



UNIVERSIDADE ESTADUAL PAULISTA
“JÚLIO DE MESQUITA FILHO”
Faculdade de Ciências e Tecnologia
Câmpus de Presidente Prudente

Uma introdução às equações diferenciais estocásticas e fluxos estocásticos de difeomorfismos em variedades diferenciáveis

Amanda Silvieri Leite de Oliveira

Orientador: Prof. Dr. Fabiano Borges da Silva

Programa: Matemática Aplicada e Computacional

Presidente Prudente, julho de 2021.

UNIVERSIDADE ESTADUAL PAULISTA

Faculdade de Ciências e Tecnologia de Presidente Prudente

Programa de Pós-Graduação em Matemática Aplicada e Computacional

**Uma introdução às equações diferenciais
estocásticas e fluxos estocásticos de difeomorfismos
em variedades diferenciáveis**

Amanda Silvieri Leite de Oliveira

Orientador: Prof. Dr. Fabiano Borges da Silva

Dissertação apresentada ao Programa de Pós-Graduação em Matemática Aplicada e Computacional da Faculdade de Ciências e Tecnologia da UNESP para obtenção do título de Mestre em Matemática Aplicada e Computacional.

Presidente Prudente, julho de 2021.

O48i

Oliveira, Amanda Silvieri Leite de

Uma introdução às equações diferenciais estocásticas e fluxos estocásticos de difeomorfismos em variedades diferenciáveis / Amanda Silvieri Leite de Oliveira. -- Presidente Prudente, 2021
126 p.

Dissertação (mestrado) - Universidade Estadual Paulista (Unesp), Faculdade de Ciências e Tecnologia, Presidente Prudente
Orientador: Fabiano Borges da Silva

1. Fluxos estocásticos. 2. Variedades diferenciáveis. 3. Fórmula de Itô. 4. Forma volume. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Faculdade de Ciências e Tecnologia, Presidente Prudente. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

CERTIFICADO DE APROVAÇÃO

TÍTULO DA DISSERTAÇÃO: Uma introdução às equações diferenciais estocásticas e fluxos estocásticos de difeomorfismos em variedades diferenciáveis

AUTORA: AMANDA SILVIERI LEITE DE OLIVEIRA

ORIENTADOR: FABIANO BORGES DA SILVA

Aprovada como parte das exigências para obtenção do Título de Mestra em MATEMÁTICA APLICADA E COMPUTACIONAL, pela Comissão Examinadora:

Prof. Dr. FABIANO BORGES DA SILVA (Participação Virtual)
Departamento de Matemática / Faculdade de Ciências de Bauru

VIDEOCONFERÊNCIA

Prof. Dr. DIEGO SEBASTIAN LEDESMA (Participação Virtual)
Instituto de Matemática Estatística e Ciência da Computação. / Universidade Estadual de Campinas

VIDEOCONFERÊNCIA

Prof. Dr. LEANDRO BATISTA MORGADO (Participação Virtual)
Departamento de Matemática / Universidade Federal de Santa Catarina

Presidente Prudente, 25 de junho de 2021

Às pessoas que estiveram presentes comigo ao longo desse caminho.

Agradecimentos

Agradeço primeiramente ao meu orientador e professor Fabiano, por todo conhecimento compartilhado, além de toda dedicação e paciência durante esta trajetória.

À minha mãe, por todo carinho, por sempre me apoiar e acreditar em mim.

Ao meu companheiro e amigo Alex, que acompanhou de perto a realização deste trabalho, suportando meus testes de apresentações e discussões, sempre disposto a dar uma opinião e me escutar, mesmo não entendendo.

Ao meu amigo Otávio, que sempre me ajudou e esteve presente desde a graduação. Ao meu amigo André, que sempre esteve presente arrumando os problemas do meu computador e do editor de texto. À minha amiga Lívia, pelo apoio nas disciplinas cursadas.

Aos meus professores, especialmente a Tatiana, por suas imensas contribuições e suporte ao longo desse período. E também aos professores Diego Sebastian Ledesma e Leandro Batista Morgado pelas sugestões e contribuições dadas a este trabalho.

Aos funcionários da Seção de Pós-Graduação do PosMac - Unesp/Presidente Prudente, especialmente a Cinthia, pela gentileza e atenção.

Agradeço também à FAPESP, Fundação de Amparo à Pesquisa do Estado de São Paulo, pela bolsa de Mestrado 2019/08076-2, que me proporcionou condições de realizar esse estudo.

E por fim, agradeço à todos os envolvidos nessa trajetória, pelos bons momentos, pela amizade e pelas conversas.

*Não sei como o mundo me vê, mas eu me sinto
como um garoto brincando na praia, contente em
achar aqui e ali, uma pedra mais lisa ou uma
concha mais bonita, mas tendo sempre diante de mim,
ainda por descobrir, “o grande oceano de verdades”.*

Isaac Newton

Lista de Figuras

1.1	Possível trajetória do movimento	22
1.2	Gráfico da função $f(x)$	31
5.1	Ilustração do campo de vetores X_1 por dois ângulos.	86
5.2	Ilustração do campo X_1 na esfera.	95
5.3	Ilustração do campo Y na esfera.	97
C.1	Problema nas derivadas direcionais de campo vetoriais, adaptado de Lee (2012).	113
C.2	Derivada direcional de um campo vetorial, adaptado de Lee (2012)	113

Resumo

Neste trabalho inicialmente apresentamos uma breve revisão sobre conteúdos de cálculo estocástico que permitam introduzir as equações diferenciais estocásticas e algumas de suas aplicações e motivações. Como objetivo principal deste trabalho estudamos as equações diferenciais estocásticas em variedades diferenciáveis e a Fórmula de Itô para ação de fluxos estocásticos em campos de vetores e em formas diferenciáveis. Por fim, estudamos sob quais condições um fluxo estocástico preserva uma forma volume seguindo a abordagem apresentada em Kunita (1982) [11].

Palavras-Chave: *Fluxos estocásticos. Variedades diferenciáveis. Fórmula de Itô. Forma volume.*

Abstract

In this work, we initially present a brief review of the contents of stochastic calculus that allow the introduction of stochastic differential equations and some of their applications and motivations. As the main objective of this work, we study the stochastic differential equations on smooth manifolds and the Itô Formula for action of stochastic flows on vector fields and differential forms. Finally, we study under which conditions a stochastic flow preserves a volume form following the approach presented in Kunita (1982) [11].

Keywords: *Stochastic flows. Smooth manifolds. Itô's formula. Volume form.*

Sumário

Lista de Figuras	7
Resumo	9
Abstract	11
Introdução	15
1 Uma introdução à análise estocástica	17
1.1 Uma introdução à Probabilidade e Processos Estocásticos	17
1.2 Movimento Browniano	21
1.2.1 Variação quadrática	23
1.2.2 Movimento browniano no \mathbb{R}^d	25
1.3 Martingale	26
1.3.1 Martingale local e semimartingale	29
1.4 Integral de Wiener	30
2 Integral estocástica, Fórmula de Itô e uma introdução às equações diferenciais estocásticas	35
2.1 Integral de Itô	36
2.2 Fórmula de Itô	41
2.2.1 Processos de Itô	43
2.2.2 Fórmula de Itô para processos de Itô	44
2.2.3 Fórmula de Itô multidimensional	46
2.3 Equações diferenciais estocásticas	47
2.3.1 Existência e unicidade de soluções	48
2.4 Integral de Stratonovich	55
3 Algumas aplicações para as EDEs	61
3.1 Crescimento populacional	61
3.2 Valor de uma ação	63
3.3 Equação de Langevin	64
4 Fluxos estocásticos	67
4.1 Diferenciabilidade com relação a variável espacial	68
4.2 $\xi_{s,t}$ é um difeomorfismo para todo $s < t$ (ω -quase sempre)	70
4.3 Fluxos com coeficientes localmente Lipschitz contínuos	73
4.4 Fórmulas de Itô generalizadas para composição de processos estocásticos	74

5	Ações de fluxos estocásticos em campos de vetores e formas diferenciais	77
5.1	Fluxos agindo em campo de vetores	79
5.1.1	Composição e decomposição de fluxos estocásticos	87
5.2	Fluxos agindo em formas diferenciais	89
5.2.1	Exemplos na esfera	94
5.3	Considerações finais	99
	Referências	101
A	Campos aleatórios que são Hölder contínuos	105
B	Ponto de compactificação	107
C	Campos de vetores	109
C.1	Diferencial de um campo	109
C.2	Colchete de Lie	111
C.3	Derivada de Lie	113
D	Formas diferenciais	115
D.1	Aplicações r -lineares	115
D.2	Formas alternadas	115
D.3	Determinantes	116
D.4	Produto exterior	117
D.5	Superfície orientada	118
D.6	Formas diferenciais	119
D.7	Operações para formas diferenciais de grau r	122
D.8	Derivada de Lie para k -formas	124
D.9	Integral de uma forma diferencial	124

Introdução

Em alguns modelos matemáticos que descrevem a evolução de um fenômeno na natureza com relação ao tempo, aparecem incertezas decorrentes de parâmetros desconhecidos ou ruídos propostos pelo ambiente. Alguns problemas de física, por exemplo, consistem no estudo de objetos que deslocam-se sobre uma determinada superfície, ou ainda num sentido mais amplo, numa variedade. Quando a evolução é em tempo contínuo, uma modelagem possível é obtida através de equações diferenciais estocásticas (EDEs), as quais podem ser tomadas no sentido das integrais de Itô ou das integrais de Stratonovich. As integrais de Itô (denotadas por $\int_0^t f(s)dB_s$) em geral são usadas quando se tem interesse nas propriedades probabilísticas, enquanto que as de Stratonovich (denotadas por $\int_0^t f(s) \circ dB_s$) são úteis para equações diferenciais estocásticas em variedades, pois as regras usuais com relação a derivação de composição de funções seguem as regras do cálculo diferencial usual. Dessa forma, as EDEs geralmente usadas em variedades diferenciáveis são da forma

$$d\xi_t = X_0(\xi_t)dt + X_1(\xi_t) \circ dB_t, \quad (0.0.0.1)$$

onde X_0 e X_1 são campos de vetores diferenciáveis em M e B_t é o movimento browniano. Usaremos M para denotar variedade diferenciável ao longo do trabalho.

Se $x \in M$ é uma condição inicial, então, a grosso modo, $\xi_t(x)$ representa uma possível trajetória aleatória em M , e sob certas condições nos campos de vetores da EDE, $\xi_t : M \rightarrow M$ é um fluxo de difeomorfismos (ω -quase sempre), e portanto, estudando a dinâmica da diferencial do fluxo $\xi_{t*} : T_x M \rightarrow T_{\xi_t(x)} M$, que é o interesse deste trabalho, é possível saber se o fluxo estocástico preserva uma forma volume na variedade diferenciável M . Um aspecto importante, é que numa variedade compacta como a esfera por exemplo, podemos garantir a recorrência do fluxo. Isto significa que um fluxo estocástico ξ_t agindo numa região compacta U da esfera, para um tempo t grande, o conjunto $\xi_t(U)$ vai interceptar o conjunto U para quase todo ω . Quem garante isto é o Teorema da Recorrência de Poincaré, que afirma que certos sistemas, após um tempo suficientemente longo, finito, retornarão para um estado muito próximo ao estado inicial, mais detalhes sobre isso podem ser vistos em Doering; Lopes (2008, p. 272, Teorema 6.26) [4].

Dessa forma, este trabalho consiste no estudo de equações diferenciais estocásticas em variedades diferenciáveis e na Fórmula de Itô para ação de fluxos estocásticos em campos de vetores e tensores alternados, também chamados de formas diferenciais, seguindo a abordagem apresentada em Kunita (1982) [11]. Com isso, apresentamos um resultado que garante sob quais condições nos campos de vetores, um fluxo estocástico de difeomorfismo $\xi_t : M \rightarrow M$ (ω -quase sempre) associado a uma EDE da forma (0.0.0.1), preserva uma forma volume sobre uma variedade diferenciável M .

No primeiro capítulo, a fim de fixar notação e fornecer um texto para o leitor não familiarizado com Cálculo Estocástico, fizemos uma revisão geral com relação aos conceitos básicos da Teoria de Probabilidade e processos estocásticos, entre estes destacamos o estudo do movimento browniano n -dimensional, dos martingales e a integral de Wiener para integrandos determinísticos.

No Capítulo 2 apresentamos a integral de Itô e Stratonovich e as equações diferenciais estocásticas. Além disso, nesse capítulo é apresentado a forma de transformação de uma integral

de Stratonovich em Itô e vice e versa, conceito usado em um resultado importante no final do trabalho. As aplicações do cálculo estocástico é o foco do Capítulo 3, e para tanto apresentamos alguns exemplos que ilustram sua relevância na Biologia, na Física e em Finanças.

Finalmente, no Capítulo 4 provamos que uma equação diferencial estocástica gera um fluxo de homeomorfismos ou difeomorfismos sob certas condições. Um dos problemas que estaremos preocupados é se $\xi_t : \mathbb{R}^d \rightarrow \mathbb{R}^d$ associado a uma EDE induz um fluxo estocástico de homeomorfismos de \mathbb{R}^d ou não. A resposta é sim, mas a verificação não é de modo algum simples, comparada com o caso determinístico. Também estudamos algumas estimativas para $|\xi_t(x) - \xi_{t'}(x')|$ fazendo o uso do cálculo estocástico e com o objetivo de usar o critério de Kolmogorov para mostrar a propriedade homeomórfica de $\xi_t(x)$.

Por fim, estudamos fluxos estocásticos em variedades analisando sob quais condições nos campos de vetores uma EDE da forma (0.0.0.1) gera um fluxo estocástico de difeomorfismo sobre uma variedade diferenciável e chegamos no objetivo principal do trabalho: Fórmula de Itô para estes fluxos agindo em campos de vetores e tensores alternados sobre uma variedade diferenciável. Na seção 5.1 obtemos a Fórmula de Itô para ξ_t atuando sobre campos vetoriais, onde as derivadas de Lie desempenham um papel importante. Como aplicação, decompomos a equação em duas mais simples e obtemos a solução como a composição das soluções correspondentes a essas duas equações. A Fórmula de Itô para ξ_t atuando em campos de tensores alternados é obtida na seção 5.2.

A fórmula pode ser usada em alguns casos para determinar se o fluxo ξ_t preserva a forma diferencial K , ou seja, $\xi_t^*K = K$ (ω -quase sempre) se, e somente se, $L_X K = 0$. Assim podemos saber sob quais condições um fluxo estocástico preserva uma forma volume. Como exemplo, estudamos o comportamento de um fluxo agindo sobre determinados campos em uma superfície esférica. Assim podemos saber, por exemplo, se um fluxo estocástico preserva volume (ou área no caso da esfera em \mathbb{R}^3) através das propriedades dos campos de vetores da EDE que o gera.

Uma introdução à análise estocástica

Dedicamos este capítulo a introdução de conceitos e uma revisão dos principais tópicos de análise estocástica que serão úteis ao longo do texto e que podem ser encontrados com mais detalhes, por exemplo, em Allen (2010) [1], Brzezniak; Zastawniak (1999) [2], Evans (2012) [5], Klebaner (2005) [10] e Oksendal (1998) [19]. É importante ressaltar que este capítulo não pretende ser uma revisão abrangente da teoria da probabilidade, mas apenas uma breve revisão de alguns conceitos da teoria da probabilidade que são importantes para a teoria dos processos estocásticos e se fazem úteis no desenvolvimento deste trabalho.

1.1 Uma introdução à Probabilidade e Processos Estocásticos

Seja Ω um conjunto qualquer (supostamente não vazio) de elementos, que será referido como espaço amostral (espaço de resultados, universo ou conjunto universal). Ele representa o conjunto de todos os possíveis resultados de uma situação aleatória, como por exemplo o lançamento de um dado ou uma moeda. Consideremos o caso de um lançamento de dois dados, um branco e um preto. Então, $\Omega = \{(1, 1), (1, 2), \dots, (2, 1), (2, 2), \dots, (6, 6)\}$ é um conjunto formado por 36 elementos, onde cada um representa um possível resultado de lançamento. Dessa forma, o elemento $(5, 1)$ representa o resultado de aparecer no dado branco a face com cinco bolinhas e uma bolinha no dado preto.

O espaço amostral pode ser representado de várias formas. Por exemplo, suponha agora que estamos observando o lançamento de uma moeda; há duas possibilidades de registro quando ela cai: cara ou coroa. Então o espaço amostral é $\{CA, CO\}$, onde CA representa a possibilidade de aparecer cara e CO representa a possibilidade de aparecer coroa. Dessa forma, podemos ter vários tipos de espaço amostral, como por exemplo:

$$\Omega = \{CARA, COROA\}, \quad \Omega = \{0, 1, 2, \dots\} \quad \Omega = \{\omega : \omega \in [0, \infty)\}.$$

Ω ainda pode representar qualquer coleção de objetos ou elementos.

Cada elemento de Ω é chamado de ponto amostral e cada subconjunto de Ω é chamado de evento. Por exemplo, no caso do lançamento dos dados $(2, 4)$ é um ponto de amostra e $\{(2, 3), (4, 5)\}$ é um possível evento, como o espaço amostral é composto por 36 elementos, existem inúmeros eventos possíveis, já que os eventos são cada subconjunto do espaço amostral. Já no caso do lançamento da moeda um ponto de amostra é CA ou CO, e os eventos podem ser $\{CA\}$, $\{CO\}$, $\{CA, CO\}$ ou \emptyset ; como o espaço amostral é composto apenas por 2 elementos, fica mais simples determinar todos os eventos possíveis. Se a moeda for justa, a probabilidade de uma cara aparecer é $\frac{1}{2}$, a probabilidade de uma coroa é $\frac{1}{2}$ e a probabilidade de qualquer outro evento é zero. Em geral, para qualquer experimento, há um espaço de probabilidade, um triplo

ordenado $(\Omega, \mathcal{F}, \mathbb{P})$, onde Ω é o espaço amostral, \mathcal{F} é a coleção de subconjuntos ou eventos em Ω e \mathbb{P} é uma medida de probabilidade definida em \mathcal{F} . O conjunto \mathcal{F} deve satisfazer três propriedades e ele é chamado de σ -álgebra. Uma σ -álgebra sobre um conjunto Ω é uma coleção de subconjuntos de Ω , incluindo o conjunto vazio, e que é fechada sobre operações contáveis de união, interseção e complemento de conjuntos. Estas álgebras são muito usadas para definir medidas em Ω . O conceito é importante em análise e probabilidade.

Definição 1. *Seja \mathcal{F} uma coleção de subconjuntos de Ω e Ω um conjunto não vazio. Então \mathcal{F} é dito ser uma σ -álgebra e o par ordenado (Ω, \mathcal{F}) é dito ser um espaço mensurável se \mathcal{F} tem as seguintes propriedades:*

- $\Omega, \emptyset \in \mathcal{F}$;
- Se $A \in \mathcal{F}$ então o complemento de A , denotado por A^c também pertence a \mathcal{F} , isto é, $A^c \in \mathcal{F}$ ou ainda

$$A^c = \{\omega : \omega \in \Omega, \omega \notin A\} \in \mathcal{F};$$

- Para qualquer sequência $\{A_n\}_{n=1}^{\infty}$ então a união $\cup_{n=1}^{\infty} A_n \in \mathcal{F}$.

Os conjuntos $A \in \mathcal{F}$ são chamados acontecimentos ou conjuntos mensuráveis. Uma medida de probabilidade \mathbb{P} é uma função $\mathbb{P} : \mathcal{F} \rightarrow [0,1]$ definida em \mathcal{F} .

Definição 2. *Seja (Ω, \mathcal{F}) um espaço mensurável e \mathbb{P} uma função com valores reais definida na σ -álgebra \mathcal{F} . A função $\mathbb{P} : \mathcal{F} \rightarrow [0,1]$ é chamada de medida de probabilidade se são válidas as seguintes propriedades:*

- $\mathbb{P}(A) \geq 0$ para todo $A \in \mathcal{F}$;
- $\mathbb{P}(\Omega)=1$;
- (Propriedade σ -aditiva) Se A_1, A_2, \dots são conjuntos disjuntos em pares, isto é, $A_i \cap A_j = \emptyset$ para $i \neq j$, que pertencem a \mathcal{F} então $\mathbb{P}(A_1 \cup A_2 \cup \dots) = \mathbb{P}(A_1) + \mathbb{P}(A_2) + \dots$.

A medida de probabilidade \mathbb{P} junto com o espaço mensurável (Ω, \mathcal{F}) define um espaço de probabilidade, o triplo ordenado $(\Omega, \mathcal{F}, \mathbb{P})$. Muitas vezes não definimos o espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$, mas nos concentramos na medida de probabilidade \mathbb{P} . Por exemplo, se Ω é um conjunto finito, \mathcal{F} pode ser considerado como todos os subconjuntos de Ω ou se Ω é o conjunto de números reais \mathbb{R} , o conjunto \mathcal{F} pode ser considerado como os conjuntos mensuráveis de Borel, onde entende-se como σ -álgebra de Borel, a σ -álgebra gerada pelos conjuntos abertos de \mathbb{R} e os elementos dessa σ -álgebra são os conjuntos de Borel. Mais precisamente, se \mathcal{C} é uma família de subconjuntos de Ω , então existe um menor σ -álgebra $\sigma(\mathcal{C})$ em Ω que contém \mathcal{C} . Este $\sigma(\mathcal{C})$ é chamado de σ -álgebra gerada por \mathcal{C} . Se $\Omega = \mathbb{R}^d$ e \mathcal{C} é a família de todos os conjuntos abertos em \mathbb{R}^d , então $\mathcal{B}^d = \sigma(\mathcal{C})$ é chamada de σ -álgebra de Borel e os elementos de \mathcal{B}^d são chamados de conjuntos de Borel.

Podemos pensar no espaço de probabilidade como sendo uma construção matemática essencial, que, entretanto, não é “diretamente observável”. Estamos, portanto, interessados em introduzir uma função X de Ω para \mathbb{R}^d , cujos valores podemos observar.

Se Ω for um espaço amostral finito, então uma função X definida em Ω anexa valores numéricos a cada $\omega \in \Omega$. Como Ω é finito, X assume apenas muitos valores finitos $x_i, i = 1, \dots, k$.

Se uma coleção de eventos \mathcal{F} for especificado, qualquer conjunto nele será chamado de conjunto mensurável. Se $\mathcal{F} = 2^\Omega$ então qualquer subconjunto de Ω é mensurável. Uma função X em Ω é chamada de mensurável com relação a \mathcal{F} ou uma variável aleatória em (Ω, \mathcal{F}) se todos os conjuntos $\{X = x_i\}, i = 1, \dots, k$, são elementos de \mathcal{F} . Isso significa que se temos a informação descrita por \mathcal{F} , ou seja, sabemos qual evento em \mathcal{F} ocorreu, então sabemos qual valor de X ocorreu. Observe que se $\mathcal{F} = 2^\Omega$, então qualquer função em Ω é uma variável aleatória.

Definição 3. Dado um espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$, então a função $X : \Omega \rightarrow \mathbb{R}^d$ é dita ser uma variável aleatória d -dimensional se para cada $B \in \mathcal{B}^d$ temos

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

Equivalentemente, dizemos que X é mensurável com relação a \mathcal{F} .

Ao invés de verificar a definição para todos os conjuntos do Borel, podemos verificar que para todo real x o conjunto $\{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{F}$. Em outras palavras, para uma variável aleatória, podemos atribuir probabilidades a conjuntos da forma $\{X \leq x\}$, e $\{a \leq X \leq b\}$.

Observe que se $A = \{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{F}$ então $B = \{\omega \in \Omega : X(\omega) > x\} \in \mathcal{F}$, pois \mathcal{F} é uma σ -álgebra. Além disso, segue das propriedades de σ -álgebra que $C = \{\omega \in \Omega : X(\omega) < x\} \in \mathcal{F}$ e $D = \{\omega \in \Omega : X(\omega) \geq x\} \in \mathcal{F}$. Por outro lado, se $C = \{\omega \in \Omega : X(\omega) < x\} \in \mathcal{F}$, concluímos que $A, B, D \in \mathcal{F}$. Portanto, podemos usar qualquer um dos conjuntos A, B, C e D para verificar a definição de variável aleatória.

Exemplo 1. Toda função constante é mensurável. De fato, consideremos $f(\omega) = c, \forall \omega \in \Omega$. Então,

$$\{\omega \in \Omega : f(\omega) > x\} = \begin{cases} \Omega, & \text{se } x < c; \\ \emptyset, & \text{se } x \geq c. \end{cases}$$

Os conjuntos que escrevemos sempre são conjuntos de qualquer σ -álgebra de Ω .

Exemplo 2. Seja $A \in \mathcal{F}$. Então a função característica de A , definida por

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

é uma função mensurável. De fato, temos que

- $\{\omega \in \Omega : f(\omega) > x\} = \emptyset$ se $x \geq 1$ e $\emptyset \in \mathcal{F}$.
- $\{\omega \in \Omega : f(\omega) > x\} = A$ se $0 \leq x < 1$ e $A \in \mathcal{F}$.
- $\{\omega \in \Omega : f(\omega) > x\} = \Omega$ se $x < 0$ e $\Omega \in \mathcal{F}$.

Logo, \mathcal{X}_A é mensurável.

Toda variável aleatória induz uma medida de probabilidade μ_X em \mathbb{R}^d dada por

$$\mu_X(B) = \mathbb{P}(X^{-1}(B)),$$

onde $\mu_X(B)$ é denominada distribuição de X . A integral

$$\mathbb{E}(X) = \int_{\Omega} X(\omega) d\mathbb{P}(\omega) = \int_{\mathbb{R}^d} x d\mu(x),$$

quando existe é chamada de esperança ou valor esperado da variável aleatória X .

A variância de X pode ser definida como

$$\text{Var}(X) = \int_{\Omega} [X - \mathbb{E}(X)]^2 d\mathbb{P},$$

ou ainda,

$$\text{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2.$$

A variância pode ser compreendida como uma medida de dispersão que nos permitem reconhecer quanto os dados estão dispersos ao redor da média.

Seja X uma variável aleatória integrável com o espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$ e seja \mathcal{G} uma σ -álgebra contida em \mathcal{F} . Então a esperança condicional de X dado \mathcal{G} é definida como uma variável aleatória $\mathbb{E}(X|\mathcal{G})$ tal que

- $\mathbb{E}(X|\mathcal{G})$ é mensurável em \mathcal{G} ;
- $\int_A \mathbb{E}(X|\mathcal{G})d\mathbb{P} = \int_A Xd\mathbb{P}, \quad \forall A \in \mathcal{G}$.

A esperança condicional tem as seguintes propriedades:

1. $\mathbb{E}(aX + bY|\mathcal{G}) = a\mathbb{E}(X|\mathcal{G}) + b\mathbb{E}(Y|\mathcal{G})$, onde $Y : \Omega \rightarrow \mathbb{R}^d$ é uma outra variável aleatória com $\mathbb{E}[Y] < \infty$ e $a, b \in \mathbb{R}$;
2. $\mathbb{E}(\mathbb{E}(X|\mathcal{G})) = \mathbb{E}(X)$;
3. $\mathbb{E}(X|\mathcal{G}) = X$, se X é mensurável com relação a \mathcal{G} ;
4. $\mathbb{E}(X \cdot Y|\mathcal{G}) = X\mathbb{E}(Y|\mathcal{G})$, se X é mensurável em \mathcal{G} , onde \cdot denota o produto interno usual em \mathbb{R}^d ;
5. $\mathbb{E}(X|\mathcal{G}) = \mathbb{E}(X)$, se X é independente de \mathcal{G} .

As demonstrações destas propriedades podem ser vistas em Oksendal (1998, p.299) [19].

Uma coleção indexada $X_t, t \in T$ de variáveis aleatórias é dito ser um processo estocástico no espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$. Em outras palavras, processos estocásticos são uma sequência de variáveis aleatórias que variam de forma imprevisível à medida que o tempo passa, como por exemplo a variação do tráfego em um cruzamento, flutuações nos mercados de ações e nas taxas de câmbio, movimentos/passeios aleatórios (movimento browniano), entre outros. Formalmente temos a seguinte definição.

Definição 4. *Um processo estocástico é uma coleção de variáveis aleatórias $\{X_t(\omega) : t \in T, \omega \in \Omega\}$, onde T é algum conjunto de índices e Ω é o espaço amostral comum das variáveis aleatórias. Para cada t fixo, $X_t(\omega)$ denota uma única variável aleatória definida em Ω . Para cada $\omega \in \Omega$, $X_t(\omega)$ corresponde a uma função definida em T que é chamada de caminho de amostra ou uma realização estocástica do processo.*

No nosso caso, t será interpretado como tempo e o conjunto de índices T será tomado como um intervalo de tempo da forma $(-\infty, +\infty)$, $[0, +\infty)$ ou $[\alpha, \beta]$; nessas condições, o processo estocástico é dito ser contínuo (ou processos estocásticos em tempo contínuo). Em outros casos, T pode ser o conjunto dos números inteiros ou dos números inteiros não-negativos, e os processos estocásticos nessas condições são ditos discretos (ou processos estocásticos em tempo discreto).

Note que cada variável aleatória $X_t = X_t(\omega)$ é uma função do “acaso” $\omega \in \Omega$, então um processo estocástico pode ser considerado uma função de duas variáveis, $t \in T$ e $\omega \in \Omega$, isto é uma função do tempo e do acaso. Para abreviar a notação, costumamos escrever simplesmente X_t em vez de $X_t(\omega)$, mas apesar do “acaso” ω não aparecer explicitamente, não devemos esquecer que o valor do processo estocástico depende dele. Esta função de t e ω não é uma função arbitrária pois está sujeita à restrição de ser, para cada t fixo, uma função mensurável de ω , isto é, uma variável aleatória.

Um processo estocástico pode ser interpretado como uma coleção de trajetórias, uma vez que podemos fixar ω , obtendo uma função apenas do tempo, que chamamos de trajetória. Assim, para cada ω fixo, X_t representa uma coleção de trajetórias.

Muitos fenômenos que evoluem com o tempo, podem ser compreendidos por uma sequência de estados, onde a transição de um determinado estado para o seguinte ocorre segundo uma certa probabilidade. No caso em que esta transição depende apenas do estado em que o fenômeno se encontra e do seguinte, ou seja, não depende dos estados passados, o processo é chamado de processo de Markov. Formalmente, uma coleção de variáveis aleatórias (X_t) é um processo

estocástico de Markov, se para qualquer sequência finita de tempos $0 < t_1 < \dots < t_{n-1} < t_n$, e $y \in \mathbb{R}$, temos que

$$\mathbb{P}\{\omega \in \Omega : X_{t_n}(\omega) \leq y \mid X_0(\omega) = x_0, \dots, X_{t_{n-1}}(\omega) = x_{n-1}\} \quad (1.1.0.1)$$

$$= \mathbb{P}\{\omega \in \Omega : X_{t_n}(\omega) \leq y \mid X_{t_{n-1}}(\omega) = x_{n-1}\}, \quad (1.1.0.2)$$

para toda sequência x_0, x_1, \dots, x_{n-1} de elementos do espaço de estados E . Mais detalhes sobre processos de Markov podem ser vistos em Allen (2010) [1], Brzezniak; Zastawniak (1999) [2] e Klebaner (2005) [10].

1.2 Movimento Browniano

Em 1828 o botânico Robert Brown descreveu o movimento de uma partícula de pólen suspensa em um fluido como sendo um movimento irregular e aleatório. Mais tarde, em 1905, Albert Einstein argumentou que este movimento era oriundo do bombardeamento de outras partículas do fluido e interpretou o movimento numa abordagem estatística, obtendo uma equação da difusão para a função de densidade de probabilidade que descreve a dinâmica da partícula que evolui com o passar do tempo.

Posteriormente, aproximadamente em 1931, Norbert Wiener apresentou uma fundamentação matemática que pudesse interpretar o movimento browniano como um processo estocástico e ainda formalizou matematicamente a teoria. Por essa razão, o movimento browniano também é chamado em vários livros de Processo de Wiener.

Este processo estocástico contínuo, que ao longo do texto chamaremos apenas de movimento browniano, tem a propriedade de Markov descrita pela equação (1.1.0.1), é um martingale (propriedade que veremos mais adiante) e possui distribuição gaussiana. Estas propriedades explicam a sua relevância em diversas aplicações nas áreas de engenharia, economia e biologia, como pode ser visto, entre outros, em Allen (2010) [1], Klebaner (2005) [10] e Oksendal (1998) [19].

Definição 5. *Formalmente, um processo estocástico $B_t : \Omega \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}$, é um movimento browniano se*

1. *Possui trajetórias contínuas e inicializadas na origem: $B_0(\omega) = 0$, para quase todo $\omega \in \Omega$;*
2. *Os incrementos $B_t - B_s$ tem distribuição normal com média zero e variância $t - s$ para $0 \leq s \leq t$, que denotamos por $N(0, t - s)$;*
3. *Os incrementos $B_v - B_u$ e $B_t - B_s$ são independentes em intervalos de tempo disjuntos $0 \leq s \leq t \leq u \leq v$.*

Logo, como o movimento browniano assume uma normal $N(0, t)$, temos que a esperança $\mathbb{E}(B_t) = 0$, e ainda, como a variância $\text{Var}(B_t) = t$, obtemos que

$$\mathbb{E}(B_t^2) = t,$$

pois

$$t = \text{Var}(B_t) = \mathbb{E}(B_t^2) - [\mathbb{E}(B_t)]^2 = \mathbb{E}(B_t^2) - 0 = \mathbb{E}(B_t^2).$$

Sabemos que o movimento browniano B_t tem distribuição $N(0, t)$ e portanto, para $t > 0$ e $a \leq b$,

$$\mathbb{P}(a \leq B_t \leq b) = \frac{1}{\sqrt{2\pi t}} \int_a^b e^{-\frac{x^2}{2t}} dx.$$

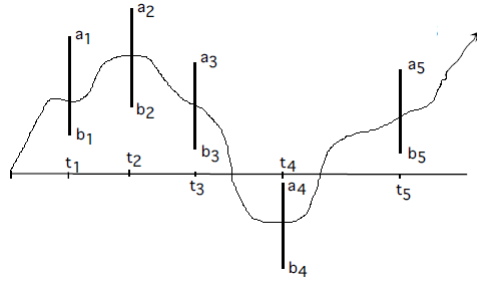


Figura 1.1: Possível trajetória do movimento

Considerando os tempos $0 < t_1 < t_2 < \dots < t_n$ e os números reais $a_i \leq b_i$, para $i = 1, \dots, n$, distribuídos de acordo com a Figura 1.1, podemos determinar qual é a probabilidade de uma trajetória de B_t passar entre os valores a_i e b_i no tempo t_i para cada $i = 1, \dots, n$.

Para isso, notemos que

$$\mathbb{P}(a_1 \leq B_{t_1} \leq b_1) = \int_{a_1}^{b_1} \frac{e^{-\frac{x_1^2}{2t_1}}}{\sqrt{2\pi t_1}} dx_1,$$

e admitindo $B_{t_1} = x_1$, para $a_1 \leq x_1 \leq b_1$ e considerando que o processo que “seguirá” tem distribuição $N(x_1, t_2 - t_1)$ no intervalo $[t_1, t_2]$, temos que a probabilidade de $a_2 \leq B_{t_2} \leq b_2$, dado que $B_{t_1} = x_1$, corresponde a equação

$$\int_{a_2}^{b_2} \frac{1}{\sqrt{2\pi(t_2 - t_1)}} e^{-\frac{|x_2 - x_1|^2}{2(t_2 - t_1)}} dx_2.$$

Logo,

$$\mathbb{P}(a_1 \leq B_{t_1} \leq b_1, a_2 \leq B_{t_2} \leq b_2) = \int_{a_1}^{b_1} \int_{a_2}^{b_2} g(x_1, t_1 | 0) g(x_2, t_2 - t_1 | x_1) dx_2 dx_1,$$

onde

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

Em geral, podemos escrever

$$\begin{aligned} & \mathbb{P}(a_1 \leq B_{t_1} \leq b_1, \dots, a_n \leq B_{t_n} \leq b_n) = \\ & = \int_{a_1}^{b_1} \dots \int_{a_n}^{b_n} g(x_1, t_1 | 0) g(x_2, t_2 - t_1 | x_1) \dots g(x_n, t_n - t_{n-1} | x_{n-1}) dx_n \dots dx_1. \end{aligned} \quad (1.2.0.1)$$

Podemos estender a equação (1.2.0.1), de acordo com o seguinte teorema, cuja demonstração pode ser vista em Evans (2012, p.39-40) [5].

Teorema 6. *Seja $B_t(\cdot)$ um movimento browniano uni-dimensional. Então, para todo n inteiro positivo, para todo tempo $0 = t_0 < t_1 < t_2 < \dots < t_n$ e para cada função $f : \mathbb{R}^d \rightarrow \mathbb{R}$, temos que*

$$\begin{aligned} \mathbb{E}[f(B_{t_1}, \dots, B_{t_n})] &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_1, \dots, x_n) g(x_1, t_1 | 0) g(x_2, t_2 - t_1 | x_1) \\ & \dots g(x_n, t_n - t_{n-1} | x_{n-1}) dx_n \dots dx_1. \end{aligned}$$

Quando $f(x_1, \dots, x_n) = \chi_{[a_1, b_1]}(x_1) \dots \chi_{[a_n, b_n]}(x_n)$, temos (1.2.0.1). Aqui $\chi_{[a_i, b_i]}(x_i)$ é definido como

$$\chi_{[a_i, b_i]}(x_i) = \begin{cases} 1, & \text{se } x_i \in [a_i, b_i] \\ 0, & \text{se } x_i \notin [a_i, b_i], \end{cases}$$

para $i = 1, \dots, n$.

Proposição 7. 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. Considere $0 \leq s \leq t$, então

$$\begin{aligned} \mathbb{E}(B_t B_s) &= \mathbb{E}((B_t)B_s) \\ &= \mathbb{E}((B_t - B_s + B_s)B_s) \\ &= \mathbb{E}([B_t - B_s] + B_s)B_s \\ &= \mathbb{E}((B_t - B_s)B_s + B_s^2) \\ &= \mathbb{E}((B_t - B_s)B_s) + \mathbb{E}(B_s^2) \end{aligned} \tag{1.2.0.2}$$

$$= \mathbb{E}((B_t - B_s)B_s) + s \tag{1.2.0.3}$$

$$= 0 \times B_s + s$$

$$= s.$$

De (1.2.0.2) para (1.2.0.3) utilizamos que $\mathbb{E}(B_s^2) = s$. Em (1.2.0.3) usamos que $(B_t - B_s)$, com $t \geq s \geq 0$, é uma $N(0, t - s)$ e portanto $\mathbb{E}(B_t - B_s) = 0$. ■

1.2.1 Variação quadrática

Considerando um processo estocástico contínuo $X_t : \Omega \rightarrow \mathbb{R}$ e uma partição no intervalo $[0, t]$ $P = \{0 = t_0^n < t_1^n < \dots < t_n^n = t\}$ e com seu módulo dado por $|P| = \max_{1 \leq i \leq n} (t_i - t_{i-1})$, temos que, para $p > 0$, a p -variação do processo X_t denotada por $[X]_t^{(p)}$ é definida como

$$[X]_t^{(p)}(\omega) = \lim_{|P| \rightarrow 0} \sum_{i=1}^n |X_{t_i}(\omega) - X_{t_{i-1}}(\omega)|^p. \tag{1.2.1.1}$$

Em particular, se $p = 2$ então $[X]_t^2$ é chamado de variação quadrática do processo. Dessa forma, a variação quadrática do movimento Browniano B_t é definida como

$$[B]_t^{(2)} = \lim_{|P| \rightarrow 0} \sum_{i=1}^n |B_{t_i} - B_{t_{i-1}}|^2. \tag{1.2.1.2}$$

Embora as somas na equação acima sejam aleatórios, seu limite não é aleatório, como mostra o resultado a seguir.

Usaremos $L^2(\Omega)$ para denotar o espaço de todas as variáveis aleatórias X com $\mathbb{E}(|X|^2) < \infty$.

Teorema 8. (Variação Quadrática) Considere um intervalo finito $[a, b]$ e suponha $P = \{a = t_0, t_1, \dots, t_{n-1}, t_n = b\}$ suas partições, com $|P| \rightarrow 0$ quando $n \rightarrow \infty$. Então

$$\sum_{k=1}^n (B_{t_k} - B_{t_{k-1}})^2 \rightarrow b - a, \tag{1.2.1.3}$$

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

Note que convergência em $L^2(\Omega)$ implica convergência em probabilidade. Por outro lado, convergência em probabilidade de uma sequência implica convergência quase certa de alguma subsequência. Portanto, existe uma subsequência $\{\tilde{P}\}$ de $\{P\}$ tal que a convergência na Equação (1.2.1.3) é uma convergência quase certa, já que $|\tilde{P}|$ tende a zero. Na verdade, a convergência quase certa na Equação (1.2.1.3) é garantida se a sequência $\{P\}$ satisfizer a condição:

$$P_1 \subset P_2 \subset \dots \subset P_n \subset \dots$$

Demonstração. Seja

$$Q_n := \sum_{k=1}^n (B_{t_k} - B_{t_{k-1}})^2.$$

Então

$$Q_n - (b - a) = \sum_{k=1}^n (B_{t_k} - B_{t_{k-1}})^2 - (t_k - t_{k-1}).$$

Logo,

$$\mathbb{E}((Q_n - (b - a))^2) = \sum_{k=1}^n \sum_{j=1}^n \mathbb{E}\{[(B_{t_k} - B_{t_{k-1}})^2 - (t_k - t_{k-1})][(B_{t_j} - B_{t_{j-1}})^2 - (t_j - t_{j-1})]\}.$$

Para $k \neq j$, o termo na soma dupla é

$$\mathbb{E}\{[(B_{t_k} - B_{t_{k-1}})^2 - (t_k - t_{k-1})]\} \mathbb{E}\{[(B_{t_j} - B_{t_{j-1}})^2 - (t_j - t_{j-1})]\},$$

pois são incrementos independentes. Como $B_t - B_s$ é $N(0, t - s)$ para todo $0 \leq s \leq t$, temos que $\mathbb{E}(B_{t_k} - B_{t_{k-1}})^2 = t_k - t_{k-1}$ e assim a soma dupla resulta em 0.

Agora, se $k = j$, temos

$$\begin{aligned} \mathbb{E}((Q_n - (b - a))^2) &= \sum_{k=1}^n \mathbb{E}\{[(B_{t_k} - B_{t_{k-1}})^2 - (t_k - t_{k-1})][(B_{t_k} - B_{t_{k-1}})^2 - (t_k - t_{k-1})]\} \\ &= \sum_{k=1}^n \mathbb{E}\{(B_{t_k} - B_{t_{k-1}})^4 - 2(B_{t_k} - B_{t_{k-1}})^2(t_k - t_{k-1}) + (t_k - t_{k-1})^2\} \\ &= \sum_{k=1}^n \mathbb{E}\left\{\left[\frac{(B_{t_k} - B_{t_{k-1}})^4}{(t_k - t_{k-1})^2} - 2\frac{(B_{t_k} - B_{t_{k-1}})^2}{(t_k - t_{k-1})} + 1\right](t_k - t_{k-1})^2\right\} \\ &= \sum_{k=1}^n \mathbb{E}\left\{\left[\underbrace{\frac{(B_{t_k} - B_{t_{k-1}})^2}{(t_k - t_{k-1})}}_{Y_k^2} - 1\right]^2 (t_k - t_{k-1})^2\right\} \\ &= \sum_{k=1}^n \mathbb{E}\{(Y_k^2 - 1)^2 (t_k - t_{k-1})^2\}. \end{aligned}$$

Ou seja,

$$\mathbb{E}((Q_n - (b - a))^2) = \sum_{k=1}^n \mathbb{E}\{(Y_k^2 - 1)^2 (t_k - t_{k-1})^2\},$$

onde

$$Y_k := \frac{(B_{t_k} - B_{t_{k-1}})}{\sqrt{(t_k - t_{k-1})}}$$

é $N(0, 1)$. Portanto, para alguma constante C

$$\mathbb{E}((Q_n - (b - a))^2) \leq C \sum_{k=1}^n (t_k - t_{k-1})^2.$$

Note que:

$$\begin{aligned} \sum_{k=1}^n (t_k - t_{k-1}) &= (t_1 - t_0) + (t_2 - t_1) + (t_3 - t_2) + \cdots + (t_n - t_{n-1}) \\ &= t_1 - t_0 + t_2 - t_1 + t_3 - t_2 + \cdots + t_n - t_{n-1} \\ &= -t_0 + t_n. \end{aligned}$$

Logo,

$$\begin{aligned} \sum_{k=1}^n (t_k - t_{k-1})^2 &= \sum_{k=0}^{n-1} \underbrace{(t_{k+1} - t_k)}_{\leq |P|} (t_{k+1} - t_k) \\ &\leq |P| \underbrace{(t_n - t_0)}_{\substack{b \\ a}} \\ &\leq |P|(b - a). \end{aligned}$$

Assim,

$$\begin{aligned} \mathbb{E}((Q_n - (b - a))^2) &\leq C \sum_{k=1}^n (t_k - t_{k-1})^2 \\ &\leq C|P|(b - a). \end{aligned}$$

onde $C|P|(b - a) \rightarrow 0$ quando $n \rightarrow \infty$. ■

Este resultado mostra que a variação quadrática de B_t , no intervalo $[0, t]$ é diferente de zero¹, e portanto a variação total de sua trajetória é infinita (Klebaner (2005) [10]), ou seja, o movimento browniano não é um processo de variação finita (mesmo em intervalos de tempo finito pequeno, o gráfico de sua trajetória tem variação infinita).

1.2.2 Movimento browniano no \mathbb{R}^d

O movimento browniano pode ser estendido para o \mathbb{R}^d da seguinte forma. Um processo estocástico $(B_t)_{t \geq 0} = (B_t^1, \dots, B_t^d)$ com valores no \mathbb{R}^d , é um movimento browniano d -dimensional se

(a) para cada $k = 1, \dots, d$, $(B_t^k)_{t \geq 0}$ é um movimento browniano unidimensional;

(b) as σ -álgebras $\mathcal{B}^k := \mathcal{F}(B_t^k | t \geq 0)$ são independentes, $k = 1, \dots, d$.

Pelos argumentos acima, podemos construir um espaço de probabilidade, e nele, d movimentos brownianos unidimensionais independentes $B^k(\cdot)$ ($k = 1, \dots, d$). Então $B^k(\cdot) = (B^1(\cdot), \dots, B^d(\cdot))$ é um movimento browniano d -dimensional. Em outras palavras, dizemos que $B_t^k = (B_t^1, \dots, B_t^d)$ é um processo de Wiener m -dimensional se B_t^1, \dots, B_t^d são processos de Wiener independentes. Desse modo, são válidas as seguintes propriedades:

(i) $\mathbb{E}(B_t^k B_s^l) = (t \wedge s) \lambda_{kl}$, ($k, l = 1, \dots, d$),

(ii) $\mathbb{E}((B_t^k - B_s^k)(B_t^l - B_s^l)) = (t - s) \lambda_{kl}$, ($k, l = 1, \dots, d$; $t \geq s \geq 0$),

em que,

$$\lambda_{kl} = \begin{cases} 1, & \text{se } k = l \\ 0, & \text{se } k \neq l. \end{cases}$$

De fato, no item (i), se $k \neq l$, então

$$\mathbb{E}(B_t^k B_s^l) = \mathbb{E}(B_t^k) \mathbb{E}(B_s^l) = 0,$$

por independência. Ainda em (i), se $k = l$, então para o movimento browniano B_t^k , temos que $\mathbb{E}(B_t^k B_s^k) = t \wedge s$, como mencionado anteriormente.

¹a variação quadrática do movimento browniano no intervalo $[0; t]$ é t quase sempre, mais detalhes podem ser vistos em diversos livros que falam sobre EDE como Klebaner (2005) [10], Kuo (2005) [12] e Oksendal (1998) [19].

Agora, no item (ii), se $k \neq l$, então

$$\begin{aligned}\mathbb{E}((B_t^k - B_s^k)(B_t^l - B_s^l)) &= \mathbb{E}((B_t^k B_t^l - B_t^k B_s^l - B_t^l B_s^k + B_s^k B_s^l)) \\ &= \mathbb{E}(B_t^k B_t^l) - \mathbb{E}(B_t^k B_s^l) - \mathbb{E}(B_t^l B_s^k) + \mathbb{E}(B_s^k B_s^l) \\ &= 0.\end{aligned}$$

E se $k = l$, então

$$\begin{aligned}\mathbb{E}((B_t^k - B_s^k)(B_t^l - B_s^l)) &= \mathbb{E}((B_t^k - B_s^k)^2) \\ &= \mathbb{E}((B_t^k)^2 - 2B_t^k B_s^k + (B_s^k)^2) \\ &= \mathbb{E}((B_t^k)^2) - 2\underbrace{\mathbb{E}(B_s^k B_t^k)}_{=s \wedge t} + \mathbb{E}((B_s^k)^2) \\ &= t - 2s + s \\ &= t - s.\end{aligned}$$

1.3 Martingale

A teoria dos martingales teve origem com o objetivo de fornecer informações sobre a aparente impossibilidade de ganhar dinheiro fazendo apostas em jogos justos. O sucesso da teoria ultrapassou em muito suas origens, e a teoria de martingale é agora uma das principais ferramentas no estudo de processos aleatórios. Nesta seção, apresentaremos brevemente alguns conceitos úteis da teoria de martingales para a dinâmica estocástica.

Definição 9. Uma filtração num espaço mensurável (Ω, \mathcal{F}) é uma família crescente $(\mathcal{F}_t), t \geq 0$ de σ -álgebras de \mathcal{F} .

Em outras palavras, para cada t , uma família \mathcal{F}_t de σ -álgebras em Ω parametrizadas pelo tempo $t \geq 0$, é chamada de filtração se $\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$ para todo $s < t$. Um espaço mensurável (Ω, \mathcal{F}) dotado de uma filtração $(\mathcal{F}_t), t \geq 0$ é chamado de espaço filtrado.

Um exemplo de filtração é a gerada por um determinado processo estocástico X_t , denotada por $\sigma(X_u : 0 \leq u \leq t)$, que corresponde a informação obtida sobre o processo até o tempo t . Heuristicamente, uma filtração é usada para modelar um fluxo de informações, isto é, conforme o tempo passa, um observador consegue saber mais detalhes sobre os acontecimentos, ou seja, partições mais “finas” de Ω . Por exemplo, seja $\mathcal{F}_t \subset \mathcal{H}_t$, onde $\mathcal{F}_t, \mathcal{H}_t$ são duas filtrações dadas que representam o conhecimento que dois observadores podem adquirir até o tempo t . Dessa maneira, o observador \mathcal{H}_t possui mais informações. De outra forma, essa σ -álgebra representa a coleção de eventos que podem ou não tem ocorrido até o tempo t .

Definição 10. Um processo estocástico X_t é dito ser adaptado em relação à filtração $\mathcal{F}_t : t \in [0, T]$ se para cada $t \in [0, T]$, a variável aleatória X_t é mensurável em \mathcal{F}_t , isto é, $\{X_t \leq \alpha\} \in \mathcal{F}_t, \forall \alpha \in \mathbb{R}$.

Dizemos que uma sequência de variáveis aleatórias discretas $\{M_n : 0 \leq n < \infty\}$ é um martingale com respeito a sequência de variáveis aleatórias $\{X_n : 1 \leq n < \infty\}$, desde que a sequência $\{M_n\}$ exista e tenha duas propriedades básicas. A primeira propriedade é que para cada $n \geq 1$ existe uma função $f_n : \mathbb{R}^d \mapsto \mathbb{R}$ tal que $M_n = f_n(X_1, X_2, \dots, X_n)$ e a segunda propriedade é que a sequência $\{M_n\}$ satisfaz a identidade fundamental de martingale:

$$\mathbb{E}(M_n | X_1, X_2, \dots, X_{n-1}) = M_{n-1}, \quad \forall n \geq 1. \quad (1.3.0.1)$$

A intuição por trás dessa definição é explicada da seguinte forma. Podemos pensar no X_i como o resultado do i -ésimo resultado de algum processo de jogo, digamos a cara ou a coroa

que alguém observaria em um lançamento de moeda. Também podemos pensar em M_n como a fortuna de um jogador que faz apostas justas em quantidades variáveis nos resultados dos lançamentos de moedas. A fórmula (1.3.0.1) nos diz que o valor esperado da fortuna do jogador no tempo n , dada todas as informações nos primeiros $n - 1$ lançamentos da moeda, é simplesmente M_{n-1} , ou seja, o valor real da fortuna do jogador antes da n -ésima rodada do jogo do lançamento da moeda. Então, se quisermos prever o próximo valor baseando-se em todos os valores prévios, não haverá melhor valor do que o atual.

A propriedade martingale (1.3.0.1) leva a uma teoria que ilumina brilhantemente o fato de que um jogador em um jogo justo não pode esperar ganhar dinheiro, por mais inteligente que ele faça suas apostas. Dessa forma, os martingales excluem a possibilidade de estratégias de ganho baseadas no histórico do jogo e, portanto, são um modelo de jogos honestos. É também uma técnica utilizada no mercado financeiro, para recuperar operações perdidas.

Definição 11. Dizemos que um processo estocástico X_t ($t \geq 0$) adaptado a filtração \mathcal{F}_t e com $\mathbb{E}|X_t| < \infty$ para todo $t \in T$ é um martingale se, e somente se,

$$\mathbb{E}(X_t | \mathcal{F}_s) = X_s, \text{ quase sempre,} \quad (1.3.0.2)$$

para $s \leq t$ em $[0, T]$.

De acordo com a definição, temos que a esperança de um martingale é constante no tempo e igual a $\mathbb{E}[M_0]$. De fato, seja M_t um martingale para $0 \leq t \leq T$, isto é, $\mathbb{E}(M_t | \mathcal{F}_s) = M_s$. Em particular, $\mathbb{E}(M_t | \mathcal{F}_0) = M_0$. Consequentemente, utilizando a propriedade 2 de esperança condicional dada na página 16 temos que

$$\mathbb{E}[M_t] = \mathbb{E}[\mathbb{E}(M_t | \mathcal{F}_0)] = \mathbb{E}[M_0].$$

Se substituirmos a igualdade na equação (1.3.0.2) por sinais de desigualdades, teremos a definição de submartingales e supermartingales, isto é,

$$\mathbb{E}(X_t | \mathcal{F}_s) \geq X_s, \text{ quase sempre,} \quad (\text{submartingales})$$

$$\mathbb{E}(X_t | \mathcal{F}_s) \leq X_s, \text{ quase sempre,} \quad (\text{supermartingales})$$

A seguir são dados alguns exemplos de processos que possuem a propriedade de martingale.

Exemplo 3. Tome $\mathcal{F}_n = \sigma(\{X_1, \dots, X_n\})$ e $S_n = \sum_{i=1}^n X_i$, onde X_n são variáveis aleatórias independentes tais que $\mathbb{E}(X_n) = 0$. Observe que $\{\mathcal{F}_n\}$ é uma filtração, pois

$$\mathcal{F}_1 = \sigma(\{X_1\}), \quad \mathcal{F}_2 = \sigma(\{X_1, X_2\}), \quad \dots, \quad \mathcal{F}_n = \sigma(\{X_1, \dots, X_n\}) \Rightarrow \mathcal{F}_1 \subset \mathcal{F}_2 \subset \dots \subset \mathcal{F}_n.$$

Além disso, $\{S_n\}$ é um martingale, pois

$$\mathbb{E}(S_n | \mathcal{F}_{n-1}) = \mathbb{E}(X_n | \mathcal{F}_{n-1}) + \mathbb{E}\left(\sum_{i=1}^{n-1} X_i | \mathcal{F}_{n-1}\right)$$

Como $\sum_{i=1}^{n-1} X_i$ é mensurável em \mathcal{F}_{n-1} , segue pela propriedade de esperança condicional que

$$\mathbb{E}\left(\sum_{i=1}^{n-1} X_i | \mathcal{F}_{n-1}\right) = \sum_{i=1}^{n-1} X_i.$$

E ainda, como X_n são variáveis aleatórias independentes,

$$\mathbb{E}(X_n | \mathcal{F}_{n-1}) = \mathbb{E}(X_n).$$

Logo,

$$\mathbb{E}(X_n | \mathcal{F}_{n-1}) + \mathbb{E}\left(\sum_{i=1}^{n-1} X_i | \mathcal{F}_{n-1}\right) = \mathbb{E}(X_n) + \sum_{i=1}^{n-1} X_i = S_{n-1}.$$

Ou seja,

$$\mathbb{E}(S_n | \mathcal{F}_{n-1}) = S_{n-1},$$

e assim, S_n é martingale.

Exemplo 4. Seja B_t o movimento browniano e considere a filtração $\mathcal{F}_t = \sigma(B_s, s \leq t)$. Então

(1) B_t é um martingale;

(2) $B_t^2 - t$ é um martingale.

De fato,

(1) Como B_t é mensurável em \mathcal{F}_t , segue pela propriedade de esperança condicional que $\mathbb{E}(B_t | \mathcal{F}_t) = B_t$, e sendo $B_{t+s} - B_t$ independente de \mathcal{F}_t segue que $\mathbb{E}(B_{t+s} - B_t | \mathcal{F}_t) = \mathbb{E}(B_{t+s} - B_t)$. Assim sendo, temos que

$$\mathbb{E}(B_{t+s} | \mathcal{F}_t) = \mathbb{E}(B_{t+s} - B_t + B_t | \mathcal{F}_t) = \mathbb{E}(B_{t+s} - B_t | \mathcal{F}_t) + \mathbb{E}(B_t | \mathcal{F}_t) = B_t.$$

(2) Usaremos para provar este item, que sendo $(B_{t+s} - B_t)$ independente de \mathcal{F}_t e g uma função contínua, então $g[B_{t+s} - B_t]$ também é independente de \mathcal{F}_t , ou seja,

$$\mathbb{E}(g[B_{t+s} - B_t] | \mathcal{F}_t) = \mathbb{E}(g[B_{t+s} - B_t]). \quad (1.3.0.3)$$

Primeiramente, notemos que

$$\begin{aligned} \mathbb{E}(B_{t+s}^2 | \mathcal{F}_t) &= \mathbb{E}[(B_t + B_{t+s} - B_t)^2 | \mathcal{F}_t] \\ &= \mathbb{E}[B_t^2 + 2B_t(B_{t+s} - B_t) + (B_{t+s} - B_t)^2 | \mathcal{F}_t] \\ &= \mathbb{E}[B_t^2 | \mathcal{F}_t] + \mathbb{E}[2B_t(B_{t+s} - B_t) | \mathcal{F}_t] + \mathbb{E}[(B_{t+s} - B_t)^2 | \mathcal{F}_t]. \end{aligned}$$

E tais esperanças podem ser calculadas da seguinte forma:

(a) $\mathbb{E}[B_t^2 | \mathcal{F}_t] = B_t^2$, pois temos a informação até o tempo t , e sendo B_t mensurável em \mathcal{F}_t , sabemos quais valores de B_t tem ocorrido. Ou seja estamos tomando $g(x) = x^2$ em (1.3.0.3).

(b) Da independência de $(B_{t+s} - B_t)$ com relação a \mathcal{F}_t e do fato que B_t é mensurável em \mathcal{F}_t , segue que

$$\begin{aligned} 2\mathbb{E}[B_t(B_{t+s} - B_t) | \mathcal{F}_t] &= B_t \mathbb{E}[B_{t+s} - B_t | \mathcal{F}_t] \\ &= B_t \mathbb{E}(B_{t+s} - B_t) \\ &= 0. \end{aligned}$$

(c) Usando $g(x) = x^2$ em (1.3.0.3), e o fato que o incremento $B_{t+s} - B_t$ tem distribuição $N(0, s)$,

$$\mathbb{E}[(B_{t+s} - B_t)^2 | \mathcal{F}_t] = \mathbb{E}[(B_{t+s} - B_t)^2] = s.$$

De (a), (b) e (c), temos que

$$\mathbb{E}[B_{t+s}^2 | \mathcal{F}_t] = B_t^2 + s.$$

Logo,

$$\mathbb{E}[B_{t+s}^2 | \mathcal{F}_t] - (t+s) = B_t^2 - t,$$

e assim,

$$\mathbb{E}[B_{t+s}^2 - (t+s) | \mathcal{F}_t] = B_t^2 - t.$$

Portanto, o processo $X_t = B_t^2 - t$ é um martingale, uma vez que

$$\mathbb{E}[X_{t+s} | \mathcal{F}_t] = X_t.$$

1.3.1 Martingale local e semimartingale

Uma variável aleatória $T : \Omega \rightarrow [0, \infty]$ munida com uma filtração é $\{\mathcal{F}_t\}$ chamada de tempo de parada se

$$[T \leq t] \in \mathcal{F}_t, (t \geq 0). \quad (1.3.1.1)$$

Em outras palavras, se o conjunto $\{\omega : T(\omega) \leq t\}$ for mensurável em \mathcal{F}_t para todo $t \geq 0$, então T é um tempo de parada. Se a filtração \mathcal{F}_t for contínua à direita, isto é, se $\mathcal{F}_t = \mathcal{F}_{t+}$ onde $\mathcal{F}_{t+} = \bigcap_{s>t} \mathcal{F}_s$ (limite lateral da filtração à direita), então a equação (1.3.1.1) dada acima corresponde a $[T < t] \in \mathcal{F}_t, (t \geq 0)$. Este fato é demonstrado em Sondermann (2006, p.39) [24].

A ideia do tempo de parada, como o nome já induz, tem o objetivo de indicar quando um determinado processo para. Por exemplo, sabemos que a quantidade de informações que uma família de variáveis aleatórias possui reflete no tamanho da sua σ -álgebra, ou seja, quanto maior for uma σ -álgebra mais informações temos sobre uma família de variáveis aleatórias. Além disso, sabemos também que conhecer um processo até um tempo t equivale a conhecer a σ -álgebra \mathcal{F}_t . Dessa forma, um tempo de parada T é uma variável aleatória que revela algum acontecimento nas trajetórias. Consequentemente, T apresenta uma situação que ocorre somente até o tempo t .

O tempo de entrada de X em G , onde X denota um processo estocástico adaptado em \mathbb{R}^d e $G \subseteq \mathbb{R}^d$, é dado pela variável aleatória

$$T_G(\omega) = \inf\{t \geq 0 : X_t(\omega) \in G\},$$

onde por conveniência definimos que $\inf\{\emptyset\} = \infty$.

Proposição 12. *Seja X um processo estocástico contínuo com \mathcal{F}_t sua filtração natural.*

1. *Se G for fechado então T_G é um \mathcal{F}_t tempo de parada.*
2. *Se G for aberto então T_G é um \mathcal{F}_{t+} tempo de parada.*

A demonstração da proposição pode ser vista em Ruffino (2009, p.42) [22].

Dado um processo estocástico X e um tempo de parada T com relação a \mathcal{F}_t , o processo que fica constante a partir de T é chamado de processo parado e é dado por $(X^T)_t := X_{t \wedge T}$.

É possível verificar que todo martingale é um martingale local. Para isso precisaremos de alguns conceitos relacionados a integração uniforme. Não entraremos em detalhes, mas o leitor interessado pode consultar, por exemplo, [Klebaner, 2005, p.186] [10].

O termo “local” em martingale local, pode ser entendido tanto no sentido de espaço como no sentido de tempo, pois podemos restringir processos adaptados a certos conjuntos mensuráveis de acordo com nossas necessidades definindo os tempos de parada adequados que podem ser abertos ou fechados de acordo com a Proposição 12.

Diz-se que um processo X tem variação finita se tiver variação limitada para cada intervalo de tempo finito (com probabilidade 1). Tais processos são muito comuns e incluem, particularmente, todas as funções continuamente diferenciáveis. Uma função $f : [a, b] \rightarrow \mathbb{R}$ é de variação limitada ou também chamada de variação finita no intervalo $[a, b]$ se

$$\sup \left\{ \sum_{i=1}^n |f(x_i) - f(x_{i-1})| \right\} < \infty,$$

em que $a = x_0 < x_1 < \dots < x_n = b$ é uma partição de $[a, b], a < b$.

Definição 13. *Um processo estocástico contínuo X é chamado de semimartingale se ele é a soma de um martingale local M e um processo de variação finita A .*

Ou seja, admite uma única decomposição, chamada de decomposição de Doob-Meyer, da forma

$$X_t = X_0 + M_t + A_t, \quad (1.3.1.2)$$

com $M_0 = A_0 = 0$ e X_0 é o valor inicial.

Para verificar que a decomposição de um semimartingale é única, precisamos do seguinte teorema cuja demonstração pode ser vista em Revuz e Yor (1999, p. 134, Proposição 1.2) [21].

Teorema 14. *Dado M um mantingale contínuo. Se M tem variação finita, então M é constante.*

Com este teorema é simples verificar que a decomposição dada na equação (1.3.1.2) é única. Suponha que existe uma outra decomposição, $X_t = X_0 + N_t + B_t$, onde N é um martingale e B um processo de variação finita. Então

$$M_t - N_t = B_t - A_t,$$

mas $M_t - N_t$ é um martingale. Como $M_0 = 0$, temos que $N_0 = 0$ e pelo teorema $M_t - N_t = 0$ e assim $M_t = N_t$ e $A_t = B_t$

Como mencionado anteriormente, como todo martingale é um martingale local podemos dizer que um processo estocástico contínuo X é chamado de semimartingale se ele é a soma de um martingale e um processo de variação finita.

1.4 Integral de Wiener

Nesta seção, queremos dar sentido à seguinte integral:

$$\int_a^b f(t)dB_t(\omega) \quad (1.4.0.1)$$

onde $B_t(\omega)$ é o movimento browniano e f é uma função determinística, ou seja, ela não depende de ω . Para cada $\omega \in \Omega$ queremos definir a integral (1.4.0.1) usando a regra de integração por partes do cálculo usual como sendo

$$\int_a^b f(t)dB_t(\omega) = f(t)B_t(\omega)|_a^b - \int_a^b B_t(\omega)df(t). \quad (1.4.0.2)$$

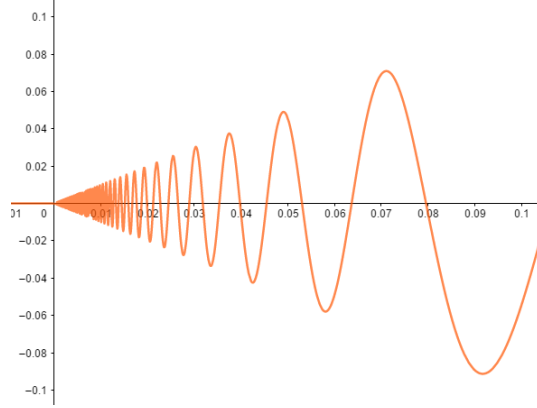
A classe de funções $f(t)$ no qual a integral $\int_a^b f(t)dB_t(\omega)$ é definida para todo $\omega \in \Omega$ é bastante limitada pois, tal integral só existe se $f(t)$ for uma função contínua de variação limitada, devido a definição da Integral de Lebesgue. Portanto, para uma função contínua de variação ilimitada, não podemos usar a equação (1.4.0.2) para definir a integral $\int_a^b f(t)dB_t(\omega)$ para cada $\omega \in \Omega$.

Note que, de fato, $f : [0, 1] \rightarrow \mathbb{R}$ dada por

$$\begin{cases} 0, & \text{se } t = 0 \\ t \operatorname{sen}(\frac{1}{t}), & \text{se } 0 < t \leq 1, \end{cases}$$

é contínua, mas não é de variação limitada, visto que a função oscila muito quando t está próximo de zero. Não entraremos em detalhes, mas isso pode ser observado no gráfico da função a seguir.

Precisamos de uma forma diferente para definir a integral $\int_a^b f(t)dB_t(\omega)$ para uma classe mais ampla de funções $f(t)$. Essa nova integral, chamada integral de Wiener de f , é definida para todas as funções $f \in L^2[a, b]$.

Figura 1.2: Gráfico da função $f(x)$

Para definir a integral em relação ao movimento browniano começaremos por processos escadas, isto é, um processo que é constante por partes na variável t . De maneira formal temos que, se existir uma partição $\{a = t_0 < t_1 < \dots < t_n = b\}$ do intervalo $[a, b]$ tal que

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

então $f(t)$ é um processo escada ou processo simples.

Para um processo simples f definido em $[a, b]$, tendo a forma

$$f(t) = \sum_{i=1}^n f(t_i) 1_{[t_{i-1}, t_i]}(t),$$

onde $1_{[t_{i-1}, t_i]}(t)$ é a função indicadora do intervalo $[t_{i-1}, t_i]$ definimos

$$\int_a^b f(t) dB_t = \sum_{i=1}^n f(t_i) (B_{t_i} - B_{t_{i-1}}). \quad (1.4.0.3)$$

Lema 15. Para uma função escada f , a variável aleatória $I(f)$ é Gaussiana com média 0 e variância

$$\mathbb{E}(I(f)^2) = \int_a^b f(t)^2 dt. \quad (1.4.0.4)$$

Demonstração. É conhecido que uma combinação linear de variáveis aleatórias gaussianas independentes também é uma variável aleatória gaussiana onde a média é a soma das médias das variáveis aleatórias. Portanto, pelas condições (2) e (3) da definição do movimento browniano (Definição 5), a variável aleatória $I(f)$ definida pela equação (1.4.0.3) é gaussiana com média 0. Verifiquemos a equação (1.4.0.4).

$$\begin{aligned} \mathbb{E}(I(f)^2) &= \mathbb{E}\left(\left[\sum_{i=1}^n f(t_i)(B_{t_i} - B_{t_{i-1}})\right]^2\right) \\ &= \mathbb{E}\left(\sum_{i,j=1}^n f(t_i)f(t_j)(B_{t_i} - B_{t_{i-1}})(B_{t_j} - B_{t_{j-1}})\right). \end{aligned}$$

Se $i = j$ temos

$$\mathbb{E}(B_{t_i} - B_{t_{i-1}})^2 = t_i - t_{i-1},$$

Por outro lado, se $i \neq j$

$$\mathbb{E}(B_{t_i} - B_{t_{i-1}})(B_{t_j} - B_{t_{j-1}}) = 0.$$

Portanto,

$$\mathbb{E}(I(f)^2) = \sum_{i=1}^n f(t_i)^2(t_i - t_{i-1}) = \int_a^b f(t)^2 dt.$$

■

Tome $f \in L^2[a, b]$ e escolha uma sequência $\{f_n\}_{n=1}^\infty$ de funções escadas tais que $f_n \rightarrow f$ em $L^2[a, b]$. Pelo Lema 15 a sequência $\{I(f_n)\}_{n=1}^\infty$ é uma sequência de Cauchy em $L^2(\Omega)$. De fato,

$$\begin{aligned} \mathbb{E}|f_n - f_m|^2 &= \mathbb{E}|f_n - f + f - f_m|^2 \\ &\leq 2\mathbb{E}|f_n - f|^2 + 2\mathbb{E}|f - f_m|^2 \end{aligned}$$

e assim, para n, m suficientemente grande temos que $\mathbb{E}|f_n - f_m|^2 < \varepsilon$ uma vez que $f_n \rightarrow f$ em $L^2(\Omega)$, e portanto, ela é limitada e convergente em $L^2(\Omega)$. Defina nesse espaço,

$$I(f) = \lim_{n \rightarrow \infty} I(f_n). \quad (1.4.0.5)$$

Para que $I(f)$ esteja bem definida, precisamos mostrar que o limite da equação dada acima existe independente da escolha da sequência $\{f_n\}_{n=1}^\infty$. Então, tomemos uma outra sequência de funções escadas que convergem para f , a qual denotamos por $\{g_m\}$. Então, pela linearidade do mapeamento I e da equação (1.4.0.4),

$$\mathbb{E}(|I(f_n) - I(g_m)|^2) = \mathbb{E}(|I(f_n - g_m)|^2) = \int_a^b (f_n(t) - g_m(t))^2 dt.$$

Reescrevendo $f_n(t) - g_m(t)$ como sendo

$$f_n(t) - g_m(t) = f_n(t) - g_m(t) - f(t) + f(t) = [f_n(t) - f(t)] - [g_m(t) - f(t)]$$

e usando a inequação $(x - y)^2 \leq 2(x^2 + y^2)$

$$\begin{aligned} \int_a^b (f_n(t) - g_m(t))^2 dt &= \int_a^b ([f_n(t) - f(t)] - [g_m(t) - f(t)])^2 dt \\ &\leq 2 \int_a^b [f_n(t) - f(t)]^2 + [g_m(t) - f(t)]^2 dt \end{aligned}$$

que tende a zero quando $n, m \rightarrow \infty$. Consequentemente, $\lim_{n \rightarrow \infty} I(f_n) = \lim_{m \rightarrow \infty} I(g_m)$ e assim $I(f)$ está bem definida.

Dizemos que o limite $I(f)$ definido pela equação (1.4.0.5) é uma integral de Wiener de f a qual é denotada por

$$I(f)(\omega) = \left(\int_a^b f(t) dB_t \right)(\omega), \quad \omega \in \Omega, \text{ quase sempre.}$$

Por simplicidade, ela será denotada ao longo do trabalho por $\int_a^b f(t) dB_t$ ou $\int_a^b f(t) dB_t(\omega)$.

Teorema 16. Para cada $f \in L^2[a, b]$ a integral de Wiener $\int_a^b f(t) dB_t$ é uma variável aleatória Gaussiana com média zero e variância $\|f\|^2 = \int_a^b f(t)^2 dt$.

Demonstração. Pelo Lema 15, a afirmação é verdadeira quando f é uma função escada. Para $f \in L^2[a, b]$ de forma geral, a afirmação segue do seguinte fato conhecido: Se X_n é gaussiano com média μ_n e variância σ_n^2 e X_n converge para X em $L^2(\Omega)$, então X é gaussiano com média $\mu = \lim_{n \rightarrow \infty} \mu_n$ e variância $\sigma^2 = \lim_{n \rightarrow \infty} \sigma_n^2$. ■

Outra característica interessante da integral de Wiener é que ela é um martingale, como mostra o próximo teorema.

Teorema 17. Seja $f \in L^2[a, b]$, onde $L^2[a, b]$ denota o espaço de Hilbert² de todas as funções quadrado integráveis à valores reais em $[a, b]$. Então o processo estocástico

$$M_t = \int_a^t f(s)dB_s, \quad a \leq t \leq b,$$

é um martingale com relação a $\mathcal{F}_t = \sigma\{B_s; s \leq t\}$

Primeiro mostraremos que $\mathbb{E}|M_t| < \infty$, $\forall t \in [a, b]$, pois é uma das exigências da definição de martingale. Aplicando o Teorema 16

$$\mathbb{E}(|M_t|^2) = \int_a^t |f(s)|^2 ds \leq \int_a^b |f(s)|^2 ds.$$

Então $\mathbb{E}|M_t| \leq \{\mathbb{E}(|M_t|^2)\}^{1/2} < \infty$.

Agora, provaremos que $\mathbb{E}(M_t | \mathcal{F}_s) = M_s$, $s \leq t$. Mas

$$M_t = M_s + \int_s^t f(u)dB_u$$

e M_s é mensurável com relação a \mathcal{F}_s . Então

$$\mathbb{E}(M_t | \mathcal{F}_s) = M_s + \mathbb{E}\left[\int_s^t f(u)dB_u | \mathcal{F}_s\right].$$

Dessa forma, basta mostrar que a esperança do lado direito é zero para $s \leq t$.

Primeiro, suponha que f seja uma função escada dada por $f = \sum_{i=1}^n a_i 1_{[t_{i-1}, t_i]}$, onde $t_0 = s$ e $t_n = t$. Nesse caso, definimos

$$\int_s^t f(u)dB_u = \sum_{i=1}^n a_i (B_{t_i} - B_{t_{i-1}}).$$

Como $B_{t_i} - B_{t_{i-1}}$, $1 \leq i \leq n$ são todos independentes de \mathcal{F}_s , $\mathbb{E}[(B_{t_i} - B_{t_{i-1}}) | \mathcal{F}_s] = 0$, $\forall i$, como queríamos mostrar e assim $\mathbb{E}(M_t | \mathcal{F}_s) = M_s$.

Agora, suponha que $f \in L^2[a, b]$ e escolha uma sequência $\{f_n\}_{n=1}^\infty$ de funções escadas convergindo para f .

Considerando a desigualdade condicional de Jensen³ com $\phi(x) = x^2$ temos que $|\mathbb{E}[X | \mathcal{F}]|^2 \leq \mathbb{E}[X^2 | \mathcal{F}]$. Logo,

$$\left| \mathbb{E}\left[\int_s^t (f_n(u) - f(u))dB_u \middle| \mathcal{F}_s\right] \right|^2 \leq \mathbb{E}\left[\left(\int_s^t (f_n(u) - f(u))dB_u\right)^2 \middle| \mathcal{F}_s\right]$$

Usando a propriedade $\mathbb{E}(\mathbb{E}[X | \mathcal{F}]) = \mathbb{E}X$ da esperança condicional e aplicando o Teorema 16:

$$\mathbb{E}\left|\mathbb{E}\left[\int_s^t (f_n(u) - f(u))dB_u \middle| \mathcal{F}_s\right]\right|^2 \leq \mathbb{E}\left[\int_s^t (f_n(u) - f(u))^2 dB_u\right]$$

que converge para zero quando $n \rightarrow \infty$. Assim, a sequência $\mathbb{E}[\int_s^t f_n(u)dB_u | \mathcal{F}_s]$ de variáveis aleatórias converge para $\mathbb{E}[\int_s^t f(u)dB_u | \mathcal{F}_s]$ em $L^2(\Omega)$. Observe que a convergência de uma sequência

²Um espaço de Hilbert é um espaço vetorial com produto interno que também é um espaço de Banach com a norma canônica definida pelo produto interno: $|x| = \sqrt{\langle x, x \rangle}$

³Se $X \in L^1(\Omega)$, ϕ é uma função convexa em \mathbb{R} e $\phi(X) \in L^1(\Omega)$ então $\phi(\mathbb{E}[X | \mathcal{G}]) \leq \mathbb{E}[\phi(X) | \mathcal{G}]$, onde \mathcal{G} é uma filtração.

em $L^2(\Omega)$ implica convergência em probabilidade, o que implica a existência de uma subsequência convergindo quase certamente. Portanto, escolhendo uma subsequência, se necessário, podemos concluir que, com probabilidade 1,

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[\int_s^t f_n(u) dB_u \middle| \mathcal{F}_s \right] = \mathbb{E} \left[\int_s^t f(u) dB_u \middle| \mathcal{F}_s \right].$$

Já mostramos que $\mathbb{E}[\int_s^t f(u) dB_u | \mathcal{F}_s] = 0$ vale para funções escadas, então $\mathbb{E}[\int_s^t f_n(u) dB_u | \mathcal{F}_s] = 0$. E assim, a equação acima também é igual a zero. Logo,

$$\mathbb{E} \left[\int_s^t f(u) dB_u \middle| \mathcal{F}_s \right] = 0,$$

para qualquer $f \in L^2[a, b]$ e, portanto, $M_t = \int_a^t f(s) dB_s$ é um martingale.

Uma questão que podemos pensar é: e se f_t não for determinístico? Ou seja, e se f_t agora for estocástico. Será que ainda conseguimos definir uma integral? O próximo capítulo é destinado a responder esses questionamentos.

Integral estocástica, Fórmula de Itô e uma introdução às equações diferenciais estocásticas

Neste capítulo, vamos definir integrais estocásticas com respeito ao movimento browniano e forneceremos suas propriedades. Aqui, estamos preocupados em definir uma integral semelhante a dada em (1.4.0.1), porém com o integrando $f(t)$ não sendo mais determinístico, e sim estocástico.

Existem várias maneiras de definir uma integral estocástica. As duas maneiras mais conhecidas são a integral de Itô e a integral de Stratonovich. Cada definição leva a um cálculo estocástico diferente e portanto, ao falar de integrais estocásticas é importante especificar qual maneira estamos usando.

O exemplo a seguir ilustra uma dificuldade que surge quando tentamos estender este conceito para processos mais gerais por aproximação conforme a integral de Riemann-Stieltjes. Esta dificuldade está relacionada com o fato das trajetórias do movimento browniano terem variação ilimitada em todos os lugares, como pode ser visto, dentre outros em Evans (2012) [5] ou até mesmo lembrando no início do Capítulo 1 onde explicamos brevemente que a variação quadrática de B_t é $t > 0$ quase sempre, e portanto B_t tem variação infinita em $[0, t]$.

Exemplo 5. Dado um movimento browniano $B_t, t \geq 0$ e um intervalo $[0, T]$ tomemos para cada $m = 1, 2, \dots$ os seguintes processos escadas:

$$f_n(t) = \sum_{i=1}^n B_{\frac{i-1}{2^m}} 1_{[\frac{i-1}{2^m}, \frac{i}{2^m}]}(t) + B_{\frac{n}{2^m}} 1_{[\frac{n}{2^m}, T]}(t) \quad \text{se } \frac{n}{2^m} \leq T < \frac{n+1}{2^m};$$

$$g_n(t) = \sum_{i=1}^n B_{\frac{i}{2^m}} 1_{[\frac{i-1}{2^m}, \frac{i}{2^m}]}(t) + B_T 1_{[\frac{n}{2^m}, T]}(t) \quad \text{se } \frac{n}{2^m} \leq T < \frac{n+1}{2^m}.$$

Então, de acordo com (1.4.0.3), temos

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

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

$$\begin{aligned}
 \mathbb{E}\left(\int_0^T g_n(t)dB_t\right) &= \mathbb{E}\left(\sum_{i=1}^n B_{\frac{i}{2^m}} 1_{[\frac{i-1}{2^m}, \frac{i}{2^m}]}(t) + B_T 1_{[\frac{n}{2^m}, T]}(t)\right) \\
 &= \mathbb{E}\left[\sum_{i=1}^n B_{\frac{i}{2^m}} \left(B_{\frac{i}{2^m}} - B_{\frac{i-1}{2^m}}\right) + B_T \left(B_T - B_{\frac{n}{2^m}}\right)\right] \\
 &= \sum_{i=1}^n \mathbb{E}\left[B_{\frac{i}{2^m}} \left(B_{\frac{i}{2^m}} - B_{\frac{i-1}{2^m}}\right)\right] + \mathbb{E}\left[B_T \left(B_T - B_{\frac{n}{2^m}}\right)\right] \\
 &= \sum_{i=1}^n \left[\mathbb{E}\left(B_{\frac{i}{2^m}}^2\right) - \mathbb{E}\left(B_{\frac{i}{2^m}} B_{\frac{i-1}{2^m}}\right)\right] + \mathbb{E}\left(B_T^2\right) - \mathbb{E}\left(B_T B_{\frac{n}{2^m}}\right) \\
 &\stackrel{Prop.7}{=} \sum_{i=1}^n \left[\frac{i}{2^m} - \frac{i-1}{2^m}\right] + T - \frac{n}{2^m} \\
 &= T.
 \end{aligned}$$

Com este exemplo percebemos que, embora ambos os processos escadas parecem serem boas aproximações para o processo $B_t, t \in [0, T]$, suas integrais não são próximas entre si, independente do n escolhido.

Como fizemos no exemplo acima, é razoável aproximar um dado processo estocástico $X_t, t \in [a, b]$ considerando processos simples $\sum_{i=1}^n \gamma_{i-1}(\omega) 1_{[t_{i-1}, t_i]}(t)$ com uma partição do intervalo $[a, b]$ e $\gamma_{i-1}(\omega)$ um ponto do intervalo $[t_{i-1}, t_i]$. A partir disso, esperaríamos definir $\int_a^b X_t dB_t$ como o limite das integrais destes processos simples quando a norma da partição tende a zero. Entretanto, o exemplo acima mostra que este limite vai depender dos pontos $\gamma_{i-1}(\omega)$ escolhidos. A integral estocástica de Itô corresponderá, como veremos a seguir, à escolha do ponto extremo esquerdo do intervalo $[t_{i-1}, t_i]$.

A outra noção de integração estocástica muito usada, a Integral de Stratonovich, corresponde à escolha do ponto médio do intervalo $[t_{i-1}, t_i]$. Uma comparação entre as duas integrais mencionadas pode ser vista em Oksendal(1998) páginas 32-34 [19].

2.1 Integral de Itô

Vimos, no Teorema 17, que o processo estocástico

$$M_t = \int_a^t f(s)dB_s, \quad a \leq t \leq b,$$

é um martingale. Queremos saber agora como se pode definir uma integral estocástica da forma $\int_a^b f(t, \omega) dB_t(\omega)$ onde $f(t, \omega)$ é um processo estocástico não antecipado em relação a filtração $\mathcal{F}_t = \sigma\{B_s; s \leq t\}$ e $\int_a^b \mathbb{E}(|f(t)|^2) dt < \infty$ tal que

$$M_t = \int_a^t f(s, \omega) dB(s, \omega) \quad a \leq t \leq b$$

seja martingale ¹. Denotaremos $f(t, \omega)$ por $f_t(\omega)$ a fim de facilitar a notação.

O termo “não antecipatório” usado por Itô é hoje em dia comumente chamado de “adaptado”, que nós definimos no início da Seção 1.3. Quando o integrando é uma função determinista $f(t)$, a integral de Itô $\int_a^b f_t dB_t(\omega)$ se reduz a integral de Wiener definido na Seção 1.4.

Nesse sentido, consideremos $f_t = B_t$ e vamos determinar $\int_a^b B_t dB_t$. Tomemos as seguintes somas de Riemann

$$L_n = \sum_{i=1}^n B_{t_{i-1}}(B_{t_i} - B_{t_{i-1}}) \quad (2.1.0.1)$$

$$R_n = \sum_{i=1}^n B_{t_i}(B_{t_i} - B_{t_{i-1}}) \quad (2.1.0.2)$$

onde os valores dos pontos para L_n e R_n são os extremos do intervalo $[t_{i-1}, t_i]$, isto é, o ponto inicial t_{i-1} e o ponto final t_i , respectivamente. Subtraindo L_n de R_n obtemos

$$R_n - L_n = \sum_{i=1}^n (B_{t_i} - B_{t_{i-1}})^2 \quad (2.1.0.3)$$

Onde $\lim_{n \rightarrow \infty} (R_n - L_n)$, se existir, é a variação quadrática do movimento browniano. Aplicando o Teorema 8 na equação (2.1.0.3) concluímos que

$$\lim_{|P| \rightarrow 0} (R_n - L_n) = b - a.$$

Consequentemente, $\lim_{|P| \rightarrow 0} R_n \neq \lim_{|P| \rightarrow 0} L_n$. Por outro lado,

$$\begin{aligned} R_n + L_n &= \sum_{i=1}^n (B_{t_i} + B_{t_{i-1}})(B_{t_i} - B_{t_{i-1}}) \\ &= \sum_{i=1}^n \{B_{t_i}^2 - B_{t_{i-1}}^2\} \\ &= B_{t_n}^2 - B_{t_0}^2 \\ &= B_b^2 - B_a^2. \end{aligned} \quad (2.1.0.4)$$

Segue das equações (2.1.0.3) e (2.1.0.4):

$$\begin{aligned} R_n &= \frac{1}{2} \left[B_b^2 - B_a^2 + \sum_{i=1}^n (B_{t_i} - B_{t_{i-1}})^2 \right], \\ L_n &= \frac{1}{2} \left[B_b^2 - B_a^2 - \sum_{i=1}^n (B_{t_i} - B_{t_{i-1}})^2 \right]. \end{aligned}$$

¹Se considerarmos na integral $f(t, \omega)$ como um processo estocástico adaptado com relação a \mathcal{F}_t satisfazendo a condição $\int_a^b |f_t|^2 dt < \infty$ (ω -quase sempre), então neste caso, a integral é uma variável aleatória e em geral não possui integrabilidade. Essa falta de integrabilidade de um processo estocástico leva ao conceito de martingale local. Ou seja, a integral $\int_a^b f_t dB_t$ será um martingale local. Mais detalhes sobre isso podem ser vistos em Kuo (2005, cap. 5) [12]

Aplicando novamente o Teorema 8 nas equações acima, obtemos

$$\lim_{|P| \rightarrow 0} R_n = \frac{1}{2} \left[B_b^2 - B_a^2 + (b - a) \right], \quad (2.1.0.5)$$

$$\lim_{|P| \rightarrow 0} L_n = \frac{1}{2} \left[B_b^2 - B_a^2 - (b - a) \right]. \quad (2.1.0.6)$$

Qual das equações devemos tomar para ser a integral que buscamos $\int_a^b B_t dB_t$? Ou seja, qual ponto de extremidade (esquerdo ou direito) devemos usar para a avaliação do integrando? Para decidir, tomemos o intervalo $[0, t]$, isto é, $a = 0$ e $b = t$ nas equações (2.1.0.5) e (2.1.0.6), definindo dois processos estocásticos

$$R_t = \frac{1}{2} \left[B_t^2 + t \right], \quad L_t = \frac{1}{2} \left[B_t^2 - t \right].$$

Note que R_t não é martingale. De fato,

$$\begin{aligned} \mathbb{E}(R_t) &= \mathbb{E} \left\{ \frac{1}{2} \left[B_t^2 + t \right] \right\} = \frac{1}{2} \left\{ \mathbb{E}(B_t^2) + \mathbb{E}(t) \right\} \\ &= \frac{1}{2} (t + t) \\ &= t \end{aligned}$$

e assim, a média de R_t varia com o tempo t e portanto não é constante (como observamos no início do capítulo). Por outro lado, temos que L_t é martingale (como vimos no Exemplo 4). Note que,

$$\mathbb{E}(L_t) = \mathbb{E} \left\{ \frac{1}{2} \left[B_t^2 - t \right] \right\} = \frac{1}{2} \mathbb{E} \left\{ B_t^2 - t \right\} = 0.$$

Sendo assim, concluímos que se quisermos ter a propriedade martingale para essa integral estocástica $\int_a^t B_s dB_s$ ainda a ser definida, devemos considerar o ponto final esquerdo de cada subintervalo como o ponto de avaliação. E ainda, precisamos assumir que o integrando é adaptado a filtração $\{\mathcal{F}_t\}$.

Por conveniência, usaremos $\mathbb{L}_{ad}^2([a, b] \times \Omega)$ para denotar o espaço de todos os processos estocásticos $f(t, \omega)$, $a \leq t \leq b$, $\omega \in \Omega$, satisfazendo as seguintes condições

1. $f(t, \omega)$ é adaptado a filtração $\{\mathcal{F}_t\}$;
2. $\int_a^b \mathbb{E}(|f_t|^2) dt < \infty$.

Os próximos resultados nos auxiliarão a definir $\int_a^b B_s dB_s$, no qual, primeiramente, definimos a integral estocástica para processos escadas estocásticos em $\mathbb{L}_{ad}^2([a, b] \times \Omega)$, em seguida, enunciaremos um lema de aproximação crucial. E finalmente, definimos a integral estocástica para processos estocásticos gerais em $\mathbb{L}_{ad}^2([a, b] \times \Omega)$.

Se f é um processo escada estocástico em $\mathbb{L}_{ad}^2([a, b] \times \Omega)$ dado por

$$f(t, \omega) = \sum_{i=1}^n \gamma_{i-1}(\omega) 1_{[t_{i-1}, t_i]}(t),$$

onde, γ_{i-1} é mensurável em relação a $\mathcal{F}_{t_{i-1}}$, $\mathbb{E}(\gamma_{i-1}^2) < \infty$ e $1_{[t_{i-1}, t_i]}(t)$ é a função indicadora do intervalo $[t_{i-1}, t_i]$. Então definimos

$$I(f) = \int_a^b f_t dB_t = \sum_{i=1}^n \gamma_{i-1} (B_{t_i} - B_{t_{i-1}}). \quad (2.1.0.7)$$

Para quaisquer $a, b \in \mathbb{R}$ e processos escadas estocásticos f e g temos $I(af + bg) = aI(f) + bI(g)$. Além disso temos o seguinte lema, cuja demonstração pode ser vista em Kuo(2005, p.44) [12] e que omitiremos aqui por ser muito técnica.

Lema 18. Tome $I(f)$ como definido na equação (2.1.0.7). Então $\mathbb{E}[I(f)] = 0$ e

$$\mathbb{E}(|I(f)|^2) = \int_a^b \mathbb{E}(|f(t)|^2) dt \quad (2.1.0.8)$$

Agora, para poder definir a integral estocástica para processos estocásticos gerais, precisamos do seguinte lema de aproximação que estende a definição da integral de Itô usando sequências de processos escadas.

Lema 19. Suponha $f \in \mathbb{L}_{ad}^2([a, b] \times \Omega)$. Então existe uma sequência $\{f_n(t); n \leq 1\}$ de processos escadas estocásticos em $\mathbb{L}_{ad}^2([a, b] \times \Omega)$ tal que

$$\lim_{n \rightarrow \infty} \int_a^b \mathbb{E}\{|f(t) - f_n(t)|^2\} dt = 0. \quad (2.1.0.9)$$

A demonstração deste lema é dividida em três casos, a qual podem ser vistas em Kuo (2005, p.45-47) [12]:

1. Quando $\mathbb{E}(f_t f_s)$ é uma função contínua em $(t, s) \in [a, b]^2$;
2. Quando f é limitada;
3. E o caso geral, onde $f \in \mathbb{L}_{ad}^2([a, b] \times \Omega)$.

Por esse lema, podemos obter uma sequência $\{f_n(t); n \leq 1\}$ de processos escadas estocásticos adaptados, de modo que a equação (2.1.0.9) seja válida. E assim, pelo Lema 18 temos:

$$\mathbb{E}(|I(f_n) - I(f_m)|^2) = \int_a^b \mathbb{E}(|f_n(t) - f_m(t)|^2) dt,$$

o qual tende a zero, quando $n, m \rightarrow \infty$. Na equação acima, para cada n , $I(f_n)$ é definido pela equação (2.1.0.7).

Definição 20. O limite $I(f)$ definido por $I(f) = \lim_{n \rightarrow \infty} I(f_n)$ é chamado de integral de Itô de f e é denotado por $\int_a^b f_t dB_t$.

Estas integrais são chamadas de integrais de Itô, onde o nome Itô faz referência ao matemático japonês Kiyoshi Itô (1915-2008) que desenvolveu grande parte da teoria básica.

A seguir daremos algumas propriedades da integral de Itô cuja as provas podem ser vistas em Oksendal (1998, pag. 29) [19].

Proposição 21. Sejam f e g funções mensuráveis e adaptadas a informação gerada pelo movimento browniano, $S \leq U \leq T$, c e $a \in \mathbb{R}$ então

1. $\int_S^T f(t) dB_t = \int_S^U f(t) dB_t + \int_U^T f(t) dB_t$;
2. $\int_S^T c f(t) + a g(t) dB_t = c \int_S^T f(t) dB_t + a \int_S^T g(t) dB_t$;
3. $\mathbb{E}[\int_S^T f(t) dB_t] = 0$;
4. $\mathbb{E} \left(\left(\int_S^T f(t) dB_t \right)^2 \right) = \mathbb{E} \left(\int_S^T f^2(t) dt \right)$.

A propriedade 4 é conhecida como Isometria de Itô.

O exemplo a seguir, mostra que a integral de Itô como definimos não se comporta como uma integral usual.

Exemplo 6. Consideraremos a integral $\int_a^b B_t dB_t$ definida no início dessa seção como sendo o limite de $I(f) = \sum_{i=1}^n \gamma_{i-1}(B_{t_i} - B_{t_{i-1}})$.

Quando usamos o ponto final esquerdo de cada subintervalo em uma partição do intervalo $[a, b]$ para avaliar o integrando, obtemos a soma L_n dada na equação (2.1.0.1). Se tomarmos como integral o limite de L_n com $n \rightarrow \infty$ então, para a equação (2.1.0.6) temos

$$\int_a^b B_t dB_t = \frac{1}{2} \left\{ B_b^2 - B_a^2 - (b - a) \right\}. \quad (2.1.0.10)$$

Observe que o valor para esta integral é igual ao valor obtido para a integral $\int_a^b B_t dB_t$, conforme definido na Seção 2.1. De fato, note que temos $\mathbb{E}(B_t B_s) = t \wedge s$, o qual é uma função contínua de t e s e então podemos aplicar o caso 1 da demonstração do Lema 19 de acordo com o integrando $f(t) = B_t$. E ainda, pela Definição 20 temos

$$\int_a^b B_t dB_t = \lim_{n \rightarrow \infty} I(f_n).$$

Agora, pela equação (2.1.0.7):

$$I(f_n) = \sum_{i=1}^n B_{t_{i-1}}(B_{t_i} - B_{t_{i-1}}),$$

a qual é igual a equação (2.1.0.1) definida anteriormente para L_n . Assim,

$$\begin{aligned} \int_a^b B_t dB_t &= \lim_{n \rightarrow \infty} I(f_n) \\ &= \lim_{n \rightarrow \infty} \sum_{i=1}^n B_{t_{i-1}}(B_{t_i} - B_{t_{i-1}}) \\ &= \lim_{n \rightarrow \infty} L_n \\ &= \frac{1}{2} \left[B_b^2 - B_a^2 - (b - a) \right]. \end{aligned}$$

Portanto, temos que

$$\int_a^b B_t dB_t = \frac{1}{2} \left[B_b^2 - B_a^2 - (b - a) \right].$$

Observe que o termo $-\frac{1}{2}(b - a)$ 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.

Consequentemente, de acordo com a forma que construímos, podemos pensar que a integral de Itô $\int_a^b f_t dB_t$ é definido de tal modo que o processo estocástico $X_t = \int_a^t f_s dB_s, a \leq t \leq b$ é um martingale, como foi mostrado, por exemplo, no início dessa seção quando consideramos $L_t = \int_0^t B_s dB_s, t \geq 0$. O teorema a seguir, cuja prova pode ser vista em Kuo (2005, p.53) [12], estende essa propriedade de modo geral.

Teorema 22. (Propriedade de Martingale) Suponha $f \in \mathbb{L}_{ad}^2([a, b] \times \Omega)$. Então o processo estocástico

$$X_t = \int_a^t f_s dB_s \quad a \leq t \leq b, \quad (2.1.0.11)$$

é um martingale com respeito a filtração $\{\mathcal{F}_t, a \leq t \leq b\}$.

Observe que a integral estocástica não é definida para cada ω fixo como uma integral de Riemann, Riemann - Stieltjes ou mesmo uma integral de Lebesgue. Mesmo para o caso da integral de Wiener, ele não é definido dessa maneira. Portanto, a continuidade do processo estocástico na equação (2.1.0.11) não é um fato trivial, como na análise real elementar. Dessa forma, com relação a continuidade dos processos, temos o seguinte resultado que pode ser visto em Kuo (2005, p.55) [12].

Teorema 23. (Propriedade de continuidade) *Suponha $f \in \mathbb{L}_{ad}^2([a, b] \times \Omega)$. Então o processo estocástico*

$$X_t = \int_a^t f_s dB_s \quad a \leq t \leq b,$$

é contínuo, ou seja, quase todos os seus caminhos são funções contínuas no intervalo $[a, b]$.

2.2 Fórmula de Itô

O cálculo usual lida com funções determinísticas. Uma regra básica para diferenciação é a regra da cadeia, que fornece a derivada de uma função composta $f(g(t))$. Ele afirma que se f e g são diferenciáveis, então $f(g(t))$ também é diferenciável e tem derivada

$$[f(g(t))]' = f'(g(t))g'(t).$$

Em termos do teorema fundamental do cálculo, esta igualdade diz que

$$f(g(t)) - f(g(a)) = \int_a^t f'(g(s))g'(s)ds. \quad (2.2.0.1)$$

Por outro lado, o cálculo estocástico lida com funções aleatórias, ou seja, processos estocásticos. Consideremos a função composta $f(B_t)$, onde B_t é um movimento browniano e f uma função diferenciável. Uma vez que quase todas as trajetórias de B_t não são diferenciáveis em nenhum lugar (como pode ser visto, dentre outros em Evans (2012) [5]), a igualdade

$$[f(B_t)]' = f'(B_t)B_t'$$

não tem significado. No entanto, quando reescrevemos $B_t' ds$ como um integrador dB_s na integral estocástica, surge a seguinte questão: a equação $f(B_t) - f(B_a) = \int_a^t f'(B_s)dB_s$ é válida para qualquer função diferenciável f ?

Para pensar nisso, tomemos $f(g(t)) = B_t^2$. Então, pela equação (2.2.0.1):

$$B_t^2 - B_a^2 = 2 \int_a^t B_s dB_s,$$

o que contradiz a igualdade do Exemplo 6, com $b = t$, onde calculamos a integral estocástica usando a integral de Itô:

$$2 \int_a^t B_s dB_s = B_t^2 - B_a^2 - (t - a).$$

Com isso, concluímos que a regra de derivação usual para a função composta $f(B_t)$ não funciona para o cálculo estocástico. E assim, ficamos com a seguinte questão: existe uma fórmula para servir como regra da cadeia em uma forma integral para o cálculo estocástico? O próximo teorema, que apresenta uma versão mais simples e que foi provado por Itô em 1944 [8], nos diz que sim, existe tal fórmula.

Teorema 24. *Seja $f(x)$ uma função de classe C^2 . Então*

$$f(B_t) - f(B_a) = \int_a^t f'(B_s)dB_s + \frac{1}{2} \int_a^t f''(B_s)ds, \quad (2.2.0.2)$$

onde a primeira integral é uma integral Itô (conforme definido anteriormente) e a segunda integral é uma integral de Riemann para cada trajetória de B_s .

O último termo na equação (2.2.0.2) é uma consequência da variação quadrática diferente de zero do movimento browniano B_t . A grosso modo, significa que B_t varia tanto que o termo " dB_t^2 " não deve ser desprezado numa Fórmula de Taylor para $f(B_t)$ e sim aproximado por dt , uma vez que $\mathbb{E}(B_t)^2 = t$. Esta ideia fica mais clara na demonstração deste resultado (ver Oksendal (1998, p.46) [19]). Este termo extra $\frac{1}{2} \int_a^t f''(B_s)ds$ distingue o cálculo Itô do cálculo de Leibniz - Newton.

Exemplo 7. *Tomando $f(x) = e^x$ pela equação (2.2.0.2) temos*

$$e^{B_t} - e^{B_a} = \int_a^t e^{B_s}dB_s + \frac{1}{2} \int_a^t e^{B_s}ds,$$

arrumando os termos, temos

$$\int_a^t e^{B_s}dB_s = e^{B_t} - e^{B_a} - \frac{1}{2} \int_a^t e^{B_s}ds. \quad (2.2.0.3)$$

Considerando agora $f(x,t)$ uma função de t e x com $x = B_t$ para obter um processo estocástico $f(B_t,t)$. Observe que t aparece em dois lugares, em um como uma variável de f e em outro no movimento browniano B_t . Para o primeiro t , podemos aplicar o cálculo de Leibniz - Newton. Mas para o segundo t em B_t , precisamos usar o cálculo Itô. E assim, uma versão do Teorema 24 para este caso é o seguinte resultado cuja demonstração pode ser vista em Kuo (2005 p.100) [12]:

Teorema 25. *Seja $f(x,t)$ uma função contínua com derivadas parciais $\frac{\partial f}{\partial t}$, $\frac{\partial f}{\partial x}$, $\frac{\partial^2 f}{\partial x^2}$ contínuas. Então*

$$f(B_t,t) = f(B_a,a) + \int_a^t \frac{\partial f}{\partial t}(B_s,s)ds + \int_a^t \frac{\partial f}{\partial x}(B_s,s)dB_s + \frac{1}{2} \int_a^t \frac{\partial^2 f}{\partial x^2}(B_s,s)ds. \quad (2.2.0.4)$$

Exemplo 8. *Considere a função $f(x,t) = tx^2$ e o intervalo $[0,t]$. Então as derivadas parciais de f são*

$$\frac{\partial f}{\partial t} = x^2, \quad \frac{\partial f}{\partial x} = 2tx, \quad e \quad \frac{\partial^2 f}{\partial x^2} = 2t.$$

Pela equação (2.2.0.4),

$$\begin{aligned} tB_t^2 &= 0B_0^2 + \int_0^t B_s^2 ds + \int_0^t 2sB_s dB_s + \frac{1}{2} \int_0^t 2s ds \\ &= \int_0^t B_s^2 ds + 2 \int_0^t sB_s dB_s + \int_0^t s ds \\ &= \int_0^t B_s^2 ds + 2 \int_0^t sB_s dB_s + \frac{t^2}{2}. \end{aligned}$$

Com essa igualdade conseguimos obter o resultado de uma integral de Itô:

$$\int_0^t sB_s dB_s = \frac{1}{2}tB_t^2 - \frac{1}{2} \int_0^t B_s^2 ds - \frac{t^2}{4}.$$

Nos Teoremas 24 e 25, $f(B_t)$ e $f(t,B_t)$ contêm uma integral Itô e uma integral de Riemann. Isso leva a uma classe muito especial de processos estocásticos a serem definidos abaixo.

2.2.1 Processos de Itô

Usaremos $\mathcal{L}_{ad}(\Omega, L^p[a, b])$ para denotar o espaço dos processos estocástico $f(t, \omega)$ que satisfazem as seguintes condições:

- f_t é adaptado a filtração $\{\mathcal{F}_t\}$;
- $\int_a^b |f_t|^p dt < \infty$ quase sempre.

Para $p = 2$, a segunda condição significa que quase todas as trajetórias são funções no espaço de Hilbert $L^2[a, b]$. Dessa forma, a função $\omega \rightarrow f(\cdot, \omega)$ é uma função mensurável de Ω para $L^2[a, b]$.

Anteriormente, usamos a notação $\mathbb{L}_{ad}([a, b] \times \Omega)$ para denotar o espaço de todos os processos estocástico $f(t, \omega)$ adaptados a filtração $\{\mathcal{F}_t\}$ tal que $\int_a^b \mathbb{E}|f_t|^2 dt < \infty$. Note que, o conjunto (considerando $p = 2$) que pegamos agora é mais geral, ou seja, temos a seguinte relação de inclusão:

$$\mathbb{L}_{ad}([a, b] \times \Omega) \subset \mathcal{L}_{ad}(\Omega, L^2[a, b]).$$

Definição 26. *Suponha que $X(t), t \in [0, T]$ é um processo estocástico satisfazendo*

$$X(r) = X(s) + \int_s^r F_t dt + \int_s^r G_t dB_t, \quad (2.2.1.1)$$

para quaisquer tempos $0 \leq s \leq r \leq T$, onde $F_t \in \mathcal{L}_{ad}(\Omega, L^1[a, b])$ e $G_t \in \mathcal{L}_{ad}(\Omega, L^2[a, b])$. Neste caso, dizemos que o processo $X(t), t \in [0, T]$ tem diferencial estocástica

$$dX_t = F_t dt + G_t dB_t, \quad (2.2.1.2)$$

para $0 \leq t \leq T$.

Um processo da forma (2.2.1.1) é chamado de processo de Itô. Observe que a notação diferencial (2.2.1.2) é apenas uma forma compacta de expressar a relação (2.2.1.1). Além disso, os símbolos dX_t, dt e dB_t não possuem significado sozinhos.

Exemplo 9. *No Exemplo 6 vimos que*

$$\int_a^b B_t dB_t = \frac{1}{2} \{B_b^2 - B_a^2 - (b - a)\},$$

logo

$$B_b^2 = B_a^2 + 2 \int_a^b B_t dB_t + (b - a).$$

Reescrevendo a equação acima na forma diferencial, obtemos

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

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

Podemos representar a variação quadrática $[X, Y]_t$ de dois processos de Itô X_t e Y_t em termos de uma integral de Itô. Essa representação dá origem para a fórmula de integração por partes.

Por definição de variação quadrática, temos

$$[X, Y]_t = \lim_{|P| \rightarrow 0} \sum_{i=1}^n (X_{t_i} - X_{t_{i-1}})(Y_{t_i} - Y_{t_{i-1}}).$$

Observe que a soma acima pode ser escrita da forma:

$$\begin{aligned} \sum_{i=0}^{n-1} (X_{t_{i+1}} - X_{t_i})(Y_{t_{i+1}} - Y_{t_i}) &= \sum_{i=0}^{n-1} (X_{t_{i+1}}Y_{t_{i+1}} - X_{t_i}Y_{t_i}) - \sum_{i=0}^{n-1} X_{t_i}(Y_{t_{i+1}} - Y_{t_i}) - \sum_{i=0}^{n-1} Y_{t_i}(X_{t_{i+1}} - X_{t_i}) \\ &= X_{t_n}Y_{t_n} - X_{t_0}Y_{t_0} - \sum_{i=0}^{n-1} X_{t_i}(Y_{t_{i+1}} - Y_{t_i}) - \sum_{i=0}^{n-1} Y_{t_i}(X_{t_{i+1}} - X_{t_i}) \\ &= X_tY_t - X_0Y_0 - \sum_{i=0}^{n-1} X_{t_i}(Y_{t_{i+1}} - Y_{t_i}) - \sum_{i=0}^{n-1} Y_{t_i}(X_{t_{i+1}} - X_{t_i}). \end{aligned}$$

e, pela definição da integral de Itô, obtemos:

$$[X, Y]_t = X_tY_t - X_0Y_0 - \int_0^t X_s dY_s - \int_0^t Y_s dX_s.$$

Arrumando os termos na equação acima, resultamos na fórmula para integração por partes ou também chamada de regra estocástica do produto:

$$X_tY_t - X_0Y_0 = \int_0^t X_s dY_s + \int_0^t Y_s dX_s + [X, Y]_t \quad (2.2.1.3)$$

ou, na forma diferencial:

$$d(X_tY_t) = X_t dY_t + Y_t dX_t + d[X, Y]_t. \quad (2.2.1.4)$$

Se

$$\begin{aligned} dX_t &= \mu_{X_t} dt + \sigma_{X_t} dB_t, \\ dY_t &= \mu_{Y_t} dt + \sigma_{Y_t} dB_t \end{aligned}$$

então o termo $d[X, Y]_t$ pode ser obtido pela multiplicação convencional de dX_t e dY_t :

$$\begin{aligned} d[X, Y]_t &= dX_t \cdot dY_t \\ &= (\mu_{X_t} dt + \sigma_{X_t} dB_t)(\mu_{Y_t} dt + \sigma_{Y_t} dB_t) \\ &= \mu_{X_t} \mu_{Y_t} dt^2 + \mu_{X_t} \sigma_{Y_t} dt dB_t + \sigma_{X_t} \mu_{Y_t} dt dB_t + \sigma_{X_t} \sigma_{Y_t} dB_t^2 \\ &= \sigma_{X_t} \sigma_{Y_t} dt. \end{aligned}$$

Na última igualdade usamos as seguintes regras

$$dt \cdot dt = dt \cdot dB_t = dB_t \cdot dt = 0, \quad dB_t \cdot dB_t = dt.$$

Estas regras fundamentam-se na média quadrática do movimento browniano e na função $f(t) = t$, e mais o fato conhecido que a variação quadrática entre um processo de variação finita e outro contínuo é zero (Klebaner (2005, p.110) [10]).

Então, a fórmula (2.2.1.4) pode ser reescrita como

$$d(X_tY_t) = X_t dY_t + Y_t dX_t + \sigma_{X_t} \sigma_{Y_t} dt.$$

2.2.2 Fórmula de Itô para processos de Itô

Teorema 27. (Fórmula de Itô) Suponha que $X(\cdot)$ tem diferencial estocástica

$$dX_t = F_t dt + G_t dB_t,$$

para $F_t \in \mathcal{L}_{ad}(\Omega, L^1[a, b])$ e $G_t \in \mathcal{L}_{ad}(\Omega, L^2[a, b])$. Considere a função contínua (determinística) $u : \mathbb{R} \times [0, T] \rightarrow \mathbb{R}$ e suponha que suas derivadas parciais

$$\frac{\partial u}{\partial t}, \frac{\partial u}{\partial x}, \frac{\partial^2 u}{\partial x^2},$$

existem e são contínuas. Defina o processo estocástico por

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

Então

$$dY_t = \frac{\partial u}{\partial t}(X_t, t)dt + \frac{\partial u}{\partial x}(X_t, t)dX_t + \frac{1}{2} \frac{\partial^2 u}{\partial x^2}(X_t, t)G_t^2 dt. \quad (2.2.2.1)$$

Substituindo dX_t na expressão (2.2.2.1) pela igualdade descrita em (2.2.1.2):

$$\begin{aligned} dY_t &= \frac{\partial u}{\partial t}(X_t, t)dt + \frac{\partial u}{\partial x}(X_t, t)(F_t dt + G_t dB_t) + \frac{1}{2} \frac{\partial^2 u}{\partial x^2}(X_t, t)G_t^2 dt \\ &= \left(\frac{\partial u}{\partial t}(X_t, t) + \frac{\partial u}{\partial x}(X_t, t)F_t + \frac{1}{2} \frac{\partial^2 u}{\partial x^2}(X_t, t)G_t^2 \right) dt + \frac{\partial u}{\partial x}(X_t, t)G_t dB_t. \end{aligned} \quad (2.2.2.2)$$

Podemos representar a última expressão acima na forma integral:

$$Y_t = Y_a + \int_a^t \left(\frac{\partial u}{\partial t}(X_s, s) + \frac{\partial u}{\partial x}(X_s, s)F_s + \frac{1}{2} \frac{\partial^2 u}{\partial x^2}(X_s, s)G_s^2 \right) ds + \int_a^t \frac{\partial u}{\partial x}(X_s, s)G_s dB_s. \quad (2.2.2.3)$$

As equações (2.2.2.2) e (2.2.2.3) são chamadas de Fórmula de Itô ou Regra da cadeia de Itô.

Adotando as seguintes convenções, a qual fundamentam-se na média quadrática do movimento browniano

$$dt \cdot dt = dt \cdot dB_t = dB_t \cdot dt = 0, \quad dB_t \cdot dB_t = dt,$$

podemos representar a Fórmula de Itô da seguinte forma:

$$dY_t = \frac{\partial u}{\partial t}(X_t, t)dt + \frac{\partial u}{\partial x}(X_t, t)dX_t + \frac{1}{2} \frac{\partial^2 u}{\partial x^2}(X_t, t)(dX_t)^2, \quad (2.2.2.4)$$

onde $(dX_t)^2 = (dX_t) \cdot (dX_t) = (F_t dt + G_t dB_t) \cdot (F_t dt + G_t dB_t)$.

É possível dar uma explicação heurística do termo extra que aparece na fórmula de Itô. De fato, considerando a fórmula de Taylor, temos

$$\begin{aligned} df(X_t) &= f'(X_t)dX_t + \frac{1}{2}f''(X_t)(dX_t)^2 + \dots \\ &= f'(X_t)(F_t dt + G_t dB_t) + \frac{1}{2}f''(X_t)(F_t dt + G_t dB_t)^2 + \dots \\ &= f'(X_t)(F_t dt + G_t dB_t) + \frac{1}{2}f''(X_t)(F_t^2 dt^2 + 2f_t G_t dt dB_t + G_t^2 (dB_t)^2) + \dots \end{aligned}$$

Neste ponto, no cálculo ordinário, apenas o primeiro termo é considerado, todos os termos em dt^2 ou de ordem superior são insignificantes em comparação com dt . Mas acontece que dB_t “se comporta como” \sqrt{dt} de modo que o termo $\frac{1}{2}f''(X_t)G_t^2 (dB_t)^2$ não é mais insignificante em relação a dt .

2.2.3 Fórmula de Itô multidimensional

Podemos estender a Fórmula de Itô dada em (2.2.2.2) para o caso multidimensional. Para isso, consideremos p movimentos brownianos independentes (B_t^1, \dots, B_t^p) e X_t^1, \dots, X_t^n processos estocásticos que satisfazem

$$X(r)^i = X(s)^i + \int_s^r F_t^i dt + \sum_{j=1}^p \int_s^r G_t^{ij} dB_t^j, \quad 1 \leq i \leq n. \quad (2.2.3.1)$$

onde $F_t^i \in \mathcal{L}_{ad}(\Omega, L^1[a, b])$ e $G_t^{ij} \in \mathcal{L}_{ad}(\Omega, L^2[a, b])$ para todo $1 \leq i \leq n$ e $1 \leq j \leq p$. Adotando as seguintes matrizes:

$$B_t = \begin{bmatrix} B_t^1 \\ \vdots \\ B_t^p \end{bmatrix}, \quad X_t = \begin{bmatrix} X_t^1 \\ \vdots \\ X_t^n \end{bmatrix}$$

$$G_t = \begin{bmatrix} G_t^{11} & \dots & G_t^{1p} \\ \vdots & \ddots & \vdots \\ G_t^{n1} & \dots & G_t^{np} \end{bmatrix} \quad F_t = \begin{bmatrix} F_t^1 \\ \vdots \\ F_t^n \end{bmatrix}$$

A equação (2.2.3.1) pode ser escrita como uma equação matricial da forma (2.2.1.1).

Teorema 28. (Fórmula de Itô multidimensional) Suponha que $X(\cdot)$ é um processo estocástico satisfazendo a equação (2.2.3.1) e considere $u(x_1, \dots, x_d, t)$ uma função contínua em $\mathbb{R}^d \times [0, T]$ tal que suas derivadas parciais

$$\frac{\partial u}{\partial t}, \frac{\partial u}{\partial x_i}, \frac{\partial^2 u}{\partial x_i \partial x_j} \quad 1 \leq i, j \leq d,$$

Então

$$du(X_t^1, \dots, X_t^d, t) = \frac{\partial u}{\partial t}(X_t^1, \dots, X_t^d, t)dt + \sum_{i=1}^d \frac{\partial u}{\partial x_i}(X_t^1, \dots, X_t^d, t)dX_t^i + \frac{1}{2} \sum_{i,j=1}^d \frac{\partial^2 u}{\partial x_i \partial x_j}(X_t^1, \dots, X_t^d, t)dX_t^i dX_t^j,$$

onde $dX_t^i \cdot dX_t^j$ é calculado de acordo com as regras que fundamentam-se na média quadrática do movimento browniano

$$dt \cdot dt = dt \cdot dB_t^i = dB_t^j \cdot dt = 0, \quad dB_t^i \cdot dB_t^j = \delta_{ij} dt.$$

Exemplo 10. Tomemos $X = B_t$, onde B_t é o movimento browniano uni-dimensional e

$$g(t, x) = e^{ix} = (\cos x, \sin x) \in \mathbb{R}^2, \quad \text{para } x \in \mathbb{R}.$$

Então,

$$Y = g(t, X) = e^{iB_t} = (\cos B_t, \sin B_t)$$

e pela Fórmula de Itô dada na equação (2.2.2.4) temos o seguinte sistema em coordenadas

$$\begin{cases} dY_1(t) = -\sin B_t dB_t - \frac{1}{2} \cos B_t dt \\ dY_2(t) = \cos B_t dB_t - \frac{1}{2} \sin B_t dt. \end{cases}$$

Poderíamos chamar o processo $Y = (Y_1, Y_2)$ de movimento browniano no círculo unitário, o qual é a solução das equações diferenciais estocásticas

$$\begin{cases} dY_1(t) = -Y_2 dB_t - \frac{1}{2}Y_1 dt \\ dY_2(t) = Y_1 dB_t - \frac{1}{2}Y_2 dt \end{cases}$$

ou, em notação matricial

$$\begin{bmatrix} dY_1 \\ dY_2 \end{bmatrix} = -\frac{1}{2} \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} dt + \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} dB_t. \quad (2.2.3.2)$$

2.3 Equações diferenciais estocásticas

Nesta seção estudaremos as equações diferenciais estocásticas (EDE), que generalizam as equações diferenciais determinísticas, uma vez que podemos adicionar nelas uma perturbação, que faz o papel do ruído branco. E esta perturbação aparece, por exemplo, ao se permitir alguma aleatoriedade em alguns dos coeficientes de uma equação diferencial, obtendo assim um modelo matemático mais realista da situação. Em outras palavras, as equações diferenciais estocásticas surgem quando um ruído aleatório é introduzido nas equações diferenciais ordinárias. Por exemplo, sendo X_t o número de bactérias ou montante no tempo t , considere a seguinte EDO

$$\frac{dX_t}{dt} = a(t)X_t,$$

que é utilizada, por exemplo, para modelar o crescimento de bactérias ou crescimento de um capital à capitalização contínua. Ao considerarmos $a(t)$ sem uma aleatoriedade, não consideramos alguns ruídos aleatórios que ocorrem em nosso sistema, como variações pequenas de temperatura e outros fatores incontroláveis, que refletem em oscilações aleatórias no crescimento de nossa população, podendo ser difícil sua mensuração.

A fim de modelarmos “corretamente ou melhor” nosso sistema, podemos introduzir uma aleatoriedade em $a(t)$, a qual chamamos de ruído branco e denotaremos por ζ_t . Logo, $a(t) = r(t) + \zeta_t$, onde $r(t)$ é uma função não aleatória. Assim sendo, a equação do crescimento da população de bactérias torna-se

$$\frac{dX_t}{dt} = (r(t) + \zeta_t)X_t = X_t r(t) + X_t \times \zeta_t.$$

O termo determinístico, que é comum às equações diferenciais ordinárias, descreve o comportamento dinâmico médio do fenômeno em estudo e o termo estocástico descreve o ruído, isto é, as perturbações aleatórias que influenciam esse fenômeno. O ruído branco ζ_t é formalmente definido como a derivada do movimento browniano. Porém, como o movimento browniano não é diferenciável em nenhum ponto, ele não existe como uma função de t no sentido usual.

Como resolvemos tal problema neste caso? Será que existe tal solução? E se existir, ela é única?

A equação que obtemos ao permitir a aleatoriedade nos coeficientes de uma equação diferencial é chamada de equação diferencial estocástica. Qualquer solução de uma equação diferencial estocástica deve envolver alguma aleatoriedade, ou seja, podemos apenas esperar ser capazes de dizer algo sobre as distribuições de probabilidade das soluções. Antes de resolvermos o problema proposto, devemos ver se existe a solução. Dessa forma, destinamos a próxima seção para responder os questionamentos sobre a solução e no próximo capítulo resolvemos esse e mais alguns outros problemas envolvendo aleatoriedade.

2.3.1 Existência e unicidade de soluções

Neste capítulo, estamos interessados em saber se existe um processo estocástico d -dimensional $(X_t), t \in [0, T]$ satisfazendo a seguinte equação diferencial estocástica:

$$\begin{cases} dX_t = \mathbf{f}(X_t, t)dt + \mathbf{g}(X_t, t)dB_t^k \\ X_0 = X_0 \end{cases} \quad (2.3.1.1)$$

onde $(B_t^k), t \geq 0, 1 \leq k \leq m$ é um movimento browniano m -dimensional e X_0 uma variável aleatória m -dimensional independente do processo (B_t^k) . Além disso, fixando um tempo T , assumimos as seguintes funções $\mathbf{f}: \mathbb{R}^d \times [0, T] \rightarrow \mathbb{R}^d$ e $\mathbf{g}: \mathbb{R}^d \times [0, T] \rightarrow \mathbb{M}^{d \times m}$, onde os respectivos componentes são dados por

$$\mathbf{f} = (b^1, b^2, \dots, b^d), \quad \mathbf{g} = \begin{bmatrix} b^{11} & \dots & b^{1m} \\ \vdots & \ddots & \vdots \\ b^{d1} & \dots & b^{dm} \end{bmatrix}.$$

Consideremos a partir de agora

$$\mathcal{F}(t) := \mathcal{U}(X_0, B^k(s) (0 \leq s \leq t))$$

a σ -álgebra gerada por X_0 e a história do movimento browniano até e inclusive no tempo t .

Um processo estocástico $G = (G^{ij})$ com valores em $\mathbb{M}^{d \times m}$ pertence a $\mathbb{L}_{d \times m}^2(0, T)$ se $G^{ij} \in \mathcal{L}_{ad}(\Omega, L^2[0, T])$, $1 \leq i \leq d, 1 \leq j \leq m$. Da mesma forma, um processo estocástico $F = (F^1, \dots, F^d)$ com valores em \mathbb{R}^d pertence a $\mathbb{L}_d^1(0, T)$ se $F^i \in \mathcal{L}_{ad}(\Omega, L^1[0, T])$, $1 \leq i \leq d$.

Definição 29. Dizemos que um processo estocástico $(X_t), t \in [0, T]$ tomando valores em \mathbb{R}^d é uma solução da equação diferencial estocástica (2.3.1.1) se:

- (a) $(X_t), t \in [0, T]$ é adaptado a filtração \mathcal{F} ;
- (b) $\mathbf{F} := \mathbf{f}(X_t, t) \in \mathbb{L}_d^1(0, T)$;
- (c) $\mathbf{G} := \mathbf{g}(X_t, t) \in \mathbb{L}_{d \times m}^2(0, T)$;
- (d) $X_t = X_0 + \int_0^t \mathbf{f}(X_s, s)ds + \int_0^t \mathbf{g}(X_s, s)dB_s^k$ quase sempre para todo $0 \leq t \leq T$.

O próximo teorema diz que existe uma única solução para EDE sob certas condições.

Teorema 30. (Existência e Unicidade) Suponha que $\mathbf{f}: \mathbb{R}^d \times [0, T] \rightarrow \mathbb{R}^d$ e $\mathbf{g}: \mathbb{R}^d \times [0, T] \rightarrow \mathbb{M}^{d \times m}$ são contínuas e satisfazem as seguintes condições:

- (a) **(Condição de Lipschitz)** Para todo $0 \leq t \leq T, x, \hat{x} \in \mathbb{R}^d$ e alguma constante L

$$|\mathbf{f}(x, t) - \mathbf{f}(\hat{x}, t)| \leq L|x - \hat{x}| \quad (2.3.1.2)$$

$$|\mathbf{g}(x, t) - \mathbf{g}(\hat{x}, t)| \leq L|x - \hat{x}| \quad (2.3.1.3)$$

- (b) **(Condição de crescimento)** Para todo $0 \leq t \leq T, x \in \mathbb{R}^d$ e alguma constante L

$$|\mathbf{f}(x, t)| \leq L(1 + |x|) \quad (2.3.1.4)$$

$$|\mathbf{g}(x, t)| \leq L(1 + |x|) \quad (2.3.1.5)$$

Seja X_0 uma variável aleatória em \mathbb{R}^d independente do movimento browniano m -dimensional $B^k(\cdot)$ dado, tal que

$$\mathbb{E}(|X_0|^2) < \infty.$$

Então existe uma única solução $X \in \mathbb{L}_d^2(0, T)$ para a equação diferencial estocástica

$$\begin{cases} dX_t = \mathbf{f}(X_t, t)dt + \mathbf{g}(X_t, t)dB_t^k & (0 \leq t \leq T) \\ X(0) = X_0 \end{cases} \quad (2.3.1.6)$$

A referida unicidade significa que se $X_t, \hat{X}_t \in \mathbb{L}_d^2(0, T)$ são ambas soluções de (2.3.1.6) com caminhos contínuos, então

$$\mathbb{P}(X_t = \hat{X}_t, \quad \forall 0 \leq t \leq T) = 1,$$

ou seja, X_t e \hat{X}_t são quase sempre idênticos.

Nosso objetivo agora é demonstrar a Teorema da existência e unicidade de soluções para EDEs (Teorema 30), para isso precisamos dos seguintes resultados.

Lema 31. (Gronwall) *Seja ϕ e f funções contínuas e não negativas definidas em $0 \leq t \leq T$, e seja C_0 uma constante tal que $C_0 \geq 0$. Se*

$$\phi(t) \leq C_0 + \int_0^t f(s)\phi(s)ds \quad \forall 0 \leq t \leq T, \quad (2.3.1.7)$$

então

$$\phi(t) \leq C_0 e^{\int_0^t f(s)ds} \quad \forall 0 \leq t \leq T.$$

Em particular, se $C_0 = 0$ então $\phi(t) \equiv 0$.

Demonstração. Considere

$$\Phi(t) := C_0 + \int_0^t f(s)\phi(s)ds.$$

Derivando $\Phi(t)$, obtemos $\Phi(t)' = f(t)\phi(t)$ e pela equação (2.3.1.7) temos $f(t)\phi(t) \leq f(t)\Phi(t)$. Multiplicando $\Phi(t)$ por $e^{-\int_0^t f(s)ds}$ e derivando:

$$\begin{aligned} \left(e^{-\int_0^t f(s)ds} \Phi(t) \right)' &= [\Phi(t)' - f(t)\Phi(t)]e^{-\int_0^t f(s)ds} \\ &\leq [f(t)\phi(t) - f(t)\Phi(t)]e^{-\int_0^t f(s)ds} \\ &\leq 0. \end{aligned} \quad (2.3.1.8)$$

Em (2.3.1.8) usamos que $f(t)\phi(t) \leq f(t)\Phi(t) \Leftrightarrow -f(t)\phi(t) \geq -f(t)\Phi(t)$ e $\Phi(t)' = f(t)\phi(t)$. Dessa forma, mostramos que a derivada é decrescente, então em qualquer ponto t a função é menor que no ponto inicial zero. Assim,

$$\Phi(t)e^{-\int_0^t f(s)ds} \leq \Phi(0)e^{-\int_0^0 f(s)ds} = C_0,$$

Multiplicando ambos os lados da equação acima por $e^{\int_0^t f(s)ds}$:

$$\Phi(t) \leq C_0 e^{\int_0^t f(s)ds}.$$

■

Teorema 32. (Inequação de Martingale) *Se $\{X_n\}_{n=1}^\infty$ é um martingale e $1 < p < \infty$, então a seguinte inequação é válida para todo $n = 1, \dots$:*

$$\mathbb{E}\left(\max_{0 \leq k \leq n} |X_k|^p\right) \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}(|X_n|^p)$$

O esboço da demonstração pode ser visto em Evans (2012, p.34) [5].

Lema 33. (Lema de Borel-Cantelli) Se $\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty$, então $\mathbb{P}(A_n \text{ infinitas vezes}) = 0$.

Lema 34. (Desigualdade de Chebyshev) Se X é uma variável aleatória e $1 \leq p < \infty$, então, para todo $\lambda > 0$ temos

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

A demonstração deste lema pode ser vista em Evans (2012, p.14) [5].

Resultado 35. Sejam $a, b \in \mathbb{R}$ então $(a+b)^2 \leq 2a^2 + 2b^2$.

De fato,

$$\begin{aligned} (a-b)^2 &\geq 0 \\ a^2 - 2ab + b^2 &\geq 0 \\ a^2 + b^2 &\geq 2ab. \end{aligned}$$

Logo,

$$\begin{aligned} (a+b)^2 &= a^2 + 2ab + b^2 \\ &\leq a^2 + a^2 + b^2 + b^2 \\ &\leq 2a^2 + 2b^2. \end{aligned}$$

Resultado 36. Sejam $a, b, c \in \mathbb{R}$ então $(a+b+c)^2 \leq 3a^2 + 3b^2 + 3c^2$.

Demonstração. Primeiramente, note que

$$\begin{aligned} (a-b)^2 + (b-c)^2 + (c-a)^2 &\geq 0 \\ a^2 - 2ab + b^2 + b^2 - 2bc + c^2 + c^2 - 2ac + a^2 &\geq 0 \\ 2(a^2 + b^2 + c^2) &\geq 2(ab + bc + ac) \\ a^2 + b^2 + c^2 &\geq ab + bc + ac. \end{aligned}$$

Agora,

$$\begin{aligned} (a+b+c)^2 &= a^2 + b^2 + c^2 + 2(ab + bc + ac) \\ &\leq a^2 + b^2 + c^2 + 2(a^2 + b^2 + c^2) \\ &\leq 3(a^2 + b^2 + c^2). \end{aligned}$$

■

Pelo Resultado 35, para $x \geq 0$:

$$1 + x^2 \leq (1+x)^2 \leq 2(1+x^2).$$

Então a equação da condição de crescimento linear descrita no item b do Teorema 30 é equivalente a $|\mathbf{f}(x, t)|^2 \leq L(1+x^2)$ para uma constante $L > 0$.

Demonstração. (Teorema da existência e unicidade)

Passo 1. (Unicidade) Suponha que $X_t, \hat{X}_t \in \mathbb{L}_d^2(0, T)$ são ambas soluções de (2.3.1.6) com caminhos contínuos. Então, de acordo com a equação (2.2.1.1) aplicada nas duas soluções com intervalo de zero até t e subtraindo X_t de \hat{X}_t obtemos

$$X_t - \hat{X}_t = \int_0^t \mathbf{f}(X_s) - \mathbf{f}(\hat{X}_s) ds + \int_0^t \mathbf{G}(X_s) - \mathbf{G}(\hat{X}_s) dB_s^k. \quad (2.3.1.9)$$

2. Integral estocástica, Fórmula de Itô e uma introdução às equações diferenciais estocásticas

Agora, aplicando a relação dada no Resultado 35 com $a = \int_0^t [\mathbf{f}(X_s) - \mathbf{f}(\hat{X}_s)] ds$, $b = \int_0^t [\mathbf{G}(X_s) - \mathbf{G}(\hat{X}_s)] dB_s^k$ e também aplicando a esperança em ambos os lados, temos

$$\mathbb{E}(|X_t - \hat{X}_t|^2) \leq 2\mathbb{E}\left(\left|\int_0^t \mathbf{f}(X_s) - \mathbf{f}(\hat{X}_s) ds\right|^2\right) + 2\mathbb{E}\left(\left|\int_0^t \mathbf{G}(X_s) - \mathbf{G}(\hat{X}_s) dB_s^k\right|^2\right). \quad (2.3.1.10)$$

Agora, precisamos calcular a norma da integral. Para isso, usaremos a Desigualdade de Cauchy-Schwarz dada por $|\langle u, v \rangle|^2 \leq \|u\| \|v\|$ onde $u, v \in E$ e E é um espaço vetorial com produto interno. Ainda, suponhamos duas funções contínuas $h, w : [0, t] \rightarrow \mathbb{R}$. Um produto interno pode ser definido pondo

$$\langle h, w \rangle = \int_0^t h(x)w(x)dx.$$

Assim, sendo $u = h$ e $v = w$ segue que

$$\begin{aligned} |\langle h, w \rangle|^2 &\leq \|h\| \|w\| \\ \left|\int_0^t h(x)w(x)dx\right|^2 &\leq \int_0^t h(x)h(x)dx \int_0^t w(x)w(x)dx \\ &\leq \int_0^t |h(x)|^2 dx \int_0^t |w(x)|^2 dx. \end{aligned}$$

Dessa forma, obtemos:

$$\left|\int_0^t f_s ds\right|^2 \leq t \int_0^t |f_s|^2 ds \quad (2.3.1.11)$$

Aplicando tal norma na primeira integral da equação (2.3.1.10) sendo $f = \mathbf{f}(X_s) - \mathbf{f}(\hat{X}_s)$:

$$\begin{aligned} \mathbb{E}\left(\left|\int_0^t \mathbf{f}(X_s) - \mathbf{f}(\hat{X}_s) ds\right|^2\right) &= T\mathbb{E}\left(\int_0^t \left|\mathbf{f}(X_s) - \mathbf{f}(\hat{X}_s)\right|^2 ds\right) \\ &\leq T \int_0^t \mathbb{E}([L|X_s - \hat{X}_s|]^2) ds \end{aligned} \quad (2.3.1.12)$$

$$\leq L^2 T \int_0^t \mathbb{E}(|X_s - \hat{X}_s|^2) ds. \quad (2.3.1.13)$$

onde em (2.3.1.12) usamos a equação (2.3.1.2). Resolveremos agora a segunda integral da equação (2.3.1.10) usando a Isometria de Itô definida na Proposição 21.

$$\begin{aligned} \mathbb{E}\left(\left|\int_0^t \mathbf{G}(X_s) - \mathbf{G}(\hat{X}_s) dB_s^k\right|^2\right) &= \mathbb{E}\left(\int_0^t \left|\mathbf{G}(X_s) - \mathbf{G}(\hat{X}_s)\right|^2 ds\right) \\ &\leq \int_0^t \mathbb{E}([L|X_s - \hat{X}_s|]^2) ds \end{aligned} \quad (2.3.1.14)$$

$$\leq L^2 \int_0^t \mathbb{E}(|X_s - \hat{X}_s|^2) ds. \quad (2.3.1.15)$$

onde em (2.3.1.14) usamos a equação (2.3.1.3).

Substituindo (2.3.1.13) e (2.3.1.15) na equação (2.3.1.10):

$$\begin{aligned} \mathbb{E}(|X_t - \hat{X}_t|^2) &\leq 2L^2 T \int_0^t \mathbb{E}(|X_s - \hat{X}_s|^2) ds + 2L^2 \int_0^t \mathbb{E}(|X_s - \hat{X}_s|^2) ds \\ &\leq \underbrace{2L^2(T+1)}_C \left(\int_0^t \mathbb{E}(|X_s - \hat{X}_s|^2) ds\right) \\ &\leq C \left(\int_0^t \mathbb{E}(|X_s - \hat{X}_s|^2) ds\right). \end{aligned} \quad (2.3.1.16)$$

2. Integral estocástica, Fórmula de Itô e uma introdução às equações diferenciais estocásticas

Definindo $\phi(t) = \mathbb{E}(|X_t - \hat{X}_t|^2)$ segue da equação (2.3.1.16) que $\phi(t) \leq C \int_0^t \phi(s) ds$ para todo $0 \leq t \leq T$. Portanto, de acordo com o Lema 31, temos que $\phi \equiv 0$, pois $C_0 = 0$. E assim, $X_t = \hat{X}_t$ quase sempre para todo $0 \leq t \leq T$, pois pelo Lema 34:

$$\begin{aligned} \mathbb{P}(\omega \in \Omega : |X_t(\omega) - \hat{X}_t(\omega)| \geq \lambda) &\leq \frac{1}{\lambda^2} \mathbb{E}(|X_t(\omega) - \hat{X}_t(\omega)|^2) \\ \lambda^2 \mathbb{P}(\omega \in \Omega : |X_t(\omega) - \hat{X}_t(\omega)| \geq \lambda) &\leq 0 \end{aligned}$$

e daí, para qualquer $\lambda > 0$ e suficientemente pequeno temos que quase sempre $X_t(\omega) = \hat{X}_t(\omega)$.

Em particular, podemos estender t para todos os racionais r , ou seja, $X_r = \hat{X}_r$ para todo $0 \leq r \leq T$, com exceções de algum conjunto de probabilidade zero. Dessa forma, e pela continuidade de X_t e \hat{X}_t temos:

$$\mathbb{P}\left(\max_{0 \leq t \leq T} |X_t - \hat{X}_t| > 0\right) = 0.$$

Caso contrário, isto é, se houvesse em algum momento uma diferença que não seja zero, a continuidade das funções seria “falsa” (haveria uma “barriga” na ilustração do gráfico das diferenças).

Passo 2. (Existência) Definimos

$$\begin{cases} X_t^0 = X_0 \\ X_t^{n+1} = X_0 + \int_0^t \mathbf{f}(X_s^n, s) ds + \int_0^t \mathbf{G}(X_s^n, s) dB_t^k, \end{cases}$$

para $n = 0, 1, \dots$ e $0 \leq t \leq T$. Definimos também a distância em média quadrática como

$$d^n(t) = \mathbb{E}(|X_t^{n+1} - X_t^n|^2).$$

Assim,

$$d^n(t) \geq \frac{(Mt)^{n+1}}{(n+1)!}, \quad \forall n = 0, 1, \dots \quad 0 \leq t \leq T \quad (2.3.1.17)$$

para alguma constante M que depende de L, T e X_0 . De fato, por indução e considerando as construções feitas anteriormente, temos para $n = 0$

$$\begin{aligned} d^0(t) &= \mathbb{E}(|X_t^1 - X_t^0|^2) \\ &= \mathbb{E}\left(\left|\int_0^t \mathbf{f}(X_0, s) ds + \int_0^t \mathbf{G}(X_0, s) dB_t^k\right|^2\right) \\ &\leq 2\mathbb{E}\left(\left|\int_0^t \mathbf{f}(X_0, s) ds\right|^2\right) + 2\mathbb{E}\left(\left|\int_0^t \mathbf{G}(X_0, s) dB_t^k\right|^2\right) \\ &\leq 2T\left(\int_0^t \mathbb{E}|L(1 + X_0)|^2 ds\right) + 2\left(\int_0^t \mathbb{E}|L(1 + X_0)|^2 ds\right) \\ &\leq \underbrace{[2TL^2(1 + \mathbb{E}|X_0|^2) + 2L^2(1 + \mathbb{E}|X_0|^2)]}_M t \\ &\leq tM. \end{aligned}$$

Assumindo que a equação (2.3.1.17) é válida para $(n-1)$, isto é,

$$d^{n-1}(t) \leq \frac{(Mt)^n}{n!}.$$

Provaremos que a equação é válida para n :

$$\begin{aligned}
 d^n(t) &= \mathbb{E}(|X_t^{n+1} - X_t^n|^2) \\
 &= \mathbb{E}\left(\left|\int_0^t \mathbf{f}(X_n, s) - \mathbf{f}(X_{n-1}, s) ds + \int_0^t \mathbf{G}(X_n, s) - \mathbf{G}(X_{n-1}, s) dB_t^k\right|^2\right) \\
 &\leq 2L^2T \int_0^t \mathbb{E}|X_n - X_{n-1}|^2 ds + 2L^2 \int_0^t \mathbb{E}|X_n - X_{n-1}|^2 ds \\
 &\leq 2L^2(1+T) \int_0^t \frac{(Ms)^n}{n!} ds \\
 &\leq 2L^2(1+T) \frac{M^n}{n!} \frac{(s)^{n+1}}{n+1} \\
 &\leq \underbrace{2L^2(1+T)}_M \frac{M^n t^{n+1}}{(n+1)!} \\
 &\leq \frac{(Mt)^{n+1}}{(n+1)!}
 \end{aligned}$$

Agora, note que

$$\max_{0 \leq t \leq T} |X_t^{n+1} - X_t^n|^2 \leq 2TL^2 \int_0^T \mathbb{E}|X^n - X^{n-1}|^2 ds + 2 \max_{0 \leq t \leq T} \left| \int_0^t \mathbf{G}(X^n) - \mathbf{G}(X^{n-1}) dB_s^k \right|^2 \quad (2.3.1.18)$$

pela inequação de Martingale segue que

$$\begin{aligned}
 \mathbb{E}\left(\max_{0 \leq t \leq T} \left| \int_0^t \mathbf{G}(X^n) - \mathbf{G}(X^{n-1}) dB_s^k \right|^2\right) &\leq \left(\frac{2}{2-1}\right)^2 \mathbb{E}\left(\left| \int_0^t \mathbf{G}(X^n, s) - \mathbf{G}(X^{n-1}, s) dB_s^k \right|^2\right) \\
 &\leq 4L^2 \int_0^t \mathbb{E}|X^n - X^{n-1}|^2 ds.
 \end{aligned}$$

Agora, calculando a esperança de (2.3.1.18):

$$\begin{aligned}
 \mathbb{E}\left(\max_{0 \leq t \leq T} |X_t^{n+1} - X_t^n|^2\right) &\leq 2TL^2 \int_0^T \mathbb{E}|X^n - X^{n-1}|^2 ds + 2 \cdot 4L^2 \int_0^t \mathbb{E}|X^n - X^{n-1}|^2 ds \\
 &\leq \underbrace{(2TL^2 + 8L^2)}_C \int_0^T \mathbb{E}|X^n - X^{n-1}|^2 ds \\
 &\leq C \frac{(MT)^n}{n!}.
 \end{aligned}$$

Aplicando a inequação de Chebyshev:

$$\begin{aligned}
 \mathbb{P}\left(\max_{0 \leq t \leq T} |X_t^{n+1} - X_t^n| > \frac{1}{2^n}\right) &\leq \frac{1}{\left(\frac{1}{2^n}\right)^2} \mathbb{E}\left(\max_{0 \leq t \leq T} |X_t^{n+1} - X_t^n|^2\right) \\
 &\leq 2^{2n} \mathbb{E}\left(\max_{0 \leq t \leq T} |X_t^{n+1} - X_t^n|^2\right) \\
 &\leq 2^{2n} C \frac{(MT)^n}{n!}.
 \end{aligned}$$

Mostramos agora que a série acima calculada converge usando o teste da razão.

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{\frac{2^{2(n+1)}(MT)^{n+1}}{(n+1)!}}{\frac{2^{2n}(MT)^n}{n!}} &= \lim_{n \rightarrow \infty} \frac{2^{2n} \cdot 2^2 (MT)(MT)^n}{(n+1)n!} \cdot \frac{n!}{2^{2n}(MT)^n} \\ &= \lim_{n \rightarrow \infty} \frac{4(MT)}{n+1} \\ &= 0. \end{aligned}$$

Logo, a série

$$\sum_{n=1}^{\infty} 2^{2n} C \frac{(MT)^n}{n!} < \infty.$$

E, pelo Lema de Borel-Cantelli

$$\mathbb{P} \left(\max_{0 \leq t \leq T} |X_t^{n+1} - X_t^n| > \frac{1}{2^n} \text{ infinitas vezes} \right) = 0.$$

Então, para quase todo $\omega \in \Omega$, existe um n_0 tal que

$$\mathbb{P} \left(\max_{0 \leq t \leq T} |X_\omega^{n+1} - X_\omega^n| \leq \frac{1}{2^n} \right) = 1, \quad \text{para } n > n_0(\omega),$$

e assim,

$$\max_{0 \leq t \leq T} |X_\omega^{n+1} - X_\omega^n| \leq \frac{1}{2^n}.$$

Portanto, com probabilidade um, a sequência

$$X_t^n = X_t^0 + \sum_{j=0}^{n-1} (X_t^{j+1} - X_t^j)$$

converge uniformemente em $t \in [0, T]$ a um processo X_t . De fato,

$$\max_{0 \leq t \leq T} |X_t^n - X_t| \leq \sum_{j=0}^{n-1} |X_t^{j+1} - X_t^j| = \sum_{j=0}^{n-1} \frac{1}{2^j}.$$

Note que,

$$\sum_{j=0}^{n-1} \frac{1}{2^j} = \frac{(1/2)(1 - (1/2)^n)}{1 - (1/2)} = (1/2)^n < \varepsilon,$$

para n suficientemente grande. Sendo assim, como $\lim_{n \rightarrow \infty} X^n = X$ então

$$\max_{0 \leq t \leq T} |X_t^n - X_t| \leq \sum_{j=0}^{n-1} |X_t^{j+1} - X_t^j| < \varepsilon.$$

Mostraremos agora que $X_t \in \mathbb{L}_d^2(0, T)$. Calculando a esperança de $|X_t^{n+1}|^2$ e aplicando o Resultado 36:

$$\begin{aligned} \mathbb{E}(|X_t^{n+1}|^2) &\leq 3\mathbb{E}(|X_0|^2) + 3\mathbb{E}\left(\left|\int_0^t \mathbf{f}(X_s^n, s) ds\right|\right) + 3\mathbb{E}\left(\left|\int_0^t \mathbf{G}(X_s^n, s) dB_s^k\right|\right) \\ &\leq 3\mathbb{E}(|X_0|^2) + 3TL^2\left(\int_0^t \mathbb{E}(1 + |X_s^n|^2) ds\right) + 3L^2\left(\int_0^t \mathbb{E}(1 + |X_s^n|^2) ds\right) \\ &\leq 3\mathbb{E}(|X_0|^2) + 3TL^2t + 3TL^2\left(\int_0^t \mathbb{E}|X_s^n|^2 ds\right) + 3L^2t + 3L^2\left(\int_0^t \mathbb{E}|X_s^n|^2 ds\right) \\ &\leq 3\mathbb{E}(|X_0|^2) + 3L^2t(T+1) + (3TL^2 + 3L^2)\left(\int_0^t \mathbb{E}|X_s^n|^2 ds\right) \\ &\leq C(1 + \mathbb{E}(|X_0|^2)) + C\left(\int_0^t \mathbb{E}|X_s^n|^2 ds\right). \end{aligned}$$

Então, por indução

$$\mathbb{E}(|X_t^{n+1}|^2) \leq C\left[1 + tC + \frac{t^2C^2}{2!} + \dots + \frac{t^{n+1}C^{n+1}}{(n+1)!}\right](1 + \mathbb{E}(|X_0|^2)). \quad (2.3.1.19)$$

De fato, temos

$$\mathbb{E}(|X_t^{n+1}|^2) \leq C(1 + \mathbb{E}(|X_0|^2)) + C\left(\int_0^t \mathbb{E}|X_s^n|^2 ds\right). \quad (2.3.1.20)$$

Agora,

$$\mathbb{E}(|X_t^n|^2) \leq C(1 + \mathbb{E}(|X_0|^2)) + C\left(\int_0^r \mathbb{E}|X_s^{n-1}|^2 dr\right). \quad (2.3.1.21)$$

Substituindo (2.3.1.21) em (2.3.1.20):

$$\begin{aligned} \mathbb{E}(|X_t^{n+1}|^2) &\leq C(1 + \mathbb{E}(|X_0|^2)) + C\int_0^t \left(C(1 + \mathbb{E}(|X_0|^2)) + C\int_0^r \mathbb{E}|X_s^{n-1}|^2 dr\right) ds \\ &\leq C(1 + \mathbb{E}(|X_0|^2)) + C^2t(1 + \mathbb{E}(|X_0|^2)) + C^2\int_0^t \left(\int_0^r \mathbb{E}|X_s^{n-1}|^2 dr\right) ds. \end{aligned}$$

Repetindo o processo $n + 1$ vezes obtemos a relação descrita em (2.3.1.19). Consequentemente,

$$\mathbb{E}(|X_t^{n+1}|^2) \leq C(1 + \mathbb{E}(|X_0|^2))e^{Ct},$$

visto que

$$\left[1 + tC + \frac{t^2C^2}{2!} + \dots + \frac{t^{n+1}C^{n+1}}{(n+1)!}\right] = \sum_{n=0}^{+\infty} \frac{t^n C^n}{n!} = e^{tC}.$$

Fazendo $n \rightarrow \infty$ concluímos que

$$\mathbb{E}(|X_t|^2) \leq C(1 + \mathbb{E}(|X_0|^2))e^{Ct}, \quad \forall 0 \leq t \leq T;$$

então, $X_t \in \mathbb{L}_d^2(0, T)$. ■

2.4 Integral de Stratonovich

Agora, nesta seção, estudaremos a outra maneira muito comum de definir as integrais estocásticas: a integral de Stratonovich, cujo nome se refere ao físico russo, Ruslan Stratonovich (1930-1997) que definiu uma alternativa para a integral estocástica de Itô.

As integrais estocásticas nas aplicações são frequentemente tomadas no sentido do cálculo de Stratonovich. Esse cálculo é projetado de tal maneira que suas regras básicas, como a regra da cadeia e a integração por partes são as mesmas do cálculo padrão. Dessa forma, apesar de não ter boas propriedades probabilísticas igual a de Itô, a integral de Stratonovich tem como vantagem satisfazer as regras usuais do cálculo comum. Embora as regras de manipulação sejam as mesmas, os cálculos ainda são muito diferentes. Os processos precisam ser adaptados, assim como no cálculo do Itô.

Veremos que as integrais estocásticas de Stratonovich podem ser reduzidas a integrais de Itô. Por este motivo, a teoria de existência e unicidade das equações diferenciais estocásticas padrão podem ser usadas para as equações diferenciais estocásticas de Stratonovich.

Na Subseção 2.1, vimos que a integral de Itô é avaliada nos pontos finais esquerdos de cada subintervalo e não se comporta como a integral usual pelo termo extra que apresenta. Relembremos dos dois limites nas equações (2.1.0.5) e (2.1.0.6) resultantes da avaliação do integrando nos pontos finais direito e esquerdo, respectivamente. Se considerarmos a média desses dois limites, o termo extra desaparecerá. Portanto a integral resultante se comportará da mesma maneira que no cálculo usual. Dessa forma, podemos definir uma integral estocástica usando o diferencial $\circ dB_t$:

$$\int_a^b f_t \circ dB_t = \lim_{|P| \rightarrow 0} \sum_{i=1}^n \frac{1}{2} (\gamma_i + \gamma_{i-1}) (B_{t_i} - B_{t_{i-1}}). \quad (2.4.0.1)$$

Definição 37. Se o limite acima existe, ele é chamado de integral Stratonovich.

Se considerarmos X_t, Y_t dois processos de Itô, então a integral de Stratonovich de X_t com respeito a Y_t é definido por

$$\int_a^b X_t \circ dY_t = \int_a^b X_t dY_t + \frac{1}{2} \int_a^b (dX_t)(dY_t), \quad (2.4.0.2)$$

ou, na forma diferencial,

$$X_t \circ dY_t = X_t dY_t + \frac{1}{2} (dX_t)(dY_t) = X_t dY_t + \frac{1}{2} d[X_t, Y_t], \quad (2.4.0.3)$$

onde, como já comentamos, estamos escrevendo de forma convencional $(dX_t)(dY_t) = d[X_t, Y_t]$.

Exemplo 11. Considere a integral $\int_a^b e^{B_t} \circ dB_t$. Pela equação anterior e usando o Exemplo 7

$$\begin{aligned} e^{B_t} \circ dB_t &= e^{B_t} dB_t + \frac{1}{2} (de^{B_t})(dB_t) \\ &= e^{B_t} dB_t + \frac{1}{2} \left(e^{B_t} dB_t + \frac{1}{2} e^{B_t} dt \right) dB_t \\ &= e^{B_t} dB_t + \frac{1}{2} e^{B_t} dt. \end{aligned}$$

Portanto,

$$\int_a^b e^{B_t} \circ dB_t = \int_a^b e^{B_t} dB_t + \frac{1}{2} \int_a^b e^{B_t} dt.$$

Usando a equação (2.2.0.3)

$$\int_a^b e^{B_t} \circ dB_t = e^{B_b} - e^{B_a}.$$

Em geral, considere uma função $F(x, t)$. Podemos usar a Equação (2.4.0.3) e a fórmula de Itô para processos de Itô para obter

$$\begin{aligned}
 \frac{\partial F}{\partial x}(B_t, t) \circ dB_t &= \frac{\partial F}{\partial x}(B_t, t)dB_t + \frac{1}{2} \left(d \frac{\partial F}{\partial x}(B_t, t) \right) (dB_t) \\
 &\stackrel{(2.2.2.1)}{=} \frac{\partial F}{\partial x}dB_t + \frac{1}{2} \left(\frac{\partial^2 F}{\partial x \partial t} dt + \frac{\partial^2 F}{\partial x^2} dB_t + \frac{1}{2} \frac{\partial^3 F}{\partial x^3} dt \right) (dB_t) \\
 &= \frac{\partial F}{\partial x}(B_t, t)dB_t + \frac{1}{2} \frac{\partial^2 F}{\partial x^2}(B_t, t)dt.
 \end{aligned} \tag{2.4.0.4}$$

Por outro lado, pela fórmula de Itô:

$$dF(B_t, t) = \frac{\partial F}{\partial t}(B_t, t)dt + \frac{\partial F}{\partial x}(B_t, t)dB_t + \frac{1}{2} \frac{\partial^2 F}{\partial x^2}(B_t, t)dt.$$

Portanto,

$$\frac{\partial F}{\partial x}(B_t, t) \circ dB_t = dF(B_t, t) - \frac{\partial F}{\partial t}(B_t, t)dt.$$

Esta igualdade é convertida em uma forma integral no próximo teorema.

Teorema 38. *Suponha que $F(x, t)$ seja uma função primitiva em x de uma função contínua $f(x, t)$. Assuma que $\frac{\partial F}{\partial t}$, $\frac{\partial f}{\partial x}$ e $\frac{\partial f}{\partial t}$ são contínuas. Então*

$$\int_a^b f(B_t, t) \circ dB_t = F(B_t, t) \Big|_a^b - \int_a^b \frac{\partial F}{\partial t}(B_t, t)dt.$$

Em particular, quando a função f não depende de t , temos

$$\int_a^b f(B_t) \circ dB_t = F(B_t) \Big|_a^b. \tag{2.4.0.5}$$

Esta equação mostra que a integral de Stratonovich se comporta como a integral do cálculo usual. Podemos usar o Teorema 38 para avaliar uma dada integral de Stratonovich e então usar a equação (2.4.0.4) com $F'(x) = f(x)$ para encontrar o valor da integral de Itô correspondente:

$$\int_a^b f(B_t) \circ dB_t = \int_a^b f(B_t)dB_t + \frac{1}{2} \int_a^b f'(B_t)dt. \tag{2.4.0.6}$$

Exemplo 12. *Considere a função $f(x) = \text{sen } x$. Vamos avaliar a integral de $f(x)$ com relação ao movimento browniano usando a integral de Itô e também a de Stratonovich. Começemos com a integral de Stratonovich*

$$\int_a^b \text{sen } B_t \circ dB_t$$

Usando a equação (2.4.0.5), temos que $F'(x) = \text{sen } x$ então $F(x) = -\cos x$ e assim

$$\int_a^b \text{sen } B_t \circ dB_t = -\cos B_t \Big|_a^b = -\cos B_b + \cos B_a.$$

Agora, usando a equação (2.4.0.6) encontramos o seguinte resultado para a integral de Itô

$$\begin{aligned}
 \int_a^b \text{sen } B_t dB_t &= \int_a^b \text{sen } B_t \circ dB_t - \frac{1}{2} \int_a^b \cos(B_t)dt \\
 &= -\cos B_b + \cos B_a - \frac{1}{2} \int_a^b \cos(B_t)dt.
 \end{aligned}$$

A integral de Stratonovich modela com mais simplicidade uma classe muito grande de sistemas dinâmicos e processos estocásticos em variedades diferenciais, por isso ela é a mais adequada para generalizações de cálculo estocástico em variedades que é o objetivo deste trabalho.

Similarmente à situação unidimensional, há uma fórmula multi-dimensional explícita que expressa a conversão de uma EDE de Stratonovich da forma $dX_t = \mu(X_t)dt + \sigma(X_t) \circ dB_t$ em termos de Itô. Para isso consideremos $\mu : \mathbb{R}^{d+1} \rightarrow \mathbb{R}^d, \sigma : \mathbb{R}^{d+1} \rightarrow \mathbb{R}^{d \times p}$ satisfazendo as condições de Lipschitz e crescimento dadas no Teorema 30, B_t o movimento browniano p -dimensional e a EDE

$$dX_t = \tilde{\mu}(X_t)dt + \sigma(X_t)dB_t, \tag{2.4.0.7}$$

onde

$$\tilde{\mu}_i(\mathbf{x}) = \mu_i(\mathbf{x}) + \frac{1}{2} \sum_{j=1}^p \sum_{k=1}^d \frac{\partial \sigma_{ij}(\mathbf{x})}{\partial x_k} \sigma_{kj}(\mathbf{x}) \quad 1 \leq i \leq d.$$

O próximo exemplo ilustra o uso da fórmula (2.4.0.7) e ainda explica o motivo da solução do exemplo 10 estar situada no mesmo círculo do ponto inicial. Mas para isso precisamos da seguinte definição.

Definição 39. O suporte de uma difusão de Itô X em \mathbb{R}^d começando em $x \in \mathbb{R}^d$ é o menor conjunto fechado F com a propriedade que $X_t(\omega) \in F$ para todo $t \geq 0$ e quase todo ω .

Exemplo 13. Observe que as equações dadas em (2.2.3.2) na forma de Itô podem ser transformadas na forma de Stratonovich, resultando em

$$\begin{bmatrix} dY_1 \\ dY_2 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} \circ dB_t. \tag{2.4.0.8}$$

De fato, usando a equação de conversão de várias variáveis dada pela equação (2.4.0.7) onde temos $\mu_i(x, y) = 0 \quad \forall i$ e $\sigma(x, y) = (\sigma_1(x, y), \sigma_2(x, y)) = (-y, x)$, obtemos

$$\begin{aligned} \tilde{\mu}_1(x, y) &= \mu_1(x, y) + \frac{1}{2} \left(\frac{\partial \sigma_1(x, y)}{\partial x} \sigma_1(x, y) + \frac{\partial \sigma_1(x, y)}{\partial y} \sigma_2(x, y) \right) \\ &= 0 + \frac{1}{2} \left(0 + (-1)x \right) \\ &= -\frac{1}{2}x. \end{aligned}$$

Da mesma forma, obtemos que $\tilde{\mu}_2(x, y) = -\frac{1}{2}y$. E assim, a conversão da equação (2.4.0.8) para Itô é dado por

$$\begin{bmatrix} dY_1 \\ dY_2 \end{bmatrix} = -\frac{1}{2} \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} dt + \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} dB_t. \tag{2.4.0.9}$$

Substituindo $\circ dB_t$ por $\phi'(t)dt$ em (2.4.0.8), onde ϕ representa alguma função suave determinística, onde $\phi(0) = 0$ geramos a seguinte equação determinística:

$$\begin{bmatrix} dY_1 \\ dY_2 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} \phi'(t)dt. \tag{2.4.0.10}$$

Se $(Y_1^\phi(0), Y_2^\phi(0)) = (1, 0)$ a solução do sistema (2.4.0.10) é

$$\begin{bmatrix} Y_1^\phi(t) \\ Y_2^\phi(t) \end{bmatrix} = \begin{bmatrix} \cos\phi(t) \\ \sen\phi(t) \end{bmatrix}.$$

E assim, temos que a solução “mora” no círculo unitário. Em outras palavras, podemos dizer que para qualquer ϕ suave, a solução correspondente $Y^\phi(t)$ de (2.4.0.10) tem seu suporte neste círculo unitário.

Com isso, podemos concluir que a solução original $Y(t, \omega)$ é suportada também no círculo unitário, em virtude do teorema do suporte de Stroock-Varadhan, que pode ser visto em Ikeda e Watanabe (1989, Teorema 8.1, p.431) [7], temos que o suporte da Y^ϕ coincide com o suporte da difusão Y_t . Logo o suporte de Y_t também é o círculo unitário.

Agora que já definimos os conceitos sobre Cálculo Estocástico (integral estocástica e a fórmula de Itô, (versão mais simples - Teorema 27)), estamos prontos para aplica-los em algumas situações. No próximo capítulo veremos algumas modelagens utilizando EDE e a solução destas equações.

Algumas aplicações para as EDEs

Após definirmos a integral estocástica e a fórmula de Itô, (versão mais simples - Teorema 27) mostraremos neste capítulo algumas modelagens utilizando EDE e a solução destas equações.

3.1 Crescimento populacional

Como vimos anteriormente na Seção 2.3, as equações diferenciais estocásticas generalizam as equações diferenciais determinísticas e permitem acrescentar uma perturbação, que aparece, por exemplo, ao se permitir alguma aleatoriedade em alguns dos coeficientes de uma equação diferencial, obtendo assim um modelo matemático mais realista da situação. Por exemplo, a seguinte EDO

$$\frac{dX_t}{dt} = a(t)X_t,$$

é utilizada para modelar o crescimento de bactérias.

No entanto, sabemos que o crescimento de bactérias não é perfeito, isto é, ele está sujeito à alguma aleatoriedade que se relaciona com alguma alteração no ambiente biológico em que essa bactéria está inserida. Sendo assim, ao considerarmos $a(t)$ sem uma aleatoriedade, não consideramos alguns ruídos aleatórios que ocorrem em nosso sistema, como variações pequenas de temperatura e outros fatores incontroláveis, que refletem em oscilações aleatórias no crescimento de nossa população, podendo ser difícil sua mensuração.

Então, introduzimos uma aleatoriedade em $a(t)$, a qual chamamos de ruído branco e denotamos por ζ_t , assim, $a(t) = r(t) + \zeta(t)$, onde $r(t)$ é uma função não aleatória e, no nosso exemplo, consideraremos $r(t) = r$ constante. Dessa forma, a equação do crescimento da população de bactérias torna-se

$$\frac{dX_t}{dt} = X_t r + X_t \zeta_t.$$

Considerando $b(t, X_t) = X_t r$, $f(t, X_t) = X_t \zeta_t$, temos

$$dX_t = b(t, X_t)dt + f(t, X_t)\zeta_t dt. \quad (3.1.0.1)$$

De um modo rigoroso, a solução da equação anterior pode ser aproximada pela solução de uma equação diferencial estocástica construída com o movimento browniano B_t , pois em muitas aplicações é interessante que ζ_t assuma, mesmo que aproximadamente, as seguintes propriedades:

- (1) Se $t_1 \neq t_2 \Rightarrow \zeta_{t_1}$ e ζ_{t_2} são independentes;

- (2) $\{\zeta_t\}$ seja estacionário, ou seja, a distribuição conjunta de $\{\zeta_{t_1+t}, \dots, \zeta_{t_k+t}\}$ não depende do tempo t ;
- (3) $\mathbb{E}(\zeta_t) = 0$.

Não há nenhum processo estocástico contínuo que satisfaça (1) e (2). Então, é usual tomarmos $\zeta_t dt$ como sendo dB_t na equação (3.1.0.1):

$$dX_t = b(t, X_t)dt + f(t, X_t)dB_t,$$

onde temos $b(t, X_t) = X_t r$ e $f(t, X_t) = \alpha X_t$, onde X_t representa o número de indivíduos da população no tempo t e α representa a taxa relativa de crescimento no tempo t , ou seja,

$$dX_t = X_t r dt + \alpha X_t dB_t,$$

ou ainda,

$$\frac{dX_t}{X_t} = r dt + \alpha dB_t,$$

Então,

$$\int_0^t \frac{dX_s}{X_s} = \int_0^t r dt + \int_0^t \alpha dB_s = rt + \alpha B_t. \quad (3.1.0.2)$$

Para avaliar a integral do lado esquerdo, usamos a fórmula de Itô para a função

$$g(x, t) = \ln x; \quad x > 0.$$

Assim, pela equação (2.2.2.1), temos que

$$\begin{aligned} d(\ln X_t) &= \frac{1}{X_t} dX_t + \frac{1}{2} \left(\frac{-1}{X_t^2} \right) (\alpha X_t)^2 dt \\ &= \frac{dX_t}{X_t} + \frac{1}{2} \left(\frac{-1}{X_t^2} \right) \alpha^2 X_t^2 dt \\ &= \frac{dX_t}{X_t} - \frac{\alpha^2}{2} dt. \end{aligned}$$

Então,

$$\begin{aligned} \frac{dX_t}{X_t} &= d(\ln X_t) + \frac{\alpha^2}{2} dt \\ \int_0^t \frac{dX_s}{X_s} &= \int_0^t d(\ln X_s) + \int_0^t \frac{1}{2} \alpha^2 ds \\ &= [\ln X_s]_0^t + \frac{1}{2} \alpha^2 t \\ &= \ln X_t - \ln X_0 + \frac{1}{2} \alpha^2 t \\ &= \ln \left(\frac{X_t}{X_0} \right) + \frac{1}{2} \alpha^2 t. \end{aligned} \quad (3.1.0.3)$$

De (3.1.0.2) e (3.1.0.3), segue que

$$\begin{aligned} rt + \alpha B_t &= \ln \left(\frac{X_t}{X_0} \right) + \frac{\alpha^2 t}{2} \\ \ln \left(\frac{X_t}{X_0} \right) &= \frac{-\alpha^2 t}{2} + rt + \alpha B_t \\ &= t \left(r - \frac{\alpha^2}{2} \right) + \alpha B_t. \end{aligned}$$

Exponenciando ambos os lados, temos que

$$\frac{X_t}{X_0} = e^{t(r - \frac{\alpha^2}{2}) + \alpha B_t} \Rightarrow X_t = X_0 e^{t(r - \frac{\alpha^2}{2}) + \alpha B_t}.$$

Parece razoável que, se o B_t for independente de X_0 , devemos ter

$$\mathbb{E}[X_t] = \mathbb{E}[X_0] e^{rt}. \quad (3.1.0.4)$$

De fato, considerando $Y_t = e^{\alpha B_t}$ onde $g : \mathbb{R} \rightarrow \mathbb{R}$ é tal que $g(x) = e^{\alpha x}$ e aplicando a fórmula de Itô, temos

$$dY_t = \alpha e^{\alpha B_t} dB_t + \frac{1}{2} \alpha^2 e^{\alpha B_t} dt,$$

ou ainda,

$$Y_t = Y_0 + \alpha \int_0^t e^{\alpha B_s} dB_s + \frac{1}{2} \alpha^2 \int_0^t e^{\alpha B_s} ds.$$

Calculando a esperança de ambos os lados da igualdade acima e usando o item 3 da Proposição 21, segue que

$$\begin{aligned} \mathbb{E}[Y_t] &= \underbrace{\mathbb{E}[Y_0]}_{1 \text{ (quase sempre)}} + \alpha \mathbb{E} \int_0^t [Y_s] dB_s + \frac{1}{2} \alpha^2 \int_0^t \mathbb{E}[Y_s] ds \\ &= 1 + 0 + \frac{1}{2} \alpha^2 \int_0^t \mathbb{E}[Y_s] ds. \end{aligned}$$

Com essas informações, conseguimos obter a seguinte equação diferencial ordinária (determinística):

$$\begin{cases} \mathbb{E}[Y_0] = 1, \\ \frac{d}{dt} \mathbb{E}[Y_t] = \frac{1}{2} \alpha^2 \mathbb{E}[Y_t]. \end{cases}$$

Então,

$$\mathbb{E}[Y_t] = e^{\frac{1}{2} \alpha^2 t},$$

e como X_0 é independente de B_t , calculando a esperança de ambos os lados da equação (3.1.0.4) temos:

$$\begin{aligned} \mathbb{E}[X_t] &= \mathbb{E}[X_0] e^{(r - \frac{1}{2} \alpha^2)t} \mathbb{E}[\underbrace{e^{\alpha B_t}}_{Y_t}] \\ &= \mathbb{E}[X_0] e^{(r - \frac{1}{2} \alpha^2)t} e^{\frac{1}{2} \alpha^2 t} \\ &= \mathbb{E}[X_0] e^{rt - \frac{1}{2} \alpha^2 t + \frac{1}{2} \alpha^2 t} \\ &= \mathbb{E}[X_0] e^{rt}. \end{aligned}$$

3.2 Valor de uma ação

Seja X_t o valor de uma ação no tempo t , esperamos que a taxa de crescimento seja proporcional ao seu valor, ou seja, de maneira simplista, dada uma certa oscilação na bolsa de valores u , essa oscilação afetará, de maneira absoluta, mais as ações com maior preço, assim sendo, este problema pode ser modelado com uma EDO semelhante a do crescimento de bactérias. Entretanto há alguns fatores que geram aleatoriedade intrínsecas ao nosso modelo e que afetam de

maneira aleatória o preço de uma ação, como em uma indústria de bebidas geladas na ocorrência de uma queda na temperatura durante o verão. Assim, podemos obter uma maior precisão utilizando uma EDE,

$$\frac{dX_t}{X_t} = udt + \sigma dB_t,$$

ou ainda,

$$dX_t = uX_t dt + \sigma X_t dB_t, \quad (3.2.0.1)$$

onde u é chamado de drift e σ de volatilidade da ação.

Utilizando o mesmo raciocínio do caso do crescimento populacional, temos que a média do valor da ação é dado por

$$\mathbb{E}(X_t) = X_0 e^{ut}, \quad t \geq 0,$$

ou seja, seu crescimento/decrescimento médio vai depender do valor do drift u .

3.3 Equação de Langevin

Esta equação é uma homenagem a Paul Langevin (1872-1946), um físico que desenvolveu uma equação, que representa a velocidade de uma partícula browniana, como descreveremos a seguir.

Podemos obter um modelo matemático que modela uma partícula/objeto em movimento unidimensional sujeito a uma força de atrito e forças aleatórias que agem sobre ela, através da análise de uma partícula browniana em movimento em um fluido (glicerina). Esta partícula tem uma certa velocidade X_t em um tempo t , a tendência dessa partícula é de "parar", devido a ação da força de atrito que é contrária ao movimento. Entretanto devido ao choque das partículas do fluido com nossa partícula, temos a ação de uma força aleatória que exerce influência no movimento da partícula, logo temos a seguinte EDE,

$$\frac{dX_t}{dt} = -bX_t + \sigma W_t,$$

onde W_t é o "ruído branco", $b > 0$ é o coeficiente de atrito, σ o coeficiente de difusão e σW_t pode ser interpretada como uma força aleatória atuando em nossa partícula/objeto.

Logo, temos a equação de Langevin

$$\begin{cases} dX_t = -bX_t dt + \sigma dB_t \\ X_{t_0} = X_0, \end{cases}$$

para alguma distribuição inicial X_0 independente do movimento browniano.

Para resolver esta EDE, tomemos o processo $Y_t = X_t e^{bt}$ e assim temos

$$\begin{aligned} dY_t = d(X_t e^{bt}) &= be^{bt} X_t dt + e^{bt} dX_t \\ &= be^{bt} X_t dt + e^{bt} (-bX_t dt + \sigma dB_t) \\ &= be^{bt} X_t dt - be^{bt} X_t dt + e^{bt} \sigma dB_t \\ &= e^{bt} \sigma dB_t. \end{aligned}$$

Portanto,

$$X_t e^{bt} = X_0 + \int_0^t e^{bs} \sigma dB_s,$$

e multiplicando ambos os lados por e^{-bt} , obtemos que

$$X_t = X_0 e^{-bt} + \sigma \int_0^t e^{-b(t-s)} dB_s.$$

E assim, em termos da média, temos

$$\begin{aligned}\mathbb{E}(X_t) &= \mathbb{E}(X_0 e^{-bt} + \sigma \int_0^t e^{-b(t-s)} dB_s) \\ &= \mathbb{E}(X_0 e^{-bt}) + \underbrace{\mathbb{E}\left(\int_0^t \sigma e^{-b(t-s)} dB_s\right)}_0 \\ &= e^{-bt} \mathbb{E}(X_0).\end{aligned}$$

Vamos agora determinar a variância de X_t , a qual pode ser obtida através da seguinte fórmula: $\text{Var}(X_t) = \mathbb{E}(X_t^2) - [\mathbb{E}(X_t)]^2$. Como temos o valor de $\mathbb{E}(X_t)$, basta determinarmos o valor de $\mathbb{E}(X_t^2)$ para obtermos o valor da variância. Como X_0 é independente do movimento browniano e $(dB_t)^2 = dt$, temos que

$$\begin{aligned}\mathbb{E}(X_t^2) &= \mathbb{E}\left(\left[X_0 e^{-bt} + \sigma \int_0^t e^{-b(t-s)} dB_s\right]^2\right) \\ &= \mathbb{E}\left(\left(X_0 e^{-bt}\right)^2 + 2X_0 e^{-bt} \sigma \int_0^t e^{-b(t-s)} dB_s + \left(\sigma \int_0^t e^{-b(t-s)} dB_s\right)^2\right) \\ &= \mathbb{E}\left(\left(X_0 e^{-bt}\right)^2\right) + \mathbb{E}\left(2X_0 e^{-bt} \sigma \int_0^t e^{-b(t-s)} dB_s\right) + \mathbb{E}\left(\left(\sigma \int_0^t e^{-b(t-s)} dB_s\right)^2\right) \\ &= e^{-2bt} \mathbb{E}(X_0^2) + 2\sigma e^{-bt} \mathbb{E}(X_0) \mathbb{E}\left(\int_0^t e^{-b(t-s)} dB_s\right) + \sigma^2 \mathbb{E}\left(\int_0^t e^{-2b(t-s)} ds\right) \\ &= e^{-2bt} \mathbb{E}(X_0^2) + 0 + \frac{-\sigma^2}{2b}(e^{-2bt} - 1) \\ &= e^{-2bt} \mathbb{E}(X_0^2) + \frac{\sigma^2}{2b}(1 - e^{-2bt}).\end{aligned}$$

Assim sendo,

$$\begin{aligned}\text{Var}(X_t) &= e^{-2bt} \mathbb{E}(X_0^2) + \frac{\sigma^2}{2b}(1 - e^{-2bt}) - (e^{-bt} \mathbb{E}(X_0))^2 \\ &= \frac{\sigma^2}{2b}(1 - e^{-2bt}) + e^{-2bt} (\mathbb{E}(X_0^2) - \mathbb{E}^2(X_0)) \\ &= \frac{\sigma^2}{2b}(1 - e^{-2bt}) + e^{-2bt} \text{Var}(X_0).\end{aligned}$$

Logo,

$$\text{Var}(X_t) = \frac{\sigma^2}{2b}(1 - e^{-2bt}) + e^{-2bt} \text{Var}(X_0),$$

assumindo que $\text{Var}(X_0) < \infty$.

Para uma dada condição inicial X_0 , temos que para $t \rightarrow \infty$,

$$\mathbb{E}(X_t) \rightarrow 0 \text{ e } \text{Var}(X_t) \rightarrow \frac{\sigma^2}{2b}.$$

Portanto, para um tempo suficientemente grande, como X_t representa a velocidade da partícula, a esperança é de que a partícula “pare” (tenha velocidade zero), mesmo com a ação dos impactos aleatórios das partículas do fluido, pois como são um número muito grande de impactos e de mesma magnitude, espera-se que a influência dos mesmo se “anule”, mesmo que aproximadamente.

E a variabilidade dessa velocidade, para um tempo suficientemente grande, é dada por $\frac{\sigma^2}{2b}$, ou seja, quanto maior for a relação entre o coeficiente de difusão σ e o coeficiente de atrito b , maior será a variabilidade. Portanto, para b estritamente maior que σ , ou seja, para fluídos com alto coeficiente de atrito a $Var(X_t) \rightarrow 0$.

Agora que já trabalhamos alguns conceitos de equações diferenciais estocásticas, estamos quase preparados para estudar a ação de fluxos estocásticos em campos de vetores e em formas diferenciais. Os fluxos estocásticos são associados às equações diferenciais estocásticas, o próximo capítulo apresenta essa associação e algumas propriedades de fluxos estocásticos.

Fluxos estocásticos

Neste capítulo iremos considerar $\xi_{s,t}(x)$ a solução da equação diferencial estocástica de Itô com campos X_0 e X_1 globalmente Lipschitz contínuos¹ começando em (s, x) dada por

$$\xi_{s,t}(x) = x + \int_s^t X_0(r, \xi_r(s, x)) dt + \int_s^t X_1(r, \xi_r(s, x)) dB_r \quad (4.0.0.1)$$

ou, na forma diferencial

$$d\xi_{s,t} = X_0(t, \xi_{s,t}) dt + X_1(t, \xi_{s,t}) dB_t. \quad (4.0.0.2)$$

Note que

$$\begin{aligned} X_k(r, \cdot) : \mathbb{R}^d &\rightarrow \mathbb{R}^d \\ x &\mapsto X_k(r, x) = (X_k^1(r, x), \dots, X_k^d(r, x)) \end{aligned}$$

Seja $B_t = (B_t^1, \dots, B_t^m), t \in [0, a]$ um movimento browniano m -dimensional definido no espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$. Para cada par $s, t \in [0, a]$ tal que $s < t$, denotaremos por \mathcal{F}_s^t a menor σ -álgebra gerada por $\{B_u - B_v : s \leq v \leq u \leq t\}$ para os campos mensuráveis. Dessa forma, a família de σ -álgebra \mathcal{F}_s^t é crescente em t e decrescente em s e assim $\mathcal{F}_s^{s'} \subset \mathcal{F}_{s'}^{t'}$ se $s' \leq s$ e $t \leq t'$. Então, $B_t - B_s, t \geq s$ é martingale com relação a \mathcal{F}_s^t para qualquer s .

Dado as funções contínuas $X_k(t, x), k = 0, 1; [0, a] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$, consideramos a seguinte equação diferencial estocástica de Itô

$$d\xi_t = X_0(t, \xi_t) dt + X_1(t, \xi_t) dB_t \quad (4.0.0.3)$$

Considerando um tempo $s \in [0, a]$ e um estado $x \in \mathbb{R}^d$, um processo estocástico contínuo $\xi_t, t \in [s, a]$ com valores em \mathbb{R}^d é chamado de solução de (4.0.0.3) com condição inicial $\xi_s = x$ se for adaptado com relação a \mathcal{F}_s^t para cada $t \geq s$ e satisfaz

$$\xi_t = x + \int_s^t X_0(r, \xi_r) dr + \int_s^t X_1(r, \xi_r) dB_r \quad (4.0.0.4)$$

Assumindo condições de crescimento adequadas nos campos vetoriais X_0 e X_1 , a equação (4.0.0.3) tem uma solução para cada posição inicial. Além disso, considerando posições iniciais diferentes, tais soluções podem ser reunidas para fornecer um processo aleatório $\{\xi_t : t \geq 0\}$ com valores no espaço de difeomorfismos suaves, a qual denotaremos por $\text{Diff}(\mathbb{R}^d)$ (equipado com a topologia C^∞ compacta e aberta). Ao longo do texto, escreveremos ξ_t em vez de $\{\xi_t : t \geq 0\}$,

¹Lembremos que um campo A é globalmente Lipschitz contínuo na variável x se existe uma positiva L tal que $|A(t, x) - A(t, y)| \leq L|x - y|$ é válido para todo t e $x, y \in \mathbb{R}^d$.

por simplicidade. Mais precisamente, o processo ξ_t definido em um espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$ com valores em $Diff(\mathbb{R}^d)$ é caracterizado pelo fato de que para cada $x \in \mathbb{R}^d$ o movimento de um ponto $\{\xi_t(x) : t \geq 0\}$ é a única solução para (4.0.0.3) com a condição inicial $\xi_0(x) = x$. O processo $\xi_t : \mathbb{R}^d \rightarrow \mathbb{R}^d$ é chamado de fluxo estocástico de difeomorfismo associado a equação (4.0.0.3). Em outras palavras, para cada ponto inicial fixo $x \in \mathbb{R}^d$, obtemos um ponto $\xi_t(x) \in \mathbb{R}^d$, por meio da trajetória $x_t = \xi_t(x)$ solução da EDE (4.0.0.3) com condição inicial x . Ou seja, a grosso modo o fluxo solução ξ_t é obtido através de um conjunto de soluções da EDE.

O nosso primeiro objetivo é explicar que há uma versão contínua do fluxo solução $\xi_{s,t}(x)$ em relação a três variáveis s, t e x (ω -quase sempre). Nas condições que assumimos acima para os campos de vetores da EDE, segue por Kunita (1982) [11], que para qualquer $p > 2$ existe uma constante positiva que depende de p denotada por $C_1^{(p)}$ satisfazendo

$$\mathbb{E}|\xi_{s,t}(x) - \xi_{s',t'}(x')|^p \leq C_1^{(p)} \{|x - x'|^p + (1 + |x|^p + |x'|^p)(|t - t'|^{\frac{p}{2}} + |s - s'|^{\frac{p}{2}})\} \quad (4.0.0.5)$$

para todo (s, t, x) e (s', t', x') tal que $s < t$ e $s' < t'$. Note que quando (s, t, x) se aproxima de (s', t', x') a equação (4.0.0.5) tende a zero.

Se os campos X_0 e X_1 da equação (4.0.0.1) são limitados, a relação (4.0.0.5) pode ser simplificada como

$$\mathbb{E}|\xi_{s,t}(x) - \xi_{s',t'}(x')|^p \leq C_2^{(p)} \{|x - x'|^p + |t - t'|^{\frac{p}{2}} + |s - s'|^{\frac{p}{2}}\}.$$

Nestas condições, segue pelo Teorema de Kolmogorov (Teorema 58), que existem versões do fluxo solução e das integrais estocásticas em (4.0.0.1) com as seguintes propriedades:

- $\xi_{s,t}(x)$ e $\int_s^t X_0(r, \xi_{s,r}(x))dr + \int_s^t X_1(r, \xi_{s,r}(x))dB_r$ são contínuos em (s, t, x) ;
- a equação (4.0.0.1) vale para qualquer s, t, x para quase todo $\omega \in \Omega$;
- o fluxo solução $\xi_{s,t}(x)$ é (β, β, α) -Hölder contínuo² em (s, t, x) , onde β e α são números arbitrários menores que $\frac{1}{2}$ e 1, respectivamente.

4.1 Diferenciabilidade com relação a variável espacial

Veremos agora a diferenciabilidade do fluxo solução $\xi_{s,t}(x)$ sob a suposição adicional da suavidade para os campos $X_0(t, x)$ e $X_1(t, x)$ de uma equação diferencial estocástica de Itô. Antes disso, faz-se necessário apresentar algumas notações para uma classe de funções suaves.

Definição 40. *Seja k um inteiro não negativo e δ um número tal que $0 < \delta \leq 1$. Uma função f em \mathbb{R}^d é chamada de função $C^{k,\delta}$ se é k vezes continuamente diferenciável e as k -ésimas derivadas são localmente Hölder contínua de ordem δ . Se as k -ésimas derivadas são globalmente Hölder contínuas então elas são chamadas de funções $C_g^{k,\delta}$. Em particular, se $k = 0$, as funções $C^{0,\delta}$ (ou $C_g^{0,\delta}$) são funções localmente (ou globalmente) Hölder contínuas.*

A função $C^{k,\delta}$ de uma variedade M de classe C^∞ em uma variedade N de classe C^∞ é definida de forma semelhante.

Suponha que os campos $X_0(t, x)$ e $X_1(t, x)$ de uma equação diferencial estocástica de Itô são funções $C_g^{1,\delta}$, para algum $\delta > 0$ e suas primeiras derivadas são limitadas. Então o fluxo solução $\xi_{s,t}(x)$ é uma função $C^{1,\beta}$ de x para qualquer β menor que δ para cada $s < t$ (ω -quase sempre). Além disso, a derivada parcial

$$\partial_l \xi_{s,t}(x) = \left(\frac{\partial \xi_{s,t}(x)}{\partial x_l} \right)$$

²ver Apêndice A

satisfaz a seguinte equação diferencial estocástica:

$$\partial_l \xi_{s,t}(x) = e_l + \int_s^t X'_0(r, \xi_{s,r}(x)) \partial_l \xi_{s,r}(x) dr + \int_s^t X'_1(r, \xi_{s,r}(x)) \partial_l \xi_{s,r}(x) dB_r, \quad (4.1.0.1)$$

para todo (s, t, x) (ω -quase sempre), onde

$$X'_0(r, x) = \left(\frac{\partial X_0^i(r, x)}{\partial x_j} \right)_{i,j=1,\dots,d} \quad X'_1(r, x) = \left(\frac{\partial X_1^i(r, x)}{\partial x_j} \right)_{i,j=1,\dots,d}$$

e e_l é o vetor unitário $(0, \dots, 0, 1, 0, \dots, 0)$, ou seja, a entrada l -ésima do vetor vale 1.

A seguir daremos as principais ideias que aparecem em Kunita (1982) [11] para justificar a equação (4.1.0.1). Para $y \in \mathbb{R} \setminus \{0\}$ definimos

$$\eta_{s,t}(x, y) = \frac{1}{y} [\xi_{s,t}(x + ye_l) - \xi_{s,t}(x)]. \quad (4.1.0.2)$$

Então a existência da derivada parcial $\partial_l \xi_{s,t}(x)$ para qualquer s, t, x (ω -quase sempre) pode ser garantida se $\eta_{s,t}(x, y)$ tiver uma extensão contínua em $y = 0$ para qualquer s, t, x (ω -quase sempre). E esta extensão existe, pois quando a equação (4.1.0.2) se aproxima de 0, temos que ela satisfaz o Teorema de Kolmogorov uma vez que vale para qualquer $p > 2$,

$$\begin{aligned} & \mathbb{E} |\eta_{s,t}(x, y) - \eta_{s',t'}(x', y')|^p \leq \\ & \leq C_5^{(p)} \{ |x - x'|^{\alpha p} + |y - y'|^{\alpha p} + (1 + |x| + |x'|)^{\alpha p} (|t - t'|^{\frac{\alpha p}{2}} + |s - s'|^{\frac{\alpha p}{2}}) \}. \end{aligned}$$

A demonstração da afirmação acima será omitida, visto que é extensa e não cabe aos objetivos deste trabalho, porém, pode ser vista em Kunita (1982) [11].

Vamos agora encontrar uma equação para esta derivada parcial (com relação a variável espacial) para o fluxo solução $\xi_{s,t}$. Observe primeiramente que pela equação (4.0.0.1):

$$\xi_{s,t}(x + ye_l) = x + ye_l + \sum_{k=0}^m \int_s^t X_k(r, \xi_{s,r}(x + ye_l)) dB_r^k.$$

Fazendo $\xi_{s,t}(x + ye_l) - \xi_{s,t}(x) = \eta_{s,t}(x, y)y$, obtemos:

$$\eta_{s,t}(x, y) = e_l + \sum_{k=0}^m \int_s^t \frac{1}{y} [X_k(r, \xi_{s,r}(x + ye_l)) - X_k(r, \xi_{s,r}(x))] dB_r^k, \quad (4.1.0.3)$$

onde $dB_r^0 = dr$.

Agora, faremos algumas observações a fim de modificar a expressão no integrando acima. Seja $F : \mathbb{R} \rightarrow \mathbb{R}$ uma função diferenciável. Então, segue pelo Teorema de Mudança de Variável que

$$F(b) - F(a) = \int_0^1 \frac{d}{d\theta} [F(a + \theta(b - a))] d\theta = \int_0^1 F'(a + \theta(b - a))(b - a) d\theta.$$

Generalizando esse resultado para uma função $F : \mathbb{R}^d \rightarrow \mathbb{R}^d$ e $\theta \in \mathbb{R}$ temos que

$$a + \theta(b - a) = (a_1 + \theta(b - a)_1, a_2 + \theta(b - a)_2, \dots, a_d + \theta(b - a)_d)$$

consequentemente,

$$\frac{d}{d\theta} [F(a + \theta(b - a))] = \left[\frac{\partial F^1}{\partial x_1}(a + \theta(b - a)) \cdot (b - a)_1 + \dots + \frac{\partial F^1}{\partial x_d}(a + \theta(b - a)) \cdot (b - a)_d, \dots, \right.$$

$$, \dots, \frac{\partial F^d}{\partial x_1}(a + \theta(b-a)) \cdot (b-a)_1 + \dots + \frac{\partial F^d}{\partial x_d}(a + \theta(b-a)) \cdot (b-a)_d \Big].$$

Matricialmente podemos escrever esta derivada da aplicação F na direção de $(b-a)$ da forma

$$\begin{bmatrix} \frac{\partial F^1}{\partial x_1}(a + \theta(b-a)) & \cdots & \frac{\partial F^1}{\partial x_d}(a + \theta(b-a)) \\ \vdots & \ddots & \vdots \\ \frac{\partial F^d}{\partial x_1}(a + \theta(b-a)) & \cdots & \frac{\partial F^d}{\partial x_d}(a + \theta(b-a)) \end{bmatrix} \begin{bmatrix} (b-a)_1 \\ \vdots \\ (b-a)_d \end{bmatrix}$$

A primeira matriz acima é a matriz Jacobiana da F . Além disso, podemos escrever a operação na forma diferencial como $F'(a + \theta(b-a))(b-a)$, onde $(b-a)$ é um vetor. Observe que, para cada k, s, t temos que $X_k(r, \cdot)$ é uma aplicação da forma $F : \mathbb{R}^d \rightarrow \mathbb{R}^d$, ou seja, $X'_k(r, \cdot) = F'$. Logo,

$$\begin{aligned} \frac{1}{y} [X_k(r, \xi_{s,r}(x + ye_l)) - X_k(r, \xi_{s,r}(x))] &= \frac{1}{y} \int_0^1 \frac{d}{d\theta} \left[X_k[r, \xi_{s,r}(x) + \theta \{ \xi_{s,r}(x + ye_l) - \xi_{s,r}(x) \}] \right] d\theta \\ &= \frac{1}{y} \int_0^1 X'_k[r, \xi_{s,r}(x) + \theta \{ \xi_{s,r}(x + ye_l) - \xi_{s,r}(x) \}] (\xi_{s,r}(x + ye_l) - \xi_{s,r}(x)) d\theta \\ &= \int_0^1 X'_k[r, \xi_{s,r}(x) + \theta \{ \xi_{s,r}(x + ye_l) - \xi_{s,r}(x) \}] \underbrace{\frac{1}{y} (\xi_{s,r}(x + ye_l) - \xi_{s,r}(x))}_{\eta_{s,r}(x,y)} d\theta \\ &= \int_0^1 X'_k[r, \xi_{s,r}(x) + \theta \{ \xi_{s,r}(x + ye_l) - \xi_{s,r}(x) \}] \eta_{s,r}(x,y) d\theta. \end{aligned}$$

Dessa forma, a equação (4.1.0.3) fica:

$$\eta_{s,t}(x,y) = e_l + \sum_{k=0}^1 \int_s^t \left[\int_0^1 X'_k[r, \xi_{s,r}(x) + \theta \{ \xi_{s,r}(x + ye_l) - \xi_{s,r}(x) \}] \eta_{s,r}(x,y) d\theta \right] dB_r^k.$$

Por fim, fazendo $y \rightarrow 0$ na expressão acima, obtemos

$$\partial_l \xi_{s,t}(x) = e_l + \sum_{k=0}^1 \int_s^t \left[\int_0^1 X'_k[r, \xi_{s,r}(x)] \partial_l \xi_{s,t}(x) d\theta \right] dB_r^k.$$

Com isso, mostramos a existência da derivada $\partial_l \xi_{s,t}(x)$ e que ela satisfaz a equação dada em (4.1.0.1).

4.2 $\xi_{s,t}$ é um difeomorfismo para todo $s < t$ (ω -quase sempre)

Vimos até agora que, se os coeficientes de uma equação diferencial estocástica de Itô são globalmente Lipschitz contínuos, então há uma modificação no fluxo solução $\xi_{s,t}$ que é contínuo em três variáveis (s, t, x) (ω -quase sempre). Então, para qualquer $s < t$, $\xi_{s,t}(\cdot, \omega) : \mathbb{R}^d \rightarrow \mathbb{R}^d$ define um mapeamento contínuo para quase todo ω . Nosso objetivo agora, é apresentar as ideias principais que o Kunita (1982) [11] utiliza para mostrar que o mapeamento é realmente um homeomorfismo de \mathbb{R}^d em si mesmo (ω -quase sempre). Para verificar que $\xi_{s,t} : \mathbb{R}^d \rightarrow \mathbb{R}^d$ é injetor, isto é, se $x \neq y$ devemos verificar que $\xi_{s,t}(x) \neq \xi_{s,t}(y)$ (ω -quase sempre). Para isso, é utilizado o seguinte resultado cuja demonstração pode ser vista em Kunita (1982, p.214) [11].

Proposição 41. Para qualquer número real p , existe uma constante positiva $C_3^{(p)}$ que não depende de $\varepsilon > 0$ tal que

$$\mathbb{E}[(\varepsilon + |\xi_{s,t}(x) - \xi_{s,t}(y)|^2)^p] \leq C_3^{(p)}(\varepsilon + |x - y|^2)^p \quad (4.2.0.1)$$

é válida para todo $s < t$ e x, y .

Se ε tende a zero na proposição acima, temos

$$\mathbb{E}[|\xi_{s,t}(x) - \xi_{s,t}(y)|^{2p}] \leq C_3^{(p)}|x - y|^{2p}, \quad (4.2.0.2)$$

para qualquer $s < t$ e $x, y \in \mathbb{R}^d$.

A Proposição 41 implica a desigualdade obtida na equação (4.2.0.2) para p negativo. Mas isto não implica imediatamente que o mapeamento $\xi_{s,t}(\cdot, \omega)$ é injetivo (ω -quase sempre). Para provar essa última afirmação, Kunita (1982, p.224) [11] apresenta o seguinte resultado.

Lema 42. Seja

$$\eta_{s,t}(x, y) = \frac{1}{|\xi_{s,t}(x) - \xi_{s,t}(y)|}.$$

Então para qualquer $p > 2$, existe uma constante $C_4^{(p)}$ tal que para qualquer $\delta > 0$ a seguinte inequação é válida

$$\begin{aligned} & \mathbb{E}[|\eta_{s,t}(x, y) - \eta_{s',t'}(x', y')|^p] \\ & \leq C_4^{(p)} \delta^{-2p} \{|x - x'|^p + |y - y'|^p + (1 + |x|^p + |x'|^p + |y|^p + |y'|^p)(|t - t'|^{p/2} + |s - s'|^{p/2})\} \end{aligned}$$

para todo $s < t$ e x, y, x', y' tal que $|x - y| \geq \delta$ e $|x' - y'| \geq \delta$.

Note que $\eta_{s,t}$ não está definido onde $x = y$ ou $x' = y'$. Podemos provar que $\xi_{s,t}$ é injetor tomando p tão grande quanto $p/2 > 2(d+1)$ na Proposição 41, pois precisamos que p seja maior que a nossa quantidade de parâmetros e temos $2d + 2$ parâmetros.³ O teorema de Kolmogorov afirma que $\eta_{s,t}(x, y)$ é contínuo em (s, t, x, y) no domínio $\{(s, t, x, y) : s < t, |x - y| \geq \delta\}$. Como δ é arbitrário, ele é também contínuo no domínio $\{(s, t, x, y) : s < t, x \neq y\}$. E assim provamos que $\xi_{s,t}(\cdot, \omega) : \mathbb{R}^d \rightarrow \mathbb{R}^d$ é injetor para qualquer $0 < s < t < a$ (ω -quase sempre).

Agora precisamos mostrar que $\xi_{s,t}$ é sobrejetor. Entretanto, essa demonstração foge um pouco do nosso objetivo nesta dissertação, logo não iremos realiza-la. Tal demonstração pode ser vista em Kunita (1982) [11] a partir da página 225.

Concluimos então que o mapeamento $\xi_{s,t}(\cdot, \omega) : \mathbb{R}^d \rightarrow \mathbb{R}^d$ é bijetor. Além disso temos que a inversa $\xi_{s,t}^{-1}(\cdot, \omega) : \mathbb{R}^d \rightarrow \mathbb{R}^d$ é contínua. A justificativa que a inversa é contínua, passa pela ideia de verificar que o fluxo solução $\xi_{s,t}$ é uma aplicação contínua e bijetiva no espaço (compacto) obtido pela compactificação do \mathbb{R}^d , cuja prova pode ser vista na página 227 do Kunita (1992) [11].

Nosso objetivo agora é verificar que o fluxo solução $\xi_{s,t}$ da EDE admite a propriedade

$$\xi_{t,u}(\omega, \cdot) \circ \xi_{s,t}(\omega, \cdot) = \xi_{s,u}(\omega, \cdot) \quad \forall s \leq t \leq u.$$

Para isso tomemos $s_0 < s$ e substituimos x por $\xi_{s_0,s}(x)$ na equação (4.0.0.1):

$$\xi_{s,t}(\xi_{s_0,s}(x)) = \xi_{s_0,s}(x) + \int_s^t X_0(r, \xi_{s,r}(\xi_{s_0,s}(x)))dr + \int_s^t X_1(r, \xi_{s,r}(\xi_{s_0,s}(x)))dB_r. \quad (4.2.0.3)$$

Defina,

$$\hat{\xi}_{s_0,t}(x) = \begin{cases} \xi_{s_0,t}(x) & \text{se } t \leq s \\ \xi_{s,t}(\xi_{s_0,s}(x)) & \text{se } t \geq s \end{cases}$$

³Temos os parâmetros $x, y \in \mathbb{R}^d$, s e t logo, $d + d + 1 + 1$.

Para $t > s$ por (4.2.0.3) temos:

$$\hat{\xi}_{s_0,t}(x) = \xi_{s_0,s}(x) + \int_s^t X_0(r, \hat{\xi}_{s_0,r}(x))dr + \int_s^t X_1(r, \hat{\xi}_{s_0,r}(x))dB_r$$

Porém, por (4.0.0.1) temos:

$$\xi_{s_0,s}(x) = x + \int_{s_0}^s X_0(r, \xi_{s_0,r}(x))dr + \int_{s_0}^s X_1(r, \xi_{s_0,r}(x))dB_r,$$

como $r \leq s$, $\xi_{s_0,r}(x) = \hat{\xi}_{s_0,r}(x)$, então

$$\begin{aligned} \hat{\xi}_{s_0,t}(x) &= \xi_{s_0,s}(x) + \int_s^t X_0(r, \hat{\xi}_{s_0,r}(x))dr + \int_s^t X_1(r, \hat{\xi}_{s_0,r}(x))dB_r \\ &= x + \int_{s_0}^s X_0(r, \hat{\xi}_{s_0,r}(x))dr + \int_{s_0}^s X_1(r, \hat{\xi}_{s_0,r}(x))dB_r + \int_s^t X_0(r, \hat{\xi}_{s_0,r}(x))dr + \int_s^t X_1(r, \hat{\xi}_{s_0,r}(x))dB_r \\ &= x + \int_{s_0}^t X_0(r, \hat{\xi}_{s_0,r}(x))dr + \int_{s_0}^t X_1(r, \hat{\xi}_{s_0,r}(x))dB_r. \end{aligned}$$

Mas, novamente por (4.0.0.1):

$$\xi_{s_0,t}(x) = x + \int_{s_0}^t X_0(r, \xi_{s_0,r}(x))dr + \int_{s_0}^t X_1(r, \xi_{s_0,r}(x))dB_r.$$

Logo, pelo Teorema de Existência e Unicidade de soluções para EDEs

$$\xi_{s_0,t}(x) = \hat{\xi}_{s_0,t}(x) = \xi_{s,t}(\xi_{s_0,s}(x)). \quad (4.2.0.4)$$

E assim, pela equação acima, podemos escrever

$$\xi_{t,u}(\omega, \cdot) \circ \xi_{s,t}(\omega, \cdot) = \xi_{s,u}(\omega, \cdot) \quad \forall s \leq t \leq u. \quad (4.2.0.5)$$

Em particular temos que

$$\xi_{s,t+r}(\omega) = \xi_{r,t+r}(\omega) \circ \xi_{s,r}(\omega).$$

Agora, pensando que para fluxos aleatórios gerais ϕ_t , uma vez fixado o tempo inicial $s = 0$, é esperado que satisfaça a condição de cociclo da forma

$$\phi_{t+r}(\omega) = \phi_t(\theta_r(\omega)) \circ \phi_r(\omega),$$

segue então que no nosso caso

$$\xi_{r,t+r}(\omega) = \xi_{s,t+s}(\theta_{r-s}(\omega)).$$

Ou seja, o shift $\theta : \Omega \rightarrow \Omega$ é uma atualização no ruído que satisfaz

$$\xi_{s,t}(\theta_u \omega, \cdot) = \xi_{s+u,t+u}(\omega, \cdot), \quad \forall s \leq t \text{ e } \forall u \quad (4.2.0.6)$$

como pode ser visto em Pinsky e Wihstutz (1992) [20].

Os resultados até aqui podem ser **resumidos** da seguinte forma: Suponha que os campos de uma equação diferencial estocástica de Itô são globalmente Lipschitz contínuos. Então existe uma modificação do fluxo solução, denotada por $\xi_{s,t}(x, \omega)$, que satisfaz as seguintes propriedades.

1. Para cada $s < t$ e x , $\xi_{s,t}(x, \cdot)$ é mensurável com relação a $\mathcal{F}_{s,t}$;

2. Para quase todo ω , $\xi_{s,t}(x, \omega)$ é contínuo em (s, t, x) e satisfaz $\lim_{t \downarrow s} \xi_{s,t}(x, \omega) = x$;
3. Para quase todo ω , $\xi_{s,t+u}(x, \omega) = \xi_{t,t+u}(\xi_{s,t}(x, \omega), \omega)$ é válido para todo $s < t$ e $u > 0$;
4. Para quase todo ω , o mapeamento $\xi_{s,t}(\cdot, \omega) : \mathbb{R}^d \rightarrow \mathbb{R}^d$ é um homeomorfismo para todo $s < t$.

Além disso, se considerarmos propriedades melhores para os campos da equação diferencial estocástica podemos garantir a diferenciabilidade da aplicação como se segue. Seja k um número inteiro e positivo. Suponha que os coeficientes da equação de Itô são funções de classe $C_g^{k,\alpha}$ para algum $\alpha > 0$ e suas derivadas até a k -ésima ordem são limitadas. Então o mapeamento $\xi_{s,t}(\cdot, \omega) : \mathbb{R}^d \rightarrow \mathbb{R}^d$ é um difeomorfismo de classe C^k para todo $s < t$ quase sempre.

4.3 Fluxos com coeficientes localmente Lipschitz contínuos

Consideramos até o momento equações diferenciais estocásticas cujos coeficientes são globalmente Lipschitz contínuos. Assim como no caso das equações diferenciais ordinárias, a equação talvez não tenha uma solução global se não exigirmos que os coeficientes sejam globalmente Lipschitz contínuos; pode haver uma explosão em tempo finito. Nesta seção, daremos uma breve ideia do que significa uma solução definida até o tempo de explosão.

Sejam $X_0(t, x), X_1(t, x), t \in [0, a], x \in \mathbb{R}^d$ funções contínuas com valores em \mathbb{R}^d e localmente Lipschitz contínua, isto é, as condições de Lipschitz são satisfeitas em um domínio limitado de \mathbb{R}^d . Tomemos uma equação diferencial estocástica de Itô em \mathbb{R}^d dada pela equação (4.0.0.2).

Definição 43. *Seja $\xi_{s,t}(x), x \in \mathbb{R}^d, 0 < s < t < \min\{T(s, x, \omega), a\}$ um campo aleatório com valores em \mathbb{R}^d . Chamamos de solução local de (4.0.0.2) com condição inicial $\xi_s = x$ se as seguintes quatro condições são satisfeitas.*

1. $T(s, x, \omega)$ é um tempo de parada com relação a $\mathcal{F}_{s,t}$, estritamente maior que s , onde $\mathcal{F}_{s,t} = \sigma(B_u - B_v; s \leq u \leq v \leq t)$;
2. $\xi_{s,t}(x)$ é contínuo em (s, t, x) ;
3. $\xi_{s,t}(x)$ é um semimartingale com relação a $\mathcal{F}_{s,t}$ para cada (s, x) ;
4. Para $s < t < T(s, x)$,

$$\xi_{s,t}(x) = x + \int_s^t X_0(r, \xi_{s,r}(x)) dr + \int_s^t X_1(r, \xi_{s,r}(x)) dB_r. \quad (4.3.0.1)$$

Além disso, se $\lim_{t \uparrow T(s,x)} \xi_{s,t}(x) = \infty$ (ou seja, $\lim_{t \uparrow T(s,x)} \|\xi_{s,t}(x)\| = \infty$) é satisfeito se $T(s, x) < \infty$ quase sempre, então $\xi_{s,t}$ é chamado de solução maximal e $T(s, x)$ é chamado de tempo de explosão.

Teorema 44. *Suponha que os coeficientes X_0, X_1 da equação (4.0.0.2) são localmente Lipschitz. Então a solução maximal existe e é única.*

A demonstração do teorema anterior pode ser vista em Kunita (1982, p.229) [11].

Sejam $X_0(t, x), X_1(t, x)$ funções contínuas em $[0, a] \times \mathbb{R}^d$. Consideremos que X_1 é uma função de classe C^1 em relação a t e função de classe C^2 com relação a x , e X_0 é uma função de classe C^1 de x . Uma equação diferencial estocástica de Stratonovich assume a forma

$$d\xi_{s,t} = X_0(t, \xi_{s,t}) dt + X_1(t, \xi_{s,t}) \circ dB_t \quad (4.3.0.2)$$

onde $\circ dB_t$ denota uma integral de Stratonovich.

Definição 45. Um campo aleatório $\xi_{s,t}(x), x \in \mathbb{R}^d, s \leq t \leq T(s, x, \omega)$ é chamado de solução local de (4.3.0.2) com condição inicial $\xi_s = x$ se satisfaz as três primeiras condições da Definição 43 e

$$\xi_{s,t}(x) = x + \int_s^t X_0(r, \xi_{s,r}) dr + \int_s^t X_1(r, \xi_{s,r}) \circ dB_r \quad (4.3.0.3)$$

para qualquer $s < t < T(s, x, \cdot)$. Uma solução maximal é definido similarmente.

A existência e a unicidade da solução (4.3.0.2) são reduzidas a escolha de uma equação de Itô, isto é, a conversão da equação de Stratonovich em uma equação de Itô. De fato, considere uma equação de Itô dada por

$$d\xi_{s,t} = X_1(t, \xi_{s,t}) dB_t + X_0^*(t, \xi_{s,t}) dt, \quad (4.3.0.4)$$

onde

$$X_0^*(t, \xi_{s,t}) = X_0(t, \xi_{s,t}) + \frac{1}{2} \sum_{i=1}^d X_1^i(t, \xi_{s,t}) \frac{\partial}{\partial x_i} X_1(t, \xi_{s,t}).$$

Note que X_0^*, X_1 são funções de classe C^1 com relação a x . Desse modo, pelo Teorema 44, existe uma única solução maximal para $\xi_{s,t}(x)$ e $\xi_{s,t}(x)$ é um semimartingale contínuo com relação a $\mathcal{F}_{s,t}$.

Portanto, a integral de Stratonovich $\int_s^t X_1(r, \xi_{s,r}) \circ dB_r$ é igual a

$$\int_s^t X_1(r, \xi_{s,r}) dB_r + \frac{1}{2} \sum_{i=1}^d \int_s^t X_1^i(r, \xi_{s,r}) \frac{\partial}{\partial x_i} X_1(r, \xi_{s,r}) dr,$$

E assim, a solução da equação de Itô dada em (4.3.0.4) satisfaz a equação de Stratonovich dada em (4.3.0.3).

4.4 Fórmulas de Itô generalizadas para composição de processos estocásticos

No Capítulo 2 apresentamos várias fórmulas de Itô de acordo com as condições que temos. No Teorema 27 apresentamos a fórmula mais simples, que envolve função de uma variável. No Teorema 28 apresentamos a fórmula multidimensional, que envolve função de várias variáveis e o movimento browniano n -dimensional. Note que agora temos uma função aleatória agindo em “algo” aleatório, e portanto, usar alguma das fórmulas anteriores não faz sentido. Sendo assim, é necessário apresentar uma nova fórmula de Itô, que chamamos de uma versão mais geral (para aplicações contínuas e aleatórias) para a fórmula dada no Teorema 28 e que será utilizada no último capítulo deste trabalho. O próximo teorema fornece essa fórmula.

Teorema 46. (Fórmula de Itô generalizada I) Considere $F_t(x), t \in [0, a], x \in \mathbb{R}^d$ um campo aleatório contínuo em (t, x) (ω -quase sempre), onde $F_t(x)$ é duas vezes continuamente diferenciável em x e para cada $x, F_t(x)$ é um semimartingale contínuo e satisfaz

$$F_t(x) = F_0(x) + \sum_{j=1}^m \int_0^t f_s^j(x) dY_s^j, \quad \forall x \in \mathbb{R}^d \quad (\omega - \text{quase sempre}),$$

onde Y_s^1, \dots, Y_s^m são semimartingales contínuos, $f_s^j(x), s \in [0, a], x \in \mathbb{R}^d$ são campos aleatórios contínuos em (s, x) , duas vezes continuamente diferenciável com relação a x e para cada x ,

$f_s^j(x)$ são processos adaptados. Consideremos $X_t = (X_t^1, \dots, X_t^d)$ semimartingales contínuos com valores em \mathbb{R}^d . Então temos

$$F_t(X_t) = F_0(X_0) + \sum_{j=1}^m \int_0^t f_s^j(X_s) dY_s^j + \sum_{i=1}^d \int_0^t \frac{\partial F_s}{\partial x_i}(X_s) dX_s^i + \\ + \sum_{i=1}^d \sum_{j=1}^m \int_0^t \frac{\partial f_s^j}{\partial x_i}(X_s) d[Y^j, X^i]_s + \frac{1}{2} \sum_{i,j=1}^d \int_0^t \frac{\partial^2 F_s}{\partial x_i \partial x_j}(X_s) d[X^i, X^j]_s.$$

Para as aplicações que faremos no capítulo final, precisaremos da fórmula acima em termos da integral de Stratonovich, fornecida no próximo teorema.

Teorema 47. (Fórmula de Itô generalizada II) Seja $F_t(x)$, $t \in [0, a]$, $x \in \mathbb{R}^d$ um campo aleatório contínuo em (t, x) (ω -quase sempre), onde para cada t , $F_t(\cdot)$ é uma função de classe C^3 de \mathbb{R}^d em \mathbb{R} (ω -quase sempre) e para cada x , $F_t(x)$ é um semimartingale e satisfaz

$$F_t(x) = F_0(x) + \sum_{j=1}^m \int_0^t f_s^j(x) \circ dY_s^j, \quad \forall x \in \mathbb{R}^d \quad (\omega - \text{quase sempre}),$$

onde Y_s^1, \dots, Y_s^m são semimartingales contínuos e $f_s^j(x)$ são campos aleatórios satisfazendo as condições impostas anteriormente em (46). Consideremos agora $X_t = (X_t^1, \dots, X_t^d)$ um semimartingale contínuo. Então

$$F_t(X_t) = F_0(X_0) + \sum_{j=1}^m \int_0^t f_s^j(X_s) \circ dY_s^j + \sum_{j=1}^m \int_0^t \frac{\partial F_s}{\partial x_i}(X_s) \circ dX_s^i.$$

A equação acima pode ser escrita na forma diferencial como

$$d(F_t(X_t)) = (dF_t)(X_t) + \sum_{i=1}^d \frac{\partial F_t}{\partial x_i}(X_t) \circ dX_t^i.$$

A demonstração dos teoremas anteriores podem ser encontradas em Kunita (1982) [11].

Apenas para ilustrar a fórmula acima, e ver como ela se assemelha a regra da cadeia usual do cálculo, consideremos dois fluxos determinísticos ξ_t, η_t tal que $\xi_t, \eta_t : \mathbb{R} \rightarrow \mathbb{R}$. Então, tomando a composta desses fluxos e aplicando num ponto x temos:

$$(\xi_t \circ \eta_t)(x) = \xi_t(\eta_t(x)) = \xi(t, \eta(t, x))$$

Assim, pela regra da cadeia

$$\frac{d}{dt} [\xi(t, \eta(t, x))] = \frac{\partial \xi}{\partial t}(t, \eta(t, x)) \cdot \frac{dt}{dt} + \frac{\partial \xi}{\partial x}(t, \eta(t, x)) \cdot \frac{d}{dt} \eta(t, x).$$

As Fórmulas de Itô Generalizadas estarão presente na maioria dos resultados que apresentaremos daqui em diante. Tais resultados nos auxiliarão no estudo da ação de fluxos estocásticos em campos de vetores e em formas diferenciais.

Ações de fluxos estocásticos em campos de vetores e formas diferenciais

Neste capítulo, estamos interessados em definir equações diferenciais estocásticas no sentido da integral de Stratonovich em variedades diferenciáveis, ou seja, queremos definir equações diferenciais estocásticas de Stratonovich em superfícies que não apresentam “bicos”. Trabalhamos com uma EDE de Stratonovich, pois ela modela uma grande classe de sistemas dinâmicos em variedades, uma vez que a sua Fórmula de Itô se ajusta melhor às ferramentas de geometria diferencial. No entanto, também podemos trabalhar com uma EDE de Itô, dependendo dos objetivos.

Como queremos uma EDE de Stratonovich, faremos uso de alguns conceitos importantes da geometria diferencial, como variedade diferenciável, campos de vetores, diferencial de uma aplicação e derivada de Lie de um campo de vetores, que podem ser vistos com mais detalhes, por exemplo, em Lee (2012) [15]. Iniciaremos expressando uma EDE de Stratonovich numa variedade M com coordenadas locais e analisando sua solução. Em seguida, fornecemos uma fórmula de Itô agindo em campo de vetores e tensores alternados, também chamados de formas diferenciais. Neste trabalho usaremos apenas dois campos vetoriais, mas os resultados são análogos para uma quantidade maior de campos vetoriais.

No capítulo anterior apresentamos a ideia de um fluxo estocástico no espaço \mathbb{R}^d . Agora, como iremos trabalhar com variedade, precisamos estender essa ideia. Seja M uma variedade diferenciável e X um campo de vetores em M . Para cada ponto $p \in M$, X tem uma única trajetória começando em p e definida para todo tempo t . Para cada t , podemos definir uma função $\xi_t : M \rightarrow M$ enviando cada $p \in M$ ao ponto obtido seguindo, para o tempo t , a trajetória começando em p . Essa função ξ_t é chamada de fluxo estocástico. De forma mais simples, considerando uma variedade diferenciável M , para cada ponto $p \in M$, ξ_t define uma única trajetória começando em p e que leva cada ponto p ao ponto obtido após um tempo t ; o fluxo indica o “caminho” percorrido por cada ponto p na variedade.

Seja M uma variedade C^∞ de dimensão d paracompacta e conexa. Considere dois campos vetoriais X_0, X_1 em M com parâmetro $t \in [0, a]$. Assumimos que X_1 é um campo vetorial de classe C^2 continuamente diferenciáveis em t , isto é, com coordenadas locais (x_1, \dots, x_d) , esse campo vetorial é expresso como

$$X_1(t) = \sum_{i=1}^d X_1^i(t, x) \frac{\partial}{\partial x_i}, \quad (5.0.0.1)$$

onde $X_1^i(t, x)$, $1 \leq i \leq d$ são funções de classe C^1 com relação a t e funções de classe C^2 com relação a x . Enquanto que o campo vetorial X_0 é tal que $X_0^i(t, x)$ é contínuo em t e de classe C^1 com relação a x . A partir desse momento, consideramos uma EDE de Stratonovich em uma variedade M como

$$d\xi_{s,t} = X_0(t, \xi_{s,t})dt + X_1(t, \xi_{s,t}) \circ dB_t, \quad (5.0.0.2)$$

onde B_t é um movimento browniano 1-dimensional definido no espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$.

Na Definição 45 consideramos $\xi_{s,t}(x)$ com $x \in \mathbb{R}^d$, como estamos em uma variedade M precisamos estender tal definição com $x \in M$, ou seja, precisamos explicar o significado de uma solução da equação (5.0.0.2) em uma variedade.

Definição 48. Uma variável aleatória $\xi_{s,t}(x)$, $x \in M$, $s \leq t < \min\{T(s, x, \omega), a\}$ com valores em M é chamado de solução local da equação (5.0.0.2) com condição inicial $\xi_s = x$ se satisfaz as seguintes condições

1. $T(s, x, \omega)$ é mensurável com relação a (s, x, ω) e para qualquer (s, x) fixo é um tempo de parada com respeito a $\mathcal{F}_{s,t}$, onde $\mathcal{F}_{s,t} = \sigma(B_u - B_v; s \leq u \leq v \leq t)$;
2. $\xi_{s,t}(x)$ é contínua em (s, t, x) ;
3. Para qualquer função $F : M \rightarrow \mathbb{R}$ de classe C^3 , $F(\xi_{s,t}(x))$ é um semimartingale contínuo com relação a $\mathcal{F}_{s,t}$ e satisfaz para todo $t < \min\{T(s, x, \omega), a\}$

$$F(\xi_{s,t}(x)) = F(x) + \int_s^t X_0(r)F(\xi_{s,r}(x))dr + \int_s^t X_1(r)F(\xi_{s,r}(x)) \circ dB_r. \quad (5.0.0.3)$$

A solução é chamada de maximal se $T(s, x) = \infty$ para todo x (ω -q.s.) e para qualquer s no caso de M ser compacto, ou se $\lim_{t \downarrow T(s, x)} \xi_{s,t}(x) = \infty$ é válido para $T(s, x) < \infty$ (ω -q.s.) no caso de M não ser compacto, onde ∞ é um ponto de compactação de M ¹.

Usando o conceito anterior de coordenadas locais, podemos expressar a equação (5.0.0.2) ou (5.0.0.3) da seguinte forma. Seja (x_1, \dots, x_d) as coordenadas locais de uma vizinhança U . Tome $F(x) = x_i$ então

$$X_0 F(x) = X_0^i(r, x) \quad \text{e} \quad X_1 F(x) = X_1^i(r, x).$$

Portanto, se $\xi_{s,t}$ satisfaz (5.0.0.3), então

$$\xi_{s,t}^i(x) \equiv x_i(\xi_{s,t}(x)) = F(\xi_{s,t}(x))$$

satisfaz:

$$\xi_{s,t}^i(x) = x_i + \int_s^t X_0^i(r, \xi_{s,r}(x))dr + \int_s^t X_1^i(r, \xi_{s,r}(x)) \circ dB_r \quad (5.0.0.4)$$

para $t < \min\{T_U, a\}$, onde $T_U = \inf\{t > s \mid \xi_{s,t}(x) \notin U\}$.

Note que a solução da equação (5.0.0.4) não depende da escolha das coordenadas locais. De fato, seja $(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_d)$ outra coordenada local da vizinhança U . Então podemos escrever o campo de vetores $X(t)$ e $Y(t)$ da seguinte forma

$$X_0(t) = \sum_{j=1}^d \bar{X}_0^j(t, x) \frac{\partial}{\partial x_j} \quad X_1(t) = \sum_{j=1}^d \bar{X}_1^j(t, x) \frac{\partial}{\partial x_j}. \quad (5.0.0.5)$$

Os coeficientes $X_0^i(t, x)$ e $\bar{X}_0^i(t, x)$ são relacionados por

$$X_0^i(t, x) = \sum_{j=1}^d \bar{X}_0^j(t, x) \frac{\partial x_i}{\partial \bar{x}_j}, \quad i = 1, \dots, d, \quad (5.0.0.6)$$

¹ver Apêndice B

o mesmo vale para X_1 pois, substituindo $X_1^i(t, x)$ na equação (5.0.0.1) pela igualdade dada em (5.0.0.6), obtemos a equação (5.0.0.5):

$$\begin{aligned} X_1(t) &= \sum_{i=1}^d \sum_{j=1}^d \bar{X}_1^j(t, x) \frac{\partial x_i}{\partial \bar{x}_j} \frac{\partial}{\partial x_i} \\ &= \sum_{j=1}^d \bar{X}_1^j(t, x) \frac{\partial}{\partial \bar{x}_j}. \end{aligned} \quad (5.0.0.7)$$

Seja $\bar{\xi}_s^t(x) = (\bar{\xi}_{s,t}^1(x), \dots, \bar{\xi}_{s,t}^d(x))$ uma outra solução de (5.0.0.4) com os coeficientes $\bar{X}_0^i(t, x)$. Aplicando a Fórmula de Itô dada na equação (5.0.0.3) para $F(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_d) = x_i(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_d)$

$$\begin{aligned} F(\bar{\xi}_{s,t}(x)) &= F(x) + \int_s^t X_0(r) F(\bar{\xi}_{s,r}(x)) dr + \int_s^t X_1(r) F(\bar{\xi}_{s,r}(x)) \circ dB_r \\ x_i(\bar{\xi}_{s,t}(x)) &= x_i + \int_s^t X_0(r) x_i(\bar{\xi}_{s,r}(x)) dr + \int_s^t X_1(r) x_i(\bar{\xi}_{s,r}(x)) \circ dB_r \\ &\stackrel{(5.0.0.5)}{=} x_i + \int_s^t \sum_{j=1}^d \bar{X}_0^j(r, \bar{\xi}_{s,r}) \frac{\partial}{\partial \bar{x}_j} x_i(\bar{\xi}_{s,r}(x)) dr + \int_s^t \sum_{j=1}^d \bar{X}_1^j(r, \bar{\xi}_{s,r}) \frac{\partial}{\partial \bar{x}_j} x_i(\bar{\xi}_{s,r}(x)) \circ dB_r \\ &\stackrel{(5.0.0.6)}{=} x_i + \int_s^t X_0^i(r, \bar{\xi}_{s,r}) dr + \int_s^t X_1^i(r, \bar{\xi}_{s,r}) \circ dB_r. \end{aligned}$$

Portanto, $x_i(\bar{\xi}_{s,t}(x))$ coincide com $\xi_{s,t}^i$.

Usando o fato que a variedade admite uma base enumerável de abertos é possível mostrar que existe uma única solução maximal. A ideia é dividir a variedade em vários abertos e usando uma base enumerável podemos construir soluções estocásticas em cada um de tal forma que essas soluções podem se concatenar afim de juntar as soluções e formar a solução maximal. Mais detalhes sobre a demonstração dessa afirmação podem ser vistas em Kunita (1982) [11], páginas 245-249.

Fornecemos até o momento uma fórmula que descreve a regra diferencial de $\xi_{s,t}$ para a variável t . Entretanto, é possível fornecer uma regra diferencial de $\xi_{s,t}$ para a variável de retorno s . Mais detalhes sobre isso podem ser visto em Kunita (1982) [11].

Além disso, seja $F : M \rightarrow \mathbb{R}$ uma função de classe C^3 e $\xi_{s,t}$ é um homeomorfismo. A equação (5.0.0.3) pode ser escrita para a inversa de $\xi_{s,t}$, como

$$F(\xi_{s,t}^{-1}(x)) = F(x) - \int_s^t X_0(r) F(\xi_{s,r}^{-1}(x)) dr - \int_s^t X_1(r) F(\xi_{s,r}^{-1}(x)) \circ dB_r. \quad (5.0.0.8)$$

5.1 Fluxos agindo em campo de vetores

Nesta seção, estamos interessados na equação diferencial estocástica ou fórmula de Itô que rege o processo avaliado pelo campo de vetores. Para simplificar a notação, consideramos a partir de agora $s = 0$. Todos os resultados podem ser estendidos para $s > 0$ de forma semelhante.

Começaremos introduzindo o conceito de diferencial de uma aplicação, conceito que pode ser visto com mais detalhes no Apêndice C. Na literatura da Geometria Diferencial, o diferencial de um campo às vezes é chamado de mapa tangente, derivada total ou simplesmente derivada de F . Como ele “empurra” vetores tangentes da variedade de domínio para o contradomínio, também pode ser chamado de *pushforward*.

Seja $\xi : M \rightarrow M$ um difeomorfismo (a princípio determinístico, que fará o papel do fluxo estocástico ξ_r em breve). A diferencial ξ_{x^*} (também chamada de *pushforward*) é por definição

uma função linear do espaço tangente $T_x M$ para o espaço tangente $T_{\xi(x)} M$ tal que dado um ponto $x \in T_x M$ a derivação de ξ_x agindo em $f \in C^\infty(M)$ denotado por ξ_x^* é dado pela regra

$$\xi_x^* X_x f = X_x(f \circ \xi) \quad \forall X_x \in T_x(M),$$

onde \circ indica uma composição e X_x denota a restrição do campo vetorial X em M no ponto x . Definimos um novo campo vetorial $\xi_* X$ por

$$(\xi_* X)_x = \xi_{*\xi^{-1}(x)} X_{\xi^{-1}(x)}.$$

Consequentemente, para qualquer função f de classe $C^\infty(M)$, onde $C^\infty(M)$ denota o espaço de todas as funções C^∞ , temos

$$\xi_* X f(x) = X(f \circ \xi)(\xi^{-1}(x)), \quad \forall x \in M. \quad (5.1.0.1)$$

Observe que podemos expressar $\xi_* X$ usando coordenadas locais. De fato, seja

$$X = \sum_{i=1}^d X^i(x) \frac{\partial}{\partial x_i}$$

a expressão de X em coordenadas locais. Então

$$\xi_* X = \sum_{i=1}^d (\xi_* X)^i(x) \frac{\partial}{\partial x_i}$$

onde

$$(\xi_* X)^i(x) = \sum_{j=1}^d X^j(\xi^{-1}(x)) \frac{\partial \xi^i}{\partial x_j}(\xi^{-1}(x))$$

e

$$\xi^i(x) = x_i(\xi(x)),$$

onde $x_i(x_1, \dots, x_d) = x_i$.

O próximo teorema fornece uma fórmula de Itô para a ação de fluxos estocásticos em campos de vetores.

Teorema 49. *Seja X_1 um campo de vetores de classe $C^{5,\alpha}$ para algum $\alpha > 0$, X_0 um campo de vetores de classe $C^{4,\alpha}$ para algum α e ξ_t a solução da EDE (5.0.0.2). Então $\xi_t^* X$ satisfaz a seguinte fórmula para algum campo de vetor X de classe C^3*

$$\xi_t^* X = X - \int_0^t L_{X_0(r)}(\xi_r^* X) dr - \int_0^t L_{X_1(r)}(\xi_r^* X) \circ dB_r \quad (5.1.0.2)$$

onde $L_{X_i(r)}(\xi_r^* X)$, $i = 0, 1$ denota a derivada de Lie do campo $\xi_r^* X$ com relação a $X_i(r)$.

Demonstração. Seja $f \in C^\infty(M)$. Então, pela equação (5.1.0.1)

$$\xi_t^* X f(x) = X(f \circ \xi_t)(\xi_t^{-1}(x)).$$

Seja

$$F_t(x) = X(f \circ \xi_t)(x) \quad \text{e} \quad M_t = \xi_t^{-1}(x) = (\xi_t^{-11}(x), \dots, \xi_t^{-1d}(x))$$

onde a última igualdade da segunda equação representa M_t em coordenadas locais. Pela fórmula de Itô generalizada com integral de Stratonovich dada no Teorema 47 obtemos

$$d(F_t(M_t)) = (dF_t)(M_t) + \sum_{i=1}^d \frac{\partial F_t}{\partial x_i}(M_t) \circ dM_t^i. \quad (5.1.0.3)$$

Agora, precisamos encontrar quem representa $(dF_t)(M_t)$ e $\sum_{i=1}^d \frac{\partial F_t}{\partial x_i}(M_t) \circ dM_t^i$. Temos pela equação (5.0.0.3):

$$f(\xi_t(x)) = f(x) + \int_0^t X_0(r)f(\xi_r(x))dr + \int_0^t X_1(r)f(\xi_r(x)) \circ dB_r$$

Aplicando o campo X de ambos os lados da equação acima:

$$X[f(\xi_t(x))] = Xf(x) + \int_0^t X[X_0(r)f(\xi_r(x))]dr + \int_0^t X[X_1(r)f(\xi_r(x))] \circ dB_r$$

ou ainda,

$$X(f \circ \xi_t)(x) = Xf(x) + \int_0^t X(X_0(r)f \circ \xi_r)dr + \int_0^t X(X_1(r)f \circ \xi_r) \circ dB_r.$$

Então,

$$\begin{aligned} (dF_t)(M_t) &= X(X_0(t)f \circ \xi_t)(M_t)dt + X(X_1(t)f \circ \xi_t)(M_t) \circ dB_t \\ &\stackrel{M_t = \xi_t^{-1}(x)}{=} X(X_0(t)f \circ \xi_t)(\xi_t^{-1}(x))dt + X(X_1(t)f \circ \xi_t)(\xi_t^{-1}(x)) \circ dB_t \\ &\stackrel{(5.1.0.1)}{=} \xi_t^* X X_0(t)f(x)dt + \xi_t^* X X_1(t)f(x) \circ dB_t. \end{aligned} \quad (5.1.0.4)$$

Descobrimos o termo $(dF_t)(M_t)$. Para achar $\sum_{i=1}^d \frac{\partial F_t}{\partial x_i}(M_t) \circ dM_t^i$ precisamos primeiramente encontrar o termo dM_t^i . Observe que dM_t^i é o mesmo que $d\xi_t^{-1i}$. Então, tomando $F(x_1, \dots, x_d) = x_i$ temos que $M_t^i = F(M_t)$ e aplicando a equação (5.0.0.8):

$$M_t^i = \xi_t^{-1i} = x_i - \int_0^t X_0(r)(\xi_r^{-1i}(x))dr - \int_0^t X_1(r)(\xi_r^{-1i}(x)) \circ dB_r$$

ou na forma diferencial

$$dM_t^i = d\xi_t^{-1i} = -X_0(t)(\xi_t^{-1i}(x))dt - X_1(t)(\xi_t^{-1i}(x)) \circ dB_t.$$

Note que a equação acima pode ser reescrita como

$$dM_t^i = -\xi_t^{*-1} X_0^i(t)(\xi_t^{-1}(x))dt - \xi_t^{*-1} X_1^i(t)(\xi_t^{-1}(x)) \circ dB_t \quad (5.1.0.5)$$

pois, como $X_0^i(t) = X_0(t)x_i$:

$$\begin{aligned} \xi_t^{*-1} X_0^i(t)(\xi_t^{-1}(x)) &= \xi_t^{*-1} X_0(t)x_i(\xi_t^{-1}(x)) \\ &\stackrel{(5.1.0.1)}{=} X_0(t) \underbrace{(x_i \circ \xi_t^{-1})}_{=\xi_t^{-1i}} \underbrace{(\xi_t(\xi_t^{-1}))}_{=x} \\ &= X_0(t)(\xi_t^{-1i}(x)). \end{aligned}$$

O mesmo vale para o campo X_1 . Para simplificar a expressão, adotamos $dB_t^0 = dt$. Portanto,

$$\begin{aligned}
\sum_{i=1}^d \frac{\partial F_t}{\partial x_i}(M_t) \circ dM_t^i &= - \sum_{i=1}^d \sum_{k=0}^1 \xi_t^{*-1} X_k^i(t) (\xi_t^{-1}(x)) \frac{\partial}{\partial x_i} (F_t) (\xi_t^{-1}(x)) \circ dB_t^k \\
&= - \sum_{k=0}^1 \xi_t^{*-1} \underbrace{\sum_{i=1}^d X_k^i(t) \frac{\partial}{\partial x_i} (F_t) (\xi_t^{-1}(x))}_{X_k(t) F_t(\xi_t^{-1}(x))} \circ dB_t^k \\
&= - \sum_{k=0}^1 \xi_t^{*-1} X_k(t) F_t(\xi_t^{-1}(x)) \circ dB_t^k \tag{5.1.0.6}
\end{aligned}$$

$$= - \sum_{k=0}^1 X_k(t) F_t(\xi_t^{-1}) (\underbrace{\xi_t(\xi_t^{-1}(x))}_{=(x)}) \circ dB_t^k \tag{5.1.0.7}$$

$$\begin{aligned}
&= - \sum_{k=0}^1 X_k(t) X(f \circ \xi_t) (\xi_t^{-1})(x) \circ dB_t^k \\
&= - \sum_{k=0}^1 X_k(t) \xi_t^* X f(x) \circ dB_t^k. \tag{5.1.0.8}
\end{aligned}$$

de (5.1.0.6) para (5.1.0.7) usamos a equação (5.1.0.1) onde $\xi_t^{-1}(x)$ é ponto e F_t é função. De (5.1.0.4) e (5.1.0.8):

$$\begin{aligned}
d(F_t(M_t)) &= \sum_{k=0}^1 \xi_t^* X X_k(t) f(x) \circ dB_t^k - \sum_{k=0}^1 X_k(t) \xi_t^* X f(x) \circ dB_t^k \\
&= \sum_{k=0}^1 \xi_t^* X X_k(t) f(x) - X_k(t) \xi_t^* X f(x) \circ dB_t^k \tag{5.1.0.9}
\end{aligned}$$

$$= \sum_{k=0}^1 \left[\xi_t^* X, X_k(t) \right] \circ dB_t^k \tag{5.1.0.10}$$

$$= - \sum_{k=0}^1 \left[X_k(t), \xi_t^* X \right] \circ dB_t^k \tag{5.1.0.11}$$

$$\begin{aligned}
&= - \sum_{k=0}^1 L_{X_k(t)} (\xi_t^* X) \circ dB_t^k \\
&= - L_{X_0(t)} (\xi_t^* X) dt - L_{X_1(t)} (\xi_t^* X) \circ dB_t.
\end{aligned}$$

onde, de (5.1.0.9) para (5.1.0.10) usamos a definição de Colchete de Lie², de (5.1.0.10) para (5.1.0.11) usamos uma propriedade do Colchete de Lie³ e na última igualdade usamos a relação bem conhecida de Derivada de Lie⁴ dada por

$$L_Y(X) = [Y, X].$$

■

O próximo exemplo ilustra o teorema anterior.

Exemplo 14. Considere o campo em \mathbb{R}^2 dado por

$$X_1(x, y) = (1, 0) = 1 \frac{\partial}{\partial x} + 0 \frac{\partial}{\partial y}.$$

²ver equação (C.2.0.1)

³ver Proposição 63

⁴ver Definição 65

e a EDE dada por

$$d(\xi_t(x, y)) = X_1(\xi_t(x, y)) \circ dB_t. \quad (5.1.0.12)$$

Podemos escrever esta EDE na forma diferencial como sistema

$$\begin{bmatrix} d\xi_t^1(x, y) \\ d\xi_t^2(x, y) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \circ dB_t$$

e que também pode ser expressa na forma integral:

$$\begin{cases} \xi_t^1(x, y) = \xi_0^1(x, y) + \int_0^t 1 \circ dB_r \\ \xi_t^2(x, y) = \xi_0^2(x, y) + \int_0^t 0 \circ dB_r \end{cases}$$

Assumindo $\xi_0^1(x, y) = x$ e $\xi_0^2(x, y) = y$, temos

$$\xi_t(x, y) = \left(x + \int_0^t 1 \circ dB_r, y + 0 \right) \quad (5.1.0.13)$$

$$= \left(x + \int_0^t 1 dB_r, y \right) \quad (5.1.0.14)$$

$$= \left(x + B_t - B_0, y \right)$$

$$= \left(x + B_t, y \right).$$

De (5.1.0.13) para (5.1.0.14) usamos a conversão de uma EDE de Stratonovich para uma EDE de Itô. Além disso, temos que a inversa de $\xi_t(x, y)$ é dada por $\xi_t^{-1}(x, y) = (x - B_t, y)$.

Não é difícil verificar que uma representação matricial para ξ_t^* é dada por

$$\left[\xi_t^* \right]_{\xi_t^{-1}(x, y)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \Big|_{\xi_t^{-1}(x, y)} \quad (5.1.0.15)$$

Vamos verificar, por exemplo, que

$$\xi_t^* \frac{\partial}{\partial x} = \frac{\partial}{\partial x}.$$

De fato, pela equação (5.1.0.2)

$$\begin{aligned} \left(\xi_t^* \frac{\partial}{\partial x} \right) f(x, y) &= \left[\frac{\partial}{\partial x} - \int_0^t L_{X_1(r)} \xi_{r^*} \frac{\partial}{\partial x} \circ dB_r \right] f(x, y) \\ &= \frac{\partial}{\partial x} f(x, y) - \int_0^t L_{X_1(r)} \xi_{r^*} \frac{\partial}{\partial x} f(x, y) \circ dB_r \\ &= \frac{\partial}{\partial x} f(x, y) - \int_0^t \left[\left(\xi_{r^*} \frac{\partial}{\partial x} \right) X_1 - X_1 \left(\xi_{r^*} \frac{\partial}{\partial x} \right) \right] f(x, y) \circ dB_r \\ &= \frac{\partial}{\partial x} f(x, y) - \int_0^t \left[\left(\xi_{r^*} \frac{\partial}{\partial x} \right) X_1 f(x, y) - X_1 \left(\xi_{r^*} \frac{\partial}{\partial x} \right) f(x, y) \right] \circ dB_r \end{aligned}$$

Calculando cada termo da integral, de acordo com a equação (5.1.0.1), obtemos

$$\begin{aligned} \left(\xi_{r^*} \frac{\partial}{\partial x} \right) X_1 f(x, y) &= \frac{\partial}{\partial x} \left(X_1 f(x, y) \circ \xi_r \right) \left(\xi_r^{-1}(x, y) \right) \\ &= \frac{\partial}{\partial x} \left(\frac{\partial}{\partial x} \xi_r^1 \right) \left(\xi_r^{-1}(x, y) \right) \end{aligned}$$

$$\begin{aligned} X_1 \left(\xi_{r^*} \frac{\partial}{\partial x} \right) f(x, y) &= \frac{\partial}{\partial x} \left[\frac{\partial}{\partial x} \left(f(x, y) \circ \xi_r \right) \left(\xi_r^{-1}(x, y) \right) \right] \\ &= \frac{\partial}{\partial x} \left(\frac{\partial}{\partial x} \xi_r^1 \right) \left(\xi_r^{-1}(x, y) \right) \end{aligned}$$

Portanto,

$$\int_0^t \left[\left(\xi_{r^*} \frac{\partial}{\partial x} \right) X_1 f(x, y) - X_1 \left(\xi_{r^*} \frac{\partial}{\partial x} \right) f(x, y) \right] \circ dB_r = \int_0^t 0 \circ dB_r = 0.$$

Logo,

$$\left(\xi_{r^*} \frac{\partial}{\partial x} \right) f(x, y) = \frac{\partial}{\partial x} f(x, y),$$

e assim

$$\xi_{r^*} \frac{\partial}{\partial x} = 1 \frac{\partial}{\partial x} + 0 \frac{\partial}{\partial y},$$

ou seja, a primeira coluna da matriz $[\xi_{r^*}]$ é dada por

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

De forma análoga,

$$\xi_{r^*} \frac{\partial}{\partial y} = 0 \frac{\partial}{\partial x} + 1 \frac{\partial}{\partial y},$$

ou seja, a segunda coluna da matriz $[\xi_{r^*}]$ é dada por

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Observe que trabalhar a fórmula (5.1.0.2) não é de modo algum fácil, uma vez que apresenta a derivada de Lie do campo $\xi_{r^*} X$ com relação a $X_i(r)$, e assim, para usá-la devemos calcular ξ_{r^*} que não é uma tarefa fácil, já que muitas vezes nem conhecemos ξ_r , imagina sua diferencial. E se conhecermos, calcular a diferencial pode ser bem trabalhoso, igual ilustrado no exemplo. Portanto, quando queremos saber a ação de fluxos estocásticos em campos de vetores usamos uma fórmula bem similar a equação (5.3.0.2) considerando ξ_t^{-1} , a próxima proposição diz isso.

Proposição 50. *Podemos obter uma equação similar a equação (5.1.0.2) considerando o fluxo estocástico inverso. Isto é, temos que ξ_t^{-1} satisfaz*

$$\xi_t^{-1} X = X + \int_0^t \xi_{r^*}^{-1} L_{X_0(r)} X dr + \int_0^t \xi_{r^*}^{-1} L_{X_1(r)} X \circ dB_r \quad (5.1.0.16)$$

Observação 51. *Como a prova deste teorema não é feita no Kunita (1992) [11], iremos mostrar aqui, mesmo sendo muito próxima da demonstração do Teorema 49.*

Demonstração. Pela equação (5.1.0.1) temos que

$$\xi_{r^*}^{-1} X f(x) = X(f \circ \xi_r^{-1})(\xi_r(x)).$$

e, pela fórmula de Itô generalizada com integral de Stratonovich dada no Teorema 47 obtemos

$$d(G_t(H_t)) = (dG_t)(H_t) + \sum_{i=1}^d \frac{\partial G_t}{\partial x_i}(H_t) \circ dH_t^i, \quad (5.1.0.17)$$

onde

$$G_t(x) = X(f \circ \xi_t^{-1})(x) \quad \text{e} \quad H_t = \xi_t(x) = (\xi_t^1(x), \dots, \xi_t^d(x)).$$

Agora, precisamos encontrar quem representa $(dG_t)(H_t)$ e $\sum_{i=1}^d \frac{\partial G_t}{\partial x_i}(H_t) \circ dH_t^i$. Temos pela equação (5.0.0.8):

$$f(\xi_t^{-1}(x)) = f(x) - \int_0^t X_0(r)f(\xi_r^{-1}(x))dr - \int_0^t X_1(r)f(\xi_r^{-1}(x)) \circ dB_r$$

Aplicando o campo X de ambos os lados da equação acima e diferenciando,

$$(dG_t)(H_t) = -\xi_t^{-1} X X_0(t)f(x)dt - \xi_t^{-1} X X_1(t)f(x) \circ dB_t. \quad (5.1.0.18)$$

Observe que dH_t^i pode ser representado como $d\xi_t^i$. Então, tomando $F(x_1, \dots, x_d) = x_i$ temos que $H_t^i = F(H_t)$ e aplicando a equação (5.0.0.3):

$$H_t^i = \xi_t^i = x_i + \int_0^t X_0(r)(\xi_r^i(x))dr + \int_0^t X_1(r)(\xi_r^i(x)) \circ dB_r$$

ou na forma diferencial

$$dH_t^i = d\xi_t^i = X_0(t)(\xi_t^i(x))dt + X_1(t)(\xi_t^i(x)) \circ dB_t.$$

Note que a equação acima, como anteriormente, pode ser reescrita como

$$dH_t^i = \xi_t^* X_0^i(t)(\xi_t(x))dt + \xi_t^* X_1^i(t)(\xi_t(x)) \circ dB_t \quad (5.1.0.19)$$

Para simplificar a expressão adotamos $dB_t^0 = dt$, como anteriormente. Portanto,

$$\begin{aligned} \sum_{i=1}^d \frac{\partial G_t}{\partial x_i}(H_t) \circ dH_t^i &= \sum_{i=1}^d \sum_{k=0}^1 \xi_t^* X_k^i(t)(\xi_t(x)) \frac{\partial}{\partial x_i}(G_t)(\xi_t(x)) \circ dB_t^k \\ &= \sum_{k=0}^1 \xi_t^* \underbrace{\sum_{i=1}^d X_k^i(t) \frac{\partial}{\partial x_i}(G_t)(\xi_t(x))}_{X_k(t)G_t(\xi_t(x))} \circ dB_t^k \\ &= \sum_{k=0}^1 \xi_t^* X_k(t)G_t(\xi_t(x)) \circ dB_t^k \end{aligned} \quad (5.1.0.20)$$

$$= \sum_{k=0}^1 X_k(t)G_t(\xi_t) \underbrace{(\xi_t^{-1}(\xi_t(x)))}_{=(x)} \circ dB_t^k \quad (5.1.0.21)$$

$$\begin{aligned} &= \sum_{k=0}^1 X_k(t)X(f \circ \xi_t^{-1})(\xi_t)(x) \circ dB_t^k \\ &= \sum_{k=0}^1 \xi_t^* X_k(t)Xf(x) \circ dB_t^k. \end{aligned} \quad (5.1.0.22)$$

de (5.1.0.20) para (5.1.0.21) usamos a equação (5.1.0.1) onde $\xi_t(x)$ é ponto e G_t é função. De (5.1.0.18) e (5.1.0.22):

$$\begin{aligned}
 d(G_t(H_t)) &= -\sum_{k=0}^1 \xi_t^{-1} X X_k(t) f(x) \circ dB_t^k + \sum_{k=0}^1 \xi_t^{-1} X_k(t) X f(x) \circ dB_t^k \\
 &= \sum_{k=0}^1 \xi_t^{-1} X_k(t) X f(x) - \xi_t^{-1} X X_k(t) f(x) \circ dB_t^k \\
 &= \sum_{k=0}^1 \xi_t^{-1} \left[X_k(t), X \right] \circ dB_t^k \\
 &= \sum_{k=0}^1 \xi_t^{-1} L_{X_k(t)} X \circ dB_t^k \\
 &= \xi_t^{-1} L_{X_0(t)} X dt + \xi_t^{-1} L_{X_1(t)} X \circ dB_t
 \end{aligned}$$

■

A seguir, daremos um exemplo para ilustrar o teorema anterior.

Exemplo 15. Vamos considerar ξ_t como sendo a solução da EDE

$$d\xi_t = X_1(\xi_t) \circ dB_t$$

onde X_1 é um campo em \mathbb{R}^3 dado por

$$X_1(x, y, z) = -y \frac{\partial}{\partial x} + x \frac{\partial}{\partial y} + 0 \frac{\partial}{\partial z}.$$

Neste caso, pelos comentários que fizemos no exemplo 13, temos que uma trajetória iniciada em (x_0, y_0, z_0) permanece no cilindro de raio $r = \|(x_0, y_0, 0)\|$ mantendo a altura z_0 . tal campo de vetores pode ser ilustrado da seguinte forma

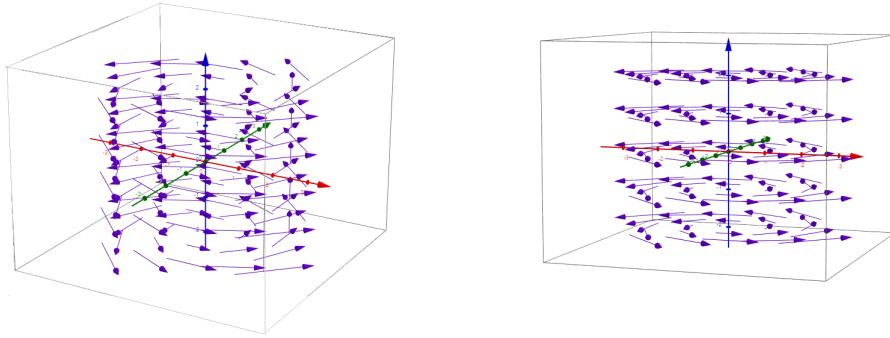


Figura 5.1: Ilustração do campo de vetores X_1 por dois ângulos.

Tomemos

$$X(x, y, z) = x \frac{\partial}{\partial x} + y \frac{\partial}{\partial y} + z \frac{\partial}{\partial z}.$$

Note que $L_{X_1} X = 0$. De fato,

$$\begin{aligned}
 L_{X_1} X &\stackrel{\text{Prop.65}}{=} [X_1, X] \\
 &\stackrel{\text{Prop.62}}{=} X_1(x) \frac{\partial}{\partial x} + X_1(y) \frac{\partial}{\partial y} + X_1(z) \frac{\partial}{\partial z} - X(-y) \frac{\partial}{\partial x} - X(x) \frac{\partial}{\partial y} - X(0) \frac{\partial}{\partial z} \\
 &= -y \frac{\partial}{\partial x} + x \frac{\partial}{\partial y} + 0 \frac{\partial}{\partial z} - (-y) \frac{\partial}{\partial x} - x \frac{\partial}{\partial y} - 0 \frac{\partial}{\partial z} \\
 &= 0.
 \end{aligned}$$

Logo,

$$\xi_t^{-1}X = X + \int_s^t \underbrace{\xi_r^{-1}L_{X_1(r)}X}_{=0} \circ dB_r = X.$$

Neste caso dizemos que o campo X é invariante pela ação do fluxo estocástico ξ_t^{-1} .

Uma observação interessante é que se X é invariante pela ação do fluxo estocástico ξ_t^{-1} então X também será invariante pela ação do fluxo estocástico ξ_t . Isto é, se $\xi_t^{-1}X = X$ então $\xi_t X = X$. De fato, de acordo com as propriedades de diferencial⁵

$$(\xi_t \circ \xi_t^{-1})_*X = (id)_*X = X,$$

mas,

$$(\xi_t \circ \xi_t^{-1})_*X = \xi_t_* \underbrace{(\xi_t^{-1}X)}_{=X} = \xi_t_*X.$$

Portanto, $\xi_t X = X$.

5.1.1 Composição e decomposição de fluxos estocásticos

Veremos agora uma propriedade interessante dos fluxos estocásticos: a composição e a decomposição. Essa característica fornece propriedades dinâmicas e geométricas de cada componente da fatoração do fluxo ξ_t e a fatoração $\xi_t = \eta_t \circ \psi_t$ é, em geral escolhida dependendo da dinâmica em que se quer estudar e da estrutura da variedade diferenciável. A seguir daremos apenas uma ideia sobre esse assunto.

Consideremos as seguintes EDEs em uma mesma variedade diferenciável compacta, para garantir que existe solução para todo $t > 0$ e que os campos vetoriais são C^∞ :

$$d\xi_t = X_0(\xi_t)dt + X_1(\xi_t) \circ dB_t; \quad (5.1.1.1)$$

$$d\eta_t = Y_0(\eta_t)dt + Y_1(\eta_t) \circ dB_t. \quad (5.1.1.2)$$

Teorema 52. Considerando as condições descritas acima o processo composição $\gamma_t = \xi_t \circ \eta_t$ é solução de

$$d\gamma_t = X_0(\gamma_t)dt + X_1(\gamma_t) \circ dB_t + \xi_t Y_0(\gamma_t)dt + \xi_t Y_1(\gamma_t) \circ dB_t.$$

Demonstração. De fato, aplicando a fórmula de Itô generalizada dada no Teorema 47 com

$$F_t(x) = f \circ \xi_t(x) \quad \text{e} \quad M_t = (\eta_t^1, \dots, \eta_t^d)$$

onde $f : M \rightarrow \mathbb{R}$ é uma função de classe C^3 e $(\eta_t^1, \dots, \eta_t^d)$ são as coordenadas da expressão η_t , temos

$$\begin{aligned} d(F_t(M_t)) &= (dF_t)(M_t) + \sum_{i=1}^d \frac{\partial F_t}{\partial x_i}(M_t) \circ dM_t^i \\ &= \sum_{k=0}^1 X_k(t) f(\xi_t \circ \eta_t) \circ dB_t^k + \sum_{k=0}^1 Y_k(t) (f \circ \xi_t)(\eta_t) \circ dB_t^k \\ &= \sum_{k=0}^1 X_k(t) f(\gamma_t) \circ dB_t^k + \sum_{k=0}^1 Y_k(t) (f \circ \xi_t)(\xi_t^{-1} \circ \gamma_t) \circ dB_t^k \\ &= \sum_{k=0}^1 X_k(t) f(\gamma_t) \circ dB_t^k + \sum_{k=0}^1 \xi_t Y_k(t) f(\gamma_t) \circ dB_t^k \\ &= X_0 f(\gamma_t) dt + X_1 f(\gamma_t) \circ dB_t + \xi_t Y_0 f(\gamma_t) dt + \xi_t Y_1 f(\gamma_t) \circ dB_t. \end{aligned}$$

⁵ver Apêndice C

Como a última igualdade vale para toda $f : M \rightarrow \mathbb{R}$ de classe C^∞ então

$$d\gamma_t = X_0(\gamma_t)dt + X_1(\gamma_t) \circ dB_t + \xi_{t*} Y_0(\gamma_t)dt + \xi_{t*} Y_1(\gamma_t) \circ dB_t.$$

■

Note que se os fluxos fossem apenas determinísticos, teríamos

$$d\gamma_t = X_0(\gamma_t)dt + \xi_{t*} Y_0(\gamma_t)dt.$$

No próximo resultado ainda estamos considerando as EDEs dadas pelas equação (5.1.1.1) e (5.1.1.2) com ξ_t e η_t seus respectivos fluxos soluções.

Teorema 53. *Considere um campo Z_k tal que vale a decomposição $X_k = Y_k + Z_k$, e seja ψ_t o processo solução da EDE*

$$d\psi_t = \eta_{t*}^{-1} Z_0(\psi_t)dt + \eta_{t*}^{-1} Z_1(\psi_t) \circ dB_t.$$

Então $\eta_t \circ \psi_t$ é solução da EDE (5.1.1.1), ou seja, temos a decomposição do fluxo estocástico $\xi_t = \eta_t \circ \psi_t$.

Demonstração. Considere a função $F_t(x) = f \circ \eta_t(x)$ e $M_t = (k_t^1, \dots, k_t^d)$. Então, aplicando a Fórmula de Itô generalizada dada no Teorema 47 com $y = \xi_t(x)$ obtemos

$$\begin{aligned} d(F_t(M_t)) &= (dF_t)(M_t) + \sum_{i=1}^d \frac{\partial F_t}{\partial x_i}(M_t) \circ dM_t^i \\ &= \sum_{j=0}^1 Y_j(t) f(\eta_t \circ k_t(x)) \circ dB_t^j + \sum_{j=0}^1 \eta_{t*}^{-1} Z_j(t) (f \circ \eta_t)(k_t(x)) \circ dB_t^j. \end{aligned}$$

Como

$$\eta_{t*}^{-1} Z_j(t) (f \circ \eta_t)(k_t(x)) = Z_j(t) f(\eta_t \circ k_t(x))$$

então

$$\begin{aligned} d(f \circ \eta_t \circ k_t) &= X_0(t) f(\eta_t \circ k_t(x)) dt + X_1(t) f(\eta_t \circ k_t(x)) \circ dB_t^k \\ d(f \circ \xi_t) &= X_0(t) f(\xi_t(x)) dt + X_1(t) f(\xi_t(x)) \circ dB_t. \end{aligned}$$

A última equação acima é válida para toda $f : M \rightarrow \mathbb{R}$ de classe C^∞ , e como

$$d(\xi_t) = X_0(\xi_t)dt + X_1(\xi_t) \circ dB_t$$

segue que pela unicidade de soluções para EDEs que $\xi_t = \eta_t \circ \psi_t$. ■

Com esse teorema conseguimos obter uma fórmula para calcular uma EDE reversa ξ_t^{-1} como mostra o próximo exemplo.

Exemplo 16. *Considere as equações*

$$d\eta_t = X_0(\eta_t)dt + X_1(\eta_t) \circ dB_t,$$

$$d\psi_t = \eta_{t*}^{-1}(-X_0)(\psi_t)dt + \eta_{t*}^{-1}(-X_1)(\psi_t) \circ dB_t.$$

Observe que aplicando o teorema acima temos que $\eta_t \circ \psi_t$ é solução de uma EDE com os campos nulos. Então

$$(\eta_t \circ \psi_t)(x) = x.$$

Ou seja, $\eta_t \circ \psi_t = id$ para todo t . Logo, $\psi_t = \eta_t^{-1}$. Isto é,

$$d\eta_t^{-1} = \eta_{t*}^{-1}(-X_0)(\eta_t^{-1})dt + \eta_{t*}^{-1}(-X_1)(\eta_t^{-1}) \circ dB_t.$$

Pela linearidade da diferencial η_{t*}^{-1} ainda podemos escrever

$$d\eta_t^{-1} = -\eta_{t*}^{-1}(X_0)(\eta_t^{-1})dt - \eta_{t*}^{-1}(X_1)(\eta_t^{-1}) \circ dB_t.$$

Exemplo 17. Considere uma EDE linear em \mathbb{R}^d :

$$d\xi_t = A\xi_t dt + CdB_t \quad (5.1.1.3)$$

onde A é um campo linear representado por uma matriz $d \times d$, C é um campo constante representado por uma matriz $d \times 1$ e B_t é o movimento browniano unidimensional. Podemos decompor a equação como

$$d\gamma_t = A\gamma_t dt, \quad d\eta_t = \gamma_t^{-1} C dB_t,$$

ou seja, se $\gamma_t(x)$ e $\eta_t(x)$ são soluções das equações acima, então $\xi_t = \gamma_t \circ \eta_t$ é uma solução para a equação (5.1.1.3), pois para multicampos a prova do Teorema 52 é análoga.

Observe que γ_t é solução de uma EDO determinística que é conhecida como $\gamma_t(x) = e^{At}x$, onde e^{At} é uma matriz e $x \in \mathbb{R}^d$. Para calcularmos γ_t^{-1} , devemos primeiro calcular γ_{t*} usando a matriz jacobiana associada a γ_t . Note que essa matriz é o próprio e^{At} , logo

$$\gamma_{t*} = e^{At}$$

e assim

$$\gamma_{t*}^{-1} = (\gamma_{t*})^{-1} = (e^{At})^{-1} = e^{-At}.$$

O resultado acima faz sentido, pois sabemos que

$$\gamma_{t*} \circ \gamma_{t*}^{-1} = id$$

e

$$e^{At} \circ e^{-At} = e^{At}(e^{-At}) = e^{(A-A)t} = e^{0 \cdot t} = id.$$

Portanto,

$$d\eta = e^{-At} C dB_t$$

Reescrevendo a equação acima na forma integral temos

$$\eta_t(x) = x + \int_0^t e^{-As} C dB_s.$$

Como $\xi_t = \gamma_t \circ \eta_t$ então

$$\xi_t = \gamma_t(\eta_t) = e^{At} \left(x + \int_0^t e^{-As} C dB_s \right).$$

Em Teoria de controle, considerando B_t um movimento browniano k -dimensional, os termos B_t^1, \dots, B_t^k e ξ_t são chamados de entrada e saída, respectivamente.

5.2 Fluxos agindo em formas diferenciais

Um fluxo estocástico de difeomorfismos determinado por uma equação diferencial estocástica em uma variedade diferenciável atua naturalmente em campos de tensores alternados, também chamados de formas diferenciais, e define um processo estocástico com valores nesses campos. Na seção anterior, obtemos uma fórmula de Itô para ξ_t agindo em campos de vetores. Nossa ideia agora é obter uma fórmula similar, mas para um fluxo agindo em formas diferenciais mais gerais. Entretanto, antes de considerarmos o caso geral, iniciaremos com uma 1-forma.

O leitor pouco familiarizado com 1-formas diferenciais pode ver mais detalhes no Apêndice D. A seguir iremos escrever apenas as propriedades básicas para estabelecer as notações.

Seja ξ um difeomorfismo de M e seja ξ_{x*} a diferencial de ξ , a qual é uma função linear de $T_x(M)$ em $T_{\xi(x)}(M)$. Denotaremos o dual da função por ξ_x^* que é uma função linear de $T_{\xi(x)}(M)^*$

(espaço cotangente) em $T_x(M)^*$, tal que para qualquer $\phi_{\xi(x)} \in T_{\xi(x)}(M)^*$ e $X_x \in T_x(M)$ a seguinte relação é satisfeita

$$\phi_{\xi(x)}(\xi_x^* X_x) = \xi_x^* \phi_{\xi(x)}(X_x). \quad (5.2.0.1)$$

Em outras palavras, temos que o dual da função é tomado da forma

$$\begin{aligned} \xi_x^* : T_{\xi(x)}(M)^* &\rightarrow T_x(M)^* \\ \phi_{\xi(x)} &\rightarrow \xi_x^* \phi_{\xi(x)} \end{aligned}$$

Seja θ uma 1-forma, denotamos por $\xi^* \theta$ uma 1-forma tal que

$$(\xi^* \theta)_x = \xi_x^* \theta_{\xi(x)}. \quad (5.2.0.2)$$

Seja X um campo vetorial completo e ξ_t seu fluxo. A derivada de Lie da uma 1-forma θ é definida por

$$L_X \theta = \lim_{t \rightarrow 0} \frac{1}{t} (\xi_t^* \theta - \theta).$$

Uma relação muito conhecida sobre a derivada de Lie de 1-formas é a seguinte

$$L_X \theta(Y) + \theta(L_X Y) = X[\theta(Y)]. \quad (5.2.0.3)$$

O próximo teorema fornece uma fórmula de Itô para a ação de fluxos estocásticos agindo em 1-forma, onde consideramos ξ_t o fluxo estocástico solução da seguinte EDE

$$d\xi_t = X_0(\xi_t)dt + X_1(\xi_t) \circ dB_t.$$

Teorema 54. *Assumindo a mesma condição de suavidade para os campos vetoriais X_0, X_1 dada no Teorema 49, temos que ξ_t^* satisfaz a seguinte fórmula para qualquer 1-forma θ de classe C^3*

$$\xi_t^* \theta = \theta + \int_0^t \xi_r^* L_{X_0(r)} \theta dr + \int_0^t \xi_r^* L_{X_1(r)} \theta \circ dB_r \quad (5.2.0.4)$$

Demonstração. Primeiro note que para um campo X em M , de acordo com a equação (5.1.0.2):

$$\xi_t^* X = X - \int_0^t L_{X_0(r)} \xi_r^* X dr - \int_0^t L_{X_1(r)} \xi_r^* X \circ dB_r$$

e portanto para uma 1-forma θ avaliada em um ponto y temos

$$\theta(\xi_t^* X)(y) = \theta X(y) - \int_0^t \theta(L_{X_0(r)} \xi_r^* X)(y) dr - \int_0^t \theta(L_{X_1(r)} \xi_r^* X)(y) \circ dB_r$$

Precisamos encontrar uma fórmula para $\theta(\xi_t^* X)$ no ponto $\xi_t(y)$, pois

$$(\xi_t^* \theta)(X)(y) = \theta(\xi_t^* X)(\xi_t(y)).$$

Para simplificar a expressão adotamos $dB_t^0 = dt$, como já feito antes. Considere a função $F_t(y) = \theta(\xi_t^* X)(y)$ de classe C^3 . Então, aplicando a Fórmula de Itô generalizada dada no Teorema 47 com $y = \xi_t(x)$ obtemos

$$F_t(\xi_t(x)) = F_s(\xi_t(x)) - \sum_{k=0}^1 \int_0^t \theta(L_{X_k} \xi_r^* X)(\xi_t(x)) \circ dB_r^k + \sum_{k=0}^1 \int_0^t X_k[\theta(\xi_r^* X)](\xi_t(x)) \circ dB_r^k \quad (5.2.0.5)$$

Note que, pela equação (5.2.0.3):

$$X_k[\theta(L_{X_k} \xi_r^* X)](\xi_t(x)) - \theta(L_{X_k} \xi_r^* X)(\xi_t(x)) = L_{X_k}[\theta \xi_r^* X](\xi_t(x))$$

e ainda, pela equação (5.2.0.1):

$$L_{X_k} \theta \xi_r^* X = \xi_r^* L_{X_k} \theta(X)$$

Portanto, a equação (5.2.0.5) resulta em

$$\theta(\xi_t^* X)(\xi_t(x)) = \theta X(\xi_t(x)) + \int_0^t \xi_r^* L_{X_0} \theta(X)(\xi_r(x)) dr + \int_0^t \xi_r^* L_{X_1} \theta(X)(\xi_r(x)) \circ dB_r$$

ou ainda,

$$\xi_r^* \theta X(\xi_t(x)) = \theta X(\xi_t(x)) + \int_0^t \xi_r^* L_{X_0} \theta(X)(\xi_r(x)) dr + \int_0^t \xi_r^* L_{X_1} \theta(X)(\xi_r(x)) \circ dB_r$$

onde na última igualdade usamos novamente a equação (5.2.0.1). ■

Agora, vamos desenvolver a mesma ideia para uma forma diferencial $K : T(M) \times \dots \times T(M) \rightarrow \mathbb{R}$. Assim, para cada x , K_x é uma forma multilinear no espaço produto

$$T_x(M) \times \dots \times T_x(M).$$

Então, dado os campos de vetores Y_1, \dots, Y_q ,

$$K_x(Y_1, \dots, Y_q) = K_x(Y_{1x}, \dots, Y_{qx}) \quad (5.2.0.6)$$

é um campo escalar.⁶

Seja ϕ um difeomorfismo de M e K uma forma diferencial, definimos uma forma diferencial $\phi^* K$ pela relação

$$(\phi^* K)_x(Y_1, \dots, Y_q) = K_{\phi(x)}(\phi_* Y_1, \dots, \phi_* Y_q). \quad (5.2.0.7)$$

A derivada de Lie de uma forma diferencial K com relação ao campo vetorial X é definido por

$$L_X K = \lim_{t \downarrow 0} \frac{1}{t} (\phi_t^* K - K).$$

Se K é uma forma diferencial igual a equação (5.2.0.6), então satisfaz

$$(L_X K)_x(Y_1, \dots, Y_q) = X(K_x)(Y_1, \dots, Y_q) - \sum_{l=1}^q K_x(Y_1, \dots, L_X Y_l, \dots, Y_q). \quad (5.2.0.8)$$

O próximo teorema indica a fórmula de Itô para ξ_t^* agindo em uma forma diferencial.

Teorema 55. *Suponha que o coeficiente X_1 da equação é um campo vetorial $C^{4,\alpha}$ para algum $\alpha > 0$ e X_0 é um campo vetorial $C^{3,\alpha}$ para algum $\alpha > 0$. Então ξ_t^* satisfaz a seguinte fórmula para alguma forma diferencial K de classe C^2*

$$\xi_t^* K = K + \int_0^t \xi_r^* \left(\frac{1}{2} L_{X_1}^2 + L_{X_0} \right) K dr + \int_0^t \xi_r^* L_{X_1} K \circ dB_r; \quad (5.2.0.9)$$

Suponha agora que o coeficiente X_1 da equação é um campo vetorial $C^{5,\alpha}$ para algum $\alpha > 0$ e X_0 é um campo vetorial $C^{4,\alpha}$ para algum $\alpha > 0$. Então ξ_t^ satisfaz a seguinte fórmula para alguma forma diferencial K de classe C^3*

$$\xi_t^* K = K + \int_0^t \xi_r^* L_{X_0(r)} K dr + \int_0^t \xi_r^* L_{X_1(r)} K \circ dB_r \quad (5.2.0.10)$$

⁶maiores detalhes sobre isso podem ser vistos no Apêndice D.

Demonstração. Por simplicidade, adotamos $dB_r^0 = dr$.

Provaremos primeiro a equação (5.2.0.10). Aplicando o Teorema 49,

$$K_x(\xi_{t*}Y_1, \dots, \xi_{t*}Y_q) = K_x(Y_1, \dots, Y_q) - \sum_{j=0}^m \left(- \sum_{l=1}^q \int_0^t K_x(\xi_{r*}Y_1, \dots, L_{X_j}\xi_{r*}Y_l, \dots, \xi_{r*}Y_q) \circ dB_r^j \right).$$

Seja

$$F_t(x) = K_x(\xi_{t*}Y_1, \dots, \xi_{t*}Y_q)$$

e aplicando a Fórmula de Itô generalizada para $F_t(\xi_t(x))$ temos

$$d(F_t(\xi_t)) = (dF_t)(\xi_t) + \sum_{i=1}^d \frac{\partial F_t}{\partial x_i}(\xi_t) \circ d\xi_t^i$$

e assim,

$$\begin{aligned} F_t(\xi_t(x)) - F_0(\xi_0(x)) &= \sum_{j=0}^1 \sum_{l=1}^q \int_0^t K_{\xi_r(x)}(\xi_{r*}Y_1, \dots, L_{X_j}\xi_{r*}Y_l, \dots, \xi_{r*}Y_q) \circ dB_r^j \\ &+ \sum_{j=0}^1 \int_0^t X_j(r)F_r(\xi_r) \circ dB_r^j \\ &\stackrel{(5.2.0.8)}{=} \sum_{j=0}^1 \int_0^t L_{X_j}K_{\xi_r(x)}(\xi_{r*}Y_1, \dots, \xi_{r*}Y_q) \circ dB_r^j \\ &= \sum_{j=0}^1 \int_0^t \xi_r^* L_{X_j}K_{(x)}(Y_1, \dots, Y_q) \circ dB_r^j. \end{aligned}$$

Com isso provamos a equação (5.2.0.10). Provaremos agora a equação (5.2.0.9).

Temos

$$\begin{aligned} \xi_t^* K - K &= \sum_{j=0}^1 \int_0^t \xi_r^* L_{X_j} K \circ dB_r^j \\ &\stackrel{(2.4.0.3)}{=} \sum_{j=0}^1 \int_0^t \xi_r^* L_{X_j} K dB_r^j + \int_0^t \frac{1}{2} d[\xi_r^* L_{X_1} K, B_r] \end{aligned}$$

Note que na segunda integral acima apareceu o termo de variação quadrática do semimartingale⁷ $\xi_r^* L_{X_1} K$, logo podemos decompor esse processo em uma soma de um martingale local com um processo de variação finita. Usando a equação (5.2.0.10) novamente e convertendo-a em uma integral de Itô

$$\begin{aligned} \xi_r^* L_{X_1} K - L_{X_1} K &= \sum_{j=0}^1 \int_0^t \xi_r^* L_{X_j} L_{X_1} K \circ dB_r^j \\ &= \underbrace{\sum_{j=0}^1 \int_0^t \xi_r^* L_{X_j} L_{X_1} K dB_r^j}_{\text{martingale}} + \underbrace{\int_0^t \frac{1}{2} d[\xi_r^* L_{X_1} K, B_r]}_A. \end{aligned}$$

um processo de variação finita

⁷ nas condições que trabalhamos o processo é um semimartingale

Substituindo $\xi_r^* L_{X_1} K$ na variação quadrática

$$\begin{aligned} \int_0^t \frac{1}{2} d \left[\sum_{j=0}^1 \int_0^t \xi_r^* L_{X_j} L_{X_1} K dB_r^j, B_r \right] &= \int_0^t \frac{1}{2} \sum_{j=0}^1 \int_0^t \xi_r^* L_{X_j} L_{X_1} K d[B_r^j, B_r] \\ &= \int_0^t \frac{1}{2} \xi_r^* L_{X_0} L_{X_1} K d[r, B_r] + \int_0^t \frac{1}{2} \xi_r^* L_{X_1} L_{X_1} K d[B_r, B_r] \\ &= \int_0^t \frac{1}{2} \xi_r^* L_{X_1}^2 K dr \end{aligned}$$

e assim

$$\begin{aligned} \xi_t^* K - K &= \sum_{j=0}^1 \int_0^t \xi_r^* L_{X_j} K dB_r^j + \int_0^t \frac{1}{2} \xi_r^* L_{X_1}^2 K dr \\ &= \int_0^t \xi_r^* L_{X_0} K dr + \int_0^t \xi_r^* L_{X_1} K dB_r + \int_0^t \frac{1}{2} \xi_r^* L_{X_1}^2 K dr \\ &= \int_0^t \xi_r^* \left(\frac{1}{2} L_{X_1}^2 + L_{X_0} \right) K dr + \int_0^t \xi_r^* L_{X_1} K dB_r. \end{aligned}$$

■

Definição 56. Dizemos que ξ_t preserva volume se $\xi_t^* K = K$.

Esta definição se justifica através da fórmula de mudança de variável. Mais precisamente, se K denota a forma volume e ξ_t é um fluxo (determinístico) então vale a fórmula

$$\text{vol}(\xi_t(U)) = \int_{\xi_t(U)} K = \int_U \xi_t^* K = \int_U K = \text{vol}(U),$$

onde U é uma região integrável. Esta fórmula também tem uma versão análoga para o caso estocástico como pode ser vista por exemplo em Ledesma; Silva (2017) [13] e Kinateder; McDonald (2002) [9].

Em vista da definição anterior, o fluxo estocástico ξ_t gerado pela EDE (5.2.0.10) preserva volume se, e somente se, $L_{X_k} K = 0$ para $k = 0, 1$. O próximo teorema confirma isso.

Teorema 57.

$$\xi_t^* K = K \Leftrightarrow L_{X_k} K = 0 \quad \text{para } k = 0, 1 \quad \omega - \text{quase sempre.}$$

Demonstração. Suponha primeiramente que $L_{X_k} K = 0$ para $k = 0, 1$. Então, segue da equação (5.2.0.9) ou da equação (5.2.0.10) que $\xi_t^* K = K$ (ω - quase sempre).

Suponha agora que $\xi_t^* K = K$, então pela equação (5.2.0.9)

$$\int_0^t \xi_r^* L_{X_1} K dB_r = - \int_0^t \xi_r^* \left(\frac{1}{2} L_{X_1}^2 + L_{X_0} \right) K(Y_1, \dots, Y_q) dr \quad \omega - \text{quase sempre.}$$

Observe que o lado esquerdo da igualdade acima depende de dB_r , logo é um martingale. Por outro lado, o lado direito da igualdade é um processo de variação finita. Portanto, ambos devem ser nulos⁸. Assim, a variação quadrática do martingale acima é

$$\int_0^t |\xi_r^* L_{X_1} K(Y_1, \dots, Y_q)|^2 dr = 0$$

para todo $Y_1, \dots, Y_q \in T_x M$. Logo,

$$\xi_r^* L_{X_1} K = 0$$

e portanto

$$L_{X_1} K = 0,$$

já que ξ_{r*} é isomorfismo, pois ξ_r é difeomorfismo. E ainda, como o processo de variação finita acima é nulo, segue que $L_{X_0} K = 0$. ■

⁸um processo martingale é de variação finita só se o martingale é nulo.

5.2.1 Exemplos na esfera

Nosso objetivo é analisar o comportamento de um fluxo que age em regiões limitadas sobre uma esfera, utilizando a teoria apresentada anteriormente. Para isso, precisamos determinar campos de vetores que pertencem a esfera e uma maneira possível de construir um campo vetorial em S^{n-1} é a seguinte.

Seja A uma matriz quadrada $n \times n$. Então a função

$$f : S^{n-1} \rightarrow \mathbb{R}^d$$

definida por

$$f(x) = Ax$$

é um campo de vetor tangente à esfera se, e somente se, o produto interno

$$\langle x, Ax \rangle = 0 \text{ para todo } x \in S^{n-1}.$$

Isso, por sua vez, é equivalente ao requisito de que A seja uma matriz antissimétrica, ou seja,

$$A + A^t = 0,$$

onde A^t é a transposta da matriz A . De fato, suponha primeiro que $\langle x, Ax \rangle = 0$ para todo $x \in S^{n-1}$, ou equivalentemente, para todo $x \in \mathbb{R}^d$. Então,

$$\begin{aligned} \langle x, (A + A^t)y \rangle &= \langle x, Ay \rangle + \langle Ax, y \rangle \\ &= \langle x, Ax \rangle + \langle x, Ay \rangle + \langle Ax, y \rangle + \langle Ay, y \rangle \\ &= \langle x, Ax \rangle + \langle x, Ay \rangle + \langle y, Ax \rangle + \langle y, Ay \rangle \\ &= \langle x, A(x+y) \rangle + \langle y, A(x+y) \rangle \\ &= \langle x+y, A(x+y) \rangle \\ &= 0, \end{aligned}$$

para todo $x, y \in \mathbb{R}^d$, então $A + A^t = 0$. Por outro lado, se $A + A^t = 0$ então

$$\begin{aligned} \langle x, Ax \rangle &= \frac{1}{2} \langle x, Ax \rangle + \frac{1}{2} \langle x, Ax \rangle \\ &= \frac{1}{2} \left(\langle x, Ax \rangle + \langle Ax, x \rangle \right) \\ &= \frac{1}{2} \left(\langle x, Ax \rangle + \langle x, A^t x \rangle \right) \\ &= \frac{1}{2} \left(\langle x, (A + A^t)x \rangle \right) \\ &= 0. \end{aligned}$$

Dizemos que o campo vetorial X obtido desta forma é um campo vetorial linear.

Exemplo 18. Vamos determinar um campo vetorial linear sobre a esfera. Consideremos a seguinte matriz anti-simétrica

$$A = \begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}.$$

Então,

$$X(x, y, z) = A \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = (-y + z, x, -x).$$

Note que, $\langle x, Ax \rangle = 0$. De fato,

$$\begin{aligned} \langle (x, y, z), (-y + z, x, -x) \rangle &= x(-y + z) + yx - zx \\ &= -yx + zx + yx - zx \\ &= 0. \end{aligned}$$

Exemplo 19. Considere $S^2 \subset \mathbb{R}^3$ onde

$$S^2 \subset \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\}$$

e a forma volume⁹

$$k = xdy \wedge dz + ydz \wedge dx + zdx \wedge dy.$$

Assim como no Exemplo 18, vamos determinar um campo linear a S^2 . Consideremos a seguinte matriz anti-simétrica

$$A = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Então,

$$X_1(x, y, z) = A \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = (-y, x, 0).$$

Note que, $\langle x, Ax \rangle = 0$. De fato,

$$\langle (x, y, z), (-y, x, 0) \rangle = -xy + xy + 0 = 0$$

A figura a seguir ilustra o campo X na esfera

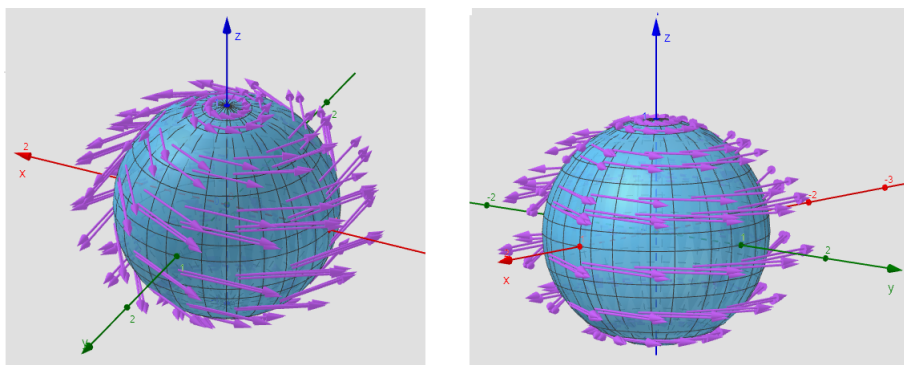


Figura 5.2: Ilustração do campo X_1 na esfera.

Pela figura, devido a rotação do campo em torno do eixo z , intuitivamente vemos que se tivermos nesse caso uma EDE da forma

$$d\psi_t = X_1(\psi_t) \circ dB_t,$$

então o fluxo estocástico ψ_t preserva área de regiões limitadas sobre a superfície esférica.

Nosso objetivo agora é mostrar que de fato $\psi_t^* k = k$ (ω -quase sempre). Para isto começamos observando que

$$L_{X_1} k = \underbrace{L_{X_1}(xdy \wedge dz)}_I + \underbrace{L_{X_1}(ydz \wedge dx)}_II + \underbrace{L_{X_1}(zdx \wedge dy)}_III \quad (5.2.1.1)$$

⁹Falar que k é uma forma volume é um abuso de linguagem, a expressão mais precisa é área, pois k é uma 2-forma em S^2 . Mais detalhes sobre isso podem ser vistos em Apêndice D.

Calculando I^{10} :

$$\begin{aligned}
 L_{X_1}(xdy \wedge dz) &= L_{X_1}(x)dy \wedge dz + x(L_{X_1}dy) \wedge dz + xdy \wedge (L_{X_1}dz) \\
 &= L_{X_1}(x)dy \wedge dz + xd(L_{X_1}y) \wedge dz + xdy \wedge d(L_{X_1}z) \\
 &= -ydy \wedge dz + xdx \wedge dz + xdy \wedge 0 \\
 &= -ydy \wedge dz + xdx \wedge dz.
 \end{aligned}$$

Calculando II :

$$\begin{aligned}
 L_{X_1}(ydz \wedge dx) &= L_{X_1}(y)dz \wedge dx + y(L_{X_1}dz) \wedge dx + ydz \wedge (L_{X_1}dx) \\
 &= L_{X_1}(y)dz \wedge dx + yd(L_{X_1}z) \wedge dx + ydz \wedge d(L_{X_1}x) \\
 &= xdz \wedge dx + y0 \wedge dx + ydz \wedge d(-y) \\
 &= xdz \wedge dx - ydz \wedge dy.
 \end{aligned}$$

Calculando III :

$$\begin{aligned}
 L_{X_1}(zdx \wedge dy) &= L_{X_1}(z)dx \wedge dy + z(L_{X_1}dx) \wedge dy + zdx \wedge (L_{X_1}dy) \\
 &= L_{X_1}(z)dx \wedge dy + zd(L_{X_1}x) \wedge dy + zdx \wedge d(L_{X_1}y) \\
 &= 0dx \wedge dy + zd(-y) \wedge dy + zdx \wedge d(x) \\
 &= \underbrace{-zdy \wedge dy}_{=0} + \underbrace{zdx \wedge dx}_{=0} \\
 &= 0.
 \end{aligned}$$

Logo,

$$\begin{aligned}
 L_{X_1}\mathbf{k} &= L_{X_1}(xdy \wedge dz) + L_{X_1}(ydz \wedge dx) + L_{X_1}(zdx \wedge dy) \\
 &= -ydy \wedge dz + xdx \wedge dz + xdz \wedge dx - ydz \wedge dy \\
 &= -ydy \wedge dz + xdx \wedge dz - xdx \wedge dz + ydy \wedge dz \\
 &= 0.
 \end{aligned}$$

E assim, temos que $\Psi_t^*\mathbf{k} = \mathbf{k}$, pois $L_{X_1}\mathbf{k} = 0$. Portanto, Ψ_t preserva volume. Note que, pela Fórmula de Cartan,

$$L_{X_1}\mathbf{k} = (\operatorname{div}X_1)\mathbf{k} \Rightarrow \operatorname{div}X_1 = 0,$$

Ou seja, o campo de vetores X_1 da EDE tem divergente nulo, logo assim como para fluxos determinísticos, o fluxo estocástico gerado por X_1 também preserva volume, visto que podemos interpretar o divergente como uma taxa de variação de área.¹¹

No próximo exemplo consideramos um campo vetorial não linear Y e queremos ver o comportamento do fluxo sobre a esfera.

Exemplo 20. Seja

$$S^2 = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\}$$

e o campo normal unitário a S^2 dado por

$$\begin{aligned}
 \eta(x, y, z) &= \frac{1}{\sqrt{x^2 + y^2 + z^2}} \left(x \frac{\partial}{\partial x} + y \frac{\partial}{\partial y} + z \frac{\partial}{\partial z} \right) \\
 &= x \frac{\partial}{\partial x} + y \frac{\partial}{\partial y} + z \frac{\partial}{\partial z}.
 \end{aligned}$$

¹⁰A regra para calcular a derivada de Lie de k -formas pode ser vista na Proposição 75

¹¹Se $\operatorname{div}Y > 0$, devemos esperar que a área esteja aumentando. Por outro lado, se $\operatorname{div}Y < 0$, esperamos que a área esteja diminuindo. E se $\operatorname{div}Y = 0$, então a área é invariante, ou seja, ela não varia. Mais detalhes sobre isso podem ver vistos por exemplo em Guidorizzi (2013) [6].

Considere também o seguinte campo de vetores

$$Y = -zx \frac{\partial}{\partial x} - zy \frac{\partial}{\partial y} + (1 - z^2) \frac{\partial}{\partial z}$$

cuja representação pode ser vista na figura a seguir.

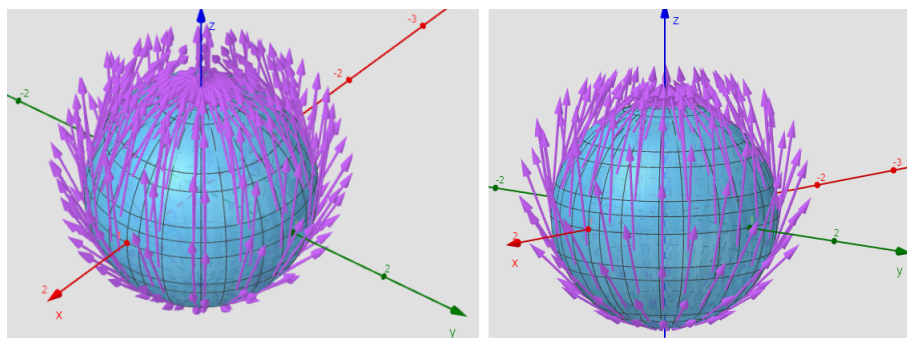


Figura 5.3: Ilustração do campo Y na esfera.

Note que

$$\begin{aligned} \langle Y(x, y, z), \eta(x, y, z) \rangle &= \langle (-zx, -zy, 1 - z^2), (x, y, z) \rangle \\ &= (-zx)x + (-zy)y + (1 - z^2)z \\ &= -zx^2 - zy^2 + z - z^3 \\ &= -z \underbrace{(x^2 + y^2 + z^2)}_{=1} + z \\ &= -z + z \\ &= 0. \end{aligned}$$

e portanto Y é de fato um campo de vetores sobre a esfera.

Considere novamente a forma volume na esfera dada

$$\mathbf{k} = xdy \wedge dz + ydz \wedge dx + zdx \wedge dy$$

e vamos calcular a derivada de Lie de \mathbf{k} com relação a Y para verificar se $L_Y \mathbf{k} = 0$ e assim conseguir determinar o comportamento do campo de vetores Y .

$$L_Y \mathbf{k} = \underbrace{L_Y(xdy \wedge dz)}_I + \underbrace{L_Y(ydz \wedge dx)}_{II} + \underbrace{L_Y(zdx \wedge dy)}_{III} \quad (5.2.1.2)$$

Calculando I:

$$\begin{aligned} L_Y(xdy \wedge dz) &= L_Y(x)dy \wedge dz + x(L_Y dy) \wedge dz + xdy \wedge (L_Y dz) \\ &= L_Y(x)dy \wedge dz + xd(L_Y y) \wedge dz + xdy \wedge d(L_Y z) \\ &= -zxdy \wedge dz + xd(-zy) \wedge dz + xdy \wedge d(1 - z^2) \\ &= -zxdy \wedge dz - zxdy \wedge dz - yx \underbrace{dz \wedge dz}_{=0} - 2xzdy \wedge dz \\ &= -4xzdy \wedge dz. \end{aligned}$$

Calculando II:

$$\begin{aligned}
L_Y(ydz \wedge dx) &= L_Y(y)dz \wedge dx + y(L_Y dz) \wedge dx + ydz \wedge (L_Y dx) \\
&= L_Y(y)dz \wedge dx + yd(L_Y z) \wedge dx + ydz \wedge d(L_Y x) \\
&= -zydz \wedge dx + yd(1 - z^2) \wedge dx + ydz \wedge d(-zx) \\
&= -zydz \wedge dx - 2zydz \wedge dx - zydz \wedge dx - \underbrace{xydz \wedge dz}_{=0} \\
&= -4zydz \wedge dx.
\end{aligned}$$

Calculando III:

$$\begin{aligned}
L_Y(zdx \wedge dy) &= L_Y(z)dx \wedge dy + z(L_Y dx) \wedge dy + zdx \wedge (L_Y dy) \\
&= L_Y(z)dx \wedge dy + zd(L_Y x) \wedge dy + zdx \wedge d(L_Y y) \\
&= (1 - z^2)dx \wedge dy + zd(-zx) \wedge dy + zdx \wedge d(-zy) \\
&= (1 - z^2)dx \wedge dy - z^2 dx \wedge dy - zxdz \wedge dy - z^2 dx \wedge dy - zydx \wedge dz \\
&= 1dx \wedge dy - z^2 dx \wedge dy - 2z^2 dx \wedge dy - zxdz \wedge dy - zydx \wedge dz \\
&= 1dx \wedge dy - 3z^2 dx \wedge dy - zxdz \wedge dy - zydx \wedge dz.
\end{aligned}$$

Logo,

$$\begin{aligned}
L_Y \mathbf{k} &= L_Y(xdy \wedge dz) + L_Y(ydz \wedge dx) + L_Y(zdx \wedge dy) \\
&= -4xzdy \wedge dz - 4zydz \wedge dx + 1dx \wedge dy - 3z^2 dx \wedge dy - zxdz \wedge dy - zydx \wedge dz \\
&= -4xzdy \wedge dz - 4zydz \wedge dx + 1dx \wedge dy - 3z^2 dx \wedge dy + zxdy \wedge dz + zydz \wedge dx \\
&= -3xzdy \wedge dz - 3zydz \wedge dx + 1dx \wedge dy - 3z^2 dx \wedge dy \\
&= -3xzdy \wedge dz - 3zydz \wedge dx + (1 - 3z^2)dx \wedge dy.
\end{aligned}$$

Como $L_Y \mathbf{k} \neq 0$, temos que se considerarmos nesse caso uma EDE da forma

$$d\xi_t = Y(\xi_t) \circ dB_t,$$

então o fluxo estocástico ξ_t não preserva área de regiões limitadas sobre a superfície esférica, ou seja, não preserva volume.

Vale observar que como o fluxo não preserva área, a função divergente do campo Y é diferente de zero, visto que (como mencionado antes) podemos interpretar o divergente como uma taxa de variação de área. E ainda, pela Fórmula de Cartan temos

$$L_Y \mathbf{k} = (\operatorname{div} Y) \mathbf{k}.$$

e assim, a função $\operatorname{div} Y$ não é nula. Por curiosidade, vamos agora determinar a função

$$\operatorname{div} Y : S^2 \rightarrow \mathbb{R}.$$

Primeiramente, note que

$$\begin{aligned}
L_Y \mathbf{k} &= -3z \underbrace{(xdy \wedge dz + ydz \wedge dx + zdx \wedge dy)}_{=\mathbf{k}} + 1dx \wedge dy \\
&= -3z \cdot \mathbf{k} + 1dx \wedge dy.
\end{aligned}$$

e assim, precisamos determinar $dx \wedge dy$ em termos da forma \mathbf{k} .

Observe que a forma volume $dx \wedge dy \wedge dz$ em \mathbb{R}^3 restrita a S^2 é identicamente nula. De fato, se $v_1, v_2, v_3 \in T_p S^2$ então

$$(dx \wedge dy \wedge dz)(v_1, v_2, v_3) \stackrel{(D.4.0.1)}{=} \det \begin{bmatrix} dx(v_1) & dx(v_2) & dx(v_3) \\ dy(v_1) & dy(v_2) & dy(v_3) \\ dz(v_1) & dz(v_2) & dz(v_3) \end{bmatrix}$$

Note que cada coluna pode ser vista como cada vetor, e ainda, como v_1, v_2 e v_3 estão no mesmo plano tangente ao ponto p , o determinante da matriz será zero¹². Logo,

$$\det \begin{bmatrix} dx(v_1) & dx(v_2) & dx(v_3) \\ dy(v_1) & dy(v_2) & dy(v_3) \\ dz(v_1) & dz(v_2) & dz(v_3) \end{bmatrix} = \det \begin{bmatrix} | & | & | \\ v_1 & v_2 & v_3 \\ | & | & | \end{bmatrix} = 0.$$

E assim, temos que para o produto interior¹³

$$i_Y(dx \wedge dy \wedge dz) = 0.$$

Por outro lado,

$$\begin{aligned} 0 &= i_Y(dx \wedge dy \wedge dz) \\ &= dx(Y) \wedge dy \wedge dz - dx \wedge dy(Y) \wedge dz + dx \wedge dy \wedge dz(Y) \\ &= -zxdy \wedge dz + zydx \wedge dz + (1 - z^2)dx \wedge dy \\ &= -z(xdy \wedge dz + ydx \wedge dz + zdx \wedge dy) + dx \wedge dy \\ &= -z\mathbf{k} + dx \wedge dy \\ dx \wedge dy &= z \cdot \mathbf{k}. \end{aligned}$$

Dessa forma,

$$L_Y \mathbf{k} = -3z \cdot \mathbf{k} + z \cdot \mathbf{k} = -2z \cdot \mathbf{k}$$

e portanto,

$$(\operatorname{div} Y)(x, y, z) = -2z.$$

Dessa forma, como $\operatorname{div} Y = -2z$ e $-1 \leq z \leq 1$ temos, que de $-1 \leq z < 0$ a área é aumentada ($\operatorname{div} Y > 0$) e de que de $0 < z \leq 1$ a área diminui ($\operatorname{div} Y < 0$).¹⁴

Temos então um exemplo de campo de vetores em S^2 em que $L_Y \mathbf{k}$ não é identicamente nula. Neste caso, se temos por exemplo uma EDE da forma

$$d\psi_t = Y(\psi_t) \circ dB_t,$$

então o fluxo estocástico ψ_t não preserva volume na esfera.

5.3 Considerações finais

Neste trabalho vimos que em alguns modelos matemáticos que descrevem a evolução de um fenômeno na natureza com relação ao tempo, aparecem incertezas decorrentes de situações que ocorrem no ambiente. Tais incertezas adicionam uma perturbação ao modelo inicial, que chamamos de ruído branco. Quando a evolução é em tempo contínuo e se permite essa perturbação,

¹²um dos vetores é combinação linear dos outros dois, e assim v_1, v_2 e v_3 são linearmente dependentes.

¹³para mais detalhes sobre isso sugerimos consultar Lee(2012, p.358) [15] e Spivak (1999, p. 228) [25]

¹⁴Essa explicação é com base no caso determinístico. O caso estocástico é bem semelhante.

uma modelagem possível é obtida através de equações diferenciais estocásticas de Stratonovich da forma

$$d\xi_t = X_0(\xi_t)dt + X_1(\xi_t) \circ dB_t \quad (5.3.0.1)$$

com condição inicial $\xi_0 = x$. X_0, X_1 são campos diferenciáveis em uma variedade diferenciável e B_t é o movimento browniano.

Uma solução para a equação acima é um processo estocástico adaptado ξ_t na variedade diferenciável M se, para toda função $f \in C^3(M)$ satisfaz

$$\begin{aligned} f(\xi_t) &= f(x) + \int_0^t X_0 f(\xi_r) dr + \int_0^t X_1 f(\xi_r) \circ dB_r \\ &= f(x) + \int_0^t \left(X_0 + \frac{1}{2} X_1^2 \right) f(\xi_r) dr + \int_0^t X_1 f(\xi_r) dB_r, \end{aligned}$$

onde $\circ dB_r$ e dB_r denotam as integrais de Stratonovich e Itô, respectivamente. Trabalhamos com uma EDE de Stratonovich, pois ela modela uma grande classe de sistemas dinâmicos em variedades, uma vez que a sua Fórmula de Itô se ajusta melhor às ferramentas de geometria diferencial. No entanto, também podemos trabalhar com uma EDE de Itô, dependendo dos objetivos.

Antes de estudarmos os fluxos estocásticos em variedades diferenciáveis, fizemos uma revisão geral de análise estocástica em espaços euclidianos, incluindo a existência e unicidade de soluções para EDE. Em seguida, estudamos EDE em variedades e verificamos que, sob certas condições nos campos de vetores, uma EDE da forma descrita na equação (5.3.0.1) gera um fluxo estocástico de difeomorfismo sobre uma variedade. E finalmente, como objetivo principal, estudamos a Fórmula de Itô para estes fluxos agindo em campos de vetores e tensores alternados sobre uma variedade. Ou seja, sendo ξ_t^* a diferencial do fluxo estocástico de difeomorfismo $\xi_t : M \rightarrow M$ e $\xi_t^* : T_{\xi_t(x)} M^* \rightarrow T_x M^*$ seu dual, X um campo de vetores e K um campo de tensores alternados em M , estudamos a fórmula de Itô para o novo campo $\xi_t^* X$ e $\xi_t^* K$ seguindo Kunita (1982) [11].

Mais precisamente, assumindo que ξ_t define um fluxo estocástico de difeomorfismo (ω -quase sempre) associado a uma EDE e que os campos X_0, X_1 assumem certas condições suficientes com relação a diferenciabilidade, temos a seguinte Fórmula de Itô para a ação de fluxos estocásticos em campos de vetores:

$$\xi_t^* X = X - \int_0^t L_{X_0(r)}(\xi_r^* X) dr - \int_0^t L_{X_1(r)}(\xi_r^* X) \circ dB_r \quad (5.3.0.2)$$

onde $L_{X_i(r)}(\xi_r^* X)$, $i = 0, 1$ denota a derivada de Lie do campo $\xi_r^* X$ com relação a $X_i(r)$.

No entanto, a fórmula acima não é muito “útil”, uma vez que apresenta a derivada de Lie do campo $\xi_r^* X$ com relação a $X_i(r)$, e assim, para usá-la devemos calcular ξ_r^* que não é uma tarefa fácil, já que muitas vezes nem conhecemos ξ_r , imagina sua diferencial. E se conhecermos, calcular a diferencial pode ser bem trabalhoso. Portanto, quando queremos saber a ação de fluxos estocásticos em campos de vetores usamos uma fórmula bem similar a equação (5.3.0.2) considerando ξ_t^{-1} dada por

$$\xi_t^{-1} X = X + \int_0^t \xi_r^{-1} L_{X_0(r)} X dr + \int_0^t \xi_r^{-1} L_{X_1(r)} X \circ dB_r \quad (5.3.0.3)$$

onde $L_{X_i(r)} X$, $i = 0, 1$ denota a derivada de Lie do campo X com relação a $X_i(r)$, duas informações conhecidas e portanto mais fácil de trabalhar.

A fórmula (5.3.0.2) pode ser estendida para a ação de fluxos estocásticos sobre um campo de tensores K , em particular para as formas diferenciais k , sobre uma variedade dada por

$$\xi_t^* k = k + \int_0^t \xi_r^* L_{X_0(r)} k dr + \int_0^t \xi_r^* L_{X_1(r)} k \circ dB_r,$$

onde $L_{X_i(r)}\mathbf{k}$, $i = 0, 1$ denota a derivada de Lie da forma diferencial \mathbf{k} com relação a $X_i(r)$.

Além disso, vimos que uma consequência direta e interessante desta fórmula é que sob certas condições nos campos de vetores um fluxo estocástico de difeomorfismo associado a uma EDE da forma (5.3.0.1) preserva uma forma volume sobre uma variedade diferenciável M , em particular, em uma esfera.

Esta abordagem pode ser estendida e aplicada em diversos contextos em que pretende-se estudar a dinâmica de fluxos estocásticos tendo em vista a geometria da variedade, como pode ser visto, por exemplo, em Liao (2000) [16], Ledesma; Silva (2017) [13] e Silva; Catuogno; Ruffino (2013) [23], abrindo uma nova possibilidade de pesquisa para seguir em um programa de doutorado.

Além disso, também estudamos a decomposição de fluxos estocásticos. Resumidamente, a ideia consiste em considerar um fluxo original ξ_t que é solução de uma EDE

$$d\xi_t = X_0(\xi_t)dt + X_1(\xi_t) \circ dB_t;$$

e supor que os campos X_k , $k = 0, 1$ possuem uma decomposição da forma

$$X_k = Y_k + Z_k.$$

Tomando η_t como sendo o fluxo solução da EDE

$$d\eta_t = Y_0(\eta_t)dt + Y_1(\eta_t) \circ dB_t.$$

e ψ_t como sendo o fluxo solução de

$$d\psi_t = \eta_{t*}^{-1}Z_0(\psi_t)dt + \eta_{t*}^{-1}Z_1(\psi_t) \circ dB_t.$$

segue que o fluxo original pode ser fatorado em uma composição de processos estocásticos da forma

$$\xi_t = \eta_t \circ \psi_t.$$

Logo, temos uma outra possibilidade de pesquisa: a decomposição obtida fornece propriedades dinâmicas e geométricas de cada componente da fatoração do fluxo ξ_t e a fatoração $\xi_t = \eta_t \circ \psi_t$ é, em geral escolhida dependendo da dinâmica em que se quer estudar e da estrutura da variedade diferenciável.

Referências

- [1] L. J. S. Allen. *An introduction to stochastic processes with applications to biology*. CRC Press, Texas, 2 edition, 2010. Citado na(s) página(s): 17, 21
- [2] Z. Brzezniak; T. Zastawniak. *Basic stochastic processes: a course through exercises*. Undergraduate Mathematics Series. Springer, London, 1999. Citado na(s) página(s): 17, 21
- [3] M. P. Carmo. *Differential forms and applications*. Springer Berlin Heidelberg, Heidelberg, 1994. Citado na(s) página(s): 115, 122
- [4] C. I. Doering; A. O. Lopes. *Equações diferenciais ordinárias*. IMPA, Rio de Janeiro, 3 edition, 2008. Citado na(s) página(s): 15
- [5] L. C. Evans. *An introduction to stochastic differential equations*. American Mathematical Society, Berkeley, 2012. Citado na(s) página(s): 17, 22, 35, 41, 50
- [6] H. L. Guidorizzi. *Um curso de cálculo, vol. 3*. Livros Tecnicos e Cientificos (LTC), Rio de Janeiro, 5 edition, 2013. Citado na(s) página(s): 96
- [7] N. Ikeda; S. Watanabe. *Stochastic differential equations and diffusion processes*, volume 24. Kodansha, Japan, 2 edition, 1989. Citado na(s) página(s): 59
- [8] K. Itô. Stochastic integral. *Proceedings of the Imperial Academy*, v. 20(n. 8):p. 519 – 524, 1944. Citado na(s) página(s): 41
- [9] K. Kinatered; P. McDonald. An ito formula for domain-valued processes driven by stochastic flows. *Probability theory and related fields*, v. 124(n. 1):p. 73 – 99, 2002. Citado na(s) página(s): 93
- [10] F. C. Klebaner. *Introduction to stochastic calculus with applications*. Imperial College Press, London, 2 edition, 2005. Citado na(s) página(s): 17, 21, 25, 29, 44
- [11] H. Kunita. *Stochastic differential equations and stochastic flow of diffeomorphisms*, volume 1097, pages 143–303. Springer, New York, 1982. Citado na(s) página(s): 9, 11, 15, 68, 69, 70, 71, 73, 75, 79, 84, 100, 105
- [12] H. Kuo. *Introduction to stochastic integration*. Springer, New York, 2005. Citado na(s) página(s): 25, 37, 38, 39, 40, 41, 42
- [13] D. S. Ledesma; F. B. Silva. Invariance of 0-currents under diffusions. *Stochastics and Dynamics*, v. 17(n. 2):p. 1750010, 2017. Citado na(s) página(s): 93, 101
- [14] J. M. Lee. *Introduction to topological manifolds*. Graduate Texts in Mathematics Series. Springer, New York, 2000. Citado na(s) página(s): 107

- [15] J. M. Lee. *Introduction to smooth manifolds*. Graduate Texts in Mathematics Series. Springer, New York, 2 edition, 2012. Citado na(s) página(s): 77, 99, 109, 110, 111, 112, 114, 115, 120, 123, 124
- [16] M. Liao. Decomposition of stochastic flows and lyapunov exponents. *Probability theory and related fields*, v. 117(n. 4):p. 589–607, 2000. Citado na(s) página(s): 101
- [17] E. L. Lima. *Análise real, v.3: Análise vetorial*. IMPA, Rio de Janeiro, 2007. Citado na(s) página(s): 115, 117
- [18] E. L. Lima. *Análise real, v.2: Funções de n variáveis*. IMPA, Rio de Janeiro, 4 edition, 2009. Citado na(s) página(s): 124
- [19] B. Oksendal. *Stochastic differential equations: an introduction with applications*. Springer-Verlag, New York, 5 edition, 1998. Citado na(s) página(s): 17, 20, 21, 25, 36, 39, 42
- [20] M. A. Pinsky; V. Wihstutz. *Diffusion processes and related problems in analysis: stochastic flows*, volume 22 of *Progress in Probability*. Birkhäuser Basel, Boston, 1992. Citado na(s) página(s): 72
- [21] D. Revuz; M. Yor. *Continuous martingales and brownian motion*. Springer-Verlag Berlin Heidelberg, New York, 3 edition, 1999. Citado na(s) página(s): 30
- [22] P. R. C. Ruffino. *Uma iniciação aos sistemas dinâmicos estocásticos*. IMPA, Rio de Janeiro, 2 edition, 2009. Citado na(s) página(s): 29
- [23] F. B. Silva; P. J. Catuogno; P. R. C. Ruffino. Decomposition of stochastic flows in manifolds with complementary distributions. *Stochastics and Dynamics*, v. 13(n. 4):p. 1350009, 2013. Citado na(s) página(s): 101
- [24] D. Sondermann. *Introduction to stochastic calculus for finance: A new didactic approach*. Springer-Verlag Berlin Heidelberg, Heidelberg, 2006. Citado na(s) página(s): 29
- [25] M. Spivak. *A comprehensive introduction to differential geometry*. Publish or Perish, Houston, 1999. Citado na(s) página(s): 99, 115, 123, 124

Campos aleatórios que são Hölder contínuos

Considere dois processos estocásticos X_t e Y_t no espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$. Dize-mos que X_t é uma versão (ou modificação) de Y_t se

$$\mathbb{P}(\omega; X_t(\omega) = Y_t(\omega)) = 1, \quad \forall t.$$

Exemplo 21. Seja $Y_t(\omega) = 0$ para todo $t, 0 \leq t \leq 1$ e ω uma variável aleatória distribuída uniformemente em $[0, 1]$. Considere

$$X_t(\omega) = \begin{cases} 1, & \text{se } t = \omega \\ 0, & \text{se } t \neq \omega \end{cases}$$

Então, para algum t fixo,

$$\mathbb{P}(X_t(\omega) \neq 0) = \mathbb{P}(\omega = t) = 0,$$

consequentemente,

$$\mathbb{P}(X_t(\omega) = 0) = \mathbb{P}(\omega \neq t) = 1$$

Portanto, todas as distribuições unidimensionais de $X_t(\omega)$ e $Y_t(\omega)$ são iguais, pois $\mathbb{P}(Y_t(\omega) = 0) = 1$. Da mesma forma, todas as distribuições finito dimensionais de X e Y são iguais. No entanto, os caminhos de amostra do processo Y e X são diferentes, pois as funções $Y_t, 0 \leq t \leq 1$ são contínuas em t , enquanto que todo caminho de amostra $X_t, 0 \leq t \leq 1$ tem um salto no ponto ω . Além disso, X_t é uma modificação de Y_t , isto é,

$$\mathbb{P}(\omega; X_t(\omega) = Y_t(\omega)) = 1, \quad \forall t, \quad 0 \leq t \leq 1.$$

O resultado a seguir é útil para mostrar que sob certas condições um processo estocástico possui uma versão contínua.

Teorema 58. (Teorema da continuidade de Kolmogorov) Suponha que o processo $X_t, t \geq 0$ satisfaça a seguinte condição: Para todo $T > 0$, existem constantes positivas α, β, D tais que

$$\mathbb{E}[|X_t - X_s|^\alpha] \leq D|t - s|^{1+\beta}; \quad 0 \leq s, t \leq T.$$

Então existe uma versão contínua de X .

Introduziremos agora um critério para campos aleatórios que são Hölder contínuos, que é uma generalização do teorema acima. Isso nos fornecerá outro método para estudar a regularidade das integrais estocásticas em relação à um determinado parâmetro. Para isso, precisamos do seguinte resultado cuja prova pode ser vista em Kunita (1982, p.196) [11].

Proposição 59. *Seja $X_\lambda(\omega)$ um campo aleatório com valor real com o parâmetro $\lambda = (\lambda_1, \dots, \lambda_d) \in \Lambda = [0, 1]^d$. Suponha que haja constantes $\gamma > 0, a_i > d, i = 1, \dots, d$ e $C > 0$ tal que*

$$\mathbb{E}[|X_\lambda - X_\mu|^\gamma] \leq C \sum_{i=1}^d |\lambda_i - \mu_i|^{a_i}, \quad \forall \lambda, \mu \in \Lambda. \quad (\text{A.0.0.1})$$

Então, X_λ tem uma modificação contínua \tilde{X}_λ .

Seja $\beta_i, i = 1, \dots, d$ números positivos arbitrários menores que $a_i(a_0 - d) \times a_0^{-1} \gamma^{-1}, i = 1, \dots, d$ respectivamente, onde $a_0 = \min\{a_1, a_2, \dots, a_d\}$. Então, para quase todo ω , existe um número inteiro positivo $m_0(\omega)$ e constante positiva K tal que

$$|\tilde{X}_\lambda(\omega) - \tilde{X}_\mu(\omega)| \leq K \sum_{i=1}^d |\lambda_i - \mu_i|^{\beta_i} \quad (\text{A.0.0.2})$$

vale para qualquer $\lambda, \mu \in \Lambda$ tal que $\sum_i |\lambda_i - \mu_i|^{\beta_i} \leq 2^{-m_0(\omega)}$.

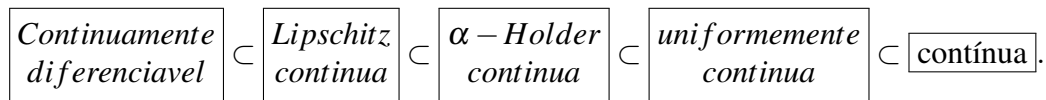
O campo aleatório satisfazendo (A.0.0.2) é chamado de $(\beta_1, \dots, \beta_d)$ -Hölder contínuo.

De forma mais simples, e apenas para situar o leitor sobre versão contínua obtida acima, uma função de valor real ou complexo f no espaço euclidiano d -dimensional satisfaz uma condição de Hölder, ou é Hölder contínua, quando existem constantes reais não negativas $C, \alpha > 0$, tais que

$$|f(x) - f(y)| \leq C|x - y|^\alpha$$

para todo x e y no domínio de f . O número α é chamado de expoente da condição de Hölder. Uma função em um intervalo que satisfaça a condição com $\alpha > 1$ é constante. Se $\alpha = 1$, então a função satisfaz uma condição de Lipschitz. Para qualquer $\alpha > 0$, a condição implica que a função é uniformemente contínua.

Temos a seguinte cadeia de inclusões estritas para funções sobre um intervalo não trivial fechado e limitado da reta real



Ponto de compactificação

Seja X um espaço de Hausdorff não compacto, mas localmente compacto. Um ponto de compactificação de X é o espaço topológico X^* definido como se segue. Seja ρ algum objeto que não pertence a X , e seja $X^* = X \cup \{\rho\}$ com a topologia τ formada pela união de conjuntos abertos de X e $U \subset X^* : X^* \setminus U$ é um subconjunto compacto de X , isto é,

$$\tau = \{\text{conjuntos abertos de } X\} \cup \{U \subset X^*\}.$$

De forma mais simples, podemos dizer que uma compactificação de um espaço topológico X é um par ordenado (X^*, ψ) , onde X^* é um espaço de Hausdorff compacto e ψ é um homeomorfismo de X sobre $\psi(X)$, com $\overline{\psi(X)} = X^*$.

Espaços de Hausdorff localmente compactos têm uma compactificação simples, conhecida como compactificação de Alexandrov. Um exemplo dessa compactificação é a seguinte. Se X é um espaço de Hausdorff não compacto, mas localmente compacto, é possível mostrar que existe um espaço de Hausdorff $X^* = X \cup \{\rho\}$, $\rho \notin X$ tal que X é homeomorfo a $\psi(X)$ e $\overline{\psi(X)} = X^*$. O ponto ρ é chamado de ponto no infinito. Por outro lado, se X já é um espaço de Hausdorff compacto, não há o que compactificar. Outro exemplo interessante de compactificação é a projeção estereográfica. Mais detalhes sobre compactificação podem ser vistos em Lee (2000) [14], capítulo 4.

Campos de vetores

C.1 Diferencial de um campo

Na literatura da Geometria Diferencial, o diferencial de um campo às vezes é chamado de mapa tangente, derivada total ou simplesmente derivada de F . Como ele “empurra” vetores tangentes da variedade de domínio para o contradomínio, também pode ser chamado de *pushforward*.

Sejam M e N variedades e $F : M \rightarrow N$ uma função diferenciável. Para cada $p \in M$ definimos uma função

$$F_*p : T_pM \rightarrow T_{F(p)}N,$$

chamada de diferencial (ou *pushforward*) de F em p dada pela fórmula

$$(F_*X)f = X(f \circ F).$$

Seja $F : M \rightarrow N$ e $G : N \rightarrow P$, então

- i. $F_* : T_pM \rightarrow T_{F(p)}M$ é linear;
- ii. $(G \circ F)_* = G_* \circ F_* : T_pM \rightarrow T_{G \circ F(p)}P$;
- iii. $(Id_M)_* = Id_{T_pM} : T_pM \rightarrow T_pM$;
- iv. Se F é um difeomorfismo, então $F_* : T_pM \rightarrow T_{F(p)}M$ é um isomorfismo.

Se $F : M \rightarrow N$ é uma função diferenciável e X é um campo de vetores em M , então para cada $p \in M$, obtemos um vetor $F_*X_p \in T_{F(p)}N$ pelo *pushforward* X_p . No entanto, isso geralmente não define um campo vetorial em N pois, se F não for sobrejetora, não há como decidir qual vetor atribuir a um ponto $q \in N - F(M)$. Por outro lado, se F não for injetiva, então para alguns pontos de N pode haver vários vetores diferentes obtidos pela aplicação de F_* a X em diferentes pontos de M . Então, suponha que exista um campo de vetores Y em N com a propriedade de que para cada $p \in M$, $F_*X_p = Y_{F(p)}$. Nesse caso, dizemos que os campos de vetores X e Y são F -relacionados. A próxima proposição, cuja demonstração pode ser vista em Lee (2012, p.182) [15], mostra como campos vetoriais F -relacionados atuam em funções suaves.

Proposição 60. *Seja $F : M \rightarrow N$ uma função diferenciável e M e N variedades. Então $X \in \mathfrak{X}(M)$ e $Y \in \mathfrak{X}(N)$ são F -relacionados a se, e somente se, para toda função de valores reais f definida em um subconjunto aberto de N ,*

$$X(f \circ F) = (Yf) \circ F,$$

onde $\mathfrak{X}(M)$, $\mathfrak{X}(N)$ denota o conjunto de todos os campos de vetores em M, N , respectivamente.

Para uma função suave $F : M \rightarrow N$ e um campo vetorial $X \in \mathfrak{X}(M)$, pode não haver nenhum campo vetorial em N que seja F -relacionado a X . No entanto, há um caso especial em que sempre há um campo vetorial, como mostra a próxima proposição, cuja demonstração pode ser vista em Lee (2012, p.183) [15].

Proposição 61. *Seja $F : M \rightarrow N$ um difeomorfismo. Para todo $X \in \mathfrak{X}(M)$, existe um único campo de vetores em N que é F -relacionado a X .*

Este único campo é denotado por F_*X e chamado de *pushforward*. Note que apenas quando F é um difeomorfismo F_*X é definido. A prova da proposição anterior mostra que F_*X é definido explicitamente pela fórmula

$$(F_*X)_{F(p)} = F_*X_p. \quad (\text{C.1.0.1})$$

A expressão acima também pode ser escrita da forma

$$(F_*X)_q = F_*F^{-1}(q)X_{F^{-1}(q)},$$

onde $q = F(p)$.

Se F^{-1} pode ser calculado explicitamente, o *pushforward* de um campo vetorial pode ser calculado diretamente a partir desta fórmula.

Exemplo 22. *Seja M e N as seguintes subvariedades abertas de \mathbb{R}^2 :*

$$M = \{(x, y) : y > 0 \text{ e } x + y > 0\},$$

$$N = \{(u, v) : u > 0 \text{ e } v > 0\},$$

e $F : M \rightarrow N$ definido por

$$F(x, y) = \left(x + y, \frac{x}{y} + 1 \right).$$

Então, F é um difeomorfismo e

$$F^{-1}(u, v) = \left(u - \frac{u}{v}, \frac{u}{v} \right).$$

A diferencial de F no ponto $(x, y) \in M$ é representado pela matriz jacobiana,

$$\begin{bmatrix} \frac{\partial(x+y)}{\partial x} & \frac{\partial(x+y)}{\partial y} \\ \frac{\partial\left(\frac{x}{y}+1\right)}{\partial x} & \frac{\partial\left(\frac{x}{y}+1\right)}{\partial y} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ \frac{1}{y} & \frac{-x}{y^2} \end{bmatrix}$$

e então, $F_*F^{-1}(u, v)$ é representado pela matriz

$$\begin{bmatrix} 1 & 1 \\ \frac{1}{\frac{u}{v}} & \frac{1}{\left(\frac{u}{v}\right)^2} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ \frac{v}{u} & \frac{v-v^2}{u} \end{bmatrix}$$

Considerando X o campo de vetor em M dado por

$$X = y^2 \frac{\partial}{\partial x} \Big|_{(x,y)},$$

então, para qualquer $(u, v) \in N$

$$X_{F^{-1}(u,v)} = \frac{u^2}{v^2} \frac{\partial}{\partial x} \Big|_{F^{-1}(u,v)}.$$

Portanto, aplicando a fórmula (C.1.0.1) com $p = (u, v)$ obtemos a fórmula para F_*X :

$$(F_*X)_{(u,v)} = \begin{bmatrix} 1 & 1 \\ v & v-v^2 \\ u & u \end{bmatrix} \cdot \begin{bmatrix} \frac{u^2}{v^2} \\ 0 \end{bmatrix} = \frac{u^2}{v^2} \frac{\partial}{\partial x} \Big|_{(u,v)} + \frac{u}{v} \frac{\partial}{\partial v} \Big|_{(u,v)}.$$

C.2 Colchete de Lie

Seja X e Y campos de vetores em uma variedade M e $f : M \rightarrow \mathbb{R}$. Um operador $[X, Y] : C^\infty(M) \rightarrow C^\infty(M)$, definido por

$$[X, Y]f = XY f - YX f \quad (\text{C.2.0.1})$$

é chamado de colchete de Lie.

Como pode ser visto em Lee (2012, p.186) [15] o colchete de Lie de qualquer par de campos vetoriais é um campo vetorial. O valor do campo vetorial $[X, Y]$ em um ponto $p \in M$ é a derivação em p dada pela fórmula

$$[X, Y]_p f = X_p(Yf) - Y_p(Xf).$$

Note que esta fórmula requer que se calcule termos envolvendo derivadas secundárias de f que sempre se cancelarão mutuamente. A próxima proposição fornece uma fórmula de coordenadas extremamente útil para o colchete de Lie, em que os cancelamentos já foram contabilizados.

Proposição 62. *Seja X, Y campos de vetores em uma variedade M , e seja*

$$X = X^i \frac{\partial}{\partial x^i} \quad e \quad Y = Y^j \frac{\partial}{\partial x^j}$$

as expressões coordenadas de X e Y em termos de coordenadas locais (x^i) para M . Então $[X, Y]$ tem a seguinte expressão coordenada

$$[X, Y] = \left(X^i \frac{\partial Y^j}{\partial x^i} - Y^j \frac{\partial X^i}{\partial x^j} \right) \frac{\partial}{\partial x^j}, \quad (\text{C.2.0.2})$$

ou, com uma notação mais simples,

$$[X, Y] = (XY^j - YX^j) \frac{\partial}{\partial x^j}. \quad (\text{C.2.0.3})$$

Demonstração. De fato,

$$\begin{aligned} [X, Y]f &= XY f - YX f \\ &= X^i \frac{\partial}{\partial x^i} \left(Y^j \frac{\partial}{\partial x^j} f \right) - Y^j \frac{\partial}{\partial x^j} \left(X^i \frac{\partial}{\partial x^i} f \right) \\ &= X^i \frac{\partial}{\partial x^i} Y^j \frac{\partial}{\partial x^j} f + X^i Y^j \frac{\partial^2}{\partial x^i \partial x^j} f - Y^j \frac{\partial}{\partial x^j} X^i \frac{\partial}{\partial x^i} f - Y^j X^i \frac{\partial^2}{\partial x^j \partial x^i} f \\ &= X^i \frac{\partial}{\partial x^i} Y^j \frac{\partial}{\partial x^j} f - Y^j \frac{\partial}{\partial x^j} X^i \frac{\partial}{\partial x^i} f \\ &= X^i \frac{\partial}{\partial x^i} Y^j \frac{\partial}{\partial x^j} f - Y^j \frac{\partial}{\partial x^j} X^i \frac{\partial}{\partial x^i} f \\ &= \left(X^i \frac{\partial}{\partial x^i} Y^j - Y^j \frac{\partial}{\partial x^j} X^i \right) \frac{\partial}{\partial x^j} f \\ &= [XY^j - YX^i]f. \end{aligned}$$

Como derivadas parciais misturadas de uma função comutam, podemos inverter os papéis dos índices i e j no somatório do segundo termo. E assim, a quinta igualdade é verdadeira. ■

Exemplo 23. Defina os campos de vetores $V, W \in \mathcal{D}(\mathbb{R}^3)$ por

$$V = x \frac{\partial}{\partial x} + \frac{\partial}{\partial y} + x(y+1) \frac{\partial}{\partial z}$$

$$W = \frac{\partial}{\partial x} + y \frac{\partial}{\partial z}$$

Pela equação (C.2.0.3) temos

$$\begin{aligned} [V, W] &= V(1) \frac{\partial}{\partial x} + V(y) \frac{\partial}{\partial z} - W(x) \frac{\partial}{\partial x} + W(1) \frac{\partial}{\partial y} - W(x(y+1)) \frac{\partial}{\partial z} \\ &= 0 \frac{\partial}{\partial x} + 1 \frac{\partial}{\partial z} - 1 \frac{\partial}{\partial x} + 0 \frac{\partial}{\partial y} - (y+1) \frac{\partial}{\partial z} \\ &= \frac{\partial}{\partial z} - 1 \frac{\partial}{\partial x} - y \frac{\partial}{\partial z} - 1 \frac{\partial}{\partial z} \\ &= -1 \frac{\partial}{\partial x} - y \frac{\partial}{\partial z} \end{aligned}$$

Note que, calculamos $V(1)$ da seguinte forma:

$$\begin{aligned} V(1) &= x \frac{\partial}{\partial x}(1) + \frac{\partial}{\partial y}(1) + x(y+1) \frac{\partial}{\partial z}(1) \\ &= x \cdot 0 + 0 + x(y+1) \cdot 0 \\ &= 0. \end{aligned}$$

Os outros cálculos foram análogos.

A próxima proposição, cuja demonstração pode ser vista em Lee (2012, p.188) [15], apresenta algumas propriedades muito úteis do colchete de Lie.

Proposição 63. (Propriedades do Colchete de Lie) O colchete de Lie satisfaz as seguintes identidades para todos os campos de vetores X, Y e Z em M :

1. Para $a, b \in \mathbb{R}$

$$[aX + bY, Z] = a[X, Z] + b[Y, Z],$$

$$[Z, aX + bY] = a[Z, X] + b[Z, Y].$$

2.

$$[X, Y] = -[Y, X].$$

3.

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0.$$

4. Para $f, g \in C^\infty(M)$,

$$[fX, gY] = fg[X, Y] + (fXg)Y - (gYf)X.$$

C.3 Derivada de Lie

No espaço euclidiano, faz muito sentido definir a derivada direcional de um campo vetorial X na direção de um vetor $Y \in T_p\mathbb{R}^d$ por

$$D_Y X_p = \lim_{t \rightarrow 0} \frac{X_{p+tY} - X_p}{t}. \tag{C.3.0.1}$$

No entanto, esta definição depende do fato de que \mathbb{R}^d é um espaço vetorial, de modo que os vetores tangentes X_{p+tY} e X_p podem ser vistos como elementos de \mathbb{R}^d .

Ao buscarmos uma maneira de dar sentido a (C.3.0.1) em uma variedade, vemos que, podemos substituir $p + tY$ por uma curva $\gamma(t)$ que começa em p e cujo vetor tangente inicial é Y . Mas mesmo com essa substituição, o quociente da diferença ainda não faz sentido porque $X_{\gamma(t)}$ e $X_{\gamma(0)}$ são elementos de diferentes espaços vetoriais ($T_{\gamma(t)}M$ e $T_{\gamma(0)}M$, respectivamente). A imagem a seguir ilustra isso.

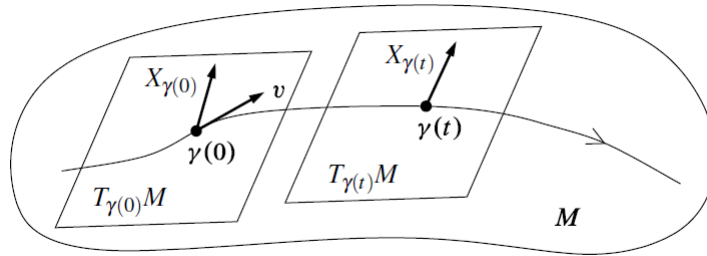


Figura C.1: Problema nas derivadas direcionais de campo vetoriais, adaptado de Lee (2012).

No espaço euclidiano há uma identificação canônica de cada espaço tangente com o próprio \mathbb{R}^d ; mas em uma variedade não há tal identificação. Portanto, não há uma maneira independente de coordenada de dar sentido à derivada direcional de X na direção do vetor Y .

Agora suponha que o próprio X seja um campo vetorial em vez de um único vetor. Nesse caso, podemos usar o fluxo de X para empurrar os valores de Y de volta para p (através do *pushforward*, também chamado de diferencial) e então diferenciar. A imagem abaixo ilustra essa ideia.

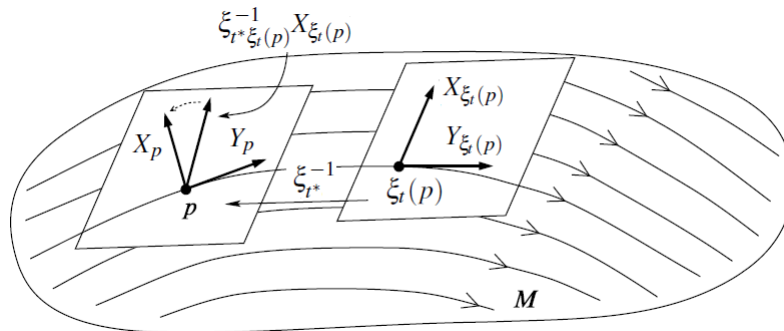


Figura C.2: Derivada direcional de um campo vetorial, adaptado de Lee (2012)

Assim, temos a seguinte definição.

Definição 64. Para quaisquer campos vetoriais X e Y em uma variedade M , seja ξ o fluxo de Y definimos um vetor $(L_Y X)_p$ para cada $p \in M$, chamado de derivada de Lie de X em relação a Y em p , por

$$(L_Y X)_p = \lim_{t \rightarrow 0} \frac{1}{t} \left[\xi_{t*}^{-1} X_{\xi_t(p)} - X_p \right]. \tag{C.3.0.2}$$

Pela ilustração anterior temos que o quociente da diferença faz sentido porque ξ_t é definido em uma vizinhança de p , e ambos $\xi_{t^* \xi_t(p)}^{-1} X_{\xi_t(p)}$ e X_p são elementos de $T_p M$.

Normalmente o fluxo é difícil ou impossível de ser anotado explicitamente, por isso, a definição de $L_Y X$ não é muito útil para cálculos. O próximo resultado, cuja demonstração pode ser vista em Lee (2012, p.229) [15], mostra que existe uma fórmula simples para calcular a derivada de Lie sem encontrar explicitamente o fluxo.

Proposição 65. *Para quaisquer campos vetoriais X e Y em uma variedade M , $L_Y X = [Y, X]$, onde $[Y, X]$ denota o colchete de Lie.*

A proposição anterior fornece uma interpretação geométrica do colchete de Lie de dois campos vetoriais: é a derivada direcional do segundo campo vetorial ao longo do fluxo do primeiro. A próxima proposição apresenta uma série de propriedades da derivada de Lie que decorrem imediatamente dos colchetes de Lie, como pode ser visto em Lee (2012, p.230) [15].

Proposição 66. *Seja M uma variedade e $V, W, X \in \mathfrak{X}(M)$, então*

i. $L_V W = -L_W V$;

ii. $L_V [W, X] = [L_V W, X] + [W, L_V X]$;

iii. $L_{[V, W]} X = L_V L_W X - L_W L_V X$;

iv. Se $g \in C^\infty(M)$ então $L_V(gW) = (Vg)W + gL_V W$;

v. Se $F : M \rightarrow N$ é um difeomorfismo, então $F_*(L_V X) = L_{F_* V} F_* X$.

Formas diferenciais

Uma forma diferencial de grau 1 é um funcional linear cujas coordenadas variam de ponto a ponto, uma forma de grau mais elevado, que chamaremos de forma exterior, é uma forma alternada com coeficientes variáveis. As formas alternadas eram chamados há muito tempo de tensores covariantes anti-simétricos. As noções que apresentaremos são voltadas para o entendimento do nosso trabalho, mais detalhes sobre esse assunto podem ser vistos por exemplo em Lima (2007) [17], Carmo (1994) [3], Lee (2012) [15] e Spivak (1999) [25].

D.1 Aplicações r -lineares

Sejam E_1, \dots, E_r, F espaços vetoriais. Uma aplicação $f : E_1 \times \dots \times E_r \rightarrow F$ é chamada de r -linear se para quaisquer $v_1 \in E_1, v_i, w_i \in E_i, v_r \in E_r$ e $\lambda \in \mathbb{R}$ satisfaz as seguintes condições

$$f(v_1, \dots, v_i + w_i, \dots, v_r) = f(v_1, \dots, v_i, \dots, v_r) + f(v_1, \dots, w_i, \dots, v_r);$$

$$f(v_1, \dots, \lambda v_i, \dots, v_r) = \lambda \cdot f(v_1, \dots, v_i, \dots, v_r).$$

O conjunto $L(E_1, \dots, E_r; F)$ das aplicações r -lineares munido das operações de soma e multiplicação é um espaço vetorial. Escrevemos $L_r(E; F)$ para representar o espaço vetorial formado pelas aplicações r -lineares $f : E \times \dots \times E \rightarrow F$. Quando $F = \mathbb{R}$, a aplicação r -linear é chamada de forma r -linear. Considerando $r = 1$ temos $L_1(E; F) = L(E; F)$ que representa o espaço das transformações lineares de E em F . Em particular, se $F = \mathbb{R}$, então $L_1(E; \mathbb{R}) = E^*$ que representa o espaço dual de E .

O produto tensorial dos funcionais lineares $f_1, f_2, \dots, f_r \in E^*$ é uma forma r -linear $f = f_1 \cdot f_2 \cdot \dots \cdot f_r \in L_r(E; \mathbb{R})$, definida por

$$f(v_1, \dots, v_i, \dots, v_r) = f_1(v_1) \cdot f_2(v_2) \cdot \dots \cdot f_r(v_r).$$

Para cada sequência $s = (i_1, \dots, i_r)$ de números em $\{1, 2, \dots, n\}$ denotamos por $\bar{e}_s = \bar{e}_{i_1} \cdot \bar{e}_{i_2} \cdot \dots \cdot \bar{e}_{i_r}$ o produto tensorial destes funcionais. Estas formas r -lineares definidas por sequências definem uma base para $L_r(E; \mathbb{R})$. Além disso, temos que se $\dim(E) = n$ então $\dim(L_r(E; \mathbb{R})) = n^r$.

D.2 Formas alternadas

Uma aplicação r -linear $f \in L_r(E; F)$ é chamada de alternada quando $f(v_1, \dots, v_i, \dots, v_r) = 0$ sempre que $v_i = v_j, i \neq j$. Ou seja, algum termo da sequência $v_1, \dots, v_i, \dots, v_r$ se repete. Uma

forma $f \in L_r(E; \mathbb{R})$ é chamada de anti-simétrica quando seu valor muda de sinal quando duas de suas variáveis trocam de posições, isto é, para quaisquer $v_1, \dots, v_r \in E$ temos

$$f(v_1, \dots, v_i, \dots, v_j, \dots, v_r) = -f(v_1, \dots, v_j, \dots, v_i, \dots, v_r). \quad (\text{D.2.0.1})$$

Teorema 67. *Toda forma anti-simétrica é alternada.*

Demonstração. Se $v_i = v_j = v$ na equação (D.2.0.1), temos

$$f(\dots, v, \dots, v, \dots) = -f(\dots, v, \dots, v, \dots) \Rightarrow f(\dots, v, \dots, v, \dots) = 0.$$

Portanto, toda forma anti-simétrica é alternada. Reciprocamente, se $f \in L_r(E; \mathbb{R})$ é alternada, então

$$\begin{aligned} 0 &= f[v_i + v_j, v_i + v_j] \\ &= \underbrace{f[v_i, v_i]}_{=0} + f[v_i, v_j] + f[v_j, v_i] + \underbrace{f[v_j, v_j]}_{=0} \\ &= f[v_i, v_j] + f[v_j, v_i]. \end{aligned}$$

onde $f[v_i, v_j] = f(\dots, v_i, \dots, v_j, \dots)$. Logo,

$$f[v_i, v_j] = -f[v_j, v_i],$$

e portanto, f é anti-simétrica. ■

Representamos o conjunto das formas r -lineares alternadas (ou anti-simétricas) no espaço vetorial E por $U_r(E)$. Consideramos $U_1(E) = L_1(E; \mathbb{R}) = E^*$, isto é, todo funcional linear é uma forma alternada.

D.3 Determinantes

Seja $E = \mathbb{R}^n$. Podemos ver uma sequência (v_1, \dots, v_n) de n vetores $v_j = (a_{1j}, \dots, a_{nj})$ como uma matriz $\mathbf{a} = [v_1, \dots, v_n] = [a_{ij}]$ do tipo $n \times n$, da qual v_j é a j -ésima coluna. Dessa maneira, uma forma n -linear f em \mathbb{R}^d é o mesmo que uma função $f : M(n \times n) \rightarrow \mathbb{R}$, cujos valores $f(\mathbf{a}) = f[v_1, \dots, v_n]$ dependem linearmente das colunas v_j da matriz $\mathbf{a} \in M(n \times n)$.

Se $\{e_1, \dots, e_n\} \subset \mathbb{R}^n$ é a base canônica, existe uma única forma n -linear alternada $\bar{e} \in U_n(\mathbb{R}^n)$ tal que $\bar{e}(e_1, \dots, e_n) = 1$. Todas as outras são múltiplos de \bar{e} , isto é, qualquer outra $f \in U_n(\mathbb{R}^n)$ temos $f = \alpha \bar{e}$.

O determinante da matriz $\mathbf{a} \in M(n \times n)$ cujas colunas são os vetores $v_j = (a_{1j}, \dots, a_{nj})$, $j = 1, \dots, n$ é definido como

$$\det \mathbf{a} = \bar{e}(v_1, \dots, v_n).$$

Se $f : M(n \times n) \rightarrow \mathbb{R}$ é qualquer função tal que $f(\mathbf{a})$ é uma função n -linear alternada das colunas de \mathbf{a} então, para cada matriz $\mathbf{a} \in M(n \times n)$ tem-se $f(\mathbf{a}) = \det \mathbf{a} \cdot f(I_n)$, onde I_n é a matriz identidade $n \times n$.

Seja $A : E \rightarrow F$ uma transformações linear. Para todo $r \geq 0$, a transformação linear $A^* : L_r(F; \mathbb{R}) \rightarrow L_r(E; \mathbb{R})$ definida da forma

$$(A^* \cdot f)(v_1, \dots, v_r) = f(A \cdot v_1, \dots, A \cdot v_r),$$

é chamada de transformação linear induzida por A , também conhecida como *pullback*. Quando $r = 1$, A^* é a adjunta $A^* : F^* \rightarrow E^*$ da transformação linear A . E ainda temos as seguintes propriedades

1. $(B \cdot A)^* = A^* \cdot B^*$;
2. $I^* = I$;
3. Quando A é invertível então $(A^*)^{-1} = (A^{-1})^*$.

Consideremos o caso particular de um operador linear $A : E \rightarrow E$, no espaço vetorial E , de dimensão n . Então $U_n(E) = 1$, de modo que a transformação linear $A^* : U_n(E) \rightarrow U_n(E)$, induzida por A , consiste na multiplicação por uma constante. Tal constante é chamada de determinante do operador A , a qual indicamos por $DetA$. Então, por definição, temos para quaisquer $v_1, \dots, v_n \in E$ e $f \in U_n(E)$

$$A^* f(v_1, \dots, v_n) = f(A \cdot v_1, \dots, A \cdot v_n) = DetA \cdot f(v_1, \dots, v_n) = DetA \cdot f(v_1, \dots, v_n).$$

Dados $A, B \in L(E)$, tomemos $f \neq 0$ em $U_n(E)$. Então,

$$(DetAB)f = (AB)^* \cdot f = B^* \cdot A^* \cdot f = DetB \cdot (A^* f) = DetB \cdot DetA \cdot f,$$

Logo, $DetAB = DetA \cdot DetB$.

Teorema 68. Para todo operador linear $A : E \rightarrow E$, tem-se $DetA = deta$, onde \mathbf{a} é a matriz de A numa base arbitrária $\{u_1, \dots, u_n\} \subset E$.

Demonstração. Consideremos o caso em que $E = \mathbb{R}^n$ e a base dada é a canônica $\{e_1, \dots, e_n\} \subset \mathbb{R}^n$. As colunas da matriz \mathbf{a} são os vetores $v_j = (a_{1j}, \dots, a_{nj}) = A \cdot e_j$. Se $\bar{e} \in U_n(\mathbb{R}^n)$ é a forma tal que $\bar{e}(e_1, \dots, e_n) = 1$, temos

$$\begin{aligned} deta &= \bar{e}(v_1, \dots, v_n) \\ &= \bar{e}(A \cdot e_1, \dots, A \cdot e_n) \\ &= (A^* \bar{e})(e_1, \dots, e_n) \\ &= DetA \cdot \bar{e}(e_1, \dots, e_n) \\ &= DetA. \end{aligned}$$

■

A demonstração do caso geral do teorema acima pode ser visto em Lima (2007,p.36) [17]. Uma matriz \mathbf{a} e sua transposta \mathbf{a}^T têm o mesmo determinante.

D.4 Produto exterior

O produto exterior de r funcionais lineares $f_1, \dots, f_r \in E^*$ é a forma r -linear $f_1 \wedge \dots \wedge f_r \in U_r(E)$ definida por

$$(f_1 \wedge \dots \wedge f_r)(v_1, \dots, v_r) = \det \begin{bmatrix} f_1(v_1) & f_1(v_2) & \cdots & f_1(v_r) \\ \vdots & \vdots & \ddots & \vdots \\ f_r(v_1) & f_r(v_2) & \cdots & f_r(v_r) \end{bmatrix}_{[r \times r]} = \det[f_i(v_j)] \quad (\text{D.4.0.1})$$

Como o determinante de uma matriz $r \times r$ é uma função r -linear alternada de suas colunas ou linhas segue que a aplicação $\Lambda : E^* \times \dots \times E^* \rightarrow U_r(E)$ definida por $\Lambda(f_1, \dots, f_r) = f_1 \wedge \dots \wedge f_r$ é também r -linear alternada.

Teorema 69. A transformação linear $A^* : U_r(E) \rightarrow U_r(E)$ preserva o produto exterior de funcionais lineares, ou seja, vale a seguinte igualdade para quaisquer $f_1, \dots, f_r \in E^*$

$$A^*(f_1 \wedge \dots \wedge f_r) = A^* f_1 \wedge \dots \wedge A^* f_r.$$

Demonstração. De fato, para quaisquer $v_1, \dots, v_r \in E$, temos

$$\begin{aligned} A^*(f_1 \wedge \dots \wedge f_r)(v_1, \dots, v_r) &= (f_1 \wedge \dots \wedge f_r)(Av_1, \dots, Av_r) \\ &= \det[f_i(Av_j)] \\ &= \det[A^*f_i(v_j)] \\ &= (A^*f_1 \wedge \dots \wedge A^*f_r)(v_1, \dots, v_r). \end{aligned}$$

■

Dados dois espaços vetoriais $U_r(E)$ e $U_s(E)$ gerados pelo produto exterior de funcionais lineares f, g , onde $f \in U_r(E)$ e $g \in U_s(E)$ existe uma única aplicação bilinear φ tal que

$$\begin{aligned} \varphi(f_1 \wedge \dots \wedge f_r, g_1 \wedge \dots \wedge g_s) &= f_1 \wedge \dots \wedge f_r \wedge g_1 \wedge \dots \wedge g_s \\ \varphi : U_r(E) \times U_s(E) &\rightarrow U_{r+s}(E) \end{aligned}$$

Ou seja, o produto exterior de uma forma $f \in U_r(E)$ por uma forma $g \in U_s(E)$, tem como resultado a forma $f \wedge g \in U_{r+s}(E)$. Temos as seguintes propriedades para o produto exterior de formas:

1. Se $f \in U_r(E)$ e $g \in U_s(E)$ então $g \wedge f = (-1)^{rs} \cdot f \wedge g$;
2. $(f \wedge g) \wedge h = f \wedge (g \wedge h)$;
3. $A^*(f \wedge g) = A^*f \wedge A^*g$.

D.5 Superfície orientada

Orientar um espaço vetorial E é escolher uma base $\{u_1, \dots, u_n\} \subset E$, chamá-la de positiva e dizer que todas as bases $\{v_1, \dots, v_n\} \subset E$ tais que

$$v_j = \sum_{i=1}^n a_{ij}u_i, \quad j = 1, \dots, n,$$

onde $\det[a_{ij}] > 0$ também são positivas.

Exemplo 24. Se M é uma superfície orientada, então para todo $p \in M$, o espaço vetorial tangente T_pM possui uma orientação natural, segundo o qual a base

$$\left\{ \frac{\partial \varphi}{\partial x_1}(x_0), \dots, \frac{\partial \varphi}{\partial x_m}(x_0) \right\} \subset T_pM,$$

associada a uma parametrização positiva $\varphi : V_0 \rightarrow V$, com $p = \varphi(x_0) \in V$, é declarada uma base positiva.

Seja E um espaço vetorial orientado, munido de produto interno. O elemento de volume de E é a forma $G \in U_n(E)$, $\dim E = n$ onde escolhe-se uma base ortonormal positiva $\{u_1, \dots, u_n\} \subset E$ e, para quaisquer $v_1, \dots, v_n \in E$ temos

$$G(v_1, \dots, v_n) = \det[a_{ij}],$$

onde $a_{ij} = \langle u_i, v_j \rangle$ e $v_j = \sum_{i=1}^n a_{ij}u_i$, isto é, v_j combinação linear de u_1, \dots, u_n . O elemento de volume é geralmente chamado de comprimento ou área de superfície para variedades de uma e duas dimensões, respectivamente.

De acordo com a definição acima, G é uma forma n -linear alternada em E .

Teorema 70. A forma n -linear alternada G não depende da escolha da base $\{u_1, \dots, u_n\}$.

Demonstração. Tomemos $g = [\langle v_i, v_j \rangle]$, onde $\langle v_i, v_j \rangle = \sum_{k=1}^n a_{ki} \cdot a_{kj}$. Logo, $g = a^T \cdot a$, $a = [a_{ij}]$. Portanto,

$$\det g = \det(a^T \cdot a) = \det a^T \cdot \det a = (\det a)^2$$

E assim,

$$\det a = \pm \sqrt{\det g}.$$

Observe que $\det g$ não depende de escolhas arbitrárias. ■

Geometricamente, $G(v_1, \dots, v_n)$ é igual ao volume do paralelepípedo n -dimensional construído sobre as arestas v_1, \dots, v_n tomando o sinal positivo ou negativo conforme a base $\{v_1, \dots, v_n\}$ seja positiva ou negativa, respectivamente. Se v_1, \dots, v_n forem linearmente dependentes, então $G(v_1, \dots, v_n) = 0$.

D.6 Formas diferenciais

Uma forma diferencial de grau r num aberto $U \subset \mathbb{R}^n$ é uma aplicação $\omega : U \rightarrow U_r(\mathbb{R}^n)$. Para cada $x \in U$, $\omega(x)$ é uma forma r -linear alternada em \mathbb{R}^n .

Denotamos $\{dx_1, \dots, dx_n\} \subset (\mathbb{R}^n)^*$ a base dual canônica $\{e_1, \dots, e_n\} \subset \mathbb{R}^n$. Uma base natural de $U_r(\mathbb{R}^n)$ consiste nas formas $dx_I = dx_{i_1} \wedge \dots \wedge dx_{i_r}$, onde $\{i_1 < \dots < i_r\}$ percorre todos os elementos de $I = \{1, 2, \dots, n\}$. Então, para cada $x \in U$, temos

$$\omega(x) = \sum_I a_I(x) dx_I,$$

onde $a_I(x) = \omega(x) \cdot (e_{i_1}, \dots, e_{i_r})$ são as coordenadas de $\omega(x)$ relativas à base composta pelos dx_I . Quando as funções $a_I : U \rightarrow \mathbb{R}$ são de classe C^k , dizemos que ω é uma forma de classe C^k .

Seja $M \subset \mathbb{R}^n$ uma superfície m -dimensional, uma forma diferencial de grau r em M é uma correspondência ω que associa a cada $x \in M$ uma forma r -linear alternada $\omega(x) \in U_r(T_x M)$. Assim, para todo $x \in M$ e toda lista de r vetores $v_1, \dots, v_r \in T_x M$, $\omega(x) \cdot (v_1, \dots, v_r)$ é um número real que depende linearmente de cada v_i e se anula quando $v_i = v_j$ com $i \neq j$.

Se $f : M \rightarrow N$ é uma aplicação de classe C^k ($k \geq 1$) entre as superfícies M, N , a cada forma diferencial ω de grau r em N corresponde uma forma $f^* \omega$, de mesmo grau de M , chamada de *pullback* de ω por f , definida por

$$(f^* \omega)(x) \cdot (v_1, \dots, v_r) = \omega(f(x)) \cdot (f'(x)v_1, \dots, f'(x)v_r)$$

para todo $x \in M$ e quaisquer $v_1, \dots, v_r \in T_x M$. A transformação linear $f'(x) : T_x M \rightarrow T_{f(x)} N$ é a derivada de f no ponto x .

A aplicação $\omega \rightarrow f^* \omega$ define uma transformação linear:

$$f^*(a\omega + b\theta) = af^* \omega + bf^* \theta, \quad a, b \in \mathbb{R}.$$

Além disso, ainda temos

$$f^*(\omega \wedge \theta) = f^* \omega \wedge f^* \theta$$

e, se $f : M \rightarrow N$ e $g : N \rightarrow P$ são aplicações de classe C^k ($k \geq 1$) então

$$(g \circ f)^* \omega = f^*(g^* \omega), \quad \forall \omega \in P.$$

Se M está contida na superfície N e $i : M \rightarrow N$ é a aplicação de inclusão, onde $i(x) = x$ então, para toda forma diferencial ω em N , seu *pullback* é a forma $i^* \omega$, chamada de forma induzida

por ω em M , às vezes representada por $\omega|_M$, pois é uma restrição de ω a M . Para obter $i^*\omega$ basta em $\omega(x)(v_1, \dots, v_r)$ restringir $x \in M$ e $v_1, \dots, v_r \in T_xM$.

Seja $\varphi : U_0 \rightarrow U \subset M$ uma parametrização local na superfície m -dimensional $M \subset \mathbb{R}^n$. Em cada ponto $x \in U$, indicaremos por $\{du_1, \dots, du_m\} \subset (T_xM)^*$ ¹ a base dual da base $\{\frac{\partial \varphi}{\partial u_1}(u), \dots, \frac{\partial \varphi}{\partial u_m}(u)\} \subset T_xM$. As formas diferenciais $du_I = du_{i_1} \wedge \dots \wedge du_{i_r}$, $I = \{i_1 < \dots < i_r\} \subset I_m$, em cada ponto $x \in U$ formam uma base de $U_r(T_xM)$ portanto, em termos de parametrização, toda forma diferencial ω de grau r em M pode ser expressa da forma

$$\omega(x) = \sum_I a_I(u) du_I, \quad x = \varphi(u).$$

Estabelecemos a convenção de que uma forma diferencial de grau zero é uma função real $g : M \rightarrow \mathbb{R}$. Se $f : M \rightarrow N$ é uma aplicação de classe C^k então o *pullback* de g por f é $f^*g = g \circ f$.

Exemplo 25. Em \mathbb{R} , as formas diferenciais de grau 1 são do tipo $\omega(x) = f(x)dx$.

Exemplo 26. Em abertos de \mathbb{R}^2 , as formas de grau 1 são da forma

$$\omega(x, y) = a(x, y)dx + b(x, y)dy$$

e as formas de grau 2 são

$$\omega(x, y) = a(x, y)dx \wedge dy.$$

Exemplo 27. Em abertos de \mathbb{R}^3 , as formas de grau 1 são da forma

$$\omega(x, y, z) = a(x, y, z)dx + b(x, y, z)dy + c(x, y, z)dz,$$

as formas de grau 2 são

$$\omega(x, y, z) = a(x, y, z)dy \wedge dz + b(x, y, z)dz \wedge dx + c(x, y, z)dx \wedge dy$$

e as formas de grau 3 são

$$\omega(x, y, z) = a(x, y, z)dx \wedge dy \wedge dz.$$

O próximo lema, cuja demonstração pode ser vista em Lee (2012, p.361) [15], fornece uma regra computacional para *pullback* de formas diferenciais semelhantes às que desenvolvemos para campos de tensores anteriormente.

Lema 71.

$$A^* \left(\sum_I a_I(x) dx_I \right) = \sum_I (a_I \circ A) d(x_I \circ A)$$

Exemplo 28. Defina $A : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ por $A(u, v) = (u, v, u^2 - v^2)$ e seja ω uma 2-forma definida por $ydx \wedge dz + xdy \wedge dz$ em \mathbb{R}^3 . O *pullback* $A^*\omega$ é calculado como se segue:

$$\begin{aligned} A^*(ydx \wedge dz + xdy \wedge dz) &= (y \circ A)d(x \circ A) \wedge d(z \circ A) + (x \circ A)d(y \circ A) \wedge d(z \circ A) \\ &= vdu \wedge d(u^2 - v^2) + u dv \wedge d(u^2 - v^2) \\ &= vdu \wedge (2udu - 2vdv) + u dv \wedge (2udu - 2vdv) \\ &= 2uv \underbrace{du \wedge du}_{=0} - 2v^2 du \wedge dv + 2u^2 dv \wedge du - 2uv \underbrace{dv \wedge dv}_{=0} \\ &= -2v^2 du \wedge dv - 2u^2 du \wedge dv \\ &= -2(v^2 + u^2) du \wedge dv. \end{aligned}$$

¹aqui a notação precisa seria $du_i(x)$.

Exemplo 29. (Elemento volume) Seja M uma superfície orientada de dimensão m . O elemento volume de M (ou também chamado de forma volume de M) é uma forma diferencial ω de grau m tal que para cada $x \in M$ e $v_1, \dots, v_m \in T_x M$, $\omega(x)(v_1, \dots, v_m) = \pm$ volume do paralelepípedo determinado por v_1, \dots, v_m . (como visto na seção D.5)

Dada uma parametrização positiva $\varphi : U_0 \rightarrow U \subset M$ definimos as funções $g_{ij} : U_0 \rightarrow \mathbb{R}$ da forma

$$g_{ij} := \left\langle \frac{\partial \varphi}{\partial u_i}(u), \frac{\partial \varphi}{\partial u_j}(u) \right\rangle \quad e \quad g = \det[g_{ij}].$$

Assim, em cada ponto $x = \varphi(u) \in U \subset M$ o volume do paralelepípedo que tem arestas $\frac{\partial \varphi}{\partial u_1}(u), \dots, \frac{\partial \varphi}{\partial u_m}(u)$ é igual a \sqrt{g} , ou seja,

$$\omega(x) \left(\frac{\partial \varphi}{\partial u_1}(u), \dots, \frac{\partial \varphi}{\partial u_m}(u) \right) = \sqrt{g}.$$

Como $\{du_1, \dots, du_m\}$ é o dual de $\{\frac{\partial \varphi}{\partial u_1}, \dots, \frac{\partial \varphi}{\partial u_m}\}$, então

$$\omega = \sqrt{g} du_1 \wedge \dots \wedge du_m.$$

Exemplo 30. (Elemento volume de uma hiperfície) Seja $M \subset \mathbb{R}^{m+1}$ uma hiperfície orientada de classe C^k , então

$$\sqrt{g} = \left| \frac{\partial \varphi}{\partial u_1} \times \dots \times \frac{\partial \varphi}{\partial u_m} \right|.$$

Para cada $x = \varphi(u)$, seja

$$N(x) = \frac{\partial \varphi}{\partial u_1}(u) \times \dots \times \frac{\partial \varphi}{\partial u_m}(u),$$

então o vetor unitário normal que define a orientação de M é dado por

$$\frac{N(x)}{|N(x)|} = \mathbf{v}(x).$$

O volume m -dimensional do paralelepípedo $\left[\frac{\partial \varphi}{\partial u_1}(u), \dots, \frac{\partial \varphi}{\partial u_m}(u) \right]$ é igual ao do paralelepípedo $(m+1)$ -dimensional $\left[\mathbf{v}(x), \frac{\partial \varphi}{\partial u_1}(u), \dots, \frac{\partial \varphi}{\partial u_m}(u) \right]$. Assim,

$$\omega(x)(v_1, \dots, v_m) = \det[\mathbf{v}(x), v_1, \dots, v_m]_{(m+1) \times (m+1)},$$

desenvolvendo o determinante com a primeira coluna da matriz, temos

$$\omega(x)(v_1, \dots, v_m) = \sum_{i=1}^{m+1} (-1)^{i+1} v_i(x) A_i,$$

onde A_i é o determinante da submatriz $[v_1, \dots, v_m]_{m \times m}$ sem a i -ésima linha. Escrevendo $dx_1 \wedge \dots \wedge \widehat{dx}_i \wedge \dots \wedge dx_{m+1}$ em vez de $dx_1 \wedge \dots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \dots \wedge dx_{m+1}$ temos

$$A_i = (dx_1 \wedge \dots \wedge \widehat{dx}_i \wedge \dots \wedge dx_{m+1})(v_1, \dots, v_m).$$

Então, para todo $x \in U$ a expressão da forma elemento volume da hiperfície M em termos das coordenadas do vetor unitário normal $\mathbf{v}(x)$ e da base canônica $\{dx_1, \dots, dx_{m+1}\} \subset (\mathbb{R}^{m+1})^*$ é dada por

$$\omega(x) = \sum_{i=1}^{m+1} (-1)^{i+1} v_i(x) dx_1 \wedge \dots \wedge \widehat{dx}_i \wedge \dots \wedge dx_{m+1}.$$

Exemplo 31. (Elemento volume da esfera) Do exemplo anterior temos que o elemento volume da esfera S^m é da forma

$$\omega(x) = \sum_{i=1}^{m+1} (-1)^{i+1} x_i dx_1 \wedge \dots \wedge \widehat{dx}_i \wedge \dots \wedge dx_{m+1},$$

pois $v_i(x) = x_i$. Esta expressão define uma forma diferencial de grau m em \mathbb{R}^{m+1} , cuja restrição a S^m é o elemento volume.

Se S é a esfera de centro $a = (a_1, \dots, a_{m+1})$ e raio r , seu elemento volume é dado por

$$\omega(x) = \sum_{i=1}^{m+1} (-1)^{i+1} \frac{x_i - a_i}{r} dx_1 \wedge \dots \wedge \widehat{dx}_i \wedge \dots \wedge dx_{m+1}.$$

D.7 Operações para formas diferenciais de grau r

Nosso objetivo agora é definir algumas operações para formas diferenciais de grau r em \mathbb{R}^d . Para facilitar, falaremos r -formas invés de formas diferenciais de grau r .

Consideremos, primeiramente, θ e φ duas r -formas:

$$\theta = \sum_I f_I dx_I, \quad \varphi = \sum_I g_I dx_I,$$

então definimos a soma como

$$\theta + \varphi = \sum_I f_I dx_I + \sum_I g_I dx_I = \sum_I (f_I + g_I) dx_I. \quad (\text{D.7.0.1})$$

Agora, seja θ uma r -forma e φ uma s -forma, representadas por

$$\theta = \sum_I f_I dx_I, \quad I = (i_1, \dots, i_r), \quad i_i < \dots < i_r,$$

$$\varphi = \sum_J g_J dx_J, \quad J = (j_1, \dots, j_s), \quad j_i < \dots < j_s$$

definimos o produto exterior $\theta \wedge \varphi$, o qual é uma $(s+r)$ -forma, como

$$\theta \wedge \varphi = \sum_{IJ} f_I g_J dx_I \wedge dx_J. \quad (\text{D.7.0.2})$$

Exemplo 32. Seja $\theta = xdx + ydy + zdz$ uma 1-forma em \mathbb{R}^3 e $\varphi = xdx \wedge dy + dx \wedge dz$ uma 2-forma em \mathbb{R}^3 . Então,

$$\begin{aligned} \theta \wedge \varphi &= (xdx + ydy + zdz) \wedge (xdx \wedge dy + dx \wedge dz) \\ &= x^2 dx \wedge dx \wedge dy + x dx \wedge dx \wedge dz + yx dy \wedge dx \wedge dy + y dy \wedge dx \wedge dz + \\ &+ zx dz \wedge dx \wedge dy + z dz \wedge dx \wedge dz \\ &= y dy \wedge dx \wedge dz + zx dz \wedge dx \wedge dy \\ &= -y dx \wedge dy \wedge dz - xz dx \wedge dz \wedge dy \\ &= (xz - y) dx \wedge dy \wedge dz. \end{aligned}$$

Note que $\theta \wedge \varphi$ é uma 3-forma.

O produto exterior das formas em \mathbb{R}^d tem as seguintes propriedades, cuja demonstrações podem ser vistas em Carmo (1994, p.5) [3].

Proposição 72. Seja θ uma k -forma, φ uma s -forma e ω uma r -forma. Então,

1. $(\theta \wedge \varphi) \wedge \omega = \theta \wedge (\varphi \wedge \omega)$,
2. $(\theta \wedge \varphi) = (-1)^{ks}(\varphi \wedge \theta)$,
3. $\theta \wedge (\varphi \wedge \omega) = \theta \wedge \varphi + \theta \wedge \omega$, se $r = s$.

Vamos agora definir uma operação na forma diferencial que generaliza a diferenciação de funções, chamada de diferencial exterior.

Considere a 0-forma $g : \mathbb{R}^d \rightarrow \mathbb{R}$. Então a diferencial

$$dg = \sum_{i=1}^n \frac{\partial g}{\partial x_i} dx_i$$

é uma 1-forma. Agora, considerando f uma 1-forma, temos

$$\begin{aligned} d(f_i dx_i) &= \frac{\partial f_j}{\partial x_i} dx_i \wedge dx_j \\ &= \sum_{i < j} \frac{\partial f_j}{\partial x_i} dx_i \wedge dx_j + \sum_{i > j} \frac{\partial f_j}{\partial x_i} dx_i \wedge dx_j \end{aligned} \quad (\text{D.7.0.3})$$

$$= \sum_{i < j} \left(\frac{\partial f_j}{\partial x_i} - \frac{\partial f_i}{\partial x_j} \right) dx_i \wedge dx_j. \quad (\text{D.7.0.4})$$

De (D.7.0.3) para (D.7.0.4) trocamos os índices i e j e usamos o fato que $dx_i \wedge dx_j = -dx_j \wedge dx_i$.

Assim, queremos generalizar este processo definindo uma operação que assume r -formas em $(r+1)$ -formas, como se segue.

Definição 73. Seja $\theta = \sum f_I dx_I$ uma r -forma em \mathbb{R}^d . A diferencial exterior $d\theta$ é definida por

$$d\theta = \sum_I df_I \wedge dx_I.$$

Exemplo 33. Seja $\theta = xyz dx + yz dy + (x+z) dz$. Então sua diferencial exterior é dada por

$$\begin{aligned} d\theta &= d(xyz) \wedge dx + d(yz) \wedge dy + d(x+z) \wedge dz \\ &= (yz dx + xz dy + xy dz) \wedge dx + (z dy + y dz) \wedge dy + (dx + dz) \wedge dz \\ &= yz dx \wedge dx + xz dy \wedge dx + xy dz \wedge dx + z dy \wedge dy + y dz \wedge dy + dx \wedge dz + dz \wedge dz \\ &= -xz dx \wedge dy - xy dx \wedge dz - y dy \wedge dz + dx \wedge dz \\ &= -xz dx \wedge dy + (1 - xy) dx \wedge dz - y dy \wedge dz. \end{aligned}$$

O seguinte teorema informa algumas propriedades da diferencial exterior. A demonstração pode ser vista em Lee (2012, p.364) [15] e Spivak (1999, p.210) [25].

Teorema 74. Para cada variedade suave M , a derivada exterior de uma forma diferencial tem as seguintes propriedades

1. Se f é uma função suave de valor real (uma 0-forma), então df é a diferencial de f , definida como de costume por $df(X) = Xf$.
2. $d(\theta \wedge \eta) = d\theta \wedge \eta + (-1)^k \theta \wedge d\eta$, onde θ é uma k -forma e η é uma s -forma.
3. $d \circ d = d(d\theta) = d^2\theta = 0$.
4. $d(\theta_1 + \theta_2) = d\theta_1 + d\theta_2$, onde θ_1 e θ_2 são k -formas.

D.8 Derivada de Lie para k -formas

Proposição 75. *Suponha M uma variedade, $V \in \mathcal{X}(M)$ e ω, η formas diferenciais. Então*

$$L_V(\omega \wedge \eta) = (L_V \omega) \wedge \eta + \omega \wedge (L_V \eta).$$

Os próximos resultados mostram que a Proposição ?? é válida para formas diferenciais e uma fórmula para as derivadas de Lie de formas diferenciais, cuja demonstrações podem ser vista em Lee (2012, p.373) e Lee (2010, p.425) [15], respectivamente.

Proposição 76. *Se V é um campo de vetor e ω é uma forma diferencial então*

$$L_V(d\omega) = d(L_V \omega).$$

Teorema 77. (Fórmula de Cartan)

$$L_X k = (div X)k.$$

D.9 Integral de uma forma diferencial

Estamos interessados em definir a integral de uma forma diferencial de grau m sobre uma superfície m -dimensional orientada. Para isso, consideremos inicialmente o caso de uma forma contínua $\omega : U \rightarrow U_m(\mathbb{R}^m)$ num aberto $U \subset \mathbb{R}^m$. Para todo $x \in U$, temos

$$\omega(x) = a(x)dx_1 \wedge \dots \wedge dx_m,$$

onde $a : U \rightarrow \mathbb{R}$ é contínua. Dado um compacto J -mensurável $X \subset U$ ², definimos

$$\int_X \omega = \int_X a(x)dx.$$

E ainda, se ω é uma forma volume, definimos

$$vol(X) := \int_X \omega,$$

onde X é compacto em M . Mais detalhes sobre esse conceito pode ser visto em Lee (2012, p.422) [15].

Vamos considerar um difeomorfismo $\phi : U \rightarrow V$ onde $U, V \subset \mathbb{R}^m$. Indicaremos por $y = (y_1, \dots, y_m)$ os pontos de V , $x = (x_1, \dots, x_m)$ os pontos de U , dy_1, \dots, dy_m as diferenciais das coordenadas de V e por dx_1, \dots, dx_m as diferenciais das coordenadas de U . Dada a forma diferencial $\omega(y) = a(y)dy_1 \wedge \dots \wedge dy_m$ em V , para todo $x \in U$, o pullback $\phi^* \omega$ tem o valor

$$(\phi^* \omega)(x) = a(\phi(x)) \cdot \det J\phi(x) \cdot dx_1 \wedge \dots \wedge dx_m,$$

onde $\det J\phi(x)$ é o determinante da matriz jacobiana de ϕ no ponto x . A demonstração da igualdade acima pode ser vista em Spivack (1965, Teorema 4-9, p.90) [25].

Em vista da definição dada, o Teorema de Mudança de variáveis significa

$$\int_X \phi^* \omega = \int_{\phi(X)} \omega,$$

desde que ϕ preserve orientação, isto é, $\det J\phi(x) > 0 \forall x \in U$.

Observe que

$$\int_X \phi^* \omega = \int_{\phi(X)} \omega = vol(\phi(X)).$$

Dessa forma, se $\phi^* \omega = \omega$ então ϕ preserva volume.

²mensurável segundo Jordan, mais detalhes podem ser vistos em Lima (2009, p.169) [18].