo de cada intervalo, conforme já ilustramos anteriormente.

Teorema 3.3.2 (Propriedades da Integral de Itô para processos gerais em $\mathbb{M}^2$)**.** *Sejam $(X_t)_{t\in[\alpha,\beta]}$ e $(Y_t)_{t\in[\alpha,\beta]}$ processos estocásticos em $\mathbb{M}^2[\alpha,\beta]$, a e b números reais, então:*

$$\int_\alpha^\beta aX_t + bY_t dB_t = a\int_\alpha^\beta X_t dB_t + b\int_\alpha^\beta Y_t dB_t; \tag{3.15}$$

$$\mathbb{E}\left(\int_\alpha^\beta X_t dB_t\right) = 0; \tag{3.16}$$

$$\textit{(Isometria de Itô)}\quad \mathbb{E}\left[\left(\int_\alpha^\beta X_t dB_t\right)^2\right] = \mathbb{E}\left(\int_\alpha^\beta X_t^2 dt\right); \tag{3.17}$$

$$\mathbb{E}\left(\int_\alpha^\beta X_t dB_t \int_\alpha^\beta Y_t dB_t\right) = \mathbb{E}\left(\int_\alpha^\beta X_t Y_t dt\right). \tag{3.18}$$

Demonstração. As três primeiras igualdades são consequências das correspondentes identidades para processos elementares, provadas no Lema 3.1.6. Com mais detalhes: quando a convergência é em L^2 podemos comutar o limite com a esperança. De fato, se $Y_n \to Y$ em $L^2(\Omega)$ então pela desigualdade de Cauchy-Schwarz temos que

$$\mathbb{E}(Y_n) - \mathbb{E}(Y) = \mathbb{E}[(Y_n - Y)1] \leq (\mathbb{E}[(Y_n - Y)^2])^{1/2} 1^{1/2} \to 0.$$

Fazendo o mesmo com $\mathbb{E}(Y) - \mathbb{E}(Y_n)$, concluímos que $\mathbb{E}(Y_n) \to \mathbb{E}(Y)$. Com isto, prova-se (3.16). Podemos também usar a continuidade do produto interno, e consequentemente da norma, e então a Isometria de Itô (3.17) segue naturalmente da análoga para processos elementares. A igualdade (3.18) resulta da identidade $2ab = (a+b)^2 - a^2 - b^2$ e da isometria (3.17). $\square$

Notemos que $\int_\alpha^\beta X_t dB_t + \int_\beta^\gamma X_t dB_t = \int_\alpha^\gamma X_t dB_t$ quando $0 \leq \alpha < \beta < \gamma$. Além disso, salientamos que a variável aleatória $\int_\alpha^\beta X_t dB_t$ é $\mathcal{F}_\beta$-mensurável. Claramente estas propriedades valem para integrandos elementares e são mantidas quando tomamos limites, portanto valem para processos gerais em $\mathbb{M}^2$.

Como primeiro exemplo calcularemos $\int_\alpha^\beta B_t dB_t$. Para isto, precisaremos do seguinte importante resultado.

Lema 3.3.3 (Variação quadrática do movimento browniano)**.** *Seja* $[\alpha, \beta]$ *um intervalo em* $[0, \infty)$*, e suponhamos que*

$$P_n = \{\alpha = t_0^n < t_1^n < \ldots < t_{m_n}^n = \beta\}$$

são partições de $[\alpha, \beta]$*, com* $|P_n| \to 0$ *quando* $n \to \infty$*. Então*

$$\sum_{k=0}^{m_n - 1} \left(B_{t_{k+1}^n} - B_{t_k^n} \right)^2 \to \beta - \alpha$$

em $L^2(\Omega)$ *quando* $n \to \infty$*.*

Obs. Frequentemente este resultado é expresso pela fórmula $(dB_t)^2 = dt$.

Demonstração. Para simplificar a notação introduzimos $\Delta_k^n(B) \equiv B_{t_{k+1}^n} - B_{t_k^n}$ e $\Delta_k^n \equiv t_{k+1}^n - t_k^n$. Seja $Q_n \equiv \sum_{k=0}^{m_n-1} \left(\Delta_k^n(B)\right)^2$, então

$$Q_n - (\beta - \alpha) = \sum_{k=0}^{m_n - 1} \left((\Delta_k^n(B))^2 - \Delta_k^n \right),$$

e assim

$$\mathbb{E}\left[(Q_n-(\beta-\alpha))^2\right]=\sum_{k=0}^{m_n-1}\sum_{j=0}^{m_n-1}\mathbb{E}\left[\left((\Delta_k^n(B))^2-\Delta_k^n\right)\left((\Delta_j^n(B))^2-\Delta_j^n\right)\right].$$

Para $k\neq j$, de acordo com a independência dos incrementos, o termo no somatório duplo é

$$\mathbb{E}\left((\Delta_k^n(B))^2-\Delta_k^n\right)\mathbb{E}\left((\Delta_j^n(B))^2-\Delta_j^n\right)=0,$$

pois $\mathbb{E}\left(\Delta_k^n(B))^2\right)=\Delta_k^n$. Portanto

$$\begin{aligned}\mathbb{E}\left[(Q_n-(\beta-\alpha))^2\right] &= \sum_{k=0}^{m_n-1}\mathbb{E}\left[\left((\Delta_k^n(B))^2-\Delta_k^n\right)^2\right]\\ &= \sum_{k=0}^{m_n-1}\mathbb{E}\left[\left(Y^2-1\right)^2\right](\Delta_k^n)^2,\end{aligned}$$

onde

$$Y=Y_k^n\equiv\frac{\Delta_k^n(B)}{\sqrt{\Delta_k^n}}\sim N(0,1).$$

Consequentemente, sendo $C\equiv\mathbb{E}\left[(Y^2-1)^2\right]$, temos que

$$\begin{aligned}\mathbb{E}\left[(Q_n-(\beta-\alpha))^2\right] &= C\sum_{k=0}^{m_n-1}(\Delta_k^n)^2\\ &\leq C|P_n|(\beta-\alpha)\to 0 \text{ quando } n\to\infty.\end{aligned}$$

□

Exemplo 3.3.4.

$$\int_\alpha^\beta B_t dB_t=\frac{1}{2}(B_\beta^2-B_\alpha^2)-\frac{1}{2}(\beta-\alpha)$$

Demonstração. Antes de mais nada, para que isto faça sentido, precisamos nos certificar que o processo $(B_t)_{t\in[\alpha,\beta]}$ está em $\mathbb{M}^2(\alpha,\beta)$.

1. A prova da mensurabilidade do movimento browniano não é difícil e usa basicamente a continuidade das trajetórias, mas não a faremos aqui[1].

[1] Veja, por exemplo, Billingsley [3], Theorem 37.2.

2. Naturalmente, $(B_t)_{t\geq 0}$ é adaptado à $(\mathcal{F}_t)_{t\geq 0}$.

3. Por Fubini, $\mathbb{E}\left(\int_\alpha^\beta |B_t|^2 dt\right) = \int_\alpha^\beta \mathbb{E}\left(|B_t|^2\right) dt = \int_\alpha^\beta t dt = \dfrac{\beta^2 - \alpha^2}{2} < \infty$

Logo, $(B_t)_{t\in[\alpha,\beta]} \in \mathbb{M}^2(\alpha, \beta)$.

Construímos processos elementares aproximando $(B_t)_{t\in[\alpha,\beta]}$ considerando partições $P_n = \{\alpha = t_0^n, \ldots, t_{m_n}^n = \beta\}$ do intervalo $[\alpha, \beta]$ tais que $|P_n| \to 0$ quando $n \to \infty$. Deste modo

$$\int_\alpha^\beta B_t dB_t = \lim_{n\to\infty} \sum_{k=0}^{m_n-1} B_{t_k^n}\left(B_{t_{k+1}^n} - B_{t_k^n}\right), \tag{3.19}$$

onde o limite é no sentido de $L^2(\Omega)$. Usando a identidade

$$a(b-a) = \frac{1}{2}(b^2 - a^2) - \frac{1}{2}(b-a)^2$$

em (3.19), obtemos

$$\begin{aligned} \int_\alpha^\beta B_t dB_t &= \frac{1}{2}\lim_{n\to\infty} \sum_{k=0}^{m_n-1}\left(B_{t_{k+1}^n}^2 - B_{t_k^n}^2\right) - \frac{1}{2}\lim_{n\to\infty}\sum_{k=0}^{m_n-1}\left(B_{t_{k+1}^n} - B_{t_k^n}\right)^2 \\ &= \frac{1}{2}(B_\beta^2 - B_\alpha^2) - \frac{1}{2}(\beta - \alpha) \end{aligned}$$

pelo Lema 3.3.3. □

O termo $-\frac{1}{2}(\beta - \alpha)$ mostra que a integral de Itô não se comporta como a integral usual. Adiante veremos a Fórmula de Itô, que explica o aparecimento deste termo extra e facilita o cálculo de algumas integrais.

3.4 Integral Indefinida

Definição 3.4.1. *Dado um processo estocástico $(X_t)_{t\in[0,T]} \in \mathbb{M}^2(0,T)$, o processo estocástico $(I_t)_{t\in[0,T]}$, definido por*

$$I_t = \int_0^t X_s dB_s,$$

é chamado de integral indefinida de $(X_t)_{t\in[0,T]}$.

Obs. Assuma $I_0 = \int_0^0 X_s dB_s = 0$ por definição.

Nesta seção, mostraremos duas importantes propriedades dos processos $(I_t)_{t\in[0,T]}$, a saber: que $(I_t)_{t\in[0,T]}$ é um martingale e que tem trajetórias contínuas (a menos de modificação).

Teorema 3.4.2. *Se $(X_t)_{t\in[0,T]} \in \mathbb{M}^2(0,T)$, então a sua integral indefinida $(I_t)_{t\in[0,T]}$ é um martingale em relação à filtração $(\mathcal{F}_t)_{t\geq 0}$ implícita na definição da classe $\mathbb{M}^2$.*

Demonstração. Recordando a definição, precisamos mostrar que:

1. $\mathbb{E}[|I_t|] < \infty$ para todo $t \in [0,T]$;
2. $(I_t)_{t\in[0,T]}$ é adaptado à filtração $(\mathcal{F}_t)_{t\in[0,T]}$;
3. $\mathbb{E}[I_t|\mathcal{F}_s] = I_s$ q.c. para todo $t \geq s$.

1. Pela desigualdade de Cauchy-Schwarz, pela Isometria de Itô, e porque $(X_t)_{t\in[0,T]} \in \mathbb{M}^2(0,T)$, segue que

$$(\mathbb{E}[|I_t|])^2 \leq \mathbb{E}[|I_t|^2] = \mathbb{E}\left[\left(\int_0^t X_s dB_s\right)^2\right] = \mathbb{E}\left(\int_0^t X_s^2 ds\right) < \infty.$$

2. Para cada t, $\int_0^t X_s dB_s$ é $\mathcal{F}_t$-mensurável pois é limite de variáveis aleatórias $\mathcal{F}_t$-mensuráveis.

3. Podemos considerar integrandos elementares, pois, se valer neste caso particular, o resultado geral se estende ao tomarmos limites. Seja então, ϕ um processo elementar em $\mathbb{M}^2[0,T]$. Para $t > s$, temos que

$$\begin{aligned}
\mathbb{E}\left[\int_0^t \phi(u)dB_u|\mathcal{F}_s\right] &= \mathbb{E}\left[\int_0^s \phi(u)dB_u|\mathcal{F}_s\right] + \mathbb{E}\left[\int_s^t \phi(u)dB_u|\mathcal{F}_s\right] \\
&= \int_0^s \phi(u)dB_u + \mathbb{E}\left[\int_s^t \phi(u)dB_u|\mathcal{F}_s\right], \qquad (3.20)
\end{aligned}$$

pois $\int_0^s \phi(u)dB_u$ é $\mathcal{F}_s$-mensurável. Portanto, resta mostrar que a segunda parcela em (3.20) se anula. Digamos que, restrito ao intervalo $[s,t]$, ϕ tenha a forma $\phi(u) = \sum_{k=0}^{n-1} \phi(t_k) I_{[t_k,t_{k+1})}(u)$ onde $\{s = t_0 < t_1 < \ldots < t_n = t\}$ é uma partição do intervalo $[s, t]$. Então, usando as propriedades da esperança condicional, obtemos

$$\begin{aligned}
\mathbb{E}\left[\int_s^t \phi(u)dB_u|\mathcal{F}_s\right] &= \sum_{k=0}^{n-1}\mathbb{E}\left[\phi(t_k)(B_{t_{k+1}}-B_{t_k})|\mathcal{F}_s\right] \\
\text{Item } \mathit{9} \text{ Teo. 2.5.9} \quad &= \sum_{k=0}^{n-1}\mathbb{E}\left[\mathbb{E}\left[\phi(t_k)(B_{t_{k+1}}-B_{t_k})|\mathcal{F}_{t_k}\right]|\mathcal{F}_s\right] \\
\text{Item } \mathit{7} \text{ Teo. 2.5.9} \quad &= \sum_{k=0}^{n-1}\mathbb{E}\left[\phi(t_k)\mathbb{E}\left[(B_{t_{k+1}}-B_{t_k})|\mathcal{F}_{t_k}\right]|\mathcal{F}_s\right] \\
\text{Item } \mathit{8} \text{ Teo. 2.5.9} \quad &= \sum_{k=0}^{n-1}\mathbb{E}\left[\phi(t_k)\mathbb{E}\left[(B_{t_{k+1}}-B_{t_k})\right]|\mathcal{F}_s\right] = 0.
\end{aligned}$$

□

Teorema 3.4.3. *Seja* $(X_t)_{t\in[0,T]} \in \mathbb{M}^2(0,T)$. *Então a integral indefinida* $(I_t)_{t\in[0,T]}$ *dada por*

$$I_t = \int_0^t X_s dB_s$$

admite uma modificação com trajetórias contínuas. Isto é, existe um processo $(J_t)_{t\in[0,T]}$ *com trajetórias contínuas quase certamente, tal que* $\mathbb{P}[I_t = J_t] = 1$, *para todo* $t \in [0,T]$.

Demonstração. Primeiramente, notemos que se ψ é um processo elementar em $\mathbb{M}^2(0,T)$ então a sua integral indefinida $(\int_0^t \psi(s)dB_s)_{t\in[0,T]}$ tem trajetórias contínuas quase certamente. De fato, suponhamos que o processo ψ seja dado por $\psi(s) = \sum_{k=0}^{n-1}\psi(t_k)I_{[t_k,t_{k+1})}(s)$. Então, se $t_m \le t < t_{m+1}$ temos que

$$\int_0^t \psi(s)dB_s = \sum_{k=0}^{m-1}\psi(t_k)\left(B_{t_{k+1}} - B_{t_k}\right) + B_t - B_{t_m}.$$

Assim, se $t \neq t_m$ a continuidade em t segue diretamente da continuidade das trajetórias do movimento browniano. Se, entretanto, $t = t_m$ então é simples verificar, através da expressão acima, que $\lim_{r\to t_m^-}\int_0^r \psi(s)dB_s = \lim_{r\to t_m^+}\int_0^r \psi(s)dB_s = \int_0^{t_m}\psi(s)dB_s$. Portanto, o resultado vale para integrandos elementares.

Seja $\{\phi_n\}_{n=1}^{\infty}$ uma sequência de processos elementares em $\mathbb{M}^2(0,T)$ tais que

$$\lim_{n\to\infty} \mathbb{E}\left(\int_0^T |X_t - \phi_n(t)|^2 dt\right) = 0.$$

Denotemos $I_t^n = \int_0^t \phi_n(t) dB_t$. Pelo que observamos acima, os processos $(I_t^n)_{t\in[0,T]}$ têm trajetórias contínuas quase certamente. Além disso, pelo teorema anterior $(I_t^n)_{t\in[0,T]}$ é um martingale pra cada n e, consequentemente, $(I_t^n - I_t^m)_{t\in[0,T]}$ é um martingale para qualquer par n, m. Portanto, pelo Corolário 2.5.20 e pela Isometria de Itô temos que, para qualquer $\varepsilon > 0$,

$$\begin{aligned} \mathbb{P}\left(\sup_{0\le t\le T} |I_t^n - I_t^m| > \varepsilon\right) &\le \frac{1}{\varepsilon^2}\mathbb{E}\left(|I_T^n - I_T^m|^2\right) \\ &= \frac{1}{\varepsilon^2}\mathbb{E}\left(\int_0^T |\phi_n(t) - \phi_m(t)|^2 dt\right) \to 0 \end{aligned}$$

quando $n, m \to \infty$. Em particular, para $\varepsilon = \frac{1}{2^k}$ existe um n_k tal que

$$\mathbb{P}\left(\sup_{0\le t\le T} |I_t^{n_k} - I_t^m| > \frac{1}{2^k}\right) < \frac{1}{k^2},$$

para todo $m \ge n_k$. Podemos escolher os n_k's de tal forma que $n_k < n_{k+1}$. Assim, obtemos um subsequência $\{n_k\}_{k=1}^{\infty}$ tal que

$$\mathbb{P}\left(\sup_{0\le t\le T} |I_t^{n_k} - I_t^{n_{k+1}}| > \frac{1}{2^k}\right) < \frac{1}{k^2}.$$

Como $\sum \frac{1}{k^2}$ converge, pelo Lema de Borel-Cantelli temos que

$$\mathbb{P}\left(\sup_{0\le t\le T} |I_t^{n_k} - I_t^{n_{k+1}}| > \frac{1}{2^k} \text{ para infinitos valores de } k\right) = 0.$$

Ou seja, para quase todo $\omega \in \Omega$ existe um número $k_0(\omega)$ tal que

$$\sup_{0\le t\le T} |I_t^{n_k}(\omega) - I_t^{n_{k+1}}(\omega)| \le \frac{1}{2^k}, \text{ para todo } k \ge k_0(\omega).$$

Para qualquer $j > k$, usando a desigualdade triangular, obtemos

$$\sup_{0\le t\le T} |I_t^{n_k}(\omega) - I_t^{n_j}(\omega)| \le \sum_{i=0}^{j-k-1} \sup_{0\le t\le T} |I_t^{n_{k+i}}(\omega) - I_t^{n_{k+i+1}}(\omega)| \le \sum_{i=0}^{j-k-1} \frac{1}{2^{k+i}} < \frac{1}{2^{k-1}}$$

para todo $k \geq k_0(\omega)$. Logo, $\{I_t^{n_k}(\omega)\}_{k=1}^{\infty}$ é uma sequência de Cauchy (uniformemente em t) e então converge t-uniformemente, para quase todo $\omega \in \Omega$. Consequentemente, o limite, denotado por $J_t(\omega)$, é t-contínuo, para $t \in [0, T]$, para quase todo $\omega \in \Omega$. Mas temos também que $I_t^{n_k} \to I_t$, em $L^2(\Omega)$, quando $k \to \infty$, para todo $t \in [0, T]$ e, consequentemente[2], para cada t, existe uma subsequência de $\{I_t^{n_k}\}_{k=1}^{\infty}$ que converge quase certamente para I_t. Portanto, devemos ter $J_t = I_t$ quase certamente, para todo $t \in [0, T]$. Isto completa a demonstração. □

3.5 Uma possível extensão da definição

A definição de Integral de Itô dada neste texto se restringe a integrandos pertencentes à classe $\mathbb{M}^2$ (veja Definição 3.1.3). Entretanto, é possível definir (de uma maneira diferente) a Integral de Itô para uma classe mais ampla que $\mathbb{M}^2$. A condição

3. $\mathbb{E}\left(\int_{\alpha}^{\beta} |X_t|^2 dt\right) < \infty$

pode ser enfraquecida e podemos exigir apenas que

3'. $\mathbb{P}\left(\int_{\alpha}^{\beta} |X_t|^2 dt < \infty\right) = 1.$

Denotemos por $\mathbb{L}^2(\alpha, \beta)$ a classe dos processos estocásticos $(X_t)_{t\in[\alpha,\beta]}$ que satisfazem *1* e *2* da Definição 3.1.3 e *3'*. Neste caso, pode-se mostrar que para cada $(X_t)_{t\in[\alpha,\beta]} \in \mathbb{L}^2(\alpha, \beta)$ existe uma sequência $\{\phi_n\}_{n=1}^{\infty}$ de processos simples em $\mathbb{L}^2(\alpha, \beta)$ tais que

$$\lim_{n\to\infty} \int_{\alpha}^{\beta} |X_t - \phi_n(t)|^2 dt = 0 \quad \text{quase certamente.}$$

Para uma tal sequência, é possível provar que $\{\int_{\alpha}^{\beta} \phi_n(t) dB_t\}_{n=1}^{\infty}$ é convergente em probabilidade. A integral $\int_{\alpha}^{\beta} X_t dB_t$ é, então, definida como sendo este limite em probabilidade. Esta construção é feita no Capítulo 4 de Friedman [6]. Para processos em $\mathbb{M}^2$ as duas definições coincidem, já que convergência em em L^2 implica convergência em probabilidade. Assim, quando acharmos conveniente, trataremos a integral $\int_{\alpha}^{\beta} X_t dB_t$ com sendo o limite *em probabilidade* da sequência $\{\int_{\alpha}^{\beta} \phi_n(t) dB_t\}_{n=0}^{\infty}$ onde as ϕ_n's são os processos elementares que aproximam o processo $(X_t)_{t\in[\alpha,\beta]}$.

[2] Veja Bartle [2], página 73.

3.6 Fórmula de Itô

De forma análoga como definimos a classe $\mathbb{M}^2$ definiremos agora a classe $\mathbb{M}^1$.

Definição 3.6.1. *Denotemos por $\mathbb{M}^1(\alpha, \beta)$ a classe de processos estocásticos $(X_t)_{t\in[\alpha,\beta]}$ que satisfazem as condições 1 e 2 da Definição 3.1.3 e*

3". $\mathbb{E}\left(\int_\alpha^\beta |X_t| dt\right) < \infty$.

Definição 3.6.2. *Seja $(X_t)_{t\in[0,T]}$ um processo estocástico satisfazendo*

$$X_r = X_s + \int_s^r F(t)dt + \int_s^r G(t)dB_t \tag{3.21}$$

para quaisquer tempos $0 \leq s \leq r \leq T$, onde $F \in \mathbb{M}^1(0, T)$ e $G \in \mathbb{M}^2(0, T)$. Neste caso, dizemos que o processo $(X_t)_{t\in[0,T]}$ tem diferencial estocástico

$$dX_t = Fdt + GdB_t. \tag{3.22}$$

Obs. 1. A notação diferencial (3.22) é apenas uma forma compacta de expressar a relação (3.21). Os símbolos dX_t, dt e dB_t não têm significado sozinhos.

Obs. 2. Se $H \in \mathbb{M}^2$ e $(X_t)_{t\in[0,T]}$ tem diferencial (3.22), então a expressão $dY_t = HdX_t$ significa $dY_t = HFdt + HGdB_t$.

Exemplo 3.6.3. No Exemplo 3.3.4 vimos que

$$\int_\alpha^\beta B_t dB_t = \frac{1}{2}(B_\beta^2 - B_\alpha^2) - \frac{1}{2}(\beta - \alpha),$$

portanto

$$B_\beta^2 = B_\alpha^2 + 2\int_\alpha^\beta B_t dB_t + (\beta - \alpha).$$

Ou seja,

$$d\big(B_t^2\big) = 2B_t dB_t + dt.$$

Note que pela regra da cadeia do cálculo usual, o termo dt não apareceria.

Enunciaremos agora o principal resultado deste capítulo, que é o análogo estocástico da regra da cadeia convencional.

Teorema 3.6.4 (Fórmula de Itô)**.** *Seja* $(X_t)_{t\in[0,T]}$ *um processo estocástico tal que*

$$dX_t = Fdt + GdB_t$$

onde $F \in \mathbb{M}^1(0,T)$ *e* $G \in \mathbb{M}^2(0,T)$*. Seja* $h : \mathbb{R} \times [0,T] \to \mathbb{R}$ *uma função contínua tal que* $\frac{\partial h}{\partial x}$, $\frac{\partial^2 h}{\partial x^2}$ *e* $\frac{\partial h}{\partial t}$ *existem e são contínuas. Defina o processo estocástico* $(Y_t)_{t\in[0,T]}$ *por*

$$Y_t \equiv h(X_t, t).$$

Então

$$dY_t = \frac{\partial h}{\partial t}(X_t,t)dt + \frac{\partial h}{\partial x}(X_t,t)dX_t + \frac{1}{2}\frac{\partial^2 h}{\partial x^2}(X_t,t)G^2dt \tag{3.23}$$

ou, equivalentemente,

$$dY_t = \left(\frac{\partial h}{\partial t}(X_t,t) + \frac{\partial h}{\partial x}(X_t,t)F + \frac{1}{2}\frac{\partial^2 h}{\partial x^2}(X_t,t)G^2\right)dt + \frac{\partial h}{\partial x}(X_t,t)GdB_t. \tag{3.24}$$

Obs. Em alusão à fórmula de Taylor, é comum também expressar esta relação da seguinte maneira:

$$dY_t = \frac{\partial h}{\partial t}(X_t,t)dt + \frac{\partial h}{\partial x}(X_t,t)dX_t + \frac{1}{2}\frac{\partial^2 h}{\partial x^2}(X_t,t)(dX_t)^2,$$

onde $(dX_t)^2$ é formalmente calculado através da tabela de multiplicação abaixo, que fundamenta-se no Lema 3.3.3.

$\times$	dt	dB_t
dt	0	0
dB_t	0	dt

(3.25)

Em termos de integrais, notemos que as fórmulas acima significam que para quaisquer tempos $0 \leq s \leq r \leq T$ vale

$$\begin{aligned} h(X_r,r) = h(X_s,s) + &\int_s^r \left(\frac{\partial h}{\partial t}(X_t,t) + \frac{\partial h}{\partial x}(X_t,t)F(t) + \frac{1}{2}\frac{\partial^2 h}{\partial x^2}(X_t,t)G^2(t)\right)dt \\ &+ \int_s^r \frac{\partial h}{\partial x}(X_t,t)G(t)dB_t. \end{aligned}$$

Demonstração. A prova será dividida em vários passos. Comecemos verificando (3.23) diretamente em dois casos particulares.

Passo 1 (Dois casos simples)**.**

(i) $d(B_t^2) = 2B_t dB_t + dt$;

(ii) $d(tB_t) = B_t dt + t dB_t$.

Obs. Note que as expressões acima são casos particulares de (3.23) quando $X_t = B_t$, o que implica $F = 0$ e $G = 1$. No primeiro caso $h(x,t) = x^2$ e no segundo caso $h(x,t) = tx$.

Demonstração. O item (i) já foi provado (Exemplo 3.6.3). Verifiquemos (ii).

$$\int_s^r t dB_t = \mathbb{P}\lim_{n\to\infty} \sum_{k=0}^{m_n-1} t_k^n \left(B_{t_{k+1}^n} - B_{t_k^n} \right), \tag{3.26}$$

onde $P_n = \{s = t_0^n < t_1^n < \ldots < t_{m_n}^n = r\}$ são partições do intervalo $[r,s]$, com $|P_n| \to 0$ e o limite considerado é em probabilidade (veja o comentário no final da Seção 3.5). Para cada ω para o qual $t \to B_t(\omega)$ é contínua temos também que

$$\int_s^r B_t(\omega) dt = \lim_{n\to\infty} \sum_{k=0}^{m_n-1} B_{t_{k+1}^n}(\omega) \left(t_{k+1}^n - t_k^n \right). \tag{3.27}$$

Note acima que podemos avaliar a função no extremo direito de cada intervalo (ao invés do esquerdo como na integral de Itô) pois se trata de uma integral de Riemann usual. Podemos escrever (3.27) como uma igualdade entre variáveis aleatórias:

$$\int_s^r B_t dt = \mathbb{P}\lim_{n\to\infty} \sum_{k=0}^{m_n-1} B_{t_{k+1}^n} \left(t_{k+1}^n - t_k^n \right), \tag{3.28}$$

onde usamos que convergência quase certa implica convergência em probabilidade. Somando as equações (3.26) e (3.28), usando a linearidade do limite, obtemos

$$\begin{aligned} \int_s^r t dB_t + \int_s^r B_t dt &= \mathbb{P}\lim_{n\to\infty} \sum_{k=0}^{m_n-1} \left[t_k^n \left(B_{t_{k+1}^n} - B_{t_k^n} \right) + B_{t_{k+1}^n} \left(t_{k+1}^n - t_k^n \right) \right] \\ &= \lim_{n\to\infty} \sum_{k=0}^{m_n-1} \left(t_{k+1}^n B_{t_{k+1}^n} - t_k^n B_{t_k^n} \right) = rB_r - sB_s, \end{aligned}$$

por soma telescópica. Ou seja, $rB_r = sB_s + \int_s^r B_t dt + \int_s^r t dB_t$, o que, na notação diferencial, fica $d(tB_t) = B_t dt + t dB_t$ conforme desejávamos. $\square$

Passo 2 (Regra do Produto de Itô)**.** *Suponhamos que* $(X_t)_{t\in[0,T]}$ *e* $(Z_t)_{t\in[0,T]}$ *sejam processos estocásticos tais que*

$$\begin{cases} dX_t &= F_1 dt + G_1 dB_t \\ dZ_t &= F_2 dt + G_2 dB_t \end{cases} \tag{3.29}$$

onde $F_i \in \mathbb{M}^1(0,T)$ *e* $G_i \in \mathbb{M}^2(0,T)$, $i = 1, 2$. *Então*

$$d(X_t Z_t) = Z_t dX_t + X_t dZ_t + G_1 G_2 dt. \tag{3.30}$$

Obs. Note que, em comparação com a correspondente fórmula do cálculo usual, há um termo extra $G_1 G_2 dt$. Encontra-se também a notação

$$d(X_t Z_t) = Z_t dX_t + X_t dZ_t + dX_t dZ_t,$$

onde $dX_t dZ_t$ é calculado de acordo com a tabela de multiplicação (3.25). Em termos de integrais, a regra do produto corresponde à seguinte regra de integração por partes

$$\int_s^r Z_t dX_t = X_r Z_r - X_s Z_s - \int_s^r X_t dZ_t - \int_s^r G_1(t) G_2(t) dt.$$

Em particular, se $f : [0,T] \to \mathbb{R}$ for uma função (determinística) diferenciável, então vale a regra usual de integração por partes

$$\int_s^r f(t) dB_t = f(r) B_r - f(s) B_s - \int_s^r B_t f'(t) dt.$$

Demonstração. A equação (3.30) é equivalente a

$$d(X_t Z_t) = (X_t F_2 + Z_t F_1 + G_1 G_2)\, dt + (X_t G_2 + Z_t G_1)\, dB_t.$$

Logo, precisamos mostrar que, para qualquer tempo $0 \le r \le T$ vale

$$X_r Z_r - X_0 Z_0 = \tag{3.31}$$

$$\int_0^r (X_t F_2(t) + Z_t F_1(t) + G_1(t) G_2(t)) dt + \int_0^r (X_t G_2(t) + Z_t G_1(t)) dB_t. \tag{3.32}$$

Inicialmente, suponhamos que F_i e G_i, $i = 1, 2$, sejam processos constantes no tempo. Neste caso, de (3.29) obtemos que

$$\begin{cases} X_t &= X_0 + F_1 t + G_1 B_t \\ Z_t &= Z_0 + F_2 t + G_2 B_t \end{cases} . \tag{3.33}$$

Usando (3.33), notemos que a expressão em (3.32) é igual a

$$\int_0^r [(X_0 + F_1 t + G_1 B_t)F_2 + (Z_0 + F_2 t + G_2 B_t)F_1 + G_1 G_2]\, dt$$
$$+ \int_0^r [(X_0 + F_1 t + G_1 B_t)G_2 + (Z_0 + F_2 t + G_2 B_t)G_1]\, dB_t$$
$$= X_0(F_2 r + G_2 B_r) + Z_0(F_1 r + G_1 B_r) + F_1 F_2 r^2$$
$$+(F_1 G_2 + G_1 F_2)\left[\int_0^r B_t dt + \int_0^r t dB_t\right] + 2G_1 G_2 \int_0^r B_t dB_t + G_1 G_2 r.$$

Pelo Passo 1, temos que $2\int_0^r B_t dB_t = B_r^2 - r$ e $\int_0^r B_t dt + \int_0^r t dB_t = rB_r$. Logo, a expressão acima se torna

$$= X_0(F_2 r + G_2 B_r) + Z_0(F_1 r + G_1 B_r) + F_1 F_2 r^2 + (F_1 G_2 + G_1 F_2) r B_r + G_1 G_2 B_r^2$$
$$= X_r Z_r - X_0 Z_0,$$

onde a última igualdade é verificada diretamente a partir de (3.33). Se F_i, G_i, $i = 1, 2$, forem processos simples, então aplicamos o resultado que acabamos de provar em cada subintervalo onde F_i, G_i são constantes e somamos as integrais resultantes, obtendo (3.30) também neste caso. O caso geral, como de costume, segue por aproximação por processos simples e passagem ao limite. $\square$

Passo 3 (Caso particular $h(x,t) = x^m$)**.** *Suponhamos que $dX_t = Fdt + GdB_t$, então, para $m = 0, 1, 2, \ldots$,*

$$d(X_t^m) = mX_t^{m-1} dX_t + \frac{1}{2} m(m-1) X_t^{m-2} G^2 dt, \tag{3.34}$$

que é fórmula (3.23) *neste caso particular.*

Demonstração. A prova é feita por indução. O resultado é claro para $m = 0, 1$, e o caso $m = 2$ segue da regra do produto. Supondo que vale o resultado para $m - 1$, isto é

$$\begin{aligned} d(X_t^{m-1}) &= (m-1)X_t^{m-2} dX_t + \frac{1}{2}(m-1)(m-2)X_t^{m-3} G^2 dt \\ &= \left((m-1)X_t^{m-2} F + \frac{1}{2}(m-1)(m-2)X_t^{m-3} G^2\right) dt \\ &\quad +(m-1)X_t^{m-2} G dB_t, \end{aligned}$$

provaremos a validade para m. Pela regra do produto,

$$\begin{aligned} d(X_t^m) &= d(X_t X_t^{m-1}) \\ &= X_t d(X_t^{m-1}) + X_t^{m-1} dX_t + (m-1)X_t^{m-2}G^2dt \\ &= X_t \left((m-1)X_t^{m-2}dX_t + \frac{1}{2}(m-1)(m-2)X_t^{m-3}G^2dt \right) \\ &\quad + X_t^{m-1}dX_t + (m-1)X_t^{m-2}G^2dt \\ &= mX_t^{m-1}dX_t + \frac{1}{2}m(m-1)X_t^{m-2}G^2dt, \end{aligned}$$

onde, na igualdade final, usamos que $m-1+\frac{1}{2}(m-1)(m-2) = \frac{1}{2}m(m-1)$. $\square$

Passo 4. *Vale a Fórmula de Itô no caso em que h é uma função da forma $h(x,t) = p(x)q(t)$, p e q polinômios.*

Demonstração. Pelo Passo 3 e pela linearidade dos operadores diferenciais, o resultado vale para polinômios na variável x, ou seja, se $dX_t = Fdt + GdB_t$ então

$$d(p(X_t)) = p'(X_t)dX_t + \frac{1}{2}p''(X_t)G^2dt.$$

Além disso, a igualdade $q(t) = q(0) + \int_0^t q'(s)ds$ significa que $dq = q'dt$. Pela regra do produto temos, então, que

$$\begin{aligned} d(h(X_t,t)) &= d(p(X_t)q(t)) \\ &= p(X_t)dq(t) + q(t)d(p(X_t)) + 0 \\ &= p(X_t)q'(t)dt + q(t)\left(p'(X_t)dX_t + \frac{1}{2}p''(X_t)G^2dt \right) \\ &= \frac{\partial h}{\partial t}(X_t,t)dt + \frac{\partial h}{\partial x}(X_t,t)dX_t + \frac{1}{2}\frac{\partial^2 h}{\partial x^2}(X_t,t)G^2dt. \end{aligned}$$

$\square$

Por linearidade, segue diretamente que:

Passo 5. *Vale a Fórmula de Itô no caso em que h é uma função da forma $h(x,t) = \sum_{k=1}^n p_k(x)q_k(t)$, p_k e q_k polinômios. Ou seja, a Fórmula de Itô é válida para funções h polinomiais nas variáveis x e t.*

Finalmente, concluímos:

Passo 6 (Caso geral)**.** *Vale a Fórmula de Itô para h qualquer satisfazendo as hipóteses do teorema.*

Demonstração. Dada $h : \mathbb{R} \times [0, T] \to \mathbb{R}$ como na hipótese do teorema, de acordo com uma versão mais forte do Teorema da Aproximação de Weierstrass[3], existe uma sequência $\{h_n\}_{n=1}^{\infty}$ de funções polinomiais tais que

$$h_n \to h,\ \frac{\partial h_n}{\partial x} \to \frac{\partial h}{\partial x},\ \frac{\partial^2 h_n}{\partial x^2} \to \frac{\partial^2 h}{\partial x^2}\ \text{ e }\ \frac{\partial h_n}{\partial t} \to \frac{\partial h}{\partial t}$$

uniformemente em subconjuntos compactos de $\mathbb{R} \times [0, T]$. Pelo Passo 5, para cada $0 \leq r \leq T$ temos que

$$\begin{aligned} h_n(X_r, r) = h_n(X_0, 0) + &\int_0^r \left(\frac{\partial h_n}{\partial t}(X_t, t) + \frac{\partial h_n}{\partial x}(X_t, t) F(t) + \frac{1}{2} \frac{\partial^2 h_n}{\partial x^2}(X_t, t) G^2(t) \right) dt \\ &+ \int_0^r \frac{\partial h_n}{\partial x}(X_t, t) G(t) dB_t. \end{aligned}$$

Assim, o resultado desejado é obtido por passagem ao limite na expressão acima, quando $n \to \infty$. □

Exemplo 3.6.5. Tomemos $X_t = B_t$ (o que implica $F \equiv 0$ e $G \equiv 1$) e $h(x, t) = e^{\lambda x - \frac{\lambda^2 t}{2}}$. Então, pela fórmula de Itô, temos que

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

Portanto $Y_t = e^{\lambda B_t - \frac{\lambda^2 t}{2}}$ satisfaz

$$\begin{cases} dY_t = \lambda Y_t dB_t \\ Y_0 = 1 \end{cases}. \tag{3.35}$$

A expressão (3.35) é um exemplo de *equação diferencial estocástica*, sobre as quais falaremos mais no capítulo seguinte. Portanto, no cálculo estocástico, a expressão $e^{\lambda B_t - \frac{\lambda^2 t}{2}}$ tem propriedade análoga à propriedade da função $e^{\lambda t}$ no cálculo usual.

[3]Proposição 19. Capítulo 8 de Lima [10].

A seguir, enunciamos, sem provar, uma forma mais geral da Fórmula de Itô.

Teorema 3.6.6 (Fórmula de Itô - Extensão)**.** *Sejam* $(X_t^{(i)})_{t\in[0,T]}$, $i = 1, \ldots, n$, *processos estocásticos tais que* $dX_t^{(i)} = F_i dt + G_i dB_t$, *onde* $F_i \in \mathbb{M}^1(0,T)$ *e* $G_i \in \mathbb{M}^2(0,T)$ *para* $i = 1, \ldots, n$*. Seja* $h : \mathbb{R}^n \times [0,T] \to \mathbb{R}$ *contínua com derivadas parciais* $\frac{\partial h}{\partial t}$, $\frac{\partial h}{\partial x_i}$ *e* $\frac{\partial^2 h}{\partial x_i x_j}$, $i, j = 1, \ldots, n$, *contínuas. Defina o processo estocástico* $(Y_t)_{t\in[0,T]}$ *por*

$$Y_t \equiv h(X_t^{(1)}, \ldots, X_t^{(n)}, t).$$

Então

$$dY_t = \frac{\partial h}{\partial t}(X_t, t)dt + \sum_{i=1}^{n} \frac{\partial h}{\partial x_i}(X_t, t)dX_t^{(i)} + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 h}{\partial x_i x_j}(X_t, t)G_i G_j dt. \quad (3.36)$$

Outra notação comumente encontrada:

$$dY_t = \frac{\partial h}{\partial t}(X_t, t)dt + \sum_{i=1}^{n} \frac{\partial h}{\partial x_i}(X_t, t)dX_t^{(i)} + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 h}{\partial x_i x_j}(X_t, t)dX_t^{(i)} dX_t^{(j)},$$

onde $dX_t^{(i)} dX_t^{(j)}$ é formalmente calculado de acordo com a tabela de multiplicação dada em (3.25).

3.7 Integral de Itô n-dimensional

Seja $(B_t)_{t\geq 0} = (B_t^{(1)}, \ldots, B_t^{(m)})_{t\geq 0}$ um movimento browniano m-dimensional definido num espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$. Consideremos uma filtração $(\mathcal{F}_t)_{t\geq 0}$ do espaço $(\Omega, \mathcal{F}, \mathbb{P})$ tal que

(i) $\sigma(B_s, 0 \leq s \leq t) \subset \mathcal{F}_t$ para todo $t \geq 0$.

(ii) $\sigma(B_{t+\lambda} - B_t, \lambda \geq 0)$ é independente de $\mathcal{F}_t$ para todo $t \geq 0$.

Definição 3.7.1. *Dizemos que um processo estocástico* $n \times m$*-dimensional* $(X_t)_{t\in[\alpha,\beta]} = (X_t^{(ij)} | \ i = 1, \ldots, n; \ j = 1, \ldots, m)_{t\in[\alpha,\beta]}$ *pertence à classe* $\mathbb{M}^2_{n\times m}(\alpha, \beta)$ *se cada componente* $(X_t^{(ij)})_{t\in[\alpha,\beta]}$ *pertence à classe* $\mathbb{M}^2(\alpha, \beta)$*. Da mesma forma, define-se* $\mathbb{M}^1_n(\alpha, \beta)$*, como sendo a classe dos processos estocásticos* n*-dimensionais cujas componentes pertencem a* $\mathbb{M}^1(\alpha, \beta)$*.*

Definição 3.7.2. *Seja* $(X_t)_{t\in[\alpha,\beta]}$ *um processo estocástico em* $\mathbb{M}^2_{n\times m}(\alpha,\beta)$. *Então, usando notação matricial, definimos*

$$\int_\alpha^\beta X_t dB_t = \int_\alpha^\beta \begin{pmatrix} X_t^{(11)} & \dots & X_t^{(1m)} \\ \vdots & \ddots & \vdots \\ X_t^{(n1)} & \dots & X_t^{(nm)} \end{pmatrix} \begin{pmatrix} dB_t^{(1)} \\ \vdots \\ dB_t^{(m)} \end{pmatrix}$$

como sendo o vetor aleatório n-dimensional cuja i-ésima componente é

$$\sum_{k=1}^m \int_\alpha^\beta X_t^{(ik)} dB_t^{(k)}.$$

Por aproximação por processos simples, prova-se as seguintes propriedades análogas ao caso unidimensional.

Teorema 3.7.3. *Seja* $(X_t)_{t\in[\alpha,\beta]} \in \mathbb{M}^2_{n\times m}(\alpha,\beta)$, *então*

$$\mathbb{E}\left(\int_\alpha^\beta X_t dB_t\right) = 0 \in \mathbb{R}^n \tag{3.37}$$

e

$$\mathbb{E}\left(\left\|\int_\alpha^\beta X_t dB_t\right\|^2\right) = \mathbb{E}\left(\int_\alpha^\beta \|X_t\|^2 dt\right) \tag{3.38}$$

onde $\|X_t\|^2 = \sum_{i=1}^n \sum_{j=1}^m |X_t^{(ij)}|^2$.

Definição 3.7.4. *Se* $(X_t)_{t\in[0,T]} = (X_t^{(1)},\dots,X_t^{(n)})_{t\in[0,T]}$ *é um processo estocástico n-dimensional tal que*

$$X_r = X_s + \int_s^r F(t)dt + \int_s^r G(t)dB_t$$

para quaisquer tempos $0 \le s \le t \le T$, *onde* $F = (F^{(1)},\dots,F^{(n)}) \in \mathbb{M}^1_n(0,T)$ *e* $G = (G^{(ij)}|\ i = 1,\dots,n;\ j = 1,\dots,m) \in \mathbb{M}^2_{n\times m}(0,T)$, *então dizemos que o processo* $(X_t)_{t\in[0,T]}$ *tem diferencial estocástico*

$$dX_t = Fdt + GdB_t.$$

Isto significa que, para $i = 1,\dots,n$,

$$dX_t^{(i)} = F^{(i)}dt + \sum_{k=1}^m G^{(ik)} dB_t^{(k)}.$$

Enunciamos agora, uma forma ainda mais geral da Fórmula de Itô.

Teorema 3.7.5 (Fórmula de Itô - Extensão 2)**.** *Suponhamos que*

$$dX_t = Fdt + GdB_t,$$

como acima. Seja $h : \mathbb{R}^n \times [0,T] \to \mathbb{R}$ *contínua com derivadas parciais* $\frac{\partial h}{\partial t}$, $\frac{\partial h}{\partial x_i}$ *e* $\frac{\partial^2 h}{\partial x_i x_j}$, $i,j = 1,\ldots,n$, *contínuas. Defina o processo estocástico unidimensional* $(Y_t)_{t\in[0,T]}$ *por*

$$Y_t \equiv h(X_t^{(1)},\ldots,X_t^{(n)},t).$$

Então

$$\begin{aligned} dY_t &= \frac{\partial h}{\partial t}(X_t,t)dt + \sum_{i=1}^{n}\frac{\partial h}{\partial x_i}(X_t,t)dX_t^{(i)} \\ &+\frac{1}{2}\sum_{i,j=1}^{n}\frac{\partial^2 h}{\partial x_i x_j}(X_t,t)\sum_{k=1}^{m}G^{(ik)}G^{(jk)}dt. \end{aligned} \tag{3.39}$$

É também encontrada a seguinte notação:

$$\begin{aligned} dY_t &= \frac{\partial h}{\partial t}(X_t,t)dt + \sum_{i=1}^{n}\frac{\partial h}{\partial x_i}(X_t,t)dX_t^{(i)} \\ &+\frac{1}{2}\sum_{i,j=1}^{n}\frac{\partial^2 h}{\partial x_i x_j}(X_t,t)d(X_t^{(i)})d(X_t^{(j)}), \end{aligned}$$

onde $d(X_t^{(i)})d(X_t^{(j)})$ é calculado através da tabela de multiplicação

$\times$	dt	$dB_t^{(l)}$
dt	0	0
$dB_t^{(k)}$	0	$\delta_{kl}dt$

(3.40)

Exemplo 3.7.6. Sejam $(B_t^{(1)})_{t\in[0,T]}$ e $(B_t^{(2)})_{t\in[0,T]}$ dois movimentos brownianos independentes então, pela Fórmula de Itô, aplicada a $h(x_1,x_2,t) = x_1x_2$, temos que

$$d(B_t^{(1)}B_t^{(2)}) = B_t^{(1)}dB_t^{(2)} + B_t^{(2)}dB_t^{(1)}.$$

Compare com o caso $B_t^{(1)} = B_t^{(2)}$, Exemplo 3.6.3. Note que não há o termo de correção dt neste caso.

Capítulo 4

Equações Diferenciais Estocásticas

4.1 Definições

Sejam $(B_t)_{t\geq 0}$ um movimento browniano m-dimensional definido num espaço de probabilidade abstrato $(\Omega, \mathcal{F}, \mathbb{P})$ e Y uma variável aleatória n-dimensional definida no mesmo espaço e independente do processo $(B_t)_{t\geq 0}$. Consideremos em $(\Omega, \mathcal{F}, \mathbb{P})$ a filtração $(\mathcal{F}_t)_{t\geq 0}$ gerada pelo processo $(B_t)_{t\geq 0}$ e pela variável aleatória Y. Ou seja, para cada tempo $t \geq 0$, $\mathcal{F}_t = \sigma(Y, B_s, 0 \leq s \leq t)$. Fixado um tempo T, suponhamos dadas funções

$$f : \mathbb{R}^n \times [0, T] \to \mathbb{R}^n$$

e

$$g : \mathbb{R}^n \times [0, T] \to \mathbb{R}^{n\times m}.$$

Neste capítulo, estaremos interessados em saber se existe um processo estocástico n-dimensional $(X_t)_{t\in[0,T]}$ satisfazendo a seguinte *equação diferencial estocástica*:

$$\begin{cases} dX_t &= f(X_t, t)dt + g(X_t, t)dB_t \\ X_0 &= Y \end{cases} . \tag{4.1}$$

Uma interpretação não rigorosa, mas bastante instrutiva, da expressão (4.1), digamos no caso $n = m = 1$ para fixar as ideias, é que, num determinado instante $t \in [0, T)$, para um pequeno acréscimo $\delta > 0$ no tempo, o processo $(X_t)_{t\in[0,T]}$ muda seu valor aproximadamente por uma quantidade aleatória que é normalmente distribuída, com média $f(X_t, t)\delta$ e variância

$g^2(X_t, t)\delta$, e que é independente do comportamento passado do processo. Isto porque os incrementos do movimento browniano são independentes e normalmente distribuídos, com média zero e variância igual ao correspondente incremento temporal. Neste sentido, na equação (4.1), o termo $g(X_t, t)dB_t$ é usado para modelar a ação de alguma perturbação aleatória (ruído) que afeta o sistema determinístico $dX_t = f(X_t, t)dt$. A associação do ruído à distribuição normal é fundamentada no Teorema Central do Limite, pois em muitas situações o ruído pode ser interpretado como resultante da soma de várias pequenas perturbações independentes.

No contexto do capítulo anterior, a equação (4.1), da forma como está escrita, é apenas uma abreviação para a equação integral correspondente. Enfatizamos isto na definição a seguir.

Definição 4.1.1. *Dizemos que um processo estocástico* $(X_t)_{t\in[0,T]}$ *tomando valores em* $\mathbb{R}^n$ *é uma solução da equação diferencial estocástica* (4.1) *se:*

1. $(X_t)_{t\in[0,T]}$ *é adaptado à filtração* $(\mathcal{F}_t)_{t\geq 0}$*;*
2. $(f(X_t, t))_{t\in[0,T]} \in \mathbb{M}^1_n(0, T)$*;*
3. $(g(X_t, t))_{t\in[0,T]} \in \mathbb{M}^2_{n\times m}(0, T)$*;*
4. $X_t = Y + \int_0^t f(X_s, s)ds + \int_0^t g(X_s, s)dB_s$ *q.c. para todo* $0 \leq t \leq T$.

A função f é chamada de *drift* e a função matricial gg^*, onde g^* é matriz transposta de g, é chamada de *coeficiente de difusão*. A condição *1* capta a ideia de que o processo $(X_t)_{t\in[0,T]}$ pode ser construído a partir do processo $(B_t)_{t\in[0,T]}$ e da condição inicial Y, através das funções f e g, de tal maneira que, para um determinado instante t, o valor de X_t dependa apenas do valor de Y e dos valores de B_s para $s \leq t$, o que é coerente com o ponto de vista dinâmico do processo.

A definição acima se refere a *solução forte*. Note que é dado um espaço $(\Omega, \mathcal{F}, \mathbb{P})$ onde está definido o movimento browniano $(B_t)_{t\geq 0}$, o qual gera uma filtração, e a solução $(X_t)_{t\in[0,T]}$ é construída a partir disto. Existe também a noção de *solução fraca*, neste caso são dadas apenas as funções f e g, e a solução consiste num espaço de probabilidade com uma filtração satisfazendo as condições usuais, onde está definido um movimento browniano $(B_t)_{t\geq 0}$, e um processo $(X_t)_{t\in[0,T]}$ adaptado a esta filtração, satisfazendo itens semelhantes aos da definição acima. Neste texto nos restringimos a soluções fortes.

Mais detalhes sobre soluções fracas podem ser vistos na Seção 5.3 de Karatzas & Shreve [8], por exemplo.

Neste texto não serão discutidos métodos para resolver equações diferenciais estocásticas, mas na seção seguinte resolveremos alguns exemplos simples, com uso da Fórmula de Itô. Na Seção 4.3 provaremos um Teorema de Existência e Unicidade de solução sob certas hipóteses acerca das funções f e g.

4.2 Alguns exemplos

Exemplo 4.2.1. Consideremos a seguinte equação diferencial ordinária (determinística) unidimensional:

$$\begin{cases} \frac{dX_t}{dt} &= a(t)X_t \\ X_0 &= x_0 \end{cases} . \tag{4.2}$$

A equação (4.2) serve, por exemplo, como um modelo simples para o crescimento populacional. Neste caso, X_t representa o tamanho da população, e o termo $a(t)$ é interpretado como a taxa relativa de crescimento no instante t. Vejamos o que acontece quando adicionamos algum tipo de ruído aleatório à função $a(\cdot)$. Suponhamos

$$a(\cdot) = r(\cdot) + \text{“ruído”}$$

onde $r(\cdot)$ é uma função conhecida. Um possível modelo para tal situação é dado pela seguinte equação diferencial estocástica:

$$\begin{cases} dX_t &= r(t)X_t dt + \sigma X_t dB_t \\ X_0 &= x_0 \end{cases} , \tag{4.3}$$

onde σ é uma constante. Simplificando um pouco a situação, suponhamos que a função $r(\cdot)$ seja constante, digamos $r(t) = \mu$ para todo $t \geq 0$. Em suma, procuremos uma solução para a seguinte equação diferencial estocástica:

$$\begin{cases} dX_t &= \mu X_t dt + \sigma X_t dB_t \\ X_0 &= x_0 \end{cases} , \tag{4.4}$$

onde μ, σ e x_0 são constantes quaisquer. Alertamos que, para encontrar uma candidata a solução de (4.4), faremos alguns cálculos pouco rigorosos. Dividindo a equação (4.4) por X_t, obtemos

$$\frac{1}{X_t} dX_t = \mu dt + \sigma dB_t. \tag{4.5}$$

Pela Fórmula de Itô (3.23), aplicada à função $h(x,t) = \log(x)$, temos que:

$$\begin{aligned} d(\log(X_t)) &= \frac{\partial h}{\partial t}(X_t,t)dt + \frac{\partial h}{\partial x}(X_t,t)dX_t + \frac{1}{2}\frac{\partial^2 h}{\partial x^2}(X_t,t)\sigma^2 X_t^2 dt \\ &= 0 + \frac{1}{X_t}dX_t + \frac{1}{2}\left(-\frac{1}{X_t^2}\right)\sigma^2 X_t^2 dt \\ &= \left(\mu - \frac{\sigma^2}{2}\right)dt + \sigma dB_t. \end{aligned}$$

Ou seja, na notação integral,

$$\begin{aligned} \log(X_t) &= \log(X_0) + \int_0^t \left(\mu - \frac{\sigma^2}{2}\right)ds + \int_0^t \sigma dB_s \\ &= \log(x_0) + \left(\mu - \frac{\sigma^2}{2}\right)t + \sigma B_t \end{aligned}$$

donde

$$X_t = x_0 e^{\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma B_t}. \tag{4.6}$$

Processos deste tipo são chamados de *movimentos brownianos geométricos* e são usados em finanças como modelos para a evolução de preços de ações. Nesse contexto, o coeficiente μ é chamado de *drift* e o coeficiente σ é chamado de *volatilidade* da ação.

Calculemos $\mathbb{E}(X_t)$. De (4.4) temos que

$$X_t = X_0 + \int_0^t \mu X_s ds + \int_0^t \sigma X_s dB_s.$$

Logo, por Fubini e por (3.16),

$$\begin{aligned} \mathbb{E}(X_t) &= \mathbb{E}(X_0) + \mathbb{E}\left(\int_0^t \mu X_s ds\right) + \mathbb{E}\left(\int_0^t \sigma X_s dB_s\right) \\ &= \mathbb{E}(X_0) + \int_0^t \mu \mathbb{E}(X_s) ds + 0. \end{aligned}$$

Portanto, $\varphi(t) = \mathbb{E}(X_t)$ satisfaz a EDO

$$d\varphi(t) = \mu\varphi(t)dt$$

que corresponde a equação (4.4) sem o termo de ruído ($\sigma = 0$). Logo,

$$\mathbb{E}(X_t) = x_0 e^{\mu t}.$$

Para encontrar a solução da equação (4.4), em determinado momento consideramos $\log(X_t)$. Isto pode causar certo desconforto, pois a princípio X_t poderia assumir valores negativos. Este problema é evitado, simplesmente definido-se o processo $(X_t)_{t\geq 0}$ pela igualdade dada em (4.6) e verificando-se, através da Fórmula de Itô, que X_t satisfaz (4.4). De fato, no exemplo seguinte, apresentamos e verificamos a solução para um caso um pouco mais geral, que tem (4.4) como caso particular.

Exemplo 4.2.2. *Suponhamos* $n = m = 1$*. Sejam dadas* $a : [0, T] \to \mathbb{R}$ *e* $b : [0, T] \to \mathbb{R}$ *funções contínuas, e* x_0 *uma constante. A equação*

$$\begin{cases} dX_t &= a(t)X_t dt + b(t)X_t dB_t \\ X_0 &= x_0 \end{cases} \tag{4.7}$$

tem sua única solução dada por

$$X_t = x_0 e^{\int_0^t a(s) - \frac{1}{2}b^2(s)ds + \int_0^t b(s)dB_s},$$

para $0 \leq t \leq T$. Para verificar isto, definimos o processo $(Y_t)_{t\geq 0}$ por

$$Y_t = \int_0^t a(s) - \frac{1}{2}b^2(s)ds + \int_0^t b(s)dB_s.$$

Notemos que vale

$$dY_t = \left(a(t) - \frac{1}{2}b^2(t)\right)dt + b(t)dB_t.$$

Agora, aplicando a Fórmula de Itô para $h(x, t) = x_0 e^x$, obtemos

$$\begin{aligned} d(X_t) &= \frac{\partial h}{\partial t}(Y_t, t)dt + \frac{\partial h}{\partial x}(Y_t, t)dY_t + \frac{1}{2}\frac{\partial^2 h}{\partial x^2}(Y_t, t)b^2(t)dt \\ &= 0 + x_0 e^{Y_t} dY_t + \frac{1}{2}x_0 e^{Y_t} b^2(t)dt \\ &= x_0 e^{Y_t}\left[\left(a(t) - \frac{1}{2}b^2(t)\right)dt + b(t)dB_t + \frac{1}{2}b^2(t)dt\right] \\ &= X_t\left[a(t)dt + b(t)dB_t\right] = a(t)X_t dt + b(t)X_t dB_t. \end{aligned}$$

Para a unicidade basta verificar que estão satisfeitas as hipóteses do Teorema 4.3.3 (seção seguinte).

Exemplo 4.2.3 (Equação de Langevin)**.**

$$\left\{\begin{array}{rcl} dX_t &=& \alpha X_t dt + \sigma dB_t \\ X_0 &=& x_0 \end{array}\right. . \tag{4.8}$$

Solução. Consideremos o processo definido por $Y_t = e^{-\alpha t} X_t$. Note que $Y_0 = x_0$. Pela Fórmula de Itô, aplicada a $h(x,t) = e^{-\alpha t} x$, temos que

$$\begin{aligned} dY_t &= \frac{\partial h}{\partial t}(X_t,t)dt + \frac{\partial h}{\partial x}(X_t,t)dX_t + \frac{1}{2}\frac{\partial^2 h}{\partial x^2}(X_t,t)\sigma^2 dt \\ &= -\alpha e^{-\alpha t} X_t dt + e^{-\alpha t} dX_t + 0 \\ &= \sigma e^{-\alpha t} dB_t. \end{aligned}$$

Ou seja,

$$e^{-\alpha t} X_t = Y_t = Y_0 + \sigma \int_0^t e^{-\alpha s} dB_s,$$

donde segue que

$$X_t = e^{\alpha t} x_0 + \sigma e^{\alpha t} \int_0^t e^{-\alpha s} dB_s.$$

□

4.3 Teorema de Existência e Unicidade

Comecemos com dois lema simples.

Lema 4.3.1. *Seja $\phi : [0,T] \to \mathbb{R}$ uma função contínua não negativa tal que $\phi(t) \le C \int_0^t \phi(s)ds$ para todo $0 \le t \le T$, onde C é uma constante qualquer. Então $\phi \equiv 0$.*

Demonstração. Seja $\psi : [0,T] \to \mathbb{R}$ definida por

$$\psi(t) = e^{-Ct} \int_0^t \phi(s)ds.$$

Então $\psi'(t) = e^{-Ct}\Big(\phi(t) - C\int_0^t \phi(s)ds\Big) \le 0$ para todo $0 \le t \le T$. Logo $\psi(t) \le \psi(0)$, ou seja, $e^{-Ct}\int_0^t \phi(s)ds \le e^{-0}\int_0^0 \phi(s)ds = 0$, se $0 \le t \le T$. Portanto $\int_0^t \phi(s)ds = 0$ para qualquer $0 \le t \le T$. Como ϕ é não negativa e $\phi(t) \le C\int_0^t \phi(s)ds = 0$ para $0 \le t \le T$, devemos ter $\phi \equiv 0$, conforme enunciado. □

Lema 4.3.2. *Se u e v são vetores em $\mathbb{R}^n$ então $\|u+v\|^2 \leq 2\|u\|^2 + 2\|v\|^2$.*

Demonstração. Note que, por Cauchy-Schwarz em $\mathbb{R}^2$, para quaisquer números reais a e b vale $(a+b)^2 \leq 2a^2+2b^2$ (basta considerar os vetores (a,b) e $(1,1)$). Portanto, pela desigualdade triangular,

$$(\|u+v\|)^2 \leq (\|u\| + \|v\|)^2 \leq 2\|u\|^2 + 2\|v\|^2.$$

□

Teorema 4.3.3 (Existência e Unicidade)**.** *Sejam $f : \mathbb{R}^n \times [0,T] \to \mathbb{R}^n$ e $g : \mathbb{R}^n \times [0,T] \to \mathbb{R}^{n\times m}$ funções mensuráveis satisfazendo às seguintes condições:*

(i) *$\|f(x,t) - f(y,t)\| + \|g(x,t) - g(y,t)\| \leq K\|x-y\|$ para todo $0 \leq t \leq T$ e $x, y \in \mathbb{R}^n$ (condição de Lipschitz);*

(ii) *$\|f(x,t)\|^2 + \|g(x,t)\|^2 \leq K^2 (1 + \|x\|^2)$ para todo $0 \leq t \leq T$ e $x \in \mathbb{R}^n$,*

onde $\|\cdot\|$ denota a correspondente norma euclidiana em cada espaço considerado e K é uma constante. Seja Y uma variável aleatória n-dimensional, independente do movimento browniano m-dimensional $(B_t)_{t\in[0,T]}$ dado, tal que

$$\mathbb{E}(\|Y\|^2) < \infty. \tag{4.9}$$

Então a equação diferencial estocástica

$$\begin{cases} dX_t &= f(X_t,t)dt + g(X_t,t)dB_t \\ X_0 &= Y \end{cases} . \tag{4.10}$$

admite uma única solução $(X_t)_{t\in[0,T]} \in \mathbb{M}^2_n(0,T)$ com trajetórias contínuas quase certamente.

Obs. A referida unicidade significa que se $(X_t)_{t\in[0,T]}$ e $(\widehat{X}_t)_{t\in[0,T]} \in \mathbb{M}^2_n(0,T)$ são ambas soluções de (4.10) com trajetórias contínuas, então

$$\mathbb{P}[X_t = \widehat{X}_t \text{ para todo } 0 \leq t \leq T] = 1.$$

Neste caso, os processos $(X_t)_{t\in[0,T]}$ e $(\widehat{X}_t)_{t\in[0,T]}$ são ditos *indistinguíveis*.

Demonstração.
1. *Unicidade.* Suponhamos que $(X_t)_{t\in[0,T]}$ e $(\widehat{X}_t)_{t\in[0,T]} \in \mathbb{M}^2_n(0,T)$ sejam soluções de (4.10) com trajetórias contínuas. Então, para todo $0 \le t \le T$,

$$X_t - \widehat{X}_t = \int_0^t f(X_s, s) - f(\widehat{X}_s, s)ds + \int_0^t g(X_s, s) - g(\widehat{X}_s, s)dB_s.$$

Pelo Lema 4.3.2, com

$$u = \int_0^t f(X_s, s) - f(\widehat{X}_s, s)ds \;\text{ e }\; v = \int_0^t g(X_s, s) - g(\widehat{X}_s, s)dB_s,$$

segue que

$$\begin{aligned}\mathbb{E}\left(\|X_t - \widehat{X}_t\|^2\right) \;&\le\; 2\mathbb{E}\left(\left\|\int_0^t f(X_s, s) - f(\widehat{X}_s, s)ds\right\|^2\right) \\ &+2\mathbb{E}\left(\left\|\int_0^t g(X_s, s) - g(\widehat{X}_s, s)dB_s\right\|^2\right).\end{aligned} \tag{4.11}$$

Para toda função $h : [0,t] \to \mathbb{R}^n$ vale $\left\|\int_0^t h(s)ds\right\|^2 \le t\int_0^t \|h(s)\|^2 ds$. De fato,

$$\begin{aligned}\left\|\int_0^t h(s)ds\right\|^2 &= \left\|\left(\int_0^t h_1(s)ds, \ldots, \int_0^t h_n(s)ds\right)\right\|^2 \\ &= \left(\int_0^t h_1(s)ds\right)^2 + \ldots + \left(\int_0^t h_n(s)ds\right)^2 \\ \text{Cauchy-Schwarz} \quad &\le t\int_0^t h_1^2(s)ds + \ldots + t\int_0^t h_n^2(s)ds = t\int_0^t \|h(s)\|^2 ds,\end{aligned}$$

onde usamos que $\int_0^t h_i(s)ds$ é o produto interno de h_i com a função constante igual a 1, em $L^2[0,t]$.

Usando este fato com $h(t) = f(X_t, t) - f(\widehat{X}_t, t)$, estimamos

$$\begin{aligned}\mathbb{E}\left(\left\|\int_0^t f(X_s, s) - f(\widehat{X}_s, s)ds\right\|^2\right) &\le T\mathbb{E}\left(\int_0^t \left\|f(X_s, s) - f(\widehat{X}_s, s)ds\right\|^2 ds\right) \\ \text{por (i) e Fubini} \qquad &\le TK^2\int_0^t \mathbb{E}\left(\|X_s - \widehat{X}_s\|^2\right) ds.\end{aligned} \tag{4.12}$$

Além disso, pelo Isometria de Itô (3.38),

$$\begin{aligned}
\mathbb{E}\left(\left\|\int_0^t g(X_s,s)-g(\widehat{X}_s,s)dB_s\right\|^2\right) &= \mathbb{E}\left(\int_0^t \left\|g(X_s,s)-g(\widehat{X}_s,s)\right\|^2 ds\right)\\
&\leq K^2\int_0^t \mathbb{E}\left(\|X_s-\widehat{X}_s\|^2\right)ds. \quad (4.13)
\end{aligned}$$

De (4.11), (4.12) e (4.13) concluímos que, para $C = 2K^2(T+1)$, vale

$$\mathbb{E}\left(\|X_t-\widehat{X}_t\|^2\right) \leq C\int_0^t \mathbb{E}\left(\|X_s-\widehat{X}_s\|^2\right)ds \qquad (4.14)$$

para $0 \leq t \leq T$. Definido $\phi(t) = \mathbb{E}\left(\|X_t-\widehat{X}_t\|^2\right)$, a equação (4.14) diz que $\phi(t) \leq C\int_0^t \phi(s)ds$ para todo $0 \leq t \leq T$. Logo, de acordo com o Lema 4.3.1, temos que $\phi \equiv 0$, ou seja $\mathbb{E}\left(\|X_t-\widehat{X}_t\|^2\right) = 0$ para todo $0 \leq t \leq T$. Disto segue que $X_t = \widehat{X}_t$ quase certamente, para todo $0 \leq t \leq T$. Tomando interseção enumerável de conjuntos com probabilidade um, obtemos

$$\mathbb{P}\left(\|X_t-\widehat{X}_t\| = 0 \text{ para todo } t \in [0,T]\cap\mathbb{Q}\right) = 1$$

e então, pela continuidade quase certa das trajetórias,

$$\mathbb{P}\left(\|X_t-\widehat{X}_t\| = 0 \text{ para todo } t \in [0,T]\right) = 1,$$

e está provada a unicidade a menos de indistinguibilidade.

2. *Existência.* A demonstração usa o *método das aproximações sucessivas* e é semelhante à prova do correspondente resultado para equações diferenciais ordinárias.

Consideremos a sequência $\{(X_t^{(k)})_{t\in[0,T]}\}_{k=0}^{\infty}$ de processos estocásticos assim definida:

$$\left\{\begin{aligned}
X_t^{(0)} &\equiv Y\\
X_t^{(k+1)} &= Y + \int_0^t f(X_s^{(k)},s)ds + \int_0^t g(X_s^{(k)},s)dB_s,
\end{aligned}\right. \qquad (4.15)$$

para $k = 0, 1, \ldots$ e $0 \leq t \leq T$.

O objetivo é mostrar que, quando $k \to \infty$, os processos $(X_t^{(k)})_{t\in[0,T]}$ convergem (uniformemente em t, q.c., ou seja, para quase todo $\omega \in \Omega$ fixado a convergência é uniforme em t) a um processo $(X_t)_{t\in[0,T]} \in \mathbb{M}_n^2$ com trajetórias contínuas que satisfaz a equação (4.10).

Afirmação 1. *Para $k = 1, 2, \ldots$, o processo $(X_t^{(k)})_{t\in[0,T]}$ está bem definido, pertence a $\mathbb{M}_n^2(0,T)$ e tem trajetórias contínuas quase certamente. Além disso,*

$$\mathbb{E}\left(\|X_t^{(k)} - X_t^{(k-1)}\|^2\right) \leq \frac{(Mt)^k}{k!} \tag{4.16}$$

para todo $t \in [0,T]$, onde M é uma constante que depende apenas de K, T e $\mathbb{E}(\|Y\|^2)$.

Prova. Por indução. Como $(X_t^{(0)})_{t\in[0,T]} \equiv Y \in \mathbb{M}_n^2$, a mensurabilidade das funções f e g e a hipótese (ii) garantem que os processos $(f(Y,t))_{t\in[0,T]}$ e $(g(Y,t))_{t\in[0,T]}$ pertencem a $\mathbb{M}_n^2$ e $\mathbb{M}_{n\times m}^2$ respectivamente e, assim, $(X_t^{(1)})_{t\in[0,T]}$ está bem definido. Além disso, pelo Lema 4.3.2,

$$\begin{aligned}\mathbb{E}\left(\|X_t^{(1)} - X_t^{(0)}\|^2\right) &= \mathbb{E}\left(\left\|\int_0^t f(Y,s)ds + \int_0^t g(Y,s)dB_s\right\|^2\right)\\ &\leq 2\mathbb{E}\left(\left\|\int_0^t f(Y,s)ds\right\|^2\right) + 2\mathbb{E}\left(\left\|\int_0^t g(Y,s)dB_s\right\|^2\right),\end{aligned}$$

e então, pela desigualdade de Cauchy-Schwarz, pela Isometria de Itô (3.38) e pela hipótese (ii) temos que

$$\begin{aligned}\mathbb{E}\left(\|X_t^{(1)} - X_t^{(0)}\|^2\right) &\leq 2T\mathbb{E}\left(\int_0^t K^2(1+\|Y\|^2)ds\right) + 2\mathbb{E}\left(\int_0^t K^2(1+\|Y\|^2)ds\right)\\ &= 2TtK^2(1+\mathbb{E}(\|Y\|^2)) + 2tK^2(1+\mathbb{E}(\|Y\|^2)) \leq Mt\end{aligned}$$

desde que $M \geq 2K^2(T+1)(1+\mathbb{E}(\|Y\|^2))$. Isto implica que

$$\begin{aligned}\mathbb{E}\left(\int_0^T \|X_t^{(1)}\|^2 dt\right) &= \int_0^T \mathbb{E}\left(\|X_t^{(1)}\|^2\right) dt\\ &\overset{\text{Lema 4.3.2}}{\leq} 2\int_0^T \mathbb{E}\left(\|X_t^{(1)} - X_t^{(0)}\|^2\right) dt + 2\int_0^T \mathbb{E}\left(\|X_t^{(0)}\|^2\right) dt\\ &\leq MT^2 + 2T\mathbb{E}\left(\|Y\|^2\right) < \infty\end{aligned}$$

e então $(X_t^{(1)})_{t\in[0,T]} \in \mathbb{M}_n^2$. A continuidade das trajetórias é garantida pela hipótese (ii) e pelo Teorema 3.4.3. Logo, vale o resultado para $k = 1$.

Agora assumindo que vale a afirmação até k, provaremos a validade para $k+1$. Como $(X_t^{(k)})_{t\in[0,T]} \in \mathbb{M}_n^2$, novamente a mensurabilidade de f e g e a hipótese (ii) garantem que $(X_t^{(k+1)})_{t\in[0,T]}$ está bem definido. Agora, estimativas semelhantes às feitas na prova da unicidade mostram que

$$\begin{aligned}
\mathbb{E}\left(\|X_t^{(k+1)} - X_t^{(k)}\|^2\right) &= \mathbb{E}\left(\left\|\int_0^t f(X_s^{(k)}, s) - f(X_s^{(k-1)}, s)ds\right.\right. \\
&\quad \left.\left. + \int_0^t g(X_s^{(k)}, s) - g(X_s^{(k-1)}, s)dB_s\right\|^2\right) \\
&\leq 2TK^2 \int_0^t \mathbb{E}\left(\|X_s^{(k)} - X_s^{(k-1)}\|^2\right) ds \\
&\quad +2K^2 \int_0^t \mathbb{E}\left(\|X_s^{(k)} - X_s^{(k-1)}\|^2\right) ds.
\end{aligned}$$

Usando a hipótese de indução, obtemos, portanto,

$$\mathbb{E}\left(\|X_t^{(k+1)} - X_t^{(k)}\|^2\right) \leq 2K^2(T+1) \int_0^t \frac{M^k s^k}{k!} ds \leq \frac{M^{k+1}t^{k+1}}{(k+1)!}$$

contanto que $M \geq 2K^2(T+1)$. Como antes, isto também implica que $(X_t^{(k+1)})_{t\in[0,T]} \in \mathbb{M}_n^2$ e, assim, a indução está completa.

Afirmação 2. *Com probabilidade um, quando $k \to \infty$, $(X_t^{(k)})_{t\in[0,T]}$ converge, uniformemente em t, a um processo $(X_t)_{t\in[0,T]} \in \mathbb{M}_n^2(0,T)$ com trajetórias contínuas que é solução da equação* (4.10).

Prova. Com estimativas análogos às já feitas anteriormente, notemos que

$$\begin{aligned}
\sup_{0\leq t\leq T} \|X_t^{(k+1)} - X_t^{(k)}\|^2 &\leq 2TK^2 \int_0^T \|X_s^{(k)} - X_s^{(k-1)}\|^2 ds \\
&\quad +2 \sup_{0\leq t\leq T} \left\|\int_0^t g(X_s^{(k)}, s) - g(X_s^{(k-1)}, s)dB_s\right\|^2.
\end{aligned}$$

Lembremos que, na expressão acima, cada componente da integral estocástica n-dimensional no segundo termo do lado esquerdo é uma soma de integrais estocásticas unidimensionais, as quais são martingales de acordo com o Teorema 3.4.2. Portanto, a desigualdade (2.39) do Teorema 2.5.19, com $p = 2$, aplicada a cada parcela implica

$$\begin{aligned}
\mathbb{E}\left(\sup_{0\leq t\leq T}\|X_t^{(k+1)} - X_t^{(k)}\|^2\right) \leq 2TK^2 \int_0^T \mathbb{E}\left(\|X_s^{(k)} - X_s^{(k-1)}\|^2\right) ds \\
+8\mathbb{E}\left(\left\|\int_0^T g(X_s^{(k)}, s) - g(X_s^{(k-1)}, s)dB_s\right\|^2\right) \\
\leq (2TK^2 + 8K^2)\int_0^T \mathbb{E}\left(\|X_s^{(k)} - X_s^{(k-1)}\|^2\right) ds.
\end{aligned}$$

Logo, por (4.16), para $C = 2T^2K^2 + 8TK^2$ vale

$$\mathbb{E}\left(\sup_{0\leq t\leq T}\|X_t^{(k+1)} - X_t^{(k)}\|^2\right) \leq C\frac{(MT)^k}{k!}.$$

Consequentemente, pela desigualdade de Tchebychev, temos que

$$\begin{aligned}
\mathbb{P}\left(\sup_{0\leq t\leq T}\|X_t^{(k+1)} - X_t^{(k)}\| > \frac{1}{2^k}\right) &\leq 4^k\mathbb{E}\left(\sup_{0\leq t\leq T}\|X_t^{(k+1)} - X_t^{(k)}\|^2\right) \\
&\leq C\frac{(4MT)^k}{k!}.
\end{aligned}$$

Como $\sum_{k=1}^{\infty}\frac{(4MT)^k}{k!} < \infty$, o Lema de Borel-Cantelli então implica que

$$\mathbb{P}\left(\sup_{0\leq t\leq T}\|X_t^{(k+1)} - X_t^{(k)}\| > \frac{1}{2^k} \text{ para infinitos valores de } k\right) = 0.$$

Assim, para quase todo $\omega \in \Omega$ existe um $k_0(\omega)$ tal que

$$\sup_{0\leq t\leq T}\|X_t^{(k+1)}(\omega) - X_t^{(k)}(\omega)\| \leq \frac{1}{2^k}$$

para $k \geq k_0(\omega)$. Portanto, com probabilidade um, a sequência

$$X_t^{(k)} = X_t^{(0)} + \sum_{i=0}^{k-1}(X_t^{(i+1)} - X_t^{(i)})$$

converge uniformemente em $t \in [0, T]$ a um processo $(X_t)_{t\in[0,T]}$. Naturalmente, o processo $(X_t)_{t\in[0,T]}$ tem trajetórias contínuas quase certamente e é adaptado à filtração $(\mathcal{F}_t)_{t\in[0,T]}$, já que $(X_t^{(k)})_{t\in[0,T]}$ tem estas propriedades para todo k. Agora, mostremos que $(X_t)_{t\in[0,T]} \in \mathbb{M}_n^2(0,T)$.

Por um resultado análogo a Lema 4.3.2 (a saber

$$\|u+v+w\|^2 \leq 3\|u\|^2 + 3\|v\|^2 + 3\|w\|^2$$

para quaisquer vetores u, v, e w em $\mathbb{R}^n$), segue que

$$\begin{aligned}
\mathbb{E}\left(\|X_t^{(k+1)}\|^2\right) &\leq 3\mathbb{E}\left(\|Y\|^2\right) + 3\mathbb{E}\left(\left\|\int_0^t f(X_s^{(k)},s)ds\right\|^2\right) \\
&\quad +3\mathbb{E}\left(\left\|\int_0^t g(X_s^{(k)},s)dB_s\right\|^2\right) \\
&\leq C\left[1+\mathbb{E}\left(\|Y\|^2\right)\right] + C\int_0^t \mathbb{E}\left(\|X_s^{(k)}\|^2\right) ds,
\end{aligned}$$

onde C é uma constante dependendo somente de K e T. Então, por indução,

$$\mathbb{E}\left(\|X_t^{(k+1)}\|^2\right) \leq \left[C + C^2 t + \ldots + C^{k+2}\frac{t^{k+1}}{(k+1)!}\right]\left[1+\mathbb{E}\left(\|Y\|^2\right)\right].$$

Consequentemente,

$$\mathbb{E}\left(\|X_t^{(k+1)}\|^2\right) \leq C\left[1+\mathbb{E}\left(\|Y\|^2\right)\right] e^{Ct}.$$

Fazendo $k\to\infty$ e usando o Lema de Fatou, concluímos que

$$\begin{aligned}
\mathbb{E}\left(\|X_t\|^2\right) &= \mathbb{E}\left(\lim_{k\to\infty}\|X_t^{(k+1)}\|^2\right) \\
&\leq \liminf_{k\to\infty}\mathbb{E}\left(\|X_t^{(k+1)}\|^2\right) \leq C\left[1+\mathbb{E}\left(\|Y\|^2\right)\right] e^{Ct}.
\end{aligned}$$

Isto implica que

$$\mathbb{E}\left(\int_0^T \|X_t\|^2 dt\right) = \int_0^T \mathbb{E}\left(\|X_t\|^2\right) dt < \infty$$

e, portanto, $(X_t)_{t\in[0,T]} \in \mathbb{M}_n^2(0,T)$.

Para concluir a demonstração, precisamos verificar que $(X_t)_{t\in[0,T]}$ é solução da equação (4.10). De (4.16) obtemos que, para $j > i \geq 0$,

$$
\begin{aligned}
\left[\mathbb{E}\left(\|X_t^{(j)} - X_t^{(i)}\|^2\right)\right]^{\frac{1}{2}} &= \|X_t^{(j)} - X_t^{(i)}\|_{L^2(\Omega)} = \left\|\sum_{k=i+1}^{j}(X_t^{(k)} - X_t^{(k-1)})\right\|_{L^2(\Omega)} \\
&\leq \sum_{k=i+1}^{j}\|X_t^{(k)} - X_t^{(k-1)}\|_{L^2(\Omega)} \\
&\leq \sum_{k=i+1}^{\infty}\left(\frac{M^k t^k}{k!}\right)^{\frac{1}{2}} \to 0 \quad \text{quando } i \to \infty. \qquad (4.17)
\end{aligned}
$$

Portanto, identificando variáveis aleatórias que são iguais quase certamente, temos que, quando $k \to \infty$, $X_t^{(k+1)}$ converge para X_t também em $L^2(\Omega)$. Como, para quase todo $\omega \in \Omega$, $X_t^{(k+1)}(\omega) \to X_t(\omega)$ quando $k \to \infty$, uniformemente em $t \in [0,T]$, pelo Lema de Fatou e por (4.17) obtemos que

$$
\mathbb{E}\left(\int_0^T \|X_t - X_t^{(i)}\|^2 dt\right) \leq \liminf_{j\to\infty} \mathbb{E}\left(\int_0^T \|X_t^{(j)} - X_t^{(i)}\|^2 dt\right) \to 0
$$

quando $i \to \infty$. Então, pela hipótese (i) e pela Isometria de Itô (3.38) temos que

$$
\mathbb{E}\left(\left\|\int_0^t g(X_s, s) - g(X_s^{(k)}, s) dB_s\right\|^2\right) \to 0
$$

quando $k \to \infty$. Usando Cauchy-Schwarz notemos também que

$$
\mathbb{E}\left(\left\|\int_0^t f(X_s, s) - f(X_s^{(k)}, s) dB_s\right\|^2\right) \to 0.
$$

Portanto, tomando-se o limite em $L^2(\Omega)$ quando $k \to \infty$ na expressão (4.15), obtemos

$$
X_t = Y + \int_0^t f(X_s, s) ds + \int_0^t g(X_s, s) dB_s.
$$

Ou seja, $(X_t)_{t\in[0,T]}$ satisfaz a equação (4.10). Isto completa a demonstração do teorema. $\square$

Referências Bibliográficas

[1] Bachman, George. Narici, Lawrence. Functional Analysis. New York: Academic Press, 1966. Dover Edition, 2000.

[2] Bartle, Robert G. The Elements of Integration and Lebesgue Measure. Wiley Classics Library ed. New York: Wiley, 1995.

[3] Billingsley, Patrick. Probability and measure. New York: Wiley-Interscience, 1995.

[4] Durrett, Richard. Probability Theory and Examples. 2nd. ed. Belmont: Duxbury, 1996.

[5] Evans, Lawrence C. An introduction to stochastic differential equations: Version 1.2. Berkeley: University of California, s.d. 139 p. Disponível em http://math.berkeley.edu/˜evans/SDE.course.pdf.

[6] Friedman, Avner. Stochastic differential equations and applications. Vol 1. New York: Academic Press, 1975.

[7] James. Barry R. Probabilidade: um curso em nível intermediário. 3 ed. Rio de Janeiro: IMPA, 2008.

[8] Karatzas, Ioannis. Shreve, Steven E. Brownian motion and stochastic calculus. 2nd ed. New York: Springer, 1996.

[9] Karlin, Samuel. Taylor, Howard M. A first course in stochastic processes. 2nd.ed. Boston: Academic Press, 1996.

[10] Lima, Elon Lages. Espaços Métricos. 4.ed. Rio de Janeiro: IMPA, 2009.

[11] Øksendal, Bernt Karsten. Stochastic differential equations: an introduction with applications. 6th ed. Berlin: Springer, 2005.

[12] Ruffino, Paulo. Uma iniciação aos sistemas dinâmicos estocásticos. 2 ed. Rio de Janeiro: IMPA, 2009.

[13] Schuss, Zeev. Theory and applications of stochastic differential equations. New York: John Wiley, 1980.

Índice Remissivo