

Volume 80, 2016

Editores

Alexandre Loureiro Madureira (Editor Chefe)

Laboratório Nacional de Computação Científica - LNCC

Petrópolis, RJ, Brasil

Amanda Liz Pacífico Manfrim Perticarrari

Universidade Estadual Paulista Júlio de Mesquita Filho - UNESP

Jaboticabal, SP, Brasil

Edson Luiz Cataldo Ferreira

Universidade Federal Fluminense - UFF

Niterói, RJ, Brasil

Eduardo V. O. Teixeira (Editor Executivo)

Universidade Federal do Ceará - UFC

Fortaleza, CE, Brasil

Jorge Manuel Vieira Capela

Universidade Estadual Paulista - UNESP

Araraquara, SP, Brasil

Sandra Augusta Santos

Universidade Estadual de Campinas - UNICAMP

Campinas, SP, Brasil

A Sociedade Brasileira de Matemática Aplicada e Computacional - SBMAC publica, desde as primeiras edições do evento, monografias dos cursos que são ministrados nos CNMAC.

Para a comemoração dos 25 anos da SBMAC, que ocorreu durante o XXVI CNMAC em 2003, foi criada a série **Notas em Matemática Aplicada** para publicar as monografias dos minicursos ministrados nos CNMAC, o que permaneceu até o XXXIII CNMAC em 2010.

A partir de 2011, a série passa a publicar, também, livros nas áreas de interesse da SBMAC. Os autores que submeterem textos à série Notas em Matemática Aplicada devem estar cientes de que poderão ser convidados a ministrarem minicursos nos eventos patrocinados pela SBMAC, em especial nos CNMAC, sobre assunto a que se refere o texto.

O livro deve ser preparado em **Latex (compatível com o Miktex versão 2.9)**, as **figuras em eps** e deve ter entre **80 e 150 páginas**. O texto deve ser redigido de forma clara, acompanhado de uma excelente revisão bibliográfica e de **exercícios de verificação de aprendizagem** ao final de cada capítulo.

Veja todos os títulos publicados nesta série na página
http://www.sbmac.org.br/p_notas.php

Introdução aos Métodos de Simulação e Análise Numérica para Equações Diferenciais Estocásticas

Felipe Carvalho Álvares da Silva

fellipe.alvares@gmail.com

Luiz Henrique Duczmal

duczmal@est.ufmg.br

Departamento de Estatística
Instituto de Ciências Exatas
Universidade Federal de Minas Gerais

Denise Burgarelli Duczmal

burgarelli@ufmg.br

Departamento de Matemática
Instituto de Ciências Exatas
Universidade Federal de Minas Gerais



Sociedade Brasileira de Matemática Aplicada e Computacional

São Carlos - SP, Brasil
2016

Coordenação Editorial: Igor Leite Freire

Coordenação Editorial da Série: Alexandre L. Madureira

Editora: SBMAC

Capa: Matheus Botossi Trindade

Patrocínio: SBMAC

Copyright ©2016 by Felipe C. Álvares da Silva, Luiz Henrique Duczmal e Denise Burgarelli Duczmal. Direitos reservados, 2016 pela SBMAC. A publicação nesta série não impede o autor de publicar parte ou a totalidade da obra por outra editora, em qualquer meio, desde que faça citação à edição original.

Catálogo elaborado pela Biblioteca do IBILCE/UNESP
Bibliotecária: Maria Luiza Fernandes Jardim Froner

Álvares, Felipe C.

Introdução aos Métodos de Simulação e Análise Numérica
para Equações Diferenciais Estocásticas - São Carlos, SP :
SBMAC, 2016, 124 p., 21.5 cm - (Notas em Matemática
Aplicada; v. 80)

e-ISBN 978-85-8215-069-6

1. Equações Diferenciais Estocásticas 2. Processo de Wiener
3. Integral de Itô 4. Discretização 5. Análise Numérica
6. Simulação Exata

I. Álvares, Felipe C. II. Duczmal, Luiz H.
III. Burgarelli, Denise . IV. Título. V. Série

CDD - 51

Conteúdo

Prefácio	xi
1 Introdução	1
1.1 Roteiro	2
2 Revisão de Probabilidade e Processos Estocásticos	3
2.1 Variáveis Aleatórias e Espaços de Probabilidade	3
2.2 Momentos e Esperança Condicional	4
2.3 Processos Estocásticos	5
3 Integral Estocástica	11
3.1 Processo de Wiener	11
3.1.1 Existência	13
3.1.2 Trajetórias	14
3.1.3 Processo de Wiener Discretizado	20
3.2 A Integral de Itô	21
3.3 Considerações Finais	30
3.4 Exercícios	31
4 Equações Diferenciais Estocásticas	33
4.1 Processos de Difusão	33
4.2 Solução de uma EDE	35
4.2.1 Soluções Fortes	36
4.2.2 Soluções Fracas	39
4.3 Fórmula de Itô	42
4.4 Transformação de Lamperti	48
4.5 Transformação de Cameron-Martin-Girsanov	50
4.6 Exercícios	54
5 Métodos de Discretização	55
5.1 Expansão de Itô-Taylor	55
5.2 Métodos de Euler-Maruyama e Milstein	58
5.3 Esquema Livre de Derivada	61
5.4 Método θ Estocástico	64
5.5 Método da Linearização Local	65
5.6 Miscelânea	68
5.7 Exercícios	70

6	Análise Numérica	71
6.1	Convergência	71
6.2	Estabilidade	74
6.2.1	Funções e Regiões de Estabilidade de Alguns Esquemas . . .	76
6.2.2	Método da Linearização Local	80
6.3	Avaliação Numérica	81
6.3.1	Convergência	81
6.3.2	Estabilidade	87
6.4	Exercícios	93
7	Simulação Exata	95
7.1	Algoritmo	95
7.2	Avaliação Numérica	102
7.3	Exercícios	107

Prefácio

O propósito do presente trabalho é realizar uma apresentação formal dos principais métodos de simulação e análise numérica no contexto das equações diferenciais estocásticas. A obra resulta da compilação de trabalhos desenvolvidos por um grupo de estudos iniciado na Universidade Federal de Minas Gerais (UFMG), pelos professores Luiz Duczmal e Denise Burgarelli, com o intuito de estudar e buscar novas aplicações para o tema. A fim de deixar o texto acessível aos leitores iniciantes, a primeira metade do livro é dedicada à discussão de conceitos fundamentais para o entendimento geral do tópico, evitando o desmoderado rigor matemático encontrado em outras obras da área; é indispensável apenas algum conhecimento em teoria da medida. Sempre que possível, noções intuitivas dos resultados são apresentadas em uma primeira instância de modo a valorizar a interpretação prática do conteúdo. Dado o envolvimento de estudantes de graduação e mestrado no projeto, acreditamos que o material é adequado tanto a alunos de matemática aplicada e computacional cursando os últimos períodos da graduação e interessados na aplicação desses modelos, quanto a alunos/pesquisadores com um conhecimento mais sólido que desejam focar seus estudos no desenvolvimento teórico de novos métodos.

Belo Horizonte, 29 de janeiro de 2015.

Felipe Álvares
Luiz Duczmal
Denise Burgarelli

Capítulo 1

Introdução

Equações diferenciais estocásticas (EDEs) constituem uma classe de modelos probabilísticos capazes de expandir as ideias já consolidadas no contexto das equações diferenciais ordinárias (EDOs), explicando de maneira mais eficaz diversos fenômenos observados na natureza e em nosso cotidiano. Como uma breve motivação, podemos citar que as equações diferenciais estocásticas tratam da união entre os campos das equações diferenciais e dos processos estocásticos, oferecendo não apenas uma crescente (e cativante) área de pesquisa mas também um grande leque de aplicações multidisciplinares. Sua grande contribuição está relacionada à modelagem formal de ruídos associados a sistemas dinâmicos observados em cenários complexos onde as fontes de perturbação são grandes o suficiente para impossibilitar o ajuste suave de trajetórias.

Assim como nos casos clássicos, a modelagem de fenômenos sob essa perspectiva exige a consolidação de uma base teórica consistente em conjunto com uma série de métodos numéricos para o tratamento das equações envolvidas. No entanto, vale ressaltar que a importância do ferramental numérico nesse contexto é ainda mais evidente se comparado aos casos clássicos. Isso se deve à complexidade das estruturas matemáticas envolvidas as quais fatalmente impossibilitam a obtenção de soluções analíticas na (quase) totalidade dos casos.

Em um primeiro contato com o tema, muitas perguntas surgem naturalmente:

- o que é uma equação diferencial estocástica?
- quais são suas características essenciais?
- como obter aproximações para os modelos?
- dado o contexto probabilístico, como avaliar propriedades como convergência e estabilidade das aproximações?

As respectivas respostas constituem basicamente a espinha dorsal deste material. Um roteiro mais detalhado das notas é apresentado a seguir visando facilitar a adequação da leitura aos objetivos de cada interessado. Infelizmente, não entramos em detalhes nas metodologias de modelagem dos fenômenos já que as possíveis discussões deste tópico seriam material suficiente para um novo projeto. Uma extensa seleção de referências disponíveis na literatura é oferecida ao final do texto para os leitores interessados em aprofundar ainda mais seus conhecimentos na área.

1.1 Roteiro

A fim de organizar a proposta deste trabalho, apresentamos a seguir um roteiro dos assuntos que serão abordados no decorrer dos próximos capítulos.

Os capítulos 2, 3 e 4 abrangem os alicerces do tema. O Capítulo 2 apresenta conceitos matemáticos subjacentes (apresentação de notações, definições e resultados básicos) à formalização teórica do universo das EDEs. Leitores mais familiarizados com a teoria dos processos estocásticos podem, em uma primeira leitura, pular este capítulo. Uma vez que o foco do texto está nas equações unidimensionais, os conceitos que se seguem levarão em consideração apenas funções em \mathbb{R} . Estes mesmos conceitos podem, sem exceção, ser generalizados para vetores (sistemas de EDEs) em \mathbb{R}^n , $n \in \mathbb{N}$ [19, 28], o que seria a continuação natural deste trabalho. O terceiro capítulo introduz os dois pilares do cálculo estocástico: processos de Wiener e integral de Itô. Começamos com um breve histórico a respeito dos processos de Wiener e com a discussão de algumas de suas propriedades fundamentais. A integral de Itô é abordada em seguida traçando um paralelo com a teoria clássica de Riemann-Stieltjes. O conceito de equação diferencial estocástica é enfim abordado no capítulo 4; uma construção intuitiva desses modelos é apresentada a partir da ideia de modelagem microscópica de processos de difusão. O capítulo é finalizado com a discussão de três resultados dentre os mais famosos do meio: fórmula de Itô, transformada de Lamperti e transformada de Girsanov.

Leitores mais experientes, interessados apenas nos métodos de simulação e análise numérica, podem iniciar seus estudos a partir do capítulo 5, o qual é reservado aos métodos de simulação baseados na discretização do domínio temporal. Apresentamos inicialmente uma versão estocástica da expansão de Taylor, permitindo assim uma associação direta entre métodos de discretização nos cenários estocástico e determinístico. A sequência trata do desenvolvimento dos métodos de Euler-Maruyama, Milstein, θ -estocástico e linearização local. O capítulo seguinte trata da análise numérica dos algoritmos abordados até então. A análise parte da discussão dos conceitos de convergência, consistência e estabilidade adaptados ao contexto das variáveis aleatórias. A fim de estimular a implementação computacional, focamos a discussão em análises experimentais dos métodos. Seguindo a proposta introdutória do livro, utilizamos experimentos bastante ilustrativos e que podem ser facilmente reproduzidos pelo leitor. O último capítulo trata da versão mais contemporânea de algoritmos: métodos de simulação exata. Esta sistemática trata basicamente da extensão do tradicional método da rejeição à simulação de trajetórias. Apresentamos apenas o modelo que originou esta perspectiva: o algoritmo EA1 de Beskos e Roberts. Referências às possíveis extensões do método são deixadas ao longo do capítulo.

Capítulo 2

Revisão de Probabilidade e Processos Estocásticos

O presente capítulo introduz conceitos básicos de probabilidade e processos estocásticos, os quais serão cruciais para o bom entendimento dos tópicos que se seguem. O objetivo principal é situar o leitor a respeito das notações e principais propriedades, destacando que uma cobertura total do tema extrapola bastante nosso enfoque. Abordagens mais completas a respeito destes tópicos podem ser obtidas em textos clássicos da área como Shiryaev [33], Billingsley [4], e Karlin e Taylor [18], por exemplo. É recomendável apenas que o leitor tenha alguma familiaridade com teoria da medida.

2.1 Variáveis Aleatórias e Espaços de Probabilidade

Seja Ω o conjunto de todos os resultados possíveis de um experimento aleatório de interesse. Ω poderia ser, por exemplo, um simples conjunto numérico ou em um subintervalo da reta, por exemplo. Na teoria da probabilidade, este conjunto é denominado espaço amostral e constitui a base de construção de todo o tópico.

Uma medida é uma extensão formal dos conceitos de comprimento, área e volume utilizados amplamente em aplicações envolvendo espaços Euclidianos com até três dimensões. Dado um espaço amostral Ω , uma medida é uma função de conjuntos cujo domínio engloba algumas das partes de Ω . Para que esta função seja bem definida, é necessário que tais partes satisfaçam algumas premissas básicas, as quais são definidas a seguir.

Definição 2.1. *Seja \mathcal{F} uma coleção de subconjuntos de um espaço amostral Ω . A família \mathcal{F} é denominada uma σ -álgebra se e somente se ela cumpre as propriedades abaixo.*

(i) $\emptyset \in \mathcal{F}$;

(ii) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$;

(iii) $A_i \in \mathcal{F}, i = 1, 2, 3, \dots \Rightarrow \cup A_i \in \mathcal{F}$.

Definição 2.2. Um espaço de probabilidade é uma tripla (Ω, \mathcal{F}, P) , onde Ω é um conjunto arbitrário, \mathcal{F} é uma σ -álgebra de conjuntos em Ω e P é uma medida de probabilidade definida em \mathcal{F} .

Um espaço de probabilidade, conforme a definição acima, é um caso particular de espaço de medida (Ω, \mathcal{F}, P) onde $P(\Omega) = 1$. Um exemplo importante é verificado quando $\Omega = [0, 1]$, $\mathcal{F} = \mathcal{B}([0, 1])$ (Borelianos em $[0, 1]$) e P é a medida de Lebesgue em $[0, 1]$.

Definição 2.3. Seja (Ω, \mathcal{F}, P) um espaço de probabilidade. Uma função $X : \Omega \rightarrow \mathbb{R}$ é uma variável aleatória se for mensurável, isto é, se

$$X^{-1}(A) = \{\omega \in \Omega; X(\omega) \in A\} \in \mathcal{F}, \quad \forall A \in \mathcal{B}(\mathbb{R}), \quad (2.1.1)$$

onde $\mathcal{B}(\mathbb{R})$ representa os Borelianos em \mathbb{R} .

Variáveis aleatórias induzem medidas de probabilidade em $\mathcal{B}(\mathbb{R})$. Para tal, basta definirmos

$$P_X(A) = P(X^{-1}(A)), \quad \forall A \in \mathcal{B}(\mathbb{R}). \quad (2.1.2)$$

Neste caso, P_X é denominada a lei ou a distribuição da variável aleatória X . Frizamos aqui que o termo $X^{-1}(A)$ denota a pré-imagem do conjunto A pela função X , a existência de uma função inversa X^{-1} neste contexto não é necessária.

Definição 2.4. Seja T um conjunto arbitrário. Dizemos que uma família de σ -álgebras $\{\mathcal{F}_t\}_{t \in T}$ é independente se, para qualquer conjunto finito de índices $\{t_1, t_2, \dots, t_j\} \subset T$, e quaisquer $A_{t_k} \in \mathcal{F}_{t_k}$, $k = 1, 2, \dots, j$, verificarmos

$$P\left(\bigcap_{k=1}^j A_{t_k}\right) = \prod_{k=1}^j P(A_{t_k}). \quad (2.1.3)$$

Além disso, se X é uma variável aleatória e \mathcal{F} é uma σ -álgebra, dizemos que X é independente de \mathcal{F} se $\sigma(X)$ e \mathcal{F} são independentes, onde $\sigma(X)$ é a σ -álgebra gerada por X .

O conceito de independência é naturalmente estendido a famílias de variáveis aleatórias. Dizemos que uma família de variáveis aleatórias $\{X_t\}_{t \in T}$ é independente se a família de σ -álgebras geradas $\{\sigma(X_t)\}_{t \in T}$ for independente.

2.2 Momentos e Esperança Condicional

Definição 2.5. Seja X uma variável aleatória definida no espaço de probabilidade (Ω, \mathcal{F}, P) . Definimos o p -ésimo momento de X como sendo

$$\mathbb{E}[X^p] = \int_{\Omega} [X(\omega)]^p P(d\omega). \quad (2.2.4)$$

Cada valor de p na definição acima revela uma característica distinta a respeito da variável aleatória X . Em particular, destacamos a esperança matemática

$$\mathbb{E}[X] = \int_{\Omega} X(\omega) P(d\omega). \quad (2.2.5)$$

Se considerarmos momentos centrados podemos obter ainda conceitos como o de variância

$$\text{Var}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2. \quad (2.2.6)$$

Os momentos de uma função podem também ser obtidos de uma maneira mais prática a partir de um dispositivo ligado à sua *função geradora de momentos*. Tal função possui o poder de especificar, de maneira alternativa, a distribuição de uma variável aleatória.

Definição 2.6. *A função geradora de momentos de uma variável aleatória X é dada por*

$$\phi_X(t) := \mathbb{E} [e^{tX}], \quad t \in \mathbb{R}, \quad (2.2.7)$$

caso a esperança exista. Esta função (caso exista) caracteriza de maneira única a distribuição de probabilidade da variável aleatória correspondente:

$$\phi_X \equiv \phi_Y \implies X \text{ e } Y \text{ são identicamente distribuídas.} \quad (2.2.8)$$

Além disso,

$$\mathbb{E} [X^p] = \frac{d^p \phi}{dt^p}(0), \quad p \in \mathbb{N}. \quad (2.2.9)$$

Definição 2.7. *Seja (Ω, \mathcal{F}, P) um espaço de probabilidade. Sejam X uma variável aleatória tal que $\mathbb{E}[|X|] < +\infty$ e $\mathcal{G} \subset \mathcal{F}$ uma σ -álgebra. A esperança condicional $\mathbb{E}[X|\mathcal{G}]$ de X com relação à \mathcal{G} é a única variável aleatória ($P - q.c.$)¹ que satisfaz*

1. $\mathbb{E}[X|\mathcal{G}]$ é \mathcal{G} -mensurável;
2. $\int_G \mathbb{E}[X(\omega)|\mathcal{G}]P(d\omega) = \int_G X(\omega)P(d\omega), \forall G \in \mathcal{G}$.

Tanto a existência quanto a unicidade da variável aleatória $\mathbb{E}[X|\mathcal{G}]$ são garantidos pelo teorema de Radon-Nikodym.

Proposição 2.1. *Sejam X, Y variáveis aleatórias e $\mathcal{H} \subset \mathcal{G} \subset \mathcal{F}$ σ -álgebras. Temos:*

- (a) $\mathbb{E}[aX + Y|\mathcal{G}] = a\mathbb{E}[X|\mathcal{G}] + \mathbb{E}[Y|\mathcal{G}], \forall a \in \mathbb{R}$;
- (b) se Y é \mathcal{G} -mensurável, então $\mathbb{E}[XY|\mathcal{G}] = Y\mathbb{E}[X|\mathcal{G}]$;
- (c) $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}]$;
- (d) se Y é independente de \mathcal{G} , então $\mathbb{E}[Y|\mathcal{G}] = \mathbb{E}[Y]$;
- (e) $\mathbb{E}[\mathbb{E}[Y|\mathcal{G}]] = \mathbb{E}[Y]$.

Assim como na definição anterior, os itens desta proposição seguem basicamente das propriedades da derivada de Radon-Nikodym. Uma demonstração formal pode ser obtida em [33].

2.3 Processos Estocásticos

Definição 2.8. *Um processo estocástico é uma família*

$$\{X_t\}_{t \in T} = \{X_t; t \in T\}$$

de variáveis aleatórias em um espaço de probabilidade (Ω, \mathcal{F}, P) , parametrizada por um conjunto de índices T .

¹ $P - q.c.$: quase certamente segundo a medida de probabilidade P .

Para nossos propósitos, o conjunto T será usualmente o intervalo $[0, +\infty)$, o que nos permite interpretar os índices $t \in T$ como instantes de tempo. Devemos ressaltar que para cada instante $t \in T$ fixo, temos uma variável aleatória

$$\omega \rightarrow X_t(\omega), \quad \omega \in \Omega.$$

Por outro lado, fixando um ponto amostral $\omega \in \Omega$, temos uma função

$$t \rightarrow X_t(\omega), \quad t \in T,$$

a qual é denominada caminho amostral (trajetória) de X_t .

Definição 2.9. *Sejam $X = \{X_t; 0 \leq t < +\infty\}$ e $Y = \{Y_t; 0 \leq t < +\infty\}$ dois processos estocásticos definidos em um mesmo espaço de probabilidade (Ω, \mathcal{F}, P) . Dizemos que*

(a) *Y é uma modificação (versão) de X se, para qualquer $t \geq 0$, temos que*

$$P(X_t = Y_t) = 1. \quad (2.3.10)$$

(b) *X e Y possuem as mesmas distribuições finito-dimensionais se, para quaisquer $n \in \mathbb{N}$, $0 \leq t_1 < t_2 < \dots < t_n < +\infty$ reais, e $A \in \mathcal{B}(\mathbb{R}^n)$, temos que*

$$P[(X_{t_1}, X_{t_2}, \dots, X_{t_n}) \in A] = P[(Y_{t_1}, Y_{t_2}, \dots, Y_{t_n}) \in A]. \quad (2.3.11)$$

(c) *X e Y são indistinguíveis se suas trajetórias coincidirem $P - q.c.$*

$$P(X_t = Y_t; \forall 0 \leq t < +\infty) = 1. \quad (2.3.12)$$

A propriedade descrita pelo item (c) é a mais forte dentre as três já que

$$(c) \Rightarrow (a) \Rightarrow (b).$$

Perceba ainda que dois processos estocásticos podem ser modificações um do outro e possuir trajetórias com propriedade antagônicas. Logo,

$$(b) \not\Rightarrow (c).$$

O contra-exemplo clássico é dado pelos processos

$$X_t \equiv 0 \quad \text{e} \quad \begin{cases} Y_t = 0, & t \neq T \\ Y_t = 1, & t = T \end{cases}, \quad (2.3.13)$$

onde T é uma variável aleatória contínua não negativa. Y é uma modificação de X , porque

$$P(X_t = Y_t) = P(T \neq t) = 1, \quad \forall t \geq 0, \quad (2.3.14)$$

no entanto,

$$P(X_t = Y_t; \forall t \geq 0) = 0. \quad (2.3.15)$$

Definição 2.10. *Seja $\mathbb{R}^{[0, \infty)}$ o espaço das funções reais $\omega = (\omega_t)$ definidas para qualquer $t \in [0, \infty)$. Um cilindro n -dimensional em $\mathbb{R}^{[0, \infty)}$ é um conjunto da forma*

$$C(A) := \{\omega \in \mathbb{R}^{[0, \infty)}; (\omega(t_1), \dots, \omega(t_n)) \in A\}, \quad (2.3.16)$$

onde $t_i \in [0, \infty)$, $i = 1, 2, \dots, n$ e A é um boreliano em \mathbb{R}^n .

A definição anterior pode ser entendida intuitivamente se considerarmos o caso particular do subconjunto de borelianos em \mathbb{R}^n da forma

$$A = B_1 \times B_2 \times \dots \times B_n,$$

sendo cada B_i um boreliano em \mathbb{R} . Com esta notação, o conjunto $C(A)$ pode ser visto como a coleção das funções reais que, nos instantes t_1, t_2, \dots, t_n , projetam suas imagens nas "janelas" B_1, B_2, \dots, B_n . Denotaremos por \mathbb{A} a álgebra gerada pelos cilindros (de qualquer dimensão finita) em $\mathbb{R}^{[0, \infty)}$ e por $\mathcal{B}(\mathbb{R}^{[0, \infty)})$ a menor σ -álgebra que contém \mathbb{A} . O conceito de cilindro pode ser aplicado na teoria dos processos estocásticos de modo a definir formalmente o conceito de *distribuição finito-dimensional* o qual foi citado superficialmente na definição 2.9.

Definição 2.11. *Seja T o conjunto das seqüências finitas $\tau = (t_1, \dots, t_n)$, $n \in \mathbb{N}$, de inteiros distintos e não negativos. Suponha que para cada τ de comprimento n , temos uma medida de probabilidade Q_τ no espaço mensurável $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$. Dizemos que a coleção $\{Q_\tau\}_{\tau \in T}$ é uma família de distribuições finito-dimensionais.*

Se P é uma medida de probabilidade em $(\mathbb{R}^{[0, \infty)}, \mathcal{B}(\mathbb{R}^{[0, \infty)}))$, então podemos definir uma família de distribuições finito-dimensionais por

$$Q_\tau(A) = P(C(A)) = P\left[\omega \in \mathbb{R}^{[0, \infty)}; (\omega(t_1), \dots, \omega(t_n)) \in A\right], \quad (2.3.17)$$

onde $A \in \mathcal{B}(\mathbb{R}^n)$ e $\tau = (t_1, \dots, t_n) \in T$. Na prática, estaremos mais interessados na recíproca deste resultado: dada uma família $\{Q_\tau\}_{\tau \in T}$, desejamos obter uma medida de probabilidade que satisfaz a relação acima. Esta propriedade é validada por um dos teoremas de extensão de Kolmogorov.

Teorema 2.1. (Kolmogorov) *Considere as seguintes condições de consistência:*

- (i) *se $\tau' = (t_{i_1}, t_{i_2}, \dots, t_{i_n})$ é uma permutação de $\tau = (t_1, t_2, \dots, t_n)$, então, para qualquer $A_i \in \mathcal{B}(\mathbb{R})$, $i = 1, 2, \dots, n$, temos que*

$$Q_\tau(A_1, A_2, \dots, A_n) = Q_{\tau'}(A_{i_1}, A_{i_2}, \dots, A_{i_n}); \quad (2.3.18)$$

- (ii) *se $\tau = (t_1, t_2, \dots, t_n)$ para $n \geq 1$, $\tau'' = (t_1, t_2, \dots, t_{n-1})$ e $A \in \mathcal{B}(\mathbb{R}^{n-1})$, então*

$$Q_\tau(A \times \mathbb{R}) = Q_{\tau''}(A). \quad (2.3.19)$$

Se $\{Q_\tau\}$ é uma família que satisfaz as condições (i) e (ii), então existe uma medida de probabilidade P em $(\mathbb{R}^{[0, \infty)}, \mathcal{B}(\mathbb{R}^{[0, \infty)}))$, tal que

$$Q_\tau(A) = P(C(A)), \quad \forall \tau \in T. \quad (2.3.20)$$

O teorema acima nos garante que qualquer processo estocástico $\{X_t; t \geq 0\}$ definido pelas *funções coordenadas* $X_t(\omega) := \omega(t)$, induz uma medida P no espaço $(\mathbb{R}^{[0, \infty)}, \mathcal{B}(\mathbb{R}^{[0, \infty)}))$. A demonstração do teorema pode ser obtida em [17].

Definição 2.12. *Sejam (Ω, \mathcal{F}, P) um espaço de probabilidade e T um conjunto arbitrário de índices. Uma filtração neste espaço é uma família de σ -álgebras $\{\mathcal{F}_t\}_{t \in T}$ tal que $\mathcal{F}_s \subset \mathcal{F}_t$, para quaisquer $s, t \in T$, $s \leq t$.*

Através das filtrações podemos rastrear uma quantidade de informações disponíveis até determinado instante de tempo ao invés de lidarmos simplesmente com o espaço amostral irrestrito. A interpretação dos processos estocásticos como um fluxo de informações ao longo do tempo torna plausível o questionamento do quanto o observador conhece a respeito do fenômeno no instante presente em comparação com o conhecimento já disponível do passado. Podemos ainda utilizar esta informação disponível para fazer melhores previsões a respeito do comportamento futuro do fenômeno. Dado um processo estocástico $\{X_t\}_{t \in T}$, a escolha mais simples de filtração é a gerada pelo próprio processo, isto é,

$$\mathcal{F}_t^X := \sigma(X_s; \forall s \in T \cap [0, t]), \quad (2.3.21)$$

a menor σ -álgebra com respeito a qual X_s é mensurável para todo $s \in T \cap [0, t]$.

Definição 2.13. *Seja (Ω, \mathcal{F}, P) um espaço de probabilidade. Dizemos que um processo estocástico $X = \{X_t; t \in T\}$ é adaptado a uma filtração $\{\mathcal{F}_t\}_{t \in T}$ se X_t for \mathcal{F}_t -mensurável, para todo $t \in T$.*

A partir do fluxo de informações estabelecido por uma dada filtração, podemos definir duas classes importantes de processos estocásticos: processos Markovianos e Martingales.

Definição 2.14. *Um processo estocásticos $X = \{X_t; 0 \leq t < +\infty\}$ definido em um espaço de probabilidade (Ω, \mathcal{F}, P) é denominado um processo Markoviano com condição inicial μ se*

$$P(X_0 \in B) = \mu(B) \quad (2.3.22)$$

e

$$P(X_{s+t} \in B | \mathcal{F}_s^X) = P(X_{s+t} \in B | X_s), \quad (2.3.23)$$

para quaisquer $s, t \geq 0$ e $B \in \mathcal{F}$.

A relação (2.3.23) é popularmente denominada *propriedade de Markov* e caracteriza basicamente a falta de memória dos processos estocásticos pertencentes à esta família. Tal propriedade assegura que um processo Markoviano pode ser predito basicamente a partir do montante de informação disponível no momento presente, deixando de lado todo o histórico anterior. Por simplicidade, assumiremos que \mathcal{F} é a σ -álgebra de Borel em \mathcal{R} e consideraremos apenas processos estocásticos associados a variáveis aleatórias X_t contínuas. Desta forma, é possível associar esta propriedade a uma função densidade de probabilidade (densidade de transição), a qual nos permitirá lidar com estes processos em situações práticas:

$$P(X_{s+t} \in B | X_s = x) = \int_B p_X(s, x, s+t, y) dy, \quad \forall B \in \mathcal{F}. \quad (2.3.24)$$

Como uma função densidade de probabilidade caracteriza de maneira única (quase certamente) uma medida de probabilidade, temos que um processo Markoviano pode ser caracterizado a partir do estabelecimento de uma densidade de transição.

A segunda classe de interesse diz respeito a sequências de variáveis aleatórias cujos estados presentes são suficientes para caracterizar os valores esperados (médias) no futuro. Esta situação caracteriza basicamente o que seria um "jogo justo", desta forma, tais processos foram nomeados em homenagem a um conjunto de estratégias de aposta bastante popular na França durante o século XVIII: *Martingales*.

Definição 2.15. *Sejam (Ω, \mathcal{F}, P) um espaço de probabilidade e $\{\mathcal{F}_t\}_{t \geq 0}$ uma filtração. Um submartingale (respectivamente, um supermartingale) com relação à filtração $\{\mathcal{F}_t\}_{t \geq 0}$ é um processo $X = \{X_t; 0 \leq t < +\infty\}$, adaptado a $\{\mathcal{F}_t\}_{t \geq 0}$, tal que*

(a) $\mathbb{E}|X_t|$, para $0 \leq t < +\infty$;

(b) $\mathbb{E}[X_t | \mathcal{F}_s] \geq X_s$ (respectivamente, $\mathbb{E}[X_t | \mathcal{F}_s] \leq X_s$) para $0 \leq s \leq t < +\infty$.

Dizemos ainda que X é um martingale com respeito a $\{\mathcal{F}_t\}_{t \geq 0}$ se for simultaneamente um supermartingale e um submartingale com respeito a $\{\mathcal{F}_t\}_{t \geq 0}$.

Capítulo 3

Integral Estocástica

O estudo das equações diferenciais estocásticas ou, mais amplamente, do cálculo estocástico é fundamentado na concepção de integrais com respeito a trajetórias aleatórias. A ideia central é generalizar a metodologia clássica de Lebesgue-Stieltjes de modo a acomodar integrais da forma

$$I(f, \omega) = \int_S^T f(t, \omega) dM_t(\omega),$$

onde f é uma função de uma trajetória aleatória e M é um processo estocástico cujas características serão discutidas adiante. O grande obstáculo observado na definição dessas integrais é a variação ilimitada dos integradores M em qualquer intervalo finito, o que impossibilita a definição tradicional de integral trajetória a trajetória (i.e., para cada realização $\omega \in \Omega$ separadamente). Conforme discutiremos a seguir, este problema só poderá ser compensado através da introdução de um novo conceito de integral.

Neste trabalho estaremos particularmente interessados na metodologia proposta por Itô [15] com respeito aos processos de Wiener (movimentos Brownianos). Desta forma, o capítulo é iniciado a partir da definição deste importante processo estocástico, discutindo algumas de suas principais propriedades. Em seguida a construção da integral de Itô é apresentada. Os leitores mais avançados podem recorrer aos trabalhos de Karatzas e Shreve [17], Protter [30] e Ikeda e Watanabe [14] para uma definição mais ampla de integral estocástica.

3.1 Processo de Wiener

Processos de Wiener definem modelos matemáticos fortemente relacionados ao fenômeno conhecido como *movimento Browniano*. Este padrão peculiar de movimento faz referência ao comportamento difusivo verificado ao longo do deslocamento de pequenas partículas em um espaço sujeito à influência de forças aleatórias conforme ilustrado na Figura 3.1, por exemplo. O estudo mais detalhado desses fenômenos teve início no século XIX com os trabalhos de Robert Brown [7], um botânico que se interessou pelo deslocamento aleatório de partículas de pólen suspensas na água. O desenvolvimento das simples observações de Brown não foi suficiente para esmiuçar física ou matematicamente as causas do curioso movimento, porém evidenciou duas características fundamentais: as trajetórias induzidas pelas partículas de pólen apresentavam um comportamento extremamente irregular (a

princípio seria impossível traçar uma tangente em qualquer ponto) e tais trajetórias não demonstravam nenhuma correlação quando provenientes de partículas distintas.



Figura 3.1: Trajetória típica de um movimento Browniano bidimensional.

O que parecia ser uma dinâmica desprezível voltou a ser foco de interesse científico algumas décadas adiante com o matemático Louis Bachelier [1] em virtude da sua descrição de um fenômeno similar testemunhado na flutuação de ativos nos mercados financeiros. A tese defendida por Bachelier em 1900 deu mais relevância ao fenômeno previamente analisado por Brown, porém a falta de uma descrição mais formal, assim como de aplicações mais concretas, não contribuíram para a conquista do interesse dos principais físicos e matemáticos da época. Com o desenvolvimento da física quântica, sobretudo a teorização dos conceitos de átomos e moléculas como constituintes da matéria, Albert Einstein [10] propôs em 1905 uma explicação física precisa para o movimento Browniano: o comportamento imprevisível das partículas suspensas era fruto de choques aleatórios com as moléculas do fluido. As conclusões de Einstein alavancaram definitivamente a importância dessa conturbação como um fenômeno físico de transporte e consequentemente induziram uma série de pesquisadores a alinhar seus estudos com o tópico.

Apesar das grandes contribuições de Bachelier e Einstein, a formalização matemática do movimento Browniano como um processo estocástico foi concretizada duas décadas mais tarde com os trabalhos de Norbert Wiener [37, 38], razão pela qual este importante fenômeno é também denominado *processo de Wiener*. Atualmente sabemos que o movimento Browniano, como um fenômeno físico, é melhor explicado por um segundo modelo (ver Capítulo 4), entretanto as terminologias em homenagem a Brown e Wiener ainda se misturam naturalmente nas obras clássicas

do cálculo estocástico.

Seguindo a metodologia desenvolvida por Wiener, um movimento Browniano unidimensional $W = \{W_t; 0 \leq t < +\infty\}$ em um espaço de probabilidade (Ω, \mathcal{F}, P) é completamente especificado pelo conjunto de hipóteses:

1. W é um processo Gaussiano em Ω ;
2. W tem trajetórias contínuas P -q.c.;
3. $\mathbb{E}[W_t] = 0$ e $\mathbb{E}[W_t W_s] \propto (t \wedge s), \forall t, s \geq 0$.

É comum utilizarmos a expressão *movimento Browniano padrão* quando queremos enfatizar a igualdade

$$\mathbb{E}[W_t W_s] = (t \wedge s), \quad \forall t, s \geq 0, \quad (3.1.1)$$

em detrimento da imprecisa relação de proporcionalidade imposta no terceiro item.

O modelo acima traduz de maneira satisfatória a natureza do fenômeno físico de transporte modelado pelos movimentos Brownianos em intervalos de tempo suficientemente pequenos. No entanto, uma formalização equivalente pode ser obtida em função dos incrementos do processo, dando origem a um modelo mais simples de se lidar na prática.

Definição 3.1. *Seja (Ω, \mathcal{F}, P) um espaço de probabilidade. Denotamos por processo de Wiener um processo estocástico $W = \{W_t, 0 \leq t < \infty\}$ tal que, para quaisquer $0 \leq s < t < +\infty$, valem:*

1. $P(W_0 = 0) = 1$;
2. $(W_t - W_s) \sim N(0, t - s)$;
3. $(W_t - W_s)$ é independente da σ -álgebra $\mathcal{F}_t = \sigma(W_u, \forall 0 \leq u \leq s)$.

O primeiro item acima trata apenas de uma convenção, podendo inclusive variar em função da aplicação de interesse. O segundo ponto garante que os incrementos do processo possuem distribuição normal centrada em zero e com variância $t - s$, ou ainda, garante que os incrementos são estacionários. Por fim, o último tópico garante que tais incrementos são independentes.

3.1.1 Existência

A existência de um processo estocástico com as propriedades da definição anterior pode ser verificada a partir do Teorema 2.1. Para tal, consideraremos inicialmente a função densidade de probabilidade Gaussiana:

$$p_G(x; \mu, \sigma^2) := \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2}, \quad x \in \mathbb{R}, \quad (3.1.2)$$

onde $\mu \in \mathbb{R}$ e $\sigma \in \mathbb{R}^+$ são constantes expressando a média e o desvio padrão da variável. Em seguida, construiremos um processo W cuja densidade de transição é dada por

$$\begin{aligned} p_W(r, x; s, y) &:= p_G(y; x, s - r) \\ &= \frac{1}{\sqrt{2\pi(s - r)}} e^{-(y-x)^2/2(s-r)}, \quad s > r \geq 0. \end{aligned} \quad (3.1.3)$$

⁰ $a \wedge b \equiv \min\{a, b\}, \quad a, b \in \mathbb{R}$

Por fim, notemos que tal densidade induz medidas finito dimensionais da forma

$$\nu_{(s_1, \dots, s_n)}(A_1, \dots, A_n) := \int_{A_1} \cdots \int_{A_n} \prod_{k=1}^n p_W(s_{k-1}, y_{k-1}; s_k, y_k) dy_n \dots dy_1, \quad (3.1.4)$$

para quaisquer

$$0 < s_1 < s_2 < \dots < s_n \quad \text{e} \quad (A_1, \dots, A_n) \in \mathcal{B}(\mathbb{R}^n).$$

Esta construção equivale à afirmação de que os incrementos $\{W_{s_k} - W_{s_{k-1}}\}_{k=1}^n$ são independentes e normalmente distribuídos com média zero e variância $s_k - s_{k-1}$ (Exercício 3). Além disso, podemos verificar (Exercício 4) que as medidas acima satisfazem as condições de consistência de Kolmogorov, logo, pelo Teorema 2.1, é possível obter uma medida de probabilidade P no espaço $(\mathbb{R}^{[0, \infty)}, \mathcal{B}(\mathbb{R}^{[0, \infty)})$ com respeito a qual o processo de Wiener está bem definido.

3.1.2 Trajetórias

Os resultados discutidos até então asseguram a existência de um processo estocástico com o cunho de um processo de Wiener, porém ainda não detalham possíveis características a respeito das trajetórias realizadas por este processo. Veremos a seguir que tais trajetórias fazem parte de uma atípica classe de funções composta por elementos simultaneamente contínuos e não-diferenciáveis.

Continuidade

Iniciaremos uma breve análise topológica apresentando uma ideia específica de continuidade, a qual nos permitirá entender com mais naturalidade algumas das propriedades destas trajetórias.

Definição 3.2. (*Continuidade de Hölder*) Uma função $f : \mathbb{R} \rightarrow \mathbb{R}$ é dita Hölder contínua com expoente $\gamma > 0$ se existir uma constante $C > 0$ tal que, para quaisquer $r, s \in \mathbb{R}$, vale:

$$|f(r) - f(s)| \leq C|r - s|^\gamma. \quad (3.1.5)$$

Fundamentalmente, uma função contínua é caracterizada pela propriedade

$$f(r) \rightarrow f(s) \quad \text{à medida que} \quad r \rightarrow s.$$

Desta forma, a ideia de continuidade de Hölder nos oferece uma noção da taxa com o qual esta convergência é observada: quanto maior o expoente γ , maior a taxa de convergência e, conseqüentemente, maior o grau de suavidade da função. Para $\gamma = 1$, por exemplo, temos o conceito de continuidade de Lipschitz.

O próximo resultado, também conhecido como critério de convergência de Kolmogorov, associa a continuidade das trajetórias de um processo estocástico a determinadas restrições impostas aos momentos de seus incrementos. Detalhes da demonstração podem ser encontrados, por exemplo, em Karatzas e Shreve [17].

Teorema 3.1. (*Kolmogorov e Čentsov*) Sejam T uma constante positiva arbitrária e $X = \{X_t; 0 \leq t \leq T\}$ um processo estocástico em (Ω, \mathcal{F}, P) que satisfaz

$$\mathbb{E}[|X_t - X_s|^\alpha] \leq C|t - s|^{1+\beta}, \quad 0 \leq s, t \leq T, \quad (3.1.6)$$

para algum trio de constantes positivas α, β e C . Então, existe uma modificação $X' = \{X'_t; 0 \leq t \leq T\}$ de X cujas trajetórias são localmente Hölder contínuas com expoente $\gamma \in \left(0, \frac{\beta}{\alpha}\right)$.

O critério de Kolmogorov pode ser aplicado facilmente às trajetórias de um processo de Wiener a fim de comprovar sua continuidade. Basta utilizarmos a tradicional função geradora de momentos da distribuição Gaussiana.

Proposição 3.1. *Existe uma modificação contínua do processo de Wiener .*

Demonstração: Dado $T \geq 0$, já verificamos que é possível obter um processo estocástico $W = \{W_t; 0 \leq t \leq T\}$ que satisfaz as condições da Definição 3.1. Como $(W_t - W_s) \sim N(0, t - s)$, para quaisquer $0 \leq s < t \leq T$, temos a seguinte função geradora de momentos:

$$\phi(\lambda) = e^{\frac{\lambda^2(t-s)}{2}}. \quad (3.1.7)$$

Se $k \in \mathbb{N}$ é uma constante par, então existe $l \in \mathbb{N}$ tal que $k = 2l$. Desta forma, para qualquer expoente par, temos:

$$\begin{aligned} \mathbb{E}[|W_t - W_s|^k] &= \mathbb{E}[|W_t - W_s|^{2l}] \\ &= \mathbb{E}[(W_t - W_s)^{2l}] \\ &= \frac{d^{2l}\phi}{d\lambda^{2l}}(0) = M_{2l}(t-s)^l, \end{aligned} \quad (3.1.8)$$

onde $M_{2l} = \prod_{k=0}^l (2k-1)$. A conclusão final segue do Teorema 3.1. ■

A proposição anterior pode ser utilizada para obter uma noção mais rigorosa a respeito da continuidade do processo de Wiener. Podemos concluir, com o suporte do Teorema 3.1, que o processo possui uma modificação Hölder contínua com expoente $\gamma \in (0, \frac{1}{2})$:

$$0 < \gamma < \frac{l-1}{2l} = \frac{1}{2} - \frac{1}{2l} \xrightarrow{l \rightarrow +\infty} \frac{1}{2}.$$

Além disso, é possível demonstrar formalmente [12] que o intervalo $(0, \frac{1}{2})$ restringe o domínio de γ , i.e., a modificação assegurada pela Proposição 3.1 é Hölder contínua com expoente γ *necessariamente* inferior a $\frac{1}{2}$. Assim, com probabilidade 1, obtemos a seguinte aproximação

$$|W_s - W_r| \leq C|s - r|^{\frac{1}{2}} \quad (3.1.9)$$

para alguma constante $C > 0$, ou ainda,

$$|W_s - W_r| = \mathcal{O}\left(\sqrt{|s - r|}\right). \quad (3.1.10)$$

Esta última relação nos dá uma estimativa do tamanho dos incrementos de um processo de Wiener com relação aos incrementos temporais correspondentes. Em notação diferencial, a equação 3.1.10 é popurlamente expressa como:

$$dW_t = \sqrt{dt}. \quad (3.1.11)$$

A continuidade das trajetórias de um processo de Wiener é uma característica fundamental já que a formalização deste processo vem da descrição de um fenômeno físico de movimento, o qual é obviamente contínuo. Devemos ressaltar, no entanto,

⁰A notação \mathcal{O} descreve o comportamento limite de seus respectivos argumentos:

$$f(x) = \mathcal{O}(g(x)) \text{ à medida que } x \rightarrow a \Leftrightarrow \limsup_{x \rightarrow a} \left| \frac{f(x)}{g(x)} \right| < +\infty$$

que os resultados demonstrados até então caracterizam apenas uma modificação contínua do processo de Wiener, não um processo contínuo na totalidade dos casos. A princípio, é possível construir realizações específicas carregadas de saltos locais (Exercício 5), porém tais construções podem ser desconsideradas por tratarem de observações com probabilidade nula.

Não-Diferenciabilidade

A simples comparação entre as ordens de grandeza de incrementos temporais e incrementos Brownianos traz à tona o desenvolvimento de uma análise intuitiva da possível diferenciabilidade das trajetórias em questão. Da relação 3.1.11, segue a seguinte estimativa

$$\frac{dW_t}{dt} \approx dt^{-1/2}, \quad (3.1.12)$$

a qual cresce indefinidamente à medida que o incremento dt se aproxima de zero. Podemos assim ter uma noção razoável de que as trajetórias de um processo de Wiener não possuem reta tangente em nenhum ponto. Sob uma perspectiva um pouco mais formal, podemos utilizar a renomada *lei do logaritmo iterado* [17] para obter um resultado mais preciso.

Teorema 3.2. (*Lei do Logaritmo Iterado*) *Seja $W = \{W_t; 0 \leq t < +\infty\}$ um processo de Wiener padrão. Então, para quase todo $\omega \in \Omega$, temos que*

$$\limsup_{t \rightarrow +\infty} \frac{W_t(\omega)}{\sqrt{2t \ln(\ln(t))}} = 1 \quad e \quad \liminf_{t \rightarrow +\infty} \frac{W_t(\omega)}{\sqrt{2t \ln(\ln(t))}} = -1. \quad (3.1.13)$$

Partindo do fato de que $\widetilde{W} = \{tW_{1/t}; 0 \leq t < +\infty\}$ (Exercício 6) é um processo de Wiener, podemos reescrever as relações acima como

$$\limsup_{t \downarrow 0} \frac{\widetilde{W}_t(\omega)}{\sqrt{2t \ln(\ln(1/t))}} = 1 \quad e \quad \liminf_{t \downarrow 0} \frac{\widetilde{W}_t(\omega)}{\sqrt{2t \ln(\ln(1/t))}} = -1. \quad (3.1.14)$$

Utilizando agora que $\widehat{W} = \{\widetilde{W}_{t+s} - \widetilde{W}_s; 0 \leq t < +\infty\}$ também é um processo de Wiener (Exercício 6), temos que para qualquer $\epsilon > 0$ existem seqüências $t_n, t'_n \downarrow 0$ (dependendo da realização ω) tais que

$$\widetilde{W}_{t_n+s}(\omega) - \widetilde{W}_s(\omega)t_n \geq (1 - \epsilon) \sqrt{\frac{2 \ln(\ln(1/t_n))}{t_n}} \quad (3.1.15)$$

e

$$\frac{\widetilde{W}_{t'_n+s}(\omega) - \widetilde{W}_s(\omega)}{t'_n} \leq (\epsilon - 1) \sqrt{\frac{2 \ln(\ln(1/t'_n))}{t'_n}}, \quad (3.1.16)$$

para n suficientemente grande. Como

$$\frac{2 \ln(\ln(1/t))}{t} \xrightarrow{t \downarrow 0} +\infty, \quad (3.1.17)$$

segue que

$$\frac{\widetilde{W}_{t_n+s}(\omega) - \widetilde{W}_s(\omega)}{t_n} \xrightarrow{t \downarrow 0} +\infty \quad e \quad \frac{\widetilde{W}_{t'_n+s}(\omega) - \widetilde{W}_s(\omega)}{t'_n} \xrightarrow{t \downarrow 0} -\infty. \quad (3.1.18)$$

Assim, o caminho amostral $W(\omega)$ não é diferenciável no ponto $t = s$. Uma vez que s é arbitrário, temos que uma trajetória qualquer do processo de Wiener não possui tangente em nenhum ponto. Aos leitores mais interessados, Breiman [6] oferece uma demonstração mais rigorosa (bastante popular na literatura) para este resultado; uma demonstração alternativa apoiada na ideia de continuidade de Hölder pode ser encontradas em Evans [11].

Variações Total e Quadrática

A não-diferenciabilidade das trajetórias de um processo de Wiener é uma característica intimamente relacionada às p -ésimas variações do processo, um conceito fundamental na transição do cálculo clássico para o cálculo estocástico.

Definição 3.3. (*p -ésima variação*) Seja $[\alpha, \beta]$ um intervalo em $[0, \infty)$ e suponhamos que

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

é uma sequência de partições de $[\alpha, \beta]$, com

$$\max_{0 \leq k \leq m_n - 1} \{t_{k+1}^n - t_k^n\} = |P_n| \rightarrow 0$$

à medida que $n \rightarrow \infty$. Se X é um processo estocástico contínuo, então, para qualquer $p > 0$, definimos a p -ésima variação de X em $[\alpha, \beta]$ por

$$\langle X, X \rangle_{\alpha, \beta}^{(p)} := \lim_{|P_n| \rightarrow 0} \sum_{k=0}^{m_n-1} \left(X_{t_{k+1}^n} - X_{t_k^n} \right)^p, \quad (3.1.19)$$

sendo o limite tomado em probabilidade.

No âmbito estocástico estaremos interessados em dois casos particulares, variações de primeira (*variação total*) e segunda (*variação quadrática*) ordem. A ideia de variação quadrática é pouco explorada na análise clássica uma vez que esta quantidade é sempre nula para funções continuamente diferenciáveis. Mais precisamente, é sempre nula para funções de variação total finita (as quais constituem a vasta maioria das funções estudadas em uma primeira análise do cálculo clássico):

$$\begin{aligned} \langle f, f \rangle_{\alpha, \beta}^{(1)} = C < +\infty &\Rightarrow \langle f, f \rangle_{\alpha, \beta}^{(2)} = \lim_{|P_n| \rightarrow 0} \sum_{k=0}^{m_n-1} \left(f_{t_{k+1}^n} - f_{t_k^n} \right)^2 \\ &\leq \lim_{|P_n| \rightarrow 0} \max_{P_n} \left| f_{t_{k+1}^n} - f_{t_k^n} \right| \sum_{k=0}^{m_n-1} \left| f_{t_{k+1}^n} - f_{t_k^n} \right| \\ &= \lim_{|P_n| \rightarrow 0} \max_{P_n} \left| f_{t_{k+1}^n} - f_{t_k^n} \right| \langle f, f \rangle_{\alpha, \beta}^{(1)} \\ &\leq C \lim_{|P_n| \rightarrow 0} \max_{P_n} \left| f_{t_{k+1}^n} - f_{t_k^n} \right| = 0. \end{aligned} \quad (3.1.20)$$

A variação quadrática de um processo pode ser interpretada, no campo da análise estocástica, como uma medida da "aleatoriedade acumulada" em um dado intervalo. Assim, quanto maior a variação quadrática de um processo, mais imprevisível (mais carregado de ruído) será o comportamento de suas respectivas trajetórias. Essa interpretação nos induz a associar os movimentos Brownianos a um processo de variação quadrática crescente, já que a cada novo instante de tempo a posição

mais recente pode ser obtida do acréscimo de um incremento Gaussiano ao histórico do trajeto. O teorema a seguir, em conjunto com a Figura 3.2, quantificam precisamente esta intuição.

Teorema 3.3. (*Varição quadrática do processo de Wiener*) *Sejam W um processo de Wiener padrão e $[\alpha, \beta] \subset \mathbb{R}^+$ um intervalo de tempo qualquer. Então,*

$$\sum_{k=0}^{m_n-1} (W_{t_{k+1}^n} - W_{t_k^n})^2 \xrightarrow{L_2} \beta - \alpha. \quad (3.1.21)$$

Demonstração: Considerando uma sequência de partições análoga à descrita na definição 3.3, temos

$$\beta - \alpha = \sum_{k=0}^{m_n-1} (t_{k+1}^n - t_k^n).$$

Assim,

$$\begin{aligned} \left\| \sum_{k=0}^{m_n-1} (W_{t_{k+1}^n} - W_{t_k^n})^2 - (\beta - \alpha) \right\|_2^2 &= \mathbb{E} \left[\left(\sum_{k=0}^{m_n-1} (W_{t_{k+1}^n} - W_{t_k^n})^2 - (\beta - \alpha) \right)^2 \right] \\ &= \mathbb{E} \left[\left(\sum_{k=0}^{m_n-1} \left((W_{t_{k+1}^n} - W_{t_k^n})^2 - (t_{k+1}^n - t_k^n) \right) \right)^2 \right] \\ &= \mathbb{E} \left[\sum_{k=0}^{m_n-1} \left((W_{t_{k+1}^n} - W_{t_k^n})^2 - (t_{k+1}^n - t_k^n) \right)^2 \right], \end{aligned} \quad (3.1.22)$$

já que

$$\left((W_{t_{k+1}^n} - W_{t_k^n})^2 - (t_{k+1}^n - t_k^n) \right)_k, \quad k = 0, 1, 2, \dots, m_n - 1$$

são variáveis independentes e têm média nula. Da relação $\mathbb{E}[X^4] = 3\sigma^2$, válida para qualquer variável aleatória $X \sim N(0, \sigma^2)$, obtemos:

$$\begin{aligned} \left\| \sum_{k=0}^{m_n-1} (W_{t_{k+1}^n} - W_{t_k^n})^2 - (\beta - \alpha) \right\|_2^2 &= 2 \sum_{k=0}^{m_n-1} (t_{k+1}^n - t_k^n)^2 \\ &\leq 2|P_n| \sum_{k=0}^{m_n-1} (t_{k+1}^n - t_k^n) \\ &= 2|P_n|(\beta - \alpha) \longrightarrow 0, \end{aligned} \quad (3.1.23)$$

o que conclui o resultado. ■

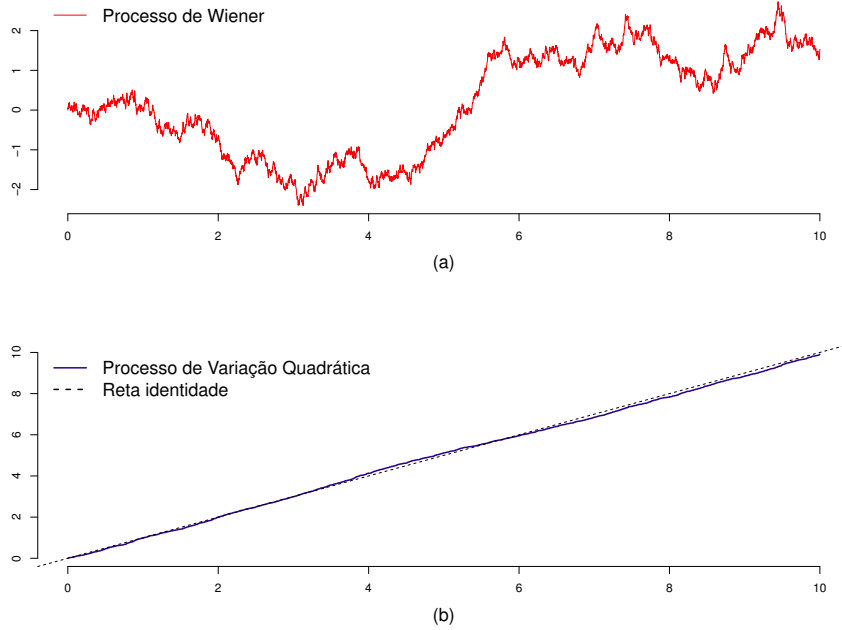


Figura 3.2: (a) Realização de uma trajetória típica do processo de Wiener. (b) Evolução do processo de variação quadrática $\langle W, W \rangle_{0,10}^2$ em comparação com o crescimento linear do tempo.

Como consequência do Teorema 3.3, temos que as trajetórias Brownianas serão caracterizadas por um processo de variação total ilimitada. Este resultado pode ser obtido imediatamente como uma adaptação do argumento 3.1.20, uma vez que as trajetórias em questão são contínuas quase certamente.

Teorema 3.4. (*Variação total do processo de Wiener*) *Sejam W um processo de Wiener padrão e $[\alpha, \beta] \subset \mathbb{R}^+$ um intervalo de tempo qualquer. Então,*

$$\sum_{k=0}^{m_n-1} |W_{t_{k+1}^n} - W_{t_k^n}| \xrightarrow{L_2} +\infty. \quad (3.1.24)$$

Demonstração: Já verificamos (3.1.20) que se uma trajetória contínua possui variação total finita, então sua respectiva variação quadrática será nula. A conclusão segue por contraposição. ■

A condição de variação total limitada é a premissa básica exigida de uma trajetória qualquer de modo a utilizá-la como elemento integrador na perspectiva de Remann-Stieltjes. Consideremos, por simplicidade, um integrando X contínuo no intervalo $[\alpha, \beta]$. Neste caso, dada uma sequência de partições

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

a possível definição de integral

$$\int_{\alpha}^{\beta} X_t dW_t = \lim_{|P_n| \rightarrow 0} \sum_{k=0}^{m_n-1} X_{t_k^n} (W_{t_{k+1}^n} - W_{t_k^n}), \quad (3.1.25)$$

com respeito a um processo de Wiener, seria inconsistente. Basta observar que, para qualquer intervalo finito $[\alpha, \beta]$,

$$\begin{aligned} \int_{\alpha}^{\beta} X_t dW_t &= \lim_{|P_n| \rightarrow 0} \sum_{k=0}^{m_n-1} X_{t_k^n} (W_{t_{k+1}^n} - W_{t_k^n}) \\ &\geq K_{\alpha, \beta} \lim_{|P_n| \rightarrow 0} \sum_{k=0}^{m_n-1} (W_{t_{k+1}^n} - W_{t_k^n}) \\ &= K_{\alpha, \beta} \langle W \rangle_{\alpha, \beta}^1 = +\infty, \end{aligned} \quad (3.1.26)$$

onde $K_{\alpha, \beta} = \min_{t \in [\alpha, \beta]} X_t$.

O maior empecilho da construção proposta acima está no modo de convergência empregado. Em particular, a ideia de convergência pontual explorada na definição de Riemann-Stieljes amarra o limite 3.1.25 à variação total do integrador (neste caso, ilimitada). Assim, uma tentativa consistente de construção partiria pela busca um modo de convergência capaz de descartar a variação total e condicionar a convergência do somatório 3.1.25 à variação quadrática do processo de Wiener. Esta é basicamente a essência da *integral de Itô*.

3.1.3 Processo de Wiener Discretizado

Para fins computacionais, é importante considerarmos o processo de Wiener discretizado, o qual nos permite simular realizações do processo de Wiener (como um processo contínuo) a partir de uma quantidade finita de pontos. Este tipo de procedimento pode ser realizado trivialmente se explorarmos as três propriedades básicas da Definição 3.1.

Inicialmente, se $W = \{W_t; 0 \leq t \leq T\}$ é um processo de Wiener padrão então sua versão discretizada $W^\delta = \{W_{t_k^\delta}; k = 0, 1, \dots, m_n\}$ também deve satisfazer a condição inicial $P(W_0^\delta = 0) = 1$.

Considerando que o intervalo de interesse foi particionado em $n + 1$ pontos igualmente espaçados, segue que os demais pontos da sequência podem ser gerados da relação:

$$W_{t_{k+1}^n} = W_{t_k^n} + \Delta \cdot Z, \quad (3.1.27)$$

onde $\Delta = \frac{T}{N}$ e Z é uma variável aleatória com distribuição normal padrão. Em outras palavras, a relação 3.1.27 indica que a posição mais atual de uma trajetória de Wiener é proveniente do acréscimo de um incremento Gaussiano ao último ponto observado. É imediato que tal relação satisfaz as propriedades 2 e 3 da Definição 3.1.

O método pode ainda ser facilmente estendido para partições adaptativas nas quais os pontos são espaçados de maneira arbitrária. Como

$$\Delta W_{t_{k+1}^n} \sim N(0, t_{k+1}^n - t_k^n), \quad (3.1.28)$$

basta ajustar o passo

$$\Delta = \Delta_k = t_{k+1}^n - t_k^n \quad (3.1.29)$$

a cada iteração.

O custo de simulação de uma trajetória específica é basicamente o custo de geração de N variáveis aleatórias Gaussianas independentes. Um algoritmo eficiente para a geração destas variáveis é o de Box-Muler [29], o qual baseia as gerações na transformação de variáveis uniformes.

3.2 A Integral de Itô

Seja W um processo de Wiener definido em um espaço de probabilidade (Ω, \mathcal{F}, P) . Uma vez caracterizado o processo W , podemos tratar da definição de integrais da forma

$$I_{\alpha, \beta}(f) = \int_{\alpha}^{\beta} f_t dW_t. \quad (3.2.30)$$

Diferentemente do cenário determinístico, a integral 3.2.30 define uma variável aleatória e não um número real. Mais geralmente,

$$I := \{I_{0,t}(f), t \geq 0\} \quad (3.2.31)$$

define um processo estocástico o qual toma valores em \mathbb{R} .

O primeiro passo para o desenvolvimento de uma definição consistente é associar a integral 3.2.30 a uma dada filtração, enriquecendo assim o conjunto de informações no qual o processo se apoiará. Consideremos uma filtração $\{\mathcal{F}_t\}_{t \geq 0}$ da forma

- (a) $\mathcal{F}_t \subset \mathcal{F}, \quad t \geq 0$;
- (b) $\mathcal{F}(W_s, 0 \leq s \leq t) \subset \mathcal{F}_t, \quad t \geq 0$;
- (c) $\mathcal{F}(W_{t+\lambda} - W_t, \lambda \geq 0)$ é independente de $\mathcal{F}_t, \quad t \geq 0$.

Caso não seja explicitada uma outra filtração, consideraremos que \mathcal{F}_t é a σ -álgebra \mathcal{F}_t^W gerada pelo processo W até o instante t . Tal convenção não passa de uma simplificação, podendo ser muito restritiva em determinados casos. De uma maneira geral, qualquer filtração \mathcal{H}_t satisfazendo (a) – (c) pode ser adotada para o desenvolvimento da teoria.

Em seguida, devemos restringir a classe de funções para as quais a integral pode ser bem definida. Devido à complexidade das estruturas envolvidas aqui, é natural esperarmos uma classe bastante restrita se comparada aos integrandos clássicos admitidos na perspectiva de Riemann-Stieltjes.

Definição 3.4. *Suponha que $\{\mathcal{H}_t\}_{t \geq 0}$ é uma filtração associada ao espaço (Ω, \mathcal{F}, P) , com respeito a qual o processo de Wiener $W = \{W_t, t \geq 0\}$ é uma martingale. Considere a classe das funções*

$$f(t, \omega) : [0, +\infty) \times \Omega \rightarrow \mathbb{R}$$

tais que

- (i) $(t, \omega) \rightarrow f(t, \omega)$ é $\mathbb{B} \times \mathcal{F}$ -mensurável, onde \mathbb{B} é a σ -álgebra de Borel do intervalo $[0, +\infty)$;
- (ii) $f(t, \omega)$ é \mathcal{H}_t -adaptada, isto é, para cada $t \geq 0, \omega \mapsto f(t, \omega)$ é \mathcal{H}_t -mensurável;
- (iii) $P \left[\int_{\alpha}^{\beta} |f(t, \omega)|^p dt < +\infty \right] = 1$;
- (iii') $\mathbb{E} \left[\int_{\alpha}^{\beta} |f(t, \omega)|^p dt \right] < +\infty$.

⁰ $\mathcal{F}(X_t, \alpha < t < \beta)$: filtração gerada pelo processo X entre os instantes α e β .

Denotamos por $\mathbb{L}_{\mathcal{H}}^p = \mathbb{L}_{\mathcal{H}}^p(\alpha, \beta)$ e $\mathbb{M}_{\mathcal{H}}^p = \mathbb{M}_{\mathcal{H}}^p(\alpha, \beta)$, $p < +\infty$, as classes de funções que satisfazem as hipóteses (i)–(iii) e (i)–(iii)', respectivamente. Por simplicidade, utilizaremos as notações \mathbb{L}^p e \mathbb{M}^p quando $\mathcal{H}_t \equiv \mathcal{F}_t^W$.

É imediato que $\mathbb{M}_{\mathcal{H}}^p \subseteq \mathbb{L}_{\mathcal{H}}^p$ já que toda variável aleatória integrável é finita quase certamente, no entanto a recíproca não é verdadeira. Durante a construção da integral de Itô estaremos particularmente interessados em funções da classe $\mathbb{M}_{\mathcal{H}}^2$. Esta escolha está relacionada à necessidade primária de modificar o modo de convergência utilizado nos conceitos clássicos de integração. Em particular, a classe $\mathbb{M}_{\mathcal{H}}^2$ engloba funções de quadrado integrável, permitindo uma associação imediata com o conceito de convergência em L_2 . Este conceito torna viável uma adaptação da aproximação 3.1.25 devido ao expoente quadrático trazido pelo novo modo de convergência, transferindo a exigência de limitação da variação total para a quadrática:

$$\begin{aligned} \mathbb{E} \left[\int_{\alpha}^{\beta} X_t dW_t \right]^2 &\approx \mathbb{E} \left[\sum_{k=0}^{m_n-1} X_{t_k^n} (W_{t_{k+1}^n} - W_{t_k^n}) \sum_{j=0}^{m_n-1} X_{t_j^n} (W_{t_{j+1}^n} - W_{t_j^n}) \right] \\ &= \mathbb{E} \left[\sum_{k=0}^{m_n-1} X_{t_k^n}^2 (W_{t_{k+1}^n} - W_{t_k^n})^2 \right] \\ &\leq M_{\alpha, \beta} \left[\sum_{k=0}^{m_n-1} (W_{t_{k+1}^n} - W_{t_k^n})^2 \right] \\ &\approx M_{\alpha, \beta} \langle W \rangle_{\alpha, \beta}^2 < +\infty, \end{aligned} \quad (3.2.32)$$

onde $M_{\alpha, \beta} = \max_{t \in [\alpha, \beta]} X_t^2$ (novamente consideramos, por simplicidade, um integrando X contínuo no intervalo $[\alpha, \beta]$).

A idéia em torno dessa construção é análoga à empregada na teoria da medida: definiremos $I(\phi)$ para uma classe elementar de funções $\phi \in \mathbb{M}_{\mathcal{H}}^2$ a qual permitirá a definição de propriedades importantes a partir de análises simples, em seguida mostraremos que cada $f \in \mathbb{M}_{\mathcal{H}}^2$ pode ser aproximada por uma sequência de funções elementares e, por fim, definiremos $I(f)$ como o limite de $I(\phi)$ à medida que $\phi \rightarrow f$.

Definição 3.5. Um processo estocástico $\phi \in \mathbb{M}^2$ é dito elementar se puder ser representado por

$$\phi(t, \omega) = \sum_{j=0}^n e_j(\omega) \mathbf{1}_{[t_j, t_{j+1})}(t), \quad (3.2.33)$$

onde $n \in \mathbb{N}$, e_1, e_2, \dots, e_n são variáveis aleatórias $\mathcal{F}_{t_1}^W, \mathcal{F}_{t_2}^W, \dots, \mathcal{F}_{t_n}^W$ -mensuráveis e $\cup_j [t_j, t_{j+1}) = [\alpha, \beta)$.

Ao exigir que e_j seja $\mathcal{F}_{t_j}^W$ -mensurável para todo j , asseguramos que o comportamento da função no intervalo $[0, t_{j+1})$ é completamente especificado pelo desenvolvimento de W até o instante t_j . Em outras palavras, ϕ desconsidera o "futuro" do processo de Wiener. Veremos adiante que esta restrição faz da integral de Itô um martingale, o que assegura à integral uma série de propriedades importantes. Outras definições de integral estocástica como a de Stratonovich [30], por exemplo, descartam esta hipótese, obtendo assim integrais com propriedades distintas.

Definição 3.6. A integral de Itô $I(\phi)$ de uma função elementar $\phi \in \mathbb{M}_{\omega}^2$ é dada pelo somatório

$$\int_s^r \phi(t, \omega) dW_t(\omega) = \sum_{j \geq 0} e_j(t_j, \omega) (W_{t_{j+1}} - W_{t_j})(\omega). \quad (3.2.34)$$

O próximo resultado complementa a definição 3.6 definindo uma associação importante entre as perspectivas de Itô e Riemann-Stieltjes. Desta forma, vários problemas de convergência, a princípio muito complicados, poderão ser transferidos para o contexto clássico a fim de explorar a teoria já consolidada.

Teorema 3.5. (*Isometria de Itô*) Se $\phi \in \mathbb{M}^2$ é elementar, então

$$\mathbb{E} \left[\left(\int_{\alpha}^{\beta} \phi(t, \omega) dW_t(\omega) \right)^2 \right] = \int_{\alpha}^{\beta} \mathbb{E}[\phi(t, \omega)^2] dt. \quad (3.2.35)$$

Demonstração: Considere, por simplicidade, a seguinte representação para os incrementos do processo de Wiener:

$$\Delta_j(\omega) = (W_{t_{j+1}} - W_{t_j})(\omega). \quad (3.2.36)$$

Para cada $i < j$, temos que os elementos aleatórios

$$\Delta_j(\omega) \quad \text{e} \quad e_i(\omega)e_j(\omega)\Delta_i(\omega)$$

são independentes. Neste caso,

$$\begin{aligned} \mathbb{E}[e_i(t_i, \omega)e_j(t_j, \omega)\Delta_i(\omega)\Delta_j(\omega)] &= \mathbb{E}[e_i(t_i, \omega)e_j(t_j, \omega)\Delta_i(\omega)] \cdot \mathbb{E}[\Delta_j(\omega)] \\ &= \mathbb{E}[e_i(t_i, \omega)e_j(t_j, \omega)\Delta_i(\omega)] \cdot 0 \\ &= 0. \end{aligned} \quad (3.2.37)$$

Por um argumento análogo, temos, para cada $i = j$, que

$$\begin{aligned} \mathbb{E}[e_i(t_i, \omega)e_j(t_j, \omega)\Delta_i(\omega)\Delta_j(\omega)] &= \mathbb{E}[e_j(t_j, \omega)^2\Delta_j(\omega)^2] \\ &= \mathbb{E}[e_j(t_j, \omega)^2] \mathbb{E}[\Delta_j(\omega)^2] \\ &= \mathbb{E}[e_j(t_j, \omega)^2] (t_{j+1} - t_j). \end{aligned} \quad (3.2.38)$$

Logo,

$$\begin{aligned} \mathbb{E} \left[\left(\int_{\alpha}^{\beta} \phi(t, \omega) dW_t(\omega) \right)^2 \right] &= \mathbb{E} \left[\sum_{i=0}^n \sum_{j=0}^n e_i(t_i, \omega)e_j(t_j, \omega)\Delta_i(\omega)\Delta_j(\omega) \right] \\ &= \sum_{i=0}^n \mathbb{E}[e_i(t_i, \omega)^2\Delta_i(\omega)^2] \\ &\quad + 2 \sum_{i=0}^n \sum_{j=i+1}^n \mathbb{E}[e_i(t_i, \omega)e_j(t_j, \omega)\Delta_i(\omega)\Delta_j(\omega)] \\ &= \sum_{i=0}^n \mathbb{E}[e_i(t_i, \omega)^2] \mathbb{E}[\Delta_i(\omega)^2] \\ &\quad + 2 \sum_{i=0}^n \sum_{j=i+1}^n \mathbb{E}[e_i(t_i, \omega)e_j(t_j, \omega)\Delta_i] \mathbb{E}[(\omega)\Delta_j(\omega)] \\ &= \sum_{i=0}^n \mathbb{E}[e_j(t_j, \omega)^2] (t_{j+1} - t_j) \\ &= \int_{\alpha}^{\beta} \mathbb{E}[\phi(t, \omega)^2] dt. \end{aligned} \quad (3.2.39)$$

■

Lema 3.1. (Propriedades da Integral de Itô – processos elementares) Sejam $\phi, \nu \in \mathbb{M}^2$ funções elementares, $c \in \mathbb{R}$ uma constante e $0 \leq \alpha < \beta < \gamma$. Então, para quase todo $\omega \in \Omega$:

- (i) $\int_0^t \phi(s, \omega) dW_s(\omega)$ é \mathcal{F}_t -mensurável;
- (ii) $\int_\alpha^\gamma \phi(s, \omega) dW_s(\omega) = \int_\alpha^\beta \phi(s, \omega) dW_s(\omega) + \int_\beta^\gamma \phi(s, \omega) dW_s(\omega)$;
- (iii) $\int_0^t (c\phi(s, \omega) + \nu(s, \omega)) dW_s(\omega) = c \cdot \int_0^t \phi(s, \omega) dW_s(\omega) + \int_0^t \nu(s, \omega) dW_s(\omega)$;
- (iv) $\mathbb{E} \left[\int_0^t \phi(s, \omega) dW_s(\omega) \right] = 0$;
- (v) $\text{Var} \left[\int_0^t \phi(s, \omega) dW_s(\omega) \right] = \mathbb{E} \left[\int_0^t \phi^2(s, \omega) ds \right]$;
- (vi) $\int_0^t \phi(s, \omega) dW_s(\omega)$ é um martingale com respeito à filtração \mathcal{F}_t^W ;
- (vii) $P \left(\sup_{r \in [0, t]} \left| \int_0^r \phi(s, \omega) dW_s(\omega) \right| \geq \epsilon \right) \leq \frac{1}{\epsilon^2} \int_0^t \mathbb{E} [\phi^2(s, \omega)] ds, \quad \forall \epsilon > 0$.

Demonstração: (i) Se ϕ é uma função elementar, então

$$\phi(t, \omega) = \sum_{j=1}^m e_j(\omega) \mathbf{1}_{[t_j, t_{j+1})}(t), \quad (3.2.40)$$

onde e_1, e_2, \dots, e_m são variáveis aleatórias \mathcal{F}_{t_j} -mensuráveis, $j = 1, 2, \dots, m$. Além disso, os incrementos $W_{t_{j+1}} - W_{t_j}$ também são $\mathcal{F}_{t_j}^W$ -mensuráveis para qualquer j . Logo, o produto $e_j \cdot (W_{t_{j+1}} - W_{t_j})$ é $\mathcal{F}_{t_j}^W$ -mensurável para qualquer j . Concluímos assim que $\int_0^t \phi(s, \omega) dW_s(\omega)$ é \mathcal{F}_t -mensurável.

(ii)–(iii) Seguem imediatamente das propriedades elementares do somatório 3.2.34.

(iv) Para ϕ elementar:

$$\begin{aligned} \mathbb{E} \left[\int_0^t \phi(s, \omega) dW_s(\omega) \right] &= \mathbb{E} \left[\sum_{j=1}^m e_j(\omega) (W_{t_{j+1}} - W_{t_j})(\omega) \right] \\ &= \sum_{j=1}^m \mathbb{E} [e_j(\omega) (W_{t_{j+1}} - W_{t_j})(\omega)] \\ &= \sum_{j=1}^m \mathbb{E} [\mathbb{E} [e_j(\omega) (W_{t_{j+1}} - W_{t_j})(\omega) | \mathcal{F}_{t_j}]] \\ &= \sum_{j=1}^m \mathbb{E} [e_j(\omega) \mathbb{E} [(W_{t_{j+1}} - W_{t_j})(\omega) | \mathcal{F}_{t_j}]] \\ &= \sum_{j=1}^m \mathbb{E} [e_j(\omega) \cdot 0] = 0. \end{aligned} \quad (3.2.41)$$

(v) Como

$$\mathbb{E} \left[\int_0^t \phi(s, \omega) dW_s(\omega) \right] = 0 \quad (3.2.42)$$

e

$$\mathbb{E} \left[\left(\int_0^t \phi(s, \omega) dW_s(\omega) \right)^2 \right] = \mathbb{E} \left[\int_0^t \phi(s, \omega)^2 ds \right], \quad (3.2.43)$$

temos que

$$\begin{aligned} \text{Var} \left[\int_0^t \phi(s, \omega) dW_s(\omega) \right] &= \mathbb{E} \left[\int_0^t \phi(s, \omega) dW_s(\omega) \right]^2 - \mathbb{E}^2 \left[\int_0^t \phi(s, \omega) dW_s(\omega) \right] \\ &= \int_0^t \mathbb{E}[\phi^2(s, \omega)] ds + 0 \\ &= \int_0^t \mathbb{E}[\phi^2(s, \omega)] ds. \end{aligned} \quad (3.2.44)$$

(vi) Já verificamos que a integral é \mathcal{F}_t^W -adaptada. Falta mostrar que

$$\mathbb{E} \left[\left| \int_0^t \phi(s, \omega) dW_s(\omega) \right| \right] < +\infty \quad (3.2.45)$$

e

$$\mathbb{E} \left[\int_0^t \phi(s, \omega) dW_s(\omega) \middle| \mathcal{F}_u^W \right] = \int_0^u \phi(s, \omega) dW_s(\omega), \quad \forall u \in (0, t). \quad (3.2.46)$$

Para a primeira relação, basta observarmos que $\phi \in \mathbb{M}^2$ e aplicar a desigualdade de Lyapunov:

$$\begin{aligned} \mathbb{E} \left[\left| \int_0^t \phi(s, \omega) dW_s(\omega) \right| \right] &\leq \sqrt{\left| \mathbb{E} \left[\int_0^t \phi(s, \omega) dW_s(\omega) \right]^2 \right|} \\ &= \sqrt{\int_0^t \mathbb{E}[\phi^2(s, \omega)] ds} < +\infty. \end{aligned} \quad (3.2.47)$$

Finalmente, utilizando um argumento similar ao desenvolvido em (iv), obtemos

$$\begin{aligned} \mathbb{E} \left[\int_0^t \phi(s, \omega) dW_s(\omega) \middle| \mathcal{F}_u^W \right] &= \mathbb{E} \left[\int_0^u \phi(s, \omega) dW_s(\omega) + \int_u^t \phi(s, \omega) dW_s(\omega) \middle| \mathcal{F}_u^W \right] \\ &= \int_0^u \phi(s, \omega) dW_s(\omega) + \mathbb{E} \left[\int_u^t \phi(s, \omega) dW_s(\omega) \middle| \mathcal{F}_u^W \right] \\ &= \int_0^u \phi(s, \omega) dW_s(\omega) + 0 \\ &= \int_0^u \phi(s, \omega) dW_s(\omega), \quad \forall u \in (0, t). \end{aligned} \quad (3.2.48)$$

(vii) Pelas propriedades (v) e (vi), a integral de Itô para processos elementares é um martingale de quadrado integrável. Logo, o resultado segue da desigualdade limite para martingales. \blacksquare

A partir dos resultados obtidos para funções elementares podemos estender a definição de integral estocástica para funções em \mathbb{M}^2 , estabelecendo uma classe razoavelmente ampla de integradores. A aproximação será apresentada em três etapas seguindo a metodologia proposta por Øksendal [28].

Lema 3.2. *Seja $g \in \mathbb{M}^2$ uma função limitada tal que $t \mapsto g(t, \omega)$ é contínua para todo $\omega \in \Omega$. Então, existe uma sequência de funções elementares $\{\phi_n\}_{n \in \mathbb{N}} \subset \mathbb{M}^2$ tal que*

$$\lim_{n \rightarrow +\infty} \int_0^t \mathbb{E} [(g - \phi_n)^2] ds = 0. \quad (3.2.49)$$

Demonstração: Seja $\{P_n\} = \{0 = t_0^n < t_1^n < \dots < t_{m_n}^n = t\}$ uma sequência de partições do intervalo $[0, t]$ tal que

$$|P_n| = \sup_k |t_{k+1}^n - t_k^n| \xrightarrow{n \uparrow +\infty} 0.$$

Para cada partição, construiremos uma função elementar da forma

$$\phi_n(t, \omega) := \sum_j g(t_j, \omega) \cdot \mathbf{1}_{[t_j, t_{j+1})}(t). \quad (3.2.50)$$

Claramente, $\phi_n \in \mathbb{M}^2$, $\forall n \in \mathbb{N}$, e

$$\int_0^t |g(s, \omega) - \phi_n(s, \omega)|^2 dt \xrightarrow{n \uparrow +\infty} 0, \quad \forall s \in [0, t], \quad (3.2.51)$$

pois $s \mapsto g(s, \omega)$ é uma aplicação contínua para todo $\omega \in \Omega$. Além disso, como g é limitada, existe uma constante $M_g > 0$ tal que $\mathbb{E} [|g(s, \omega)|^2] \leq M_g$, $\forall s \in (0, t)$. Assim,

$$\int |g(s, \omega) - \phi_n(s, \omega)|^2 dt \leq 4 \int_0^t |g(s, \omega)|^2 dt \leq 4M_g t, \quad \forall s \in (0, t). \quad (3.2.52)$$

Pelo Teorema da Convergência Dominada:

$$\mathbb{E} \left[\int_0^t (g - \phi_n)^2 dt \right] \xrightarrow{n \uparrow +\infty} 0. \quad (3.2.53)$$

A permutação entre os operadores de integração e esperança é válida já que os integrandos são sempre não negativos (Teorema de Tonelli):

$$\lim_{n \rightarrow +\infty} \int_0^t \mathbb{E} [(g - \phi_n)^2] ds = 0. \quad (3.2.54)$$

■

Lema 3.3. *Seja $h \in \mathbb{M}^2(0, T)$ uma função limitada. Então, existe uma sequência de funções limitadas $\{g_n\}_{n \in \mathbb{N}} \subset \mathbb{M}^2$, tal que $g_n(\cdot, \omega)$ é contínua para quaisquer $\omega \in \Omega$ e $n \in \mathbb{N}$, e*

$$\lim_{n \rightarrow +\infty} \mathbb{E} \left[\int_0^t (h - g_n)^2 ds \right] = 0. \quad (3.2.55)$$

Demonstração: Como h é uma função limitada, podemos obter uma constante $\bar{M}_h > 0$ tal que $|h(s, \omega)| \leq \bar{M}_h$, para quaisquer $s \in [0, t]$ e $\omega \in \Omega$. Seja $\{\phi_n\}_{n \in \mathbb{N}}$ uma sequência de núcleos de Dirac, isto é, uma sequência de funções contínuas não-negativas $\phi_n : \mathbb{R} \rightarrow \mathbb{R}$ tal que

$$\phi_n(x) = 0 \text{ para } x \in \left(-\infty, -\frac{1}{n}\right) \cup [0, +\infty) \text{ e } \int_{\mathbb{R}} \phi_n(x) dx = 1, \forall n \in \mathbb{N}. \quad (3.2.56)$$

Definindo

$$g_n(s, \omega) = \int_0^s \phi_n(u-s)h(u, \omega)du, \quad (3.2.57)$$

teremos uma sequência de funções \mathcal{F}_t^W -adaptadas pois a integral acima depende da função h avaliada apenas em pontos inferiores a t . Além disso, temos que $|g_n| \leq M$, para qualquer n ,

$$\begin{aligned} |g_n(s, \omega)| &\leq \int_0^s \phi_n(r-s)|h(r, \omega)dr \\ &\leq M_h \int_0^t \phi_n(r-s)dr \\ &\leq M_h \int_{\mathbb{R}} \phi_n(r-s)dr = M_h. \end{aligned} \quad (3.2.58)$$

Assim,

$$\mathbb{E} \left[\int_0^t |g_n(s, \omega)|^2 ds \right] \leq M_h t, \quad s \in (0, t), \quad (3.2.59)$$

assegurando que $g_n \in \mathbb{M}^2$, $\forall n \in \mathbb{N}$. É possível verificar também [13] que

$$\int_0^t |h(s, \omega) - g_n(s, \omega)|^2 ds \xrightarrow{n \uparrow +\infty} 0, \quad (3.2.60)$$

já que ϕ_n são identidades aproximadas. Para concluir o resultado, basta então mostrar que a integral acima é limitada a fim de habilitarmos o Teorema da Convergência Dominada. De fato,

$$\int_0^t |h(s, \omega) - g_n(s, \omega)|^2 ds \leq 4 \int_0^t |h(s, \omega)|^2 ds \leq 4M_h^2 t. \quad (3.2.61)$$

De maneira análoga ao lema anterior, a ordem dos operadores de integração e esperança pode ser permutada obtendo:

$$\lim_{n \rightarrow +\infty} \mathbb{E} \left[\int_0^T (h - g_n)^2 dt \right] = 0. \quad (3.2.62)$$

■

Lema 3.4. *Seja $f \in \mathbb{M}^2$ uma função arbitrária. Então, existe uma sequência de funções limitadas $\{h_n\}_n \subset \mathbb{M}^2$ tal que*

$$\lim_{n \rightarrow +\infty} \int_0^t \mathbb{E}[(f - h_n)^2] dt = 0. \quad (3.2.63)$$

Demonstração: Basta definir a sequência h_n como o truncamento de f

$$h_n(t, \omega) = \begin{cases} -n, & f(t, \omega) < -n \\ f(t, \omega), & |f(t, \omega)| \leq n \\ n, & f(t, \omega) > n \end{cases}, \quad (3.2.64)$$

e aplicar o teorema da convergência dominada. ■

Definição 3.7. *Seja $f \in \mathbb{M}^2$. A integral de Itô de f com respeito a um processo de Wiener W é dada pelo limite em $L^2(\Omega, \mathcal{F}, P)$*

$$\int_{\alpha}^{\beta} f(t, \omega) dW_t(\omega) = \lim_{n \rightarrow +\infty} \int_{\alpha}^{\beta} \phi_n(t, \omega) dW_t(\omega), \quad (3.2.65)$$

onde $\{\phi_n\}_{n \in \mathbb{N}} \subset \mathbb{M}^2$ é uma sequência de funções elementares tal que

$$\mathbb{E} \left[\int_{\alpha}^{\beta} |f(t, \omega) - \phi_n(t, \omega)|^2 dt \right] \rightarrow 0. \quad (3.2.66)$$

Como uma consequência imediata desta construção, as integrais de Itô para funções em \mathbb{M}^2 herdam as propriedades verificadas para processos elementares. Esta generalização pode ser verificada formalmente adaptando os argumentos utilizados na verificação do Lema 3.1.

Lema 3.5. *(Propriedades da Integral de Itô) Sejam $f, g \in \mathbb{M}^2$ funções quaisquer, $c \in \mathbb{R}$ uma constante e $0 \leq \alpha < \beta < \gamma$. Então, para quase todo $\omega \in \Omega$:*

- (i) $\int_0^t f(s, \omega) dW_s(\omega)$ é \mathcal{F}_t -mensurável;
- (ii) $\int_{\alpha}^{\gamma} f(t, \omega) dW_t(\omega) = \int_{\alpha}^{\beta} f(t, \omega) dW_t(\omega) + \int_{\beta}^{\gamma} f(t, \omega) dW_t(\omega)$;
- (iii) $\int_0^t (cf(s, \omega) + g(s, \omega)) dW_s(\omega) = c \cdot \int_0^t f(s, \omega) dW_s(\omega) + \int_0^t g(s, \omega) dW_s(\omega)$;
- (iv) $\mathbb{E} \left[\int_0^t f(s, \omega) dW_s(\omega) \right] = 0$;
- (v) $\text{Var} \left[\int_0^t f(s, \omega) dW_s(\omega) \right] = \int_0^t f^2(s, \omega) ds$;
- (vi) $\int_0^t f(s, \omega) dW_s(\omega)$ é um martingale com respeito à filtração \mathcal{F}_t^W
- (vii) $P \left(\sup_{s \in [0, t]} \left| \int_0^s f(r, \omega) dW_r(\omega) \right| \geq \epsilon \right) \leq \frac{1}{\epsilon^2} \int_0^t \mathbb{E} [f^2(s, \omega)] ds, \quad \forall \epsilon > 0.$

A integral de Itô, assim como o processo de Wiener, admite uma modificação com trajetórias contínuas. Este resultado, apesar de complexo, não é surpreendente se analisarmos a construção de Itô como uma extensão da metodologia de Riemann-Stieltjes. No contexto determinístico, sabemos que se $g : \mathbb{R} \rightarrow \mathbb{R}$ é uma função contínua de variação total limitada, então a integral de Stieltjes como uma função da forma

$$F(t) = \int_{t_0}^t f(\omega) dg(\omega), \quad (3.2.67)$$

está bem definida e também é contínua. No caso estocástico, o integrador W admite uma modificação contínua e o modo de convergência empregado permite uma definição similar de integral. Desta forma, é bastante razoável esperarmos que a análise se repita para

$$I_t(f) = \int_{t_0}^t f(t, \omega) dW_t(\omega). \quad (3.2.68)$$

O próximo resultado confirma esta intuição.

Teorema 3.6. *Existe uma modificação contínua da integral de Itô*

$$I_t(f) = \int_{t_0}^t f(t, \omega) dW_t(\omega), t \in [t_0, T], f \in \mathbb{M}^2. \quad (3.2.69)$$

Demonstração: Sejam $f \in \mathbb{M}^2$ uma função qualquer e $\{f_n\}_{n \in \mathbb{N}} \subset \mathbb{M}^2$ uma sequência de funções elementares tal que $f_n \rightarrow f$. Defina

$$I_t(f) := \int_{t_0}^t f(t, \omega) dW_t(\omega) \quad \text{e} \quad I_t^n(f) := \int_{t_0}^t f_n(t, \omega) dW_t(\omega). \quad (3.2.70)$$

Notemos primeiramente que

$$\begin{aligned} I_t^n(f) &= e_k(t_k^n, \omega)(W_t(\omega) - W_{t_k^n}(\omega)) \\ &\quad + \sum_{j=0}^{k-1} e_j(t_j^n, \omega)(W_{t_{j+1}^n}(\omega) - W_{t_j^n}(\omega)), \quad \forall n \in \mathbb{N}, \end{aligned} \quad (3.2.71)$$

onde

$$0 = t_0^n < t_1^n < \dots < t_{m_n}^n = T \quad \text{e} \quad t_k^n \leq t < t_{k+1}^n,$$

define uma função contínua quase certamente (este primeiro resultado segue, assim como no caso determinístico, da continuidade do integrador W). Neste caso, é suficiente mostrarmos que $I_t^n(f)$ converge uniformemente para $I_t(f)$. Como a sequência $\{f_n\}_{n \in \mathbb{N}}$ é arbitrária, podemos construí-la de modo a controlar convenientemente a velocidade de convergência. Em particular, consideremos que

$$\int_{t_0}^t \mathbb{E} [|f(s, \omega) - f_n(s, \omega)|^2] ds \leq \frac{1}{n^4}, \quad n \in \mathbb{N}. \quad (3.2.72)$$

Já verificamos que $I_t(f)$ e $I_t^n(f)$, $\forall n \in \mathbb{N}$, são exemplares de martingales de quadrado integrável. Portanto, a diferença

$$I_t(f) - I_t^n(f), \quad \forall n \in \mathbb{N},$$

também é um martingale de quadrado integrável. Pela desigualdade limite de martingales:

$$\begin{aligned} P \left(\sup_{s \in [t_0, T]} |I_t(f) - I_t^n(f)| \geq \frac{1}{n} \right) &\leq n^2 \int_{t_0}^t \mathbb{E} [|f(s, \omega) - f_n(s, \omega)|^2] ds \\ &\leq \frac{1}{n^2}, \quad n \in \mathbb{N}. \end{aligned} \quad (3.2.73)$$

Desta forma, a série

$$\sum_{n=1}^{+\infty} P \left(\sup_{s \in [t_0, T]} |I_t(f) - I_t^n(f)| \geq \frac{1}{n} \right)$$

é convergente e, pelo lema de Borel-Cantelli,

$$P \left(\sup_{s \in [t_0, T]} |I_t(f) - I_t^n(f)| \geq \frac{1}{n} \text{ infinitas vezes} \right) = 0. \quad (3.2.74)$$

Assim, $I_t^n(f)$ converge uniformemente para $I_t(f)$ (com probabilidade 1), conforme queríamos mostrar. ■

Os resultados discutidos até aqui formalizam a integral de Itô como uma estrutura matemática bem definida, porém não disponibilizam um algoritmo prático para o cálculo analítico dessas integrais. Diferentemente da teoria clássica de Riemann, não há um Teorema Fundamental do Cálculo Estocástico. Considere, por exemplo, o seguinte problema

$$I(\omega) = \int_{\alpha}^{\beta} W_t(\omega) dW_t(\omega), \quad \beta > \alpha \geq 0. \quad (3.2.75)$$

Um raciocínio natural para tentar solucionar esta integral seria nos guiar pelo Teorema Fundamental do Cálculo, o que nos levaria à seguinte expressão:

$$\int_{\alpha}^{\beta} W_t dW_t = \frac{W_{\beta}^2 - W_{\alpha}^2}{2}. \quad (3.2.76)$$

Todavia, podemos utilizar a descrição de Itô como contra-argumento para a última igualdade. Particionando o intervalo $[\alpha, \beta]$ e aproximando W por funções elementares, chegamos à seguinte aproximação para a integral:

$$\begin{aligned} \sum_k W_{t_k} \cdot (W_{t_{k+1}} - W_{t_k}) &= \sum_k (W_{t_k} W_{t_{k+1}} - W_{t_k}^2) \\ &= \frac{1}{2} \sum_k (W_{t_{k+1}}^2 - W_{t_k}^2) - \frac{1}{2} \sum_k (W_{t_{k+1}} - W_{t_k})^2 \\ &= \frac{1}{2} (W_s^2 - W_r^2) - \frac{1}{2} \sum_k (W_{t_{k+1}} - W_{t_k})^2. \end{aligned} \quad (3.2.77)$$

Passando ao limite (em L_2), temos que o último termo acima corresponde à variação quadrática do processo de Wiener. Desta forma,

$$\begin{aligned} \int_{\alpha}^{\beta} W_t dW_t &= \lim_{\Delta t_k \rightarrow 0} \sum_k W_{t_k} \cdot (W_{t_{k+1}} - W_{t_k}) \\ &= \frac{1}{2} (W_{\alpha}^2 - W_{\beta}^2) - \frac{1}{2} \langle W \rangle_{\alpha, \beta} \\ &= \frac{1}{2} (W_{\alpha}^2 - W_{\beta}^2) - \frac{1}{2} (\beta - \alpha). \end{aligned} \quad (3.2.78)$$

Na prática, não será sempre possível avaliar uma integral estocástica diretamente pela definição, o que traz uma excelente motivação para o desenvolvimento de uma metodologia numérica consistente.

3.3 Considerações Finais

Um dos pontos fundamentais na construção da Integral estocástica é a maneira com a qual definimos a ideia de função elementar. A construção apresentada aqui segue as diretrizes de Itô: dada uma função $f \in \mathbb{M}^2$, particione o domínio de integração em n subintervalos e monte funções elementares a partir dos valores de f nos extremos esquerdos destes intervalos. É possível verificar que a utilização de outros pontos

de referência, como o ponto médio ou o extremo direito dos intervalos, conduz a definições distintas de integral.

A título de exemplo, voltemos à seguinte integral

$$\int_{\alpha}^{\beta} W_t(\omega) dW_t(\omega).$$

Tomando como referência pontos da forma $\xi_k^n = (1 - \lambda)t_k^n + \lambda t_{k+1}^n$, teremos como resultado limite (Exercício 7):

$$\int_{\alpha}^{\beta} W_t(\omega) dW_t(\omega) = \frac{1}{2} W_{\beta}^2 + \left(\lambda - \frac{1}{2} \right) (\beta - \alpha). \quad (3.3.79)$$

Logo, cada escolha de λ resulta em um diferente ponto de convergência para o limite; um resultado totalmente distinto se comparado ao cálculo clássico. Em particular, a escolha $\lambda = 0$ nos leva à integral de Itô enquanto a escolha $\lambda = \frac{1}{2}$ resulta na *integral de Stratonovich*. Não apenas o resultado limite nestes casos é distinto, temos que as propriedades de cada uma dessas integrais são também distintas. Podemos verificar, por exemplo, que a integral de Stratonovich não define um martingale, no entanto ela cumpre o Teorema Fundamental do Cálculo. Uma discussão mais completa do tópico pode ser encontrada em Ikeda e Watanabe [14].

3.4 Exercícios

1. Sejam U_1 e U_2 duas variáveis aleatórias independentes uniformemente distribuídas no intervalo $[0, 1]$. Mostre que Z_1 e Z_2 dadas por

$$Z_1 := \cos(2\pi U_2) \sqrt{-2 \ln(U_1)} \quad \text{e} \quad Z_2 := \sin(2\pi U_2) \sqrt{-2 \ln(U_1)},$$

são variáveis independentes com distribuição Gaussiana padrão.

2. Mostre que o processo de Wiener $W = \{W_t; 0 \leq t \leq T\}$ é uma martingale com relação à escolha natural de filtração $\mathcal{F}_t = \sigma(W_s; 0 \leq s \leq t)$.
3. Mostre que a equação 3.1.4 é equivalente à afirmação de que os incrementos $\{W_{s_k} - W_{s_{k-1}}\}_{k=1}^n$ são independentes e normalmente distribuídos com média zero e variância $s_k - s_{k-1}$.
4. Mostre que as medidas finito dimensionais definidas pela equação 3.1.6 satisfazem as condições de consistência de Kolmogorov (Teorema 2.1).
5. Construa uma modificação descontínua do processo de Wiener.
6. Sejam $W = \{W_t; 0 \leq t < +\infty\}$ um processo de Wiener e s um instante de tempo fixo. Mostre que $\widetilde{W}_t = tW(1/t)$ e $\widehat{W}_t = W(t+s) - W(s)$ também são processos de Wiener.
7. Mostre que

$$\sum_{k=0}^{m_n-1} W_{\xi_k^n} (W_{t_{k+1}^n} - W_{t_k^n}) \xrightarrow{L_2} \frac{1}{2} W_{\beta}^2 + \left(\lambda - \frac{1}{2} \right) (\beta - \alpha),$$

onde

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

é uma sequência de partições do intervalo $[\alpha, \beta]$ e $\xi_k^n = (1 - \lambda)t_k^n + \lambda t_{k+1}^n$.

8. Mostre que

$$\int_0^t f(s) dW_t(\omega) = f(t)W_t - \int_0^t f'(s)W_s ds,$$

para qualquer função $f : [0, t] \rightarrow \mathbb{R}$ continuamente diferenciável.

Capítulo 4

Equações Diferenciais Estocásticas

Este capítulo é reservado à apresentação do conceito de equação diferencial no cenário estocástico. As ideias de solução de uma EDE, assim como questões concernentes à existência e unicidade, serão abordadas de modo a evidenciar os principais pontos de proximidade e discrepância com relação ao contexto determinístico. As últimas seções são dedicadas à discussão de alguns resultados célebres como a fórmula de Itô e a transformação de Girsanov.

4.1 Processos de Difusão

O conceito de *integral estocástica*, conforme formalizado por Itô [15], teve como principal motivação a investigação do comportamento microscópico (propriedades locais) de uma classe específica de modelos: os *processos de difusão*. Restringindo, por simplicidade, nossa análise ao caso das variáveis aleatórias X_t contínuas, temos que os processos de difusão podem ser caracterizados diretamente a partir de funções densidade de transição conforme estabelecido na definição abaixo.

Definição 4.1. (*Processo de Difusão*) Um processo Markoviano X com função densidade de transição $p(r, x; s, y)$ é denominado um processo de difusão se existirem funções $f, g : \mathbb{R}^+ \times \mathbb{R} \rightarrow \mathbb{R}$ tais que, para quaisquer $x \in \mathbb{R}$, $s > r \geq 0$ e $\epsilon > 0$, valem as relações:

$$\int_{|x-y|>\epsilon} p(r, x; s, y) dy = o(s-r), \quad (4.1.1)$$

$$\int_{|x-y|\leq\epsilon} (y-x)p(r, x; s, y) dy = f(r, x)(s-r) + o(s-r), \quad (4.1.2)$$

$$\int_{|x-y|\leq\epsilon} (y-x)^2 p(r, x; s, y) dy = g^2(r, x)(s-r) + o(s-r). \quad (4.1.3)$$

Em resumo, um processo de difusão pode ser entendido como um modelo Markoviano cujas trajetórias são quase certamente contínuas. Este aspecto é facilmente percebido através da condição (4.1.1), a qual assegura a ausência de saltos instantâneos nas trajetórias do movimento:

$$P(X_s - X_r > \epsilon | X_r = x) = o(s-r). \quad (4.1.4)$$

Uma análise mais precisa nos indica que um processo dessa natureza é definido fundamentalmente pelos coeficientes

$$\begin{cases} f(r, x) = \lim_{s \rightarrow r} \mathbb{E} \left[\frac{X_s - X_r}{s - r} \middle| X_r = x \right] \\ g^2(r, x) = \lim_{s \rightarrow r} \mathbb{E} \left[\frac{(X_s - X_r)^2}{s - r} \middle| X_r = x \right] \end{cases}, \quad (4.1.5)$$

usualmente denotados por *coeficiente de tendência (ou deriva)* e *coeficiente de difusão*, respectivamente. Notemos que o coeficiente de tendência f define localmente a velocidade média do movimento aleatório modelado por X enquanto o coeficiente de difusão g fornece uma medida aproximada da taxa de variação das flutuações médias de $X_s - X_r$ em torno de sua média. Uma análise mais rica dessa dinâmica sob o ponto de vista físico pode ser obtida no excelente trabalho de Nelson [25].

As interpretações de velocidade e taxa de variação nos desvios, atribuídas respectivamente aos coeficientes de tendência e difusão, podem ser associadas a uma variável aleatória auxiliar de modo a especificar uma aproximação local para o processo de difusão X . Para tal, basta tomarmos uma variável ξ_s , independente de X_t para qualquer $t \leq r < s$, tal que

$$\xi_s \sim N(0, s - r)$$

e, em seguida, definirmos

$$X_s - X_r \approx f(r, X_r)(s - r) + g(r, X_r)\xi_s. \quad (4.1.6)$$

Observemos que a variável auxiliar ξ_s introduz a aleatoriedade necessária para a modelagem do fenômeno e sua combinação com os coeficientes f e g assegura prontamente as condições (4.1.2)-(4.1.3):

$$\begin{aligned} \mathbb{E} \left[\frac{X_s - X_r}{s - r} \middle| X_r = x \right] &\approx f(r, x) + g(r, x) \mathbb{E} \left[\frac{\xi_s}{s - r} \right] \\ &= f(r, x) \xrightarrow{s \rightarrow r} f(r, x); \end{aligned} \quad (4.1.7)$$

$$\begin{aligned} \mathbb{E} \left[\frac{(X_s - X_r)^2}{s - r} \middle| X_r = x \right] &\approx f^2(r, x)(s - r) + 2f(r, x)g(r, x)\mathbb{E}[\xi_s] + g^2(r, x)\mathbb{E} \left[\frac{\xi_s^2}{s - r} \right] \\ &= f^2(r, x)(s - r) + g^2(r, x) \xrightarrow{s \rightarrow r} g^2(r, x). \end{aligned} \quad (4.1.8)$$

A aproximação descrita acima pode ser estendida a um conjunto enumerável de instantes temporais se tomarmos $\xi = \{\xi_t, t = 1, 2, 3, \dots\}$ como uma família de variáveis aleatórias independentes, com média nula e variância finita:

$$X_{t_n} - X_{t_{n-1}} \approx f(t_{n-1}, X_{t_{n-1}})(s - r) + g(t_{n-1}, X_{t_{n-1}})\xi_{t_n}, \quad (4.1.9)$$

$$\xi_{t_n} \sim N(0, t_n - t_{n-1}), \quad \xi_{t_{n_i}} \perp \xi_{t_{n_j}}, \forall i \neq j, \quad \xi_{t_n} \perp X_{t_{n-i}}, \forall i = 1, 2, \dots, n.$$

A escolha de uma seqüência como ξ não é feita por mero acaso. Seqüências dessa natureza são denominadas processos de *ruído branco Gaussiano* e são amplamente exploradas em aplicações envolvendo a modelagem de choques aleatórios. Traçando um paralelo com os processos de Wiener, temos que

$$\Delta W_{t_n} = W_{t_n} - W_{t_{n-1}} \sim N(0, t_n - t_{n-1}), \quad \forall n \geq 1 \quad (4.1.10)$$

⁰ $X \perp Y \Rightarrow X$ e Y são variáveis aleatórias independentes

ou seja, seus incrementos constituem um exemplar típico de ruído branco Gaussiano. Desta forma, é possível remodelar a aproximação (4.1.9) da seguinte forma

$$X_{t_n} - X_{t_{n-1}} \approx f(t_{n-1}, X_{t_{n-1}})(s - r) + g(t_{n-1}, X_{t_{n-1}})\Delta W_{t_n}, \quad (4.1.11)$$

onde W é um processo de Wiener. Podemos agora, passando ao limite, substituir as diferenças acima por diferenciais infinitesimais:

$$dX_t = f(t, X_t)dt + g(t, X_t)dW_t, \quad (4.1.12)$$

obtendo uma *equação diferencial estocástica*. Concluimos assim que uma EDE define um modelo de equação cujas soluções são caracterizadas por processos de difusão.

Exemplo 4.1. (*Movimento Browniano*) A formalização matemática dos processos de Wiener teve como principal motivação o desenvolvimento de uma descrição precisa dos movimentos Brownianos. A princípio, acreditava-se que o movimento errante de uma partícula suspensa em alguma forma de fluido sofreria a influência exclusiva de choques aleatórios frutos da interação da partícula com as moléculas do meio, dando origem à EDE

$$dX_t = \sigma dW_t, \quad \sigma \in \mathbb{R}^+, \quad (4.1.13)$$

cuja solução corresponde ao processo de Wiener definido anteriormente. Entretanto, uma análise mais criteriosa dos movimentos Brownianos como um fenômeno físico de transporte indica que o deslocamento destas partículas é também influenciado por um fenômeno de fricção proveniente da viscosidade do meio. Este efeito é proporcional à velocidade de deslocamento e, portanto, sugere a EDE

$$dX_t = -\lambda X_t dt + \sigma dW_t, \quad \lambda, \sigma \in \mathbb{R}^+, \quad (4.1.14)$$

conhecida como *Equação de Langevin*, a qual tem como solução formal um processo estocástico formalmente denominado processo de Ornstein-Uhlenbeck ou movimento Browniano geométrico; a Figura 4.1 ilustra o comportamento do processo para diferentes valores de λ . Apesar da descrição mais precisa do fenômeno físico movimento Browniano fornecida pela Equação (4.1.14), essa nomenclatura ainda se mistura muito à dos processos de Wiener por uma questão histórica.

4.2 Solução de uma EDE

Uma vez estabelecida a noção intuitiva de EDE como um modelo para processos de difusão, o próximo passo é traçar algumas das propriedades fundamentais acerca das soluções destes modelos. Aos leitores interessados, Jacod e Memin [16] e Stroock e Varadhan [36] apresentam discussões mais profundas em torno dessas propriedades.

O desenvolvimento de soluções para uma determinada EDE passa, assim como no caso determinístico, pela especificação de uma condição inicial. Desta forma, modelos descritos por EDEs também podem ser expressos como problemas de valor inicial (PVI):

$$dX_t = f(t, X_t)dt + g(t, X_t)dW_t, \quad t \in [0, T], \quad (4.2.15)$$

$$P(X_0 = \xi_0) = 1. \quad (4.2.16)$$

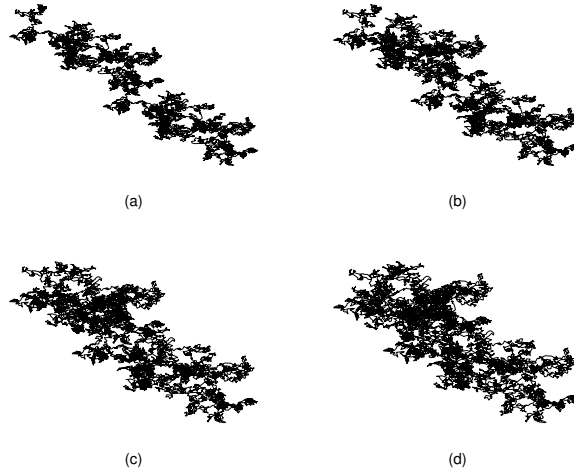


Figura 4.1: Movimentos Brownianos para distintos coeficientes de viscosidade: (a) $\lambda = 0$, $\sigma = 0$, (b) $\lambda = 0.05$, $\sigma = 0$, (c) $\lambda = 0.10$, $\sigma = 0$, (d) $\lambda = 0.15$, $\sigma = 0$.

Notemos que se $g \equiv 0$, então nosso problema se reduz a um PVI determinístico

$$dX_t = f(t, X_t)dt, \quad t \in [0, T], \quad (4.2.17)$$

$$P(X_0 = \xi_0) = 1, \quad (4.2.18)$$

cujas soluções são trajetórias contínuas partindo do ponto inicial observado. Desta forma, é razoável esperarmos alguma relação de proximidade entre as ideias de solução nos casos determinístico e estocástico. De fato, tal relação existe mas é não trivial; ao contrário do caso clássico, a aleatoriedade presente nos modelos estocásticos nos permite estabelecer dois conceitos de solução: soluções fortes (unicidade de trajetórias) e soluções fracas (unicidade em distribuição).

4.2.1 Soluções Fortes

Seja $W = \{W_t; 0 \leq t < +\infty\}$ um processo de Wiener padrão definido no espaço de probabilidade (Ω, \mathcal{F}, P) . Associaremos a este espaço a filtração aumentada

$$\mathcal{F}_t := \sigma(\mathcal{G}_t \cup \mathcal{N}), \quad (4.2.19)$$

onde

$$\mathcal{G}_t = \mathcal{F}(\xi_0, W_s; 0 \leq s \leq t) \quad \text{e} \quad \mathcal{N} = \{N \subseteq \Omega; \exists G \in \mathcal{G}_\infty \Rightarrow N \subseteq G, P(G) = 0\},$$

e assumiremos que ele é rico o suficiente a ponto de admitir que a construção da variável ξ_0 independa de \mathcal{F}_t^W , $\forall t$.

Definição 4.2. (Solução Forte) Uma solução forte para o problema (4.2.15)-(4.2.16), no espaço (Ω, \mathcal{F}, P) e com respeito a uma realização (trajetória fixa) do processo de Wiener W e à condição inicial ξ_0 , é um processo estocástico X de trajetórias contínuas, tal que

- i) X é \mathcal{F}_t -adaptado;

ii) $P(X_0 = \xi_0) = 1;$

iii) $P \left[\int_0^T (|f(t, X_t)| + |g(t, X_t)|^2) dt < +\infty \right] = 1, \forall T \in (0, +\infty).$

Dizemos ainda que esta solução é única se para quaisquer processos X e Y que satisfaçam (i) – (iii), vale

$$P(X_t = Y_t; 0 \leq t < +\infty) = 1, \tag{4.2.20}$$

isto é, se quaisquer duas soluções forem indistinguíveis.

As duas últimas condições da definição acima têm como missão assegurar que o problema é bem posto. Mais precisamente, (ii) exige que a solução se comporte segundo a variável aleatória ξ_0 no instante inicial de realização, enquanto a condição (iii) provê um cenário ideal para as implementações da integral de Riemann (associada à componente de tendência) e da isometria de Itô. A hipótese mais forte é imposta por (i); tal suposição nos permite entender o processo estocástico X como um sistema dinâmico alimentado por W e ξ_0 e cuja evolução é descrita pelo par ordenado (f, g) (ver Figura 4.2). Assim, temos que uma solução forte X em um instante t qualquer depende exclusivamente da condição inicial ξ_0 e da realização $\{W_s; 0 \leq s \leq t\}$.

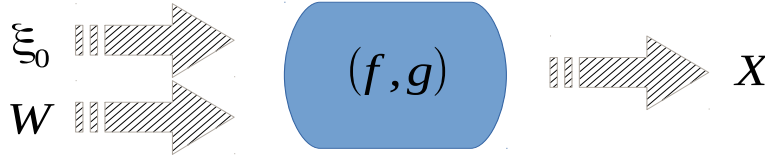


Figura 4.2: Solução forte X como um sistema dinâmico alimentado por uma condição inicial e um processo de Wiener. Dadas as realizações de W e ξ , o trajeto é especificado de maneira determinística pelos coeficiente f e g .

Teorema 4.1. (*Existência e Unicidade*) *Sejam f e g os coeficientes de tendência e de difusão, respectivamente, de uma dada equação diferencial estocástica. Suponha que esses coeficientes são Lipschitz contínuos e satisfazem uma restrição de crescimento linear*

$$|f(t, x) - f(t, y)| \leq K|x - y| \quad e \quad |g(t, x) - g(t, y)| \leq K|x - y|, \tag{4.2.21}$$

$$|f(t, x)|^2 \leq K^2(1 + |x|) \quad e \quad |g(t, x)|^2 \leq K^2(1 + |x|), \tag{4.2.22}$$

para todo $t \in [0, +\infty)$, $x \in \mathbb{R}$, $y \in \mathbb{R}$, onde K é uma constante positiva. Dado um espaço de probabilidade (Ω, \mathcal{F}, P) e uma realização $W(\omega)$ do processo de Wiener neste espaço, seja ξ_0 uma variável aleatória independente de \mathcal{F}_T^W , e com segundo momento finito

$$\mathbb{E} [|\xi_0|^2] < +\infty.$$

Considere também que a filtração (4.2.19) é adotada. Neste caso, existe um processo estocástico

$$X = \{X_t, 0 < t \leq T\}, \quad T \in (0, +\infty), \quad (4.2.23)$$

\mathcal{F}_t -adaptado, o qual representa a única solução forte do problema (4.2.15)-(4.2.16) relativo às observações de W e ξ_0 . Além disso,

$$\int_0^t \mathbb{E} [|X_s|^2] ds < +\infty, \quad \forall t \in (0, T]. \quad (4.2.24)$$

Demonstração: Começando pelo problema da unicidade, suponhamos que X e \widehat{X} são duas soluções fortes para o problema (4.2.15)-(4.2.16). Neste caso, as duas soluções coincidem no instante inicial e, para qualquer $t \in (0, T]$,

$$\mathbb{E} [|X_t - \widehat{X}_t|^2] = \mathbb{E} \left[\left| \int_0^t [f(s, X_s) - f(s, \widehat{X}_s)] ds + \int_0^t [g(s, X_s) - g(s, \widehat{X}_s)] dW_s \right|^2 \right]. \quad (4.2.25)$$

Da relação

$$(m + n)^2 \leq 2(m^2 + n^2), \quad \forall m, n \in \mathbb{R}, \quad (4.2.26)$$

obtemos

$$\mathbb{E} [|X_t - \widehat{X}_t|^2] \leq 2\mathbb{E} \left[\int_0^t [f(s, X_s) - f(s, \widehat{X}_s)] ds \right]^2 + 2\mathbb{E} \left[\int_0^t [g(s, X_s) - g(s, \widehat{X}_s)] dW_s \right]^2. \quad (4.2.27)$$

Aplicando agora a desigualdade de Cauchy-Schwarz em conjunto com a isometria de Itô:

$$\begin{aligned} \mathbb{E} [|X_t - \widehat{X}_t|^2] &\leq 2\mathbb{E} \left[\left(\int_0^t [f(s, X_s) - f(s, \widehat{X}_s)]^2 ds \right) \left(\int_0^t 1^2 ds \right) \right] \\ &\quad + 2 \int_0^t \mathbb{E} [g(s, X_s) - g(s, \widehat{X}_s)]^2 ds \\ &\leq 2t \int_0^t \mathbb{E} [f(s, X_s) - f(s, \widehat{X}_s)]^2 ds + 2 \int_0^t \mathbb{E} [g(s, X_s) - g(s, \widehat{X}_s)]^2 ds. \end{aligned} \quad (4.2.28)$$

Como os coeficientes de tendência e de difusão são Lipschitz contínuos, segue que

$$\begin{aligned} \mathbb{E} [|X_t - \widehat{X}_t|^2] &\leq (1 + 2t) \int_0^t [|f(s, X_s) - f(s, \widehat{X}_s)|^2 + |g(s, X_s) - g(s, \widehat{X}_s)|^2] ds \\ &\leq (1 + 2t) K^2 \int_0^t |X_s - \widehat{X}_s|^2 ds \\ &=: A_t \int_0^t |X_s - \widehat{X}_s|^2 ds. \end{aligned} \quad (4.2.29)$$

Desta forma, a função $\nu(t) := \mathbb{E} [|X_t - \widehat{X}_t|^2]$ satisfaz

$$\phi(t) \leq A_t \int_0^t \phi(s) ds, \quad \phi(0) = 0, \quad A_t > 0. \quad (4.2.30)$$

Portanto,

$$\phi(t) \equiv 0, \quad \forall t \in (0, T]. \quad (4.2.31)$$

Uma vez que t é arbitrário, segue que

$$P\left(X_t \neq \widehat{X}_t, \quad \forall t \in D\right) = \bigcap_{t \in D} P(\phi(t) \equiv 0) = 0 \quad (4.2.32)$$

para qualquer subconjunto denso D do intervalo $[0, T]$. Como as soluções são contínuas e coincidem em um subconjunto denso e enumerável de $[0, T]$, eles devem coincidir, quase certamente, no intervalo $[0, T]$ por completo. Temos assim que X e \widehat{X} são soluções indistinguíveis.

A demonstração da existência de soluções passa por um dispositivo similar às iterações de Picard-Lindelöf para equações diferenciais ordinárias: seja $\{\widetilde{X}_n\}_{n \geq 0}$ uma sequência tal que

$$\widetilde{X}_0 = \xi_0 \quad e \quad \widetilde{X}_{n+1} = \widetilde{X}_0 + \int_0^t f(s, \widetilde{X}_s) ds + \int_0^t g(s, \widetilde{X}_s) dW_s, \quad n \geq 1; \quad (4.2.33)$$

o objetivo é mostrar que \widetilde{X}_n converge para algum processo X o qual satisfaz o PVI (4.2.15)-(4.2.16). Destacamos que este dispositivo não explora de maneira contundente as características singulares da integral de Itô como um martingale, o que garante uma argumento bastante similar ao utilizado na versão determinística. Neste caso, omitiremos esse desenvolvimento (os detalhes podem ser encontrados em Friedman [12] p. 99–102). ■

A formalização do conceito de solução forte segue basicamente os princípios clássicos na teoria das equações diferenciais. A mesma similaridade entre as duas perspectivas será verificada no desenvolvimento de soluções a partir de métodos de discretização. A construção de uma solução aproximada para uma EDE passa inicialmente pela simulação de uma trajetória do processo de Wiener em conjunto com um ponto de partida; uma vez obtida tais instâncias, a evolução do restante da trajetória dependerá única e exclusivamente dos coeficientes f e g , conforme ilustrado na Figura 4.2. Podemos assim esperar extensões naturais de algoritmos como os de Euler e Runge-Kutta, por exemplo.

4.2.2 Soluções Fracas

Soluções fracas estão associadas a processos cujos coeficientes de tendência e de difusão são especificados, porém não estão associados a nenhuma realização específica do processo de Wiener.

Definição 4.3. (*Solução fraca*) Uma solução fraca para a equação (4.2.15) é uma tripla $(X, W), (\Omega, \mathcal{F}, P), \{\mathcal{H}_t\}_{t \geq 0}$, onde

- (a) (Ω, \mathcal{F}, P) é um espaço de probabilidade, e $\{\mathcal{H}_t\}_{t \geq 0}$ é uma filtração de sub- σ -álgebras de \mathcal{F} satisfazendo as hipóteses usuais;
- (b) X é um processo estocástico contínuo e W é um movimento Browniano, ambos \mathcal{H}_t -adaptados;
- (c) X satisfaz

$$X_t = X_0 + \int_0^t f(s, X_s) ds + \int_0^t g(s, X_s) dW_s; \quad 0 \leq t < +\infty. \quad (4.2.34)$$

Diferentemente do caso anterior, a filtração na definição (4.3) não é necessariamente a filtração aumentada gerada pelo processo de Wiener em união com a condição inicial. Desta forma, dado um instante t qualquer, temos que a variável aleatória X_t não é necessariamente obtida de uma função mensurável do trajeto $\{W_s, 0 \leq s \leq t\}$. Essa realação nos impede, a princípio, de pensar em unicidade de trajetórias conforme abordamos no caso das soluções fortes. Além disso, temos que qualquer informação, além de X_0 e $\{W_s, 0 \leq s \leq t\}$, necessária para a determinação de X_t deve ser independente de $\{W_u(\omega) - W_t(\omega), t \leq u < +\infty\}$, uma vez que X_t é incapaz de antecipar o futuro de W (ambos são \mathcal{H}_t -adaptados).

Considerando as observações acima, devemos esperar que o cenário imposto pela definição 4.3 não faz da existência de soluções fracas uma condição suficiente para a existência de soluções fortes dada uma realização de W . Todavia, é claro que a existência de soluções fortes culmina na existência de soluções fracas.

Exemplo 4.2. (*Equação de Tanaka*) Seja W um processo de Wiener padrão e defina um segundo processo \widetilde{W} por

$$\widetilde{W}_t := \int_0^t \operatorname{sgn}(W_s) dW_s, \quad (4.2.35)$$

onde $\operatorname{sgn}(\cdot)$ expressa a função sinal:

$$\operatorname{sgn}(x) = \begin{cases} +1, & x \geq 0 \\ -1, & x < 0 \end{cases}. \quad (4.2.36)$$

Em notação diferencial, temos que \widetilde{W} satisfaz a seguinte EDE:

$$d\widetilde{W}_t = \operatorname{sgn}(W_t) dW_t, \quad (4.2.37)$$

popularmente conhecida como equação de Tanaka, em homenagem ao matemático japonês Hiroshi Tanaka.

Pela simetria da distribuição Gaussiana, devemos esperar que os incrementos infinitesimais de W e \widetilde{W} descrevam o mesmo comportamento, ou seja, devemos esperar que \widetilde{W} também seja um processo de Wiener. De fato,

$$\langle \widetilde{W}, \widetilde{W} \rangle_{0,t}^2 = t, \quad (4.2.38)$$

reforçando nossa intuição: o processo de Wiener pode ser utilizado como uma solução fraca para 4.2.37. Logo, também é válida a relação

$$W_t = \operatorname{sgn}(\widetilde{W}_t) d\widetilde{W}_t. \quad (4.2.39)$$

Entretanto, podemos verificar que o mesmo problema não admite soluções fortes. A equação de Tanaka satisfaz ([17] fórmula 3.6.13)

$$\begin{aligned} \widetilde{W}_t &= \int_0^t \operatorname{sgn}(W_s) dW_s \\ &= |W_t| - \lim_{\epsilon \downarrow 0} (2\epsilon)^{-1} \int_0^t \mathbf{1}_{[0,\epsilon)}(|W_s|) ds, \end{aligned} \quad (4.2.40)$$

o que nos mostra que

$$\mathcal{F}_t^{\widetilde{W}} \subset \mathcal{F}_t^{|W|}. \quad (4.2.41)$$

Caso a equação de Tanaka admitisse soluções fortes, teríamos, da realção 4.2.39, que W seria $\mathcal{F}_t^{\tilde{W}}$ -adaptado

$$\mathcal{F}_t^W \subset \mathcal{F}_t^{\tilde{W}}. \quad (4.2.42)$$

Comparando as expressões (4.2.41)-(4.2.42), obtemos

$$\mathcal{F}_t^W \subset \mathcal{F}_t^{|W|}, \quad (4.2.43)$$

o que é um absurdo, já que tal inclusão nos permitiria determinar completamente o comportamento do processo W a partir de trajetórias de $|W|$.

A impossibilidade de lidar com trajetórias únicas nos leva à busca por um conceito mais flexível de unicidade. Em particular, podemos concentrar tal busca no comportamento macroscópico do problema, isto é, buscaremos unicidade nas distribuições de probabilidade associadas às soluções de uma EDE.

Definição 4.4. Dizemos que o problema (4.2.15)-(4.2.16) possui unicidade no sentido da distribuição de probabilidade se, para qualquer dupla de soluções fracas $(X, B), (\Omega, \mathcal{F}, P), \{\mathcal{H}_t\}_{t \geq 0}$, e $(\hat{X}, \hat{B}), (\hat{\Omega}, \hat{\mathcal{F}}, \hat{P}), \{\hat{\mathcal{H}}_t\}_{t \geq 0}$, com a mesma distribuição inicial,

$$P(X_0 \in B) = \hat{P}(\hat{X}_0 \in B); \quad \forall B \in \mathcal{B}(\mathbb{R}), \quad (4.2.44)$$

os processos X, \hat{X} induzirem o mesmo conjunto de distribuições finito-dimensionais.

Uma vez estabelecidas as noções intuitivas de solução fraca, é preciso, por uma questão de consistência, estabelecer critérios capazes de assegurar tanto a existência quanto a unicidade destas soluções. Seria plausível, por exemplo, nos ater às mesmas restrições impostas para o estudo de soluções fortes já que tais soluções cumprem as propriedades listadas na definição 4.3. Entretanto, tais hipóteses seriam excessivamente restritivas para os propósitos deste estudo. O próximo resultado lista um conjunto simples de propriedades que nos habilitará a trabalhar com uma classe específica de equações. A demonstração do teorema (ver [17], p. 303) faz uso de alguns resultados avançados de probabilidade e por isso não será abordada aqui.

Teorema 4.2. Considere a equação diferencial estocástica

$$dX_t = f(t, X_t)dt + dW_t; \quad 0 \leq t \leq T, \quad (4.2.45)$$

onde T é uma constante positiva, $W = \{W_t; 0 \leq t \leq T\}$ é um processo de Wiener e $f(t, x)$ é uma função Borel-mensurável.

(a) Se f satisfaz

$$|f(t, x)| \leq K(1 + |x|); \quad 0 \leq t \leq T, x \in \mathbb{R} \quad (4.2.46)$$

para alguma constante K . Então, para qualquer medida μ em $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, a equação (4.2.45) possui uma solução fraca com distribuição inicial μ .

(b) Dadas as condições do item (a), assumamos que $(X^{(i)}, B^{(i)}), (\Omega^{(i)}, \mathcal{F}^{(i)}, P^{(i)}), \{\mathcal{H}_t^{(i)}\}_{t \in [0, T]}$; $i = 1, 2$, são duas soluções fracas para (4.2.45) com a mesma distribuição inicial. Se

$$P^{(i)} \left(\int_0^T |f(t, X_t)|^2 dt < +\infty \right) = 1; \quad i = 1, 2, \quad (4.2.47)$$

então $(X^{(1)}, B^{(1)})$ e $(X^{(2)}, B^{(2)})$ possuem a mesma distribuição com respeito às suas respectivas medidas de probabilidade.

Modelos da forma 4.2.45 serão úteis quando lidarmos com os algoritmos de simulação exata de soluções. Veremos que esta classe específica de equações admite a geração de soluções a partir da distribuição de probabilidade induzida pelo modelo, uma metodologia consistente devido à noção de unicidade de soluções fracas.

4.3 Fórmula de Itô

A fórmula de Itô descreve uma importante transformação utilizada para determinar o diferencial de funções de um processo de difusão, atuando basicamente como uma versão estocástica da regra da cadeia. A fim de minuciar o elo entre as duas metodologias, iniciaremos a discussão com dois exemplos simples que ajudarão a elucidar o resultado formal.

Exemplo 4.3. *Sejam $W = \{W_t; r \leq t \leq s\}$, $s > r \geq 0$, um processo de Wiener padrão e $Y = \{u(t, W_t); r \leq t \leq s\}$ um processo obtido de W a partir da transformação $u(t, x) = tx$. Determinar o diferencial estocástico dY_t é equivalente a construir uma EDE:*

$$dY_t = f(t, Y_t)dt + g(t, Y_t)dW_t, \quad t \in [r, s]. \quad (4.3.48)$$

Para tal, consideremos inicialmente as seguintes integrais

$$\int_r^s t dW_t \quad e \quad \int_r^s W_t dt, \quad (4.3.49)$$

as quais podem ser obtidas como limites (em L_2) de somas da forma

$$\sum_k t_k \cdot (W_{t_{k+1}} - W_{t_k}) \quad e \quad \sum_k W_{t_{k+1}} \cdot (t_{k+1} - t_k). \quad (4.3.50)$$

Notemos que as aproximações construídas a partir do segundo somatório foram baseadas nos extremos direitos das partições, um resultado legítimo já que a continuidade da aplicação $t \mapsto W_t(\omega)$ nos habilita a trabalhar com a concepção de Riemann. Assim,

$$\begin{aligned} \int_r^s t dW_t + \int_r^s W_t dt &= \lim \left[\sum_k t_k \cdot (W_{t_{k+1}} - W_{t_k}) + \sum_k W_{t_{k+1}} \cdot (t_{k+1} - t_k) \right] \\ &= \lim \sum [t_k \cdot (W_{t_{k+1}} - W_{t_k}) + W_{t_{k+1}} \cdot (t_{k+1} - t_k)] \\ &= \lim \sum (t_{k+1} W_{t_{k+1}} - t_k W_{t_k}) = sW_s - rW_r. \end{aligned} \quad (4.3.51)$$

Reescrevendo a última expressão como uma EDE, obtemos

$$d(tW_t) = W_t dt + t dW_t, \quad t \in [r, s] \quad (4.3.52)$$

ou ainda,

$$\begin{aligned} dY_t &= W_t dt + t dW_t \\ &= \frac{\partial u}{\partial t}(t, W_t) dt + \frac{\partial u}{\partial x}(t, W_t) dW_t, \quad t \in [r, s]. \end{aligned} \quad (4.3.53)$$

O exemplo acima indica uma perfeita harmonia entre os diferenciais de transformações determinísticas e estocásticas, entretanto é possível mostrar que tal resultado trata-se de uma coincidência pontual; por mais suave que possa ser a transformação u , a mesma irregularidade que impede a diferenciabilidade das trajetórias de um processo de Wiener também será responsável por introduzir perturbações significativas no comportamento esperado de funções de um processo de difusão. O próximo exemplo ilustra bem esta situação.

Exemplo 4.4. *Seja W conforme o exemplo anterior e considere a seguinte transformação $Z = \{u(t, W_t); r \leq t \leq s\}$, onde $u(t, x) = x^2$. Já verificamos anteriormente (Equação 3.2.78) que*

$$\int_r^s W_t dW_t = \frac{1}{2}(W_s^2 - W_r^2) - \frac{1}{2}(s - r). \quad (4.3.54)$$

Logo,

$$W_s^2 = W_r^2 + 2 \int_r^s W_t dW_t + \int_r^s dt. \quad (4.3.55)$$

Em notação diferencial, obtemos

$$d(W_t^2) = dt + 2W_t dW_t, \quad t \in [r, s] \quad (4.3.56)$$

ou ainda,

$$\begin{aligned} dZ_t &= dt + 2W_t dW_t, \quad t \in [r, s] \\ &= \left(\frac{\partial u}{\partial t}(t, W_t) + \frac{1}{2} \frac{\partial^2 u}{\partial x^2}(t, W_t) \right) dt + \frac{\partial u}{\partial x}(t, W_t) dW_t. \end{aligned} \quad (4.3.57)$$

Diferentemente do resultado esperado segundo a sistemática clássica, um novo elemento foi acrescido ao coeficiente de tendência da EDE resultante do processo transformado no Exemplo 4.4. Este termo adicional surge naturalmente de modo a acomodar o efeito da irregularidade das trajetórias no processo de diferenciação da trajetória transformada. Em um cenário mais amplo, a análise informal dos incrementos infinitesimais da transformação nos permitirá especificar mais precisamente a origem desse novo termo. Para tal, construiremos um argumento apoiado em expansões via séries de Taylor.

Sejam X um processo de difusão, $u \in C^{1,2}([0, T] \times \mathbb{R})$ uma função qualquer e $U = u(t, X)$ uma transformação do processo X por u . Se $t_1 < t_2$ são dois instantes de tempo suficientemente próximos, podemos assumir que

$$t_2 - t_1 \approx dt \quad \text{e} \quad X_{t_2} - X_{t_1} \approx dX_t. \quad (4.3.58)$$

Desta forma, considerando a expansão de Taylor de segunda ordem para U :

$$\begin{aligned} U_{t_2} &\approx U_{t_1} + \frac{\partial u}{\partial t}(t_1, X_{t_1}) \cdot dt + \frac{\partial u}{\partial x}(t_1, X_{t_1}) \cdot dX_t + \frac{1}{2} \frac{\partial^2 u}{\partial t^2}(t_1, X_{t_1}) \cdot (dt)^2 \\ &\quad + \frac{\partial^2 u}{\partial t \partial x}(t_1, X_{t_1}) \cdot dt dX_t + \frac{1}{2} \frac{\partial^2 u}{\partial x^2}(t_1, X_{t_1}) \cdot (dX_t)^2. \end{aligned} \quad (4.3.59)$$

Termos da ordem $(dt)^a$ com $a > 1$ são muito próximos de zero, logo não exercem uma influência significativa na expressão acima e podem ser negligenciados. Além disso, sabendo que $dX_t = f(t, X_t)dt + g(t, X_t)dW_t$ e utilizando a aproximação $dW_t = \sqrt{dt}$, obtemos

$$dX_t = o(dt^{1/2}), \quad dt dX_t = o(dt^{3/2}) \quad \text{e} \quad (dX_t)^2 = o(dt). \quad (4.3.60)$$

Portanto,

$$U_{t_2} \approx U_{t_1} + \frac{\partial u}{\partial t}(t_1, X_{t_1}) \cdot dt + \frac{\partial u}{\partial x}(t_1, X_{t_1}) \cdot dX_t + \frac{1}{2} \frac{\partial^2 u}{\partial x^2}(t_1, X_{t_1}) \cdot (dX_t)^2, \quad (4.3.61)$$

ou, de uma maneira mais sucinta,

$$dU_t = u_t dt + \frac{1}{2} u_{xx} (dX_t)^2 + u_x dX_t. \quad (4.3.62)$$

A equação acima corresponde à *fórmula de Itô*, uma expressão amplamente utilizada no cálculo estocástico. Destacamos que o procedimento empregado para a obtenção da fórmula é bastante intuitivo e válido como motivação para o entendimento do tópico, porém não conta com o devido rigor matemático. Alguns das considerações empregadas não são triviais e podem instigar alguns questionamentos importantes a respeito da precisão do resultado. A seguir construiremos a fórmula de Itô de um ponto de vista mais formal: deduziremos inicialmente uma versão estocástica da regra do produto, em seguida obteremos a fórmula para transformações polinomiais e, por fim, estenderemos o resultado a uma classe mais ampla de transformações.

Teorema 4.3. (*Regra do Produto*) *Sejam X e Y processos de difusão satisfazendo*

$$\begin{cases} dX_t &= f_1(t, X_t)dt + g_1(t, X_t)dW_t \\ dY_t &= f_2(t, Y_t)dt + g_2(t, Y_t)dW_t \end{cases}, \quad t \in [0, T].$$

Então,

$$d(X_t Y_t) = X_t dY_t + Y_t dX_t + g_1(t, X_t)g_2(t, Y_t)dt. \quad (4.3.63)$$

Demonstração: Observemos inicialmente que

$$\begin{aligned} d(X_t Y_t) &= X_t dY + Y_t dX + g_1(t, X_t)g_2(t, Y_t)dt \\ &= X_t (f_2(t, Y_t)dt + g_2(t, Y_t)dW_t) + Y_t (f_1(t, X_t)dt + g_1(t, X_t)dW_t) \\ &\quad + g_1(t, X_t)g_2(t, Y_t)dt \\ &= (X_t f_2(t, Y_t) + Y_t f_1(t, X_t) + g_1(t, X_t)g_2(t, Y_t))dt \\ &\quad + (X_t g_2(t, Y_t) + Y_t g_1(t, X_t))dW_t, \end{aligned} \quad (4.3.64)$$

Verificar a igualdade acima é equivalente a demonstrar que, $\forall [r, s] \subseteq [0, T]$,

$$\begin{aligned} X_s Y_s &= X_r Y_r + \int_r^s (X_t f_2(t, Y_t) + Y_t f_1(t, X_t) + g_1(t, X_t)g_2(t, Y_t))dt \\ &\quad + \int_r^s (X_t g_2(t, Y_t) + Y_t g_1(t, X_t))dW_t. \end{aligned} \quad (4.3.65)$$

Suponha inicialmente que f_i e g_i são constantes em $[r, s]$. Logo,

$$\begin{cases} dX_t &= f_1 dt + g_1 dW_t \Rightarrow X_s = X_r + (s-r)f_1 + (W_s - W_r)g_1 \\ dY_t &= f_2 dt + g_2 dW_t \Rightarrow Y_s = Y_r + (s-r)f_2 + (W_s - W_r)g_2 \end{cases}. \quad (4.3.66)$$

Sem perda de generalidade, considere $r = 0$. Neste caso, dado $t \in [r, s]$, temos

$$X_t = X_r + t f_1 + W_t g_1; \quad (4.3.67)$$

$$Y_t = Y_r + t f_2 + W_t g_2. \quad (4.3.68)$$

Substituindo (4.3.67)-(4.3.68) na relação integral (4.3.65), obtemos:

$$\begin{aligned}
& \int_r^s [X_t f_2 + Y_t f_1 + g_1 g_2] dt + \int_r^s [X_t g_2 + Y_t g_1] dW_t \\
&= \int_r^s \{ [X_r + t f_1 + W_t g_1] f_2 + [Y_r + t f_2 + W_t g_2] f_1 + g_1 g_2 \} dt \\
&\quad + \int_r^s \{ [X_r + t f_1 + W_t g_1] g_2 + [Y_r + t f_2 + W_t g_2] g_1 \} dW_t \\
&= X_r (s f_2 + W_s g_2) + Y_r (s f_1 + W_s g_1) + s^2 f_1 f_2 + s g_1 g_2 \\
&\quad + (f_1 g_2 + f_2 g_1) \left(\int_r^s t dW_t + \int_r^s W_t dt \right) + 2 g_1 g_2 \int_r^s W_t dW_t. \quad (4.3.69)
\end{aligned}$$

As integrais estocásticas à direita da última equação podem ser tratadas analiticamente (Exemplos 4.3 e 4.4):

$$\begin{aligned}
& \int_r^s [X_t f_2 + Y_t f_1 + g_1 g_2] dt + \int_r^s [X_t g_2 + Y_t g_1] dW_t \\
&= X_r f_2 s + X_r g_2 W_s + Y_r f_1 s + Y_r g_1 W_s + f_1 g_2 s W_s + g_1 f_2 s W_s \\
&\quad + g_1 g_2 W_s^2 + f_1 f_2 s^2 \\
&= (X_r + s f_1 + W_s g_1)(Y_r + s f_2 + W_s g_2) - X_r Y_r \\
&= X_s Y_s - X_r Y_r, \quad (4.3.70)
\end{aligned}$$

concluindo o resultado para funções constantes. Consideremos agora que f_i e g_i são funções simples em $\mathbb{L}_{\mathcal{H}}^1(0, T)$ e $\mathbb{L}_{\mathcal{H}}^2(0, T)$, respectivamente. O resultado anterior pode então ser aplicado individualmente nos intervalos onde f_i e g_i são constantes:

$$\int_{t_n}^{t_{n+1}} [X_t f_2 + Y_t f_1 + g_1 g_2] dt + \int_{t_n}^{t_{n+1}} [X_t g_2 + Y_t g_1] dW_t = X_{t_{n+1}} Y_{t_{n+1}} - X_{t_n} Y_{t_n}, \quad (4.3.71)$$

onde $\tau = \{r = t_0 < t_1 < \dots < t_N = s\}$ é uma partição do intervalo $[r, s]$. Para o caso geral basta utilizarmos o fato de que $f_i, g_i, i = 1, 2$, podem ser aproximadas por funções simples. Assim, podemos concatenar as integrais obtidas acima e passar ao limite (em L_2). Da definição da integral de Itô, obtemos

$$\begin{aligned}
X_s Y_s &= X_r Y_r + \int_r^s (X_t f_2(t, Y_t) + Y_t f_1(t, X_t) + g_1(t, X_t) g_2(t, Y_t)) dt \\
&\quad + \int_r^s (X_t g_2(t, Y_t) + Y_t g_1(t, X_t)) dW_t, \quad (4.3.72)
\end{aligned}$$

como queríamos mostrar. ■

Rearranjando os termos da expressão obtida no teorema anterior, percebemos que a regra do produto pode ser entendida como uma versão estocástica do método de integração por partes:

$$\int_r^s X_t dY_t = X_t Y_t|_r^s - \int_r^s Y_t dX_t - \int_r^s g_1(t, X_t) g_2(t, Y_t) dt. \quad (4.3.73)$$

Note que se os diferenciais fossem determinísticos, isto é, se $g_1 \equiv g_2 \equiv 0$, então voltaríamos ao método clássico de integração por partes:

$$\int_s^r X_t dY_t = X_t Y_t|_s^r - \int_s^r Y_t dX_t. \quad (4.3.74)$$

Teorema 4.4. *Sejam $u : \mathbb{R} \rightarrow \mathbb{R}$ uma função da forma $u(x) = x^n$, $n \in \mathbb{N}$, e X um processo de difusão. Se $U = u(X)$, então U satisfaz a seguinte EDE:*

$$\begin{aligned} dU_t &= d(X_t^n) \\ &= nX_t^{n-1}dX_t + \frac{1}{2}n(n-1)X_t^{n-2}g^2(t, X_t)dt. \end{aligned} \quad (4.3.75)$$

Demonstração: Defina o seguinte conjunto

$$\mathcal{X} := \left\{ n \in \mathbb{N}; d(X_t^n) = nX_t^{n-1}dX_t + \frac{1}{2}n(n-1)X_t^{n-2}g^2(t, X_t)dt \right\}. \quad (4.3.76)$$

Nosso objetivo é verificar por indução que $\mathcal{X} = \mathbb{N}$. Para $n = 1$, temos:

$$1 \cdot X_t^{1-1}dX_t + \frac{1}{2} \cdot 1 \cdot (1-1) \cdot X_t^{1-2}g^2(t, X_t)dt = dX_t = d(X_t^1) \Rightarrow 1 \in \mathcal{X} \quad (4.3.77)$$

Para $n = 2$, podemos aplicar a regra do produto de Itô:

$$\begin{aligned} d(X_t^2) &= d(X_t X_t) \\ &= X_t dX_t + X_t dX_t + g(t, X_t) \cdot g(t, X_t)dt \\ &= 2X_t dX_t + g(t, X_t)^2 dt \\ &= 2 \cdot X_t^{2-1} + \frac{1}{2} \cdot 2 \cdot (2-1)g^2(t, X_t)dt \Rightarrow 2 \in \mathcal{X}. \end{aligned} \quad (4.3.78)$$

Suponha agora que $k-1 \in \mathcal{X}$, $k \geq 3$. Assim:

$$\begin{aligned} d(X_t^{k-1}) &= (k-1)X_t^{k-2}dX_t + \frac{1}{2}(k-1)(k-2)X_t^{k-3}g^2(t, X_t)dt \\ &= (k-1)X_t^{k-2}(f(t, X_t)dt + g(t, X_t)dW_t) + \frac{1}{2}(k-1)(k-2)X_t^{k-3}g^2(t, X_t)dt \\ &= \left[f(t, X_t)(k-1)X_t^{k-2} + \frac{1}{2}(k-1)(k-2)X_t^{k-3}g^2(t, X_t) \right] dt \\ &\quad + (k-1)X_t^{k-2}g(t, X_t)dW_t. \end{aligned} \quad (4.3.79)$$

Aplicando a regra do produto e a hipótese de indução:

$$\begin{aligned} d(X_t^k) &= d(X_t X_t^{k-1}) \\ &= X_t d(X_t^{k-1}) + X_t^{k-1}dX_t + (k-1)X_t^{k-2}g^2(t, X_t)dt. \end{aligned} \quad (4.3.80)$$

Podemos agora eliminar o termo $d(X_t^{m-1})$ da equação acima de modo a obter:

$$\begin{aligned} d(X_t^k) &= X_t \left[(k-1)X_t^{k-2}dX_t + \frac{1}{2}(k-1)(k-2)X_t^{k-3}g^2(t, X_t)dt \right] \\ &\quad + (k-1)X_t^{k-2}g^2(t, X_t)dt + X_t^{k-1}dX_t \\ &= dX_t \left[(k-1)X_t^{k-1} + X_t^{k-1} \right] + dt \left[\frac{1}{2}(k-1)(k-2) + (k-1) \right] X_t^{k-2}g^2(t, X_t) \\ &= kX_t^{k-1}dX_t + \frac{1}{2}(k-1)kX_t^{k-2}g^2(t, X_t)dt \Rightarrow k \in \mathcal{X}. \end{aligned} \quad (4.3.81)$$

Concluimos assim que $\mathcal{X} = \mathbb{N}$, isto é, a relação (4.3.63) é válida para potências $u(x) = x^n$, $n \in \mathbb{N}$. ■

Corolário 4.1. *Seja $X = \{X_t; 0 \leq t \leq T\}$, $T > 0$, um processo de difusão o qual cumpre as condições do Teorema 4.4. Se Q é um polinômio qualquer, então o diferencial estocástico $d(Q(X_t))$ é dado por*

$$dQ(X_t) = Q'(X_t)dX_t + \frac{1}{2}Q''(X_t)g^2(t, X_t)dt. \quad (4.3.82)$$

Demonstração: Como o diferencial estocástico é uma aplicação linear, basta aplicar a proposição anterior a cada uma das parcelas de Q e somar as expressões obtidas. ■

Teorema 4.5. (*Fórmula de Itô*) *Seja $X = \{X_t; 0 \leq t < T\}$, $T > 0$, um processo de difusão satisfazendo a EDE:*

$$dX_t = f(t, X_t)dt + g(t, X_t)dW_t. \quad (4.3.83)$$

Considere uma função $u \in C^{1,2}([0, T] \times \mathbb{R})$ e defina o processo estocástico

$$Y := \{u(t, X_t); 0 \leq t < T\}. \quad (4.3.84)$$

Neste caso, Y é um processo de difusão satisfazendo

$$\begin{aligned} dY &= \frac{\partial u}{\partial t}(t, X_t)dt + \frac{\partial u}{\partial x}(t, X_t)dX_t + \frac{1}{2} \frac{\partial^2 u}{\partial X^2}(t, X_t)(dX_t)^2 \\ &= \left(\frac{\partial u}{\partial t}(t, X_t) + \frac{\partial u}{\partial x}(t, X_t)f(t, X_t) + \frac{1}{2} \frac{\partial^2 u}{\partial x^2}(t, X_t)g^2(t, X_t) \right) dt \\ &\quad + \frac{\partial u}{\partial x}(t, X_t)g(t, X_t)dW_t. \end{aligned} \quad (4.3.85)$$

Demonstração: Suponha inicialmente que $u : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ é uma função da forma

$$u(t, x) = v(x)w(t), \quad (4.3.86)$$

onde v e w são funções polinomiais. Por simplicidade, considere $v_t := v(X_t)$ e $w_t := w(t)$. Aplicando a regra do produto:

$$\begin{aligned} du_t &= d(v_t w_t) \\ &= v_t dw_t + w_t dv_t + 0 \cdot g(t, X_t)dt = v_t dw_t + w_t dv_t. \end{aligned} \quad (4.3.87)$$

Pelo corolário (4.1), podemos reescrever esta equação do seguinte modo:

$$\begin{aligned} du_t &= v_t w'_t dt + w_t \left[v'_t + \frac{1}{2} v''_t g^2(t, X_t) dt \right] \\ &= \frac{\partial u}{\partial t}(t, X_t)dt + \frac{\partial u}{\partial x}(t, X_t)dX_t + \frac{1}{2} \frac{\partial^2 u}{\partial X^2}(t, X_t)g^2(t, X_t)dt \\ &= \left[\frac{\partial u}{\partial t}(t, X_t) + \frac{\partial u}{\partial x}(t, X_t)f(t, X_t) + \frac{1}{2} \frac{\partial^2 u}{\partial x^2}(t, X_t)g^2(t, X_t) \right] dt \\ &\quad + \frac{\partial u}{\partial x}(t, X_t)g(t, X_t)dW_t. \end{aligned} \quad (4.3.88)$$

Logo, a fórmula de Itô vale para funções do tipo $u(x, t) = v(x)w(t)$, onde v e w são polinômios. Como o diferencial estocástico é uma aplicação linear, podemos estender esta regra para funções do tipo

$$u(x, t) = \sum_{i=1}^m v^i(x)w^i(t), \quad (4.3.89)$$

onde v^i e w^i são polinômios com respeito a x e t , respectivamente. Desta forma, temos que a regra é válida para funções separáveis com comportamento polinomial nas variáveis x e t . Por fim, dada uma função qualquer $u \in C^{1,2}([0, T] \times \mathbb{R})$, podemos aproximá-la por uma sequência de polinômios $\{u_n\}_{n \in \mathbb{N}}$ tal que

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

em conjuntos compactos, à medida que $n \rightarrow +\infty$. A existência desses polinômios é garantida pelo teorema da aproximação de Stone-Weierstrass. Pelo passo anterior, temos que:

$$\begin{aligned} u_n(T, X_T) &= u_n(0, X_0) \\ &+ \int_0^T \left[\frac{\partial u_n}{\partial t}(t, X_t) + \frac{\partial u_n}{\partial x}(t, X_t) f(t, X_t) + \frac{1}{2} \frac{\partial^2 u_n}{\partial x^2}(t, X_t) g^2(t, X_t) \right] dt \\ &+ \int_0^T \frac{\partial u_n}{\partial x}(t, X_t) g(t, X_t) dW_t. \end{aligned} \quad (4.3.90)$$

Passando ao limite, concluímos o resultado. ■

4.4 Transformação de Lamperti

A *transformação de Lamperti* é uma aplicação particular da fórmula de Itô que nos permite estabilizar o coeficiente de difusão de algumas EDEs. Em outras palavras, se X satisfaz a equação

$$dX_t = f(t, X_t)dt + g(t, X_t)dW_t, \quad (4.4.91)$$

e Y corresponde à sua transformada de Lamperti, então

$$dY_t = \hat{f}(t, Y_t)dt + dW_t. \quad (4.4.92)$$

Notemos que todos os resultados (simulação e inferência) obtidos a partir desta transformação serão dados em função do processo Y , logo para que a metodologia seja praticável é indispensável que a transformação seja inversível.

Teorema 4.6. (*Transformação de Lamperti*) *Seja X um processo satisfazendo a EDE*

$$dX_t = f(t, X_t)dt + g(t, X_t)dW_t, \quad (4.4.93)$$

e assuma que a transformação

$$Y_t := S(t, X_t) = \int_{\xi}^x \frac{1}{g(t, u)} du \Big|_{x=X_t}, \quad \forall t \in [0, T], \quad (4.4.94)$$

onde ξ é um ponto qualquer do espaço de estados de X , é inversível e que $\frac{1}{g(t, x)}$ satisfaz as condições usuais para diferenciação sob o sinal da integral. Neste caso, Y satisfaz a seguinte EDE:

$$\begin{aligned} dY_t &= \left(\frac{f(t, X_t)}{g(t, X_t)} - \frac{1}{2} \frac{\partial g}{\partial x}(t, X_t) - \int_{\xi}^{X_t} \frac{1}{g^2(t, u)} \frac{\partial g}{\partial t}(t, u) du \right) dt + dW_t \\ &= \hat{f}(t, X_t)dt + dW_t = \hat{f}(t, S^{-1}(Y_t)) dt + dW_t. \end{aligned} \quad (4.4.95)$$

Demonstração: Basta observar que, ao aplicar a fórmula de Itô, obtemos um diferencial da forma

$$dY_t = \hat{f}(t, X_t)dt + \frac{\partial S}{\partial x}(t, X_t)g(t, X_t)dW_t. \quad (4.4.96)$$

É imediato que a escolha

$$S(t, X_t) = \int_{\xi}^x \frac{1}{g(t, u)} du \Big|_{x=X_t} \quad (4.4.97)$$

soluciona o problema. Para o cálculo de \hat{f} , basta utilizar as relações:

$$\frac{\partial S}{\partial t}(t, x) = \int_{\xi}^x \frac{\partial}{\partial t} \left(\frac{1}{g(t, u)} \right) du; \quad \frac{\partial S}{\partial x}(t, x) = \frac{1}{g(t, x)}; \quad (4.4.98)$$

$$\frac{\partial^2 S}{\partial x^2}(t, x) = -\frac{1}{g^2(t, x)} \frac{\partial g}{\partial x}(t, x). \quad (4.4.99)$$

■

As condições para a validade das hipóteses do teorema acima não são únicas. Um conjunto bem amplo de condições é dado por Luschgy [22]:

1. $g \in C^1([0, T] \times \mathbb{R})$;
2. existe um $K > 0$ tal que $\forall(t, x) \in [0, T] \times \mathbb{R}$, valem

$$|f(t, x)| \leq K(1 + |x|) \quad \text{e} \quad 0 \leq g(t, x) \leq K(1 + |x|). \quad (4.4.100)$$

A condição $g(t, x) \geq 0$ garante que a transformação S é estritamente crescente, portanto, inversível. A exigência de continuidade e as restrições de crescimento nos permitem derivar sob o sinal da integral.

Esta ferramenta é útil tanto para fins teóricos quanto experimentais. Lembremos que a maior dificuldade no trato de uma EDE está relacionada à integração da componente estocástica. A transformação de Lamperti é capaz de contornar este problema já que a perspectiva de Itô aplicada a integrandos elementares pode ser tratada trivialmente. Quanto aos métodos numéricos, mostraremos mais adiante que a transformação de Lamperti não só melhora a taxa de convergência de determinados algoritmos, como também estende determinados métodos a classes mais abrangentes de processos de Itô.

Exemplo 4.5. *Seja X um processo de difusão satisfazendo*

$$dX_t = \lambda X_t dt + \mu X_t dW_t, \quad 0 \leq t \leq T, \quad (4.4.101)$$

$$P(X_0 = 1) = 1, \quad (4.4.102)$$

onde $\lambda, \mu \in \mathbb{R}$. *Este processo é de grande relevância no contexto do cálculo estocástico e é usualmente denotado por movimento Browniano Geométrico. Neste exemplo, a transformação de Lamperti pode ser utilizada a fim de determinar analiticamente a solução da equação. Notemos que*

$$Y_t = S(t, X_t) = \int_1^x \frac{dx}{\mu x} \Big|_{x=X_t} = \frac{\log X_t}{\mu}, \quad (4.4.103)$$

e

$$X_t = S^{-1}(t, Y_t) = e^{\mu Y_t}. \quad (4.4.104)$$

Assim, Y_t será a solução da EDE

$$\begin{aligned} dY_t &= \left(\frac{\lambda X_t}{\mu X_t} - \frac{1}{2}\mu \right) dt + dW_t \\ &= \left(\frac{\lambda}{\mu} - \frac{1}{2}\mu \right) dt + dW_t. \end{aligned} \quad (4.4.105)$$

Para qualquer $t \in [0, T]$ podemos obter uma expressão fechada para Y avaliada neste instante:

$$\begin{aligned} Y_t &= \int_0^t \left(\frac{\lambda}{\mu} - \frac{1}{2}\mu \right) ds + \int_0^t dW_s \\ &= \left(\frac{\lambda}{\mu} - \frac{1}{2}\mu \right) t + W_t, \end{aligned} \quad (4.4.106)$$

ou seja,

$$Y_t \sim N \left[\left(\frac{\lambda}{\mu} - \frac{1}{2}\mu \right) t, t \right]. \quad (4.4.107)$$

Podemos agora obter o processo original:

$$X_t = e^{\mu Y_t} \therefore X_t \sim LN \left[\left(\frac{\lambda}{\mu} - \frac{1}{2}\mu \right) t, \mu^2 t \right], \quad (4.4.108)$$

onde LN denota a distribuição log-normal.

4.5 Transformação de Cameron-Martin-Girsanov

As transformações apresentadas até aqui são aplicadas em um processo estocástico X e devolvem um segundo processo \tilde{X} cujo diferencial estocástico apresenta um comportamento distinto do primeiro. Uma outra forma de se pensar em transformações seria com respeito às medidas induzidas pelos processos de Itô e não necessariamente nos processos em si.

Como motivação para este estudo, suponhamos que

$$Z_1, Z_2, Z_3, \dots, Z_n$$

sejam variáveis aleatórias independentes e normalmente distribuídas com $\mathbb{E}Z_i = 0$ e $\mathbb{E}Z_i^2 = 1$, $i = 1, 2, \dots, n$, em um espaço de probabilidade (Ω, \mathcal{F}, P) . O estudo de transformações invariantes relacionadas a variáveis com estas características é tradicional na teoria da Probabilidade. Por exemplo, dado um vetor $\underline{\mu} = (\mu_1, \mu_2, \dots, \mu_n) \in \mathbb{R}^n$, consideremos a seguinte função

$$f(Z_1(\omega), Z_2(\omega), \dots, Z_n(\omega)) = \exp \left\{ \sum_{i=1}^n \mu_i Z_i(\omega) - \frac{1}{2} \sum_{i=1}^n \mu_i^2 \right\}. \quad (4.5.109)$$

Como f é uma função \mathcal{F} -mensurável não-negativa, segue que a função Q definida por

$$Q(d\omega) := fP(d\omega)$$

é uma medida. Mais precisamente, como estamos lidando com uma sequência de variáveis Gaussianas padronizadas e independentes, temos

$$\begin{aligned}
 Q(d\omega) &= Q(Z_1 \in dz_1, Z_2 \in dz_2, \dots, Z_n \in dz_n) \\
 &= \exp \left\{ \sum_{i=1}^n \mu_i z_i(\omega) - \frac{1}{2} \sum_{i=1}^n \mu_i^2 \right\} P(Z_1 \in dz_1, Z_2 \in dz_2, \dots, Z_n \in dz_n) \\
 &= \exp \left\{ \sum_{i=1}^n \mu_i z_i(\omega) - \frac{1}{2} \sum_{i=1}^n \mu_i^2 \right\} \cdot \frac{1}{(2\pi)^{n/2}} \exp \left\{ -\frac{1}{2} \sum_{i=1}^n z_i^2 \right\} dz_1 dz_2 \dots dz_n \\
 &= (2\pi)^{-n/2} \exp \left\{ -\frac{1}{2} \sum_{i=1}^n (z_i - \mu_i)^2 \right\} dz_1 dz_2 \dots dz_n, \tag{4.5.110}
 \end{aligned}$$

o que mostra que Q é uma medida de probabilidade sob a qual as translações $Z_i - \mu_i$, $i = 1, 2, \dots, n$, são normalmente distribuídas com $\mathbb{E}_Q(Z_i - \mu_i) = 0$ e $\mathbb{E}_Q[(Z_i - \mu_i)^2] = 1$. Portanto,

$$Q \left[\left(\underline{Z} - \underline{\mu} \right)^{-1} \in A \right] = P \left[\left(\underline{Z} \right)^{-1} \in A \right], \quad \forall A \in \mathcal{F}, \tag{4.5.111}$$

onde $\underline{Z} = (Z_1, Z_2, \dots, Z_n)$ e $\underline{\mu} = (\mu_1, \mu_2, \dots, \mu_n)$.

O estudo dos processos de Wiener evidencia uma relação forte entre estes processos e as variáveis aleatórias Gaussianas. Desta forma, é natural procurarmos transformações similares que garantam a processos desta natureza a propriedade de invariância.

Teorema 4.7. (Cameron-Martin-Girsanov) *Seja $u \in \mathbb{L}^2$, e assumamos que*

$$M_t := \exp \left\{ \int_0^t u(s, \omega) dW_s - \frac{1}{2} \int_0^t u^2(s, \omega) dW_s \right\} \tag{4.5.112}$$

é uma martingale. Defina o processo \hat{W} por

$$\hat{W}_t := W_t - \int_0^t u(s, \omega) ds, \quad 0 \leq t < +\infty. \tag{4.5.113}$$

Para cada $T \in [0, +\infty)$ fixado, o processo $\{\hat{W}_t, 0 \leq t \leq T\}$ é um processo de Wiener em $(\Omega, \mathcal{F}_T, Q_T)$, onde

$$Q_T(A) = \mathbb{E}[\mathbf{1}_A M_T], \quad \forall A \in \mathcal{F}_T. \tag{4.5.114}$$

A demonstração deste teorema utiliza a caracterização de Lévy do movimento Browniano, cuja descrição exige resultados matemáticos não abordados neste texto (descrições formais podem ser obtidas em Friedman [12] e Karatzas e Shreve [17]). Mesmo omitindo o formalismo matemático, é possível analisar o teorema a fim de entender o papel das hipóteses e a importância do resultado. Notemos inicialmente que para qualquer evento $A \in \mathcal{F}_T$, a função $\mathbf{1}_A(\omega)$ é \mathcal{F}_T -mensurável. Assim, se $\{M_t\}_{t=0}^T$ é um martingale, então, para qualquer $t \in [0, T]$, vale

$$\begin{aligned}
 \int_{\Omega} \mathbf{1}_A(\omega) M_T(\omega) P(d\omega) &= \mathbb{E}[\mathbf{1}_A M_T] \\
 &= \mathbb{E}[\mathbb{E}[\mathbf{1}_A M_T | \mathcal{F}_t]] = \mathbb{E}[\mathbf{1}_A \mathbb{E}[M_T | \mathcal{F}_t]] \\
 &= \mathbb{E}[\mathbf{1}_A M_t] = \int_{\Omega} \mathbf{1}_A(\omega) M_t(\omega) P(d\omega), \tag{4.5.115}
 \end{aligned}$$

o que garante que

$$M_T P(d\omega) = M_t P(d\omega), \quad \forall t \in [0, T], \quad (4.5.116)$$

em \mathcal{F}_t . Em outras palavras, estamos garantindo que a família de medidas

$$\{Q_T; 0 \leq T < +\infty\}$$

satisfaz a condição de consistência

$$Q_T(A) = Q_t(A); \quad \forall A \in \mathcal{F}_t, 0 \leq t \leq T. \quad (4.5.117)$$

Além disso,

$$\mathbb{E}[M_t] = \mathbb{E}[M_0] = \mathbb{E}[1] = 1, \quad 0 \leq t \leq T, \quad (4.5.118)$$

pois todo martingale possui esperança constante. Logo,

$$Q_T(\Omega) = \mathbb{E}_{Q_T}[\mathbf{1}_\Omega] = \mathbb{E}_P[M_T] = 1, \quad (4.5.119)$$

o que nos mostra que Q_T é de fato uma medida de probabilidade para cada $T \in [0, +\infty)$.

A discussão acima deixa claro que a hipótese crucial do teorema é que M_t seja um martingale. A validade desta hipótese pode ser associada à integrabilidade de um dos componentes de M_t .

Teorema 4.8. (Condição de Novikov [27]) *Sejam W um processo de Wiener em (Ω, \mathcal{F}, P) e u um processo em \mathbb{L}^2 . Se*

$$\mathbb{E} \left[\exp \left(\frac{1}{2} \int_0^T u^2(t, \omega) dt \right) \right] < +\infty; \quad 0 \leq T < +\infty, \quad (4.5.120)$$

então

$$M_t = \exp \left\{ \int_0^t u(s, \omega) dW_s - \frac{1}{2} \int_0^t u^2(s, \omega) dW_s \right\} \quad (4.5.121)$$

é um martingale.

Dentre as inúmeras aplicações do Teorema 4.7, destaca-se a obtenção da densidade (derivada de Radon-Nykodim) das trajetórias de um processo de difusão com respeito às trajetórias de um segundo processo. Ao obter esta densidade estamos construindo um ferramental para a realização simulações a partir de métodos como rejeição ou amostragem por importância, por exemplo, e\ou inferir parâmetros desconhecidos do modelo a partir de métodos baseados em verossimilhança.

Teorema 4.9. (Teorema de Girsanov) *Sejam $u : [0, +\infty) \times \mathbb{R} \rightarrow \mathbb{R}$, M_t e \hat{W}_t conforme o teorma (4.7). Denote por P a medida induzida pelo processo de difusão correspondente à solução da EDE*

$$dX_t = f(t, X_t)dt + g(t, X_t)dW_t, \quad 0 \leq t \leq T. \quad (4.5.122)$$

Se existir uma função $\hat{f} \in \mathbb{L}^1(0, T)$ tal que

$$g(t, X_t)u(t, X_t) = \hat{f}(t, X_t) - f(t, X_t), \quad (4.5.123)$$

então, com respeito à medida Q_T dada por

$$dQ_T(\omega) = M_T dP(\omega), \quad (4.5.124)$$

X possui o diferencial estocástico

$$dX_t = \hat{f}(t, X_t)dt + g(t, X_t)d\hat{W}_t, \quad 0 \leq t \leq T. \quad (4.5.125)$$

Demonstração: Pelo teorema (4.7), Q_T é uma medida de probabilidade em \mathcal{F}_T e \hat{W}_t é um movimento Browniano com respeito a Q_T . Desta forma, a EDE (4.5.125) está bem definido. Além disso,

$$\begin{aligned}
 dX_t &= f(t, X_t)dt + g(t, X_t)dW_t \\
 &= f(t, X_t)dt + g(t, X_t) \left(d\hat{W}_t + u(t, X_t)dt \right) \\
 &= (f(t, X_t) + g(t, X_t)u(t, X_t)) dt + g(t, X_t)d\hat{W}_t \\
 &= \hat{f}(t, X_t)dt + g(t, X_t)d\hat{W}_t.
 \end{aligned} \tag{4.5.126}$$

■

Exemplo 4.6. (*Densidade com respeito à medida de Wiener*) Considere um processo estocástico X satisfazendo a equação diferencial estocástica

$$dX_t = f(t, X_t)dt + dW_t, \quad 0 \leq t \leq T, \tag{4.5.127}$$

e denote por Q_T a respectiva medida de probabilidade induzida por X . Nosso objetivo é obter a derivada de Radon-Nykodim de Q_T com respeito à medida \mathbb{W} induzida por:

$$dY_t = dW_t, \quad 0 \leq t \leq T. \tag{4.5.128}$$

\mathbb{W} é denominada medida de Wiener. Definindo

$$\begin{aligned}
 M_T(\omega) &= \exp \left(\int_0^T f(t, Y_t(\omega))dW_t - \frac{1}{2} \int_0^T f^2(t, Y_t(\omega))dt \right) \\
 &= \exp \left(\int_0^T f(t, W_t)dW_t - \frac{1}{2} \int_0^T f^2(t, W_t)dt \right),
 \end{aligned} \tag{4.5.129}$$

segue, do teorema (4.9), que

$$\frac{dQ_T}{d\mathbb{W}}(\omega) = M_T(W_t(\omega)) = \exp \left(\int_0^T f(t, W_t)dW_t - \frac{1}{2} \int_0^T f^2(t, W_t)dt \right). \tag{4.5.130}$$

Exemplo 4.7. (*Verossimilhança de uma trajetória*) Considere um processo de difusão X o qual induz uma medida de probabilidade Q_T e cujo diferencial estocástico

$$dX_t = f(X_t, \theta)dt + g(X_t)dW_t, \quad 0 \leq t \leq T, \tag{4.5.131}$$

não depende explicitamente do tempo t , porém é função de um parâmetro desconhecido θ . Nosso objetivo é construir a derivada de Radon-Nykodim de Q_T com respeito à medida de probabilidade P induzida por

$$dY_t = g(Y_t)dW_t, \quad 0 \leq t \leq T. \tag{4.5.132}$$

Definindo

$$\begin{aligned}
 M_T(\omega) &= \exp \left(\int_0^T \frac{f(Y_t(\omega), \theta)}{g(Y_t(\omega))}dW_t - \frac{1}{2} \int_0^T \frac{f^2(Y_t(\omega), \theta)}{g^2(Y_t(\omega))}dt \right) \\
 &= \left(\int_0^T \frac{f(Y_t, \theta)}{g^2(Y_t)}dY_t - \frac{1}{2} \int_0^T \frac{f^2(Y_t, \theta)}{g^2(Y_t)}dt \right),
 \end{aligned} \tag{4.5.133}$$

segue, do teorema (4.9), que

$$\frac{dQ_T}{dP}(\omega) = M_T(Y_t(\omega)) = \left(\int_0^T \frac{f(Y_t, \theta)}{g^2(Y_t)} dY_t - \frac{1}{2} \int_0^T \frac{f^2(Y_t, \theta)}{g^2(Y_t)} dt \right). \quad (4.5.134)$$

Este procedimento é bastante utilizado em se tratando de inferência em processos de difusão [5].

4.6 Exercícios

1. Mostre que $X_t = \frac{W_t}{1+t}$, onde W é um processo de Wiener com $W_0 = 0$, é solução da EDE

$$dX_t = -\frac{1}{1+t}X_t dt + \frac{1}{1+t}dW_t; \quad X_0 = 0.$$

2. Sejam $a, b \in \mathbb{R}$ constantes quaisquer e considere a seguinte EDE:

$$dX_t = \frac{b - X_t}{1-t} dt + dW_t; \quad 0 \leq t < 1, X_0 = a.$$

Mostre que

$$X_t = a(1-t) + bt + (1-t) \int_0^t \frac{dW_s}{1-s}; \quad 0 \leq s < 1$$

satisfaz a equação e que $\lim_{t \rightarrow 1} X_t = b$ (q.c.). O processo X está relacionado a um importante processo denominado ponte Browniana, o qual será crucial para a metodologia desenvolvida no Capítulo 7.

3. Considere a seguinte EDE:

$$dX_t = \mu X_t dt + \sigma \sqrt{X_t} dW_t, \quad X_0 = x,$$

onde $\mu, \sigma > 0$. Utilize a transformação de Lamperti para determinar uma nova EDE com coeficiente de difusão constante.

4. Considere a EDE $X_t = W_t + \text{sent}$, onde W é um processo de Wiener com respeito à medida \mathbb{P} . Dado que X_t é um processo de Wiener com respeito a uma segunda medida \mathbb{Q} , utilize o Teorema de Girsanov para estabelecer $\frac{d\mathbb{Q}}{d\mathbb{P}}$.

Capítulo 5

Métodos de Discretização

O primeiro passo para a construção de um algoritmo numérico para equações diferenciais é a discretização do domínio de interesse, em nosso caso, um intervalo da forma $[0, T]$. Esta discretização é feita através da escolha de um conjunto de N pontos $\tau = \{0 = t_0 < t_1 < t_2 < \dots < t_{N-2} < t_{N-1} = T\}$, o qual definimos como uma *partição* do intervalo. É comum denotarmos por δ a maior distância entre dois pontos consecutivos da partição τ , isto é,

$$\delta = \max_n \{t_{n+1} - t_n\}, \quad n = 0, 1, 2, \dots, N - 2.$$

Esta notação é importante pois a escolha dos pontos da partição depende do algoritmo em questão e diversos teoremas de convergência baseiam-se no maior espaçamento da partição. Neste trabalho consideraremos apenas partições uniformes, ou seja, partições onde os N pontos estão igualmente espaçados a uma distância de

$$\Delta = \frac{T}{N - 1}.$$

Desta forma, a representação t_n denotará o instante de tempo $t = n\Delta$.

Antes de entrarmos nos métodos de discretização, apresentaremos uma versão estocástica da expansão de Taylor. Este resultado é importante para o entendimento de alguns resultados que serão discutidos. Ao final do capítulo citaremos alguns métodos numéricos populares, porém não entraremos em maiores detalhes.

5.1 Expansão de Itô-Taylor

a teoria dos polinômios de Taylor oferece um ferramental muito importante para o cálculo numérico pois nos permite aproximar localmente uma função

$$f : I \rightarrow \mathbb{R},$$

n vezes diferenciável, através de uma combinação linear de suas derivadas. Mais precisamente, dado qualquer ponto x_0 no interior de I , é possível obter um $h \in \mathbb{R}$ tal que $x_0 + h \in I$ e

$$f(x_0 + h) = f(x_0) + \sum_{i=1}^n \frac{1}{i!} f^{(i)}(x_0) \cdot h^i + o(h^n). \quad (5.1.1)$$

Desta forma, se h é suficientemente pequeno, então o termo $o(h^n)$ é negligenciável, ou seja

$$f(x_0 + h) \approx f(x_0) + \sum_{i=1}^n \frac{1}{i!} f^{(i)}(x_0) \cdot h^i. \quad (5.1.2)$$

Como as derivadas acima podem ser aproximadas pelo método das diferenças finitas, segue que a relação acima de fato nos dá uma ferramenta para obter aproximações locais de f .

No contexto estocástico, o comportamento local de uma função aleatória pode ser estudado através da aplicação da fórmula de Itô. Este dispositivo nos permite estender o resultado acima de modo a obter aproximações locais para um processo estocástico X_t solução da equação diferencial estocástica

$$dX_t = f(t, X_t)dt + g(t, X_t)dW_t, \quad X_0 = x_0, t \in [0, T]. \quad (5.1.3)$$

Para simplificar a notação, definimos inicialmente os seguintes operadores para a fórmula de Itô:

$$\begin{aligned} \mathcal{M} &= g \cdot \frac{\partial}{\partial x}; \\ \mathcal{L} &= \frac{\partial}{\partial t} + f \cdot \frac{\partial}{\partial x} + \frac{1}{2}g^2 \cdot \frac{\partial^2}{\partial x^2}; \end{aligned} \quad (5.1.4)$$

onde f e g são, respectivamente, os coeficientes de tendência e difusão da equação diferencial estocástica 5.1.3. Neste caso, dado um processo estocástico $Y_t = z(t, X_t)$, $z \in C^{1,2}([0, T] \times \mathbb{R})$, temos que

$$\begin{aligned} dY_t &= \left(\frac{\partial z}{\partial t}(t, X_t) + f(t, X_t) \frac{\partial z}{\partial x}(t, X_t) + \frac{1}{2}g^2(t, X_t) \frac{\partial^2 z}{\partial x^2}(t, X_t) \right) dt \\ &\quad + g(t, X_t) \frac{\partial z}{\partial x}(t, X_t) dW_t \\ &= (\mathcal{L} \circ z)(t, X_t)dt + (\mathcal{M} \circ z) dW_t. \end{aligned} \quad (5.1.5)$$

Se o objetivo é avaliar o comportamento local de X_t em torno de algum ponto $s \in (0, T)$, então consideramos um incremento h suficientemente pequeno, de modo que $t = s + h \in (0, T]$ e tentamos obter uma aproximação para

$$X_t = X_s + \int_s^t f(u, X_u)du + \int_s^t g(u, X_u)dW_u. \quad (5.1.6)$$

Aplicando a fórmula de Itô aos coeficientes f e g , obtemos

$$\begin{aligned} X_t &= X_s + \int_s^t \left[f(s, X_s) + \int_s^u (\mathcal{L} \circ f)(s_1, X_{s_1})ds_1 + \int_s^u (\mathcal{M} \circ f)(s_1, X_{s_1})dW_{s_1} \right] du \\ &\quad + \int_s^t \left[g(s, X_s) + \int_s^u (\mathcal{L} \circ g)(s_1, X_{s_1})ds_1 + \int_s^u (\mathcal{M} \circ g)(s_1, X_{s_1})dW_{s_1} \right] dW_u \\ &= X_s + f(s, X_s) \cdot h + g(s, X_s)(W_t - W_s) \\ &\quad + \int_s^t \left[\int_s^u (\mathcal{L} \circ f)(s_1, X_{s_1})ds_1 + \int_s^u (\mathcal{M} \circ f)(s_1, X_{s_1})dW_{s_1} \right] du \\ &\quad + \int_s^t \left[\int_s^u (\mathcal{L} \circ g)(s_1, X_{s_1})ds_1 + \int_s^u (\mathcal{M} \circ g)(s_1, X_{s_1})dW_{s_1} \right] dW_u. \end{aligned} \quad (5.1.7)$$

Como $(W_t - W_s) \sim N(0, h)$, podemos expressar a relação acima em termos de uma variável aleatória $Z \sim N(0, 1)$:

$$X_t = X_s + f(s, X_s) \cdot h + g(s, X_s) \cdot (\sqrt{h}Z) + R_1(t, X_t), \quad (5.1.8)$$

obtendo assim a primeira forma da expansão de Itô-Taylor. Era de se esperar, de maneira análoga à expansão de Taylor determinística, que $R(t, X_t)$ não dependesse explicitamente de termos lineares com respeito ao passo de discretização h , no entanto, a integral

$$\int_s^t \int_s^u (\mathcal{M} \circ g)(s_1, X_{s_1}) dW_{s_1} dW_u \quad (5.1.9)$$

garante a presença de um termo desta ordem no resto da expansão. Este fato pode ser explicado (de uma maneira simplista) através da variação quadrática do movimento Browniano. Basta observar que a integral acima dará origem a um incremento Browniano quadrático, o qual pode ser aproximado por um incremento linear com respeito ao tempo.

Caso f e g possuam as derivadas necessárias, podemos aplicar a fórmula de Itô aos demais coeficientes do resto $R_1(t, X_t)$ da equação (5.1.8) de modo a torná-lo mais negligenciável. Em particular, podemos extrair o termo de ordem h da integral dupla acima. Para tal, obtemos o diferencial de $(\mathcal{M} \circ g)$:

$$d[(\mathcal{M} \circ g)(t, X_t)] = (\mathcal{L} \circ \mathcal{M} \circ g)(t, X_t) dt + (\mathcal{M}^2 \circ g)(t, X_t) dW_t. \quad (5.1.10)$$

Assim,

$$\begin{aligned} (\mathcal{M} \circ g)(t, X_t) &= (\mathcal{M} \circ g)(s, X_s) + \int_s^t (\mathcal{L} \circ \mathcal{M} \circ g)(u, X_u) du \\ &\quad + \int_s^t (\mathcal{M}^2 \circ g)(u, X_u) dW_u \\ &= g(s, X_s) \frac{\partial g}{\partial x}(s, X_s) + \hat{R}_1(t, X_t). \end{aligned} \quad (5.1.11)$$

Substituindo o resultado acima na integral (5.1.9):

$$\begin{aligned} \int_s^t \int_s^u (\mathcal{M} \circ g)(s_1, X_{s_1}) dW_{s_1} dW_u &= \int_s^t \int_s^u g(s, X_s) \frac{\partial g}{\partial x}(s, X_s) dW_{s_1} dW_u \\ &\quad + \int_s^t \int_s^u \hat{R}_1(s_1, X_{s_1}) dW_{s_1} dW_u \\ &= g(s, X_s) \frac{\partial g}{\partial x}(s, X_s) \int_s^t \int_s^u dW_{s_1} dW_u \\ &\quad + R_2(t, X_t). \end{aligned} \quad (5.1.12)$$

A integral estocástica dupla que aparece na relação acima pode ser tratada analiticamente através das fórmulas obtidas na seção 4.3:

$$\begin{aligned} \int_s^t \int_s^u dW_{s_1} dW_u &= \int_s^t (W_u - W_s) dW_u \\ &= \int_s^t W_u dW_u - W_s \int_s^t dW_s = \left[\frac{W_u^2 - u}{2} \right]_s^t - W_s(W_t - W_s) \\ &= \frac{1}{2} [W_t^2 - W_s^2 - (t - s)] - W_s(W_t - W_s) \\ &= \frac{1}{2} [(W_t - W_s)^2 - h]. \end{aligned} \quad (5.1.13)$$

Finalmente, podemos obter a versão estocástica para a aproximação de Taylor de primeira ordem:

$$\begin{aligned} X_t &= X_s + f(s, X_s) \cdot h + g(s, X_s) \cdot (\sqrt{h}Z) \\ &\quad + \frac{1}{2}g(s, X_s) \frac{\partial g}{\partial x}(s, X_s) (Z^2 - h) + R_2(t, X_t). \end{aligned} \quad (5.1.14)$$

Assim como no caso determinístico, a série de Taylor estocástica pode continuar a ser expandida a partir do resto R_2 . Esta expansão não é trivial devido às integrais estocásticas múltiplas que surgem em decorrência da aplicação iterada da fórmula de Itô. Uma discussão mais profunda desta expansão pode ser encontrada em Kloeden e Platen [19].

5.2 Métodos de Euler-Maruyama e Milstein

A aproximação discreta mais simples para um processo de difusão é caracterizada pelo *método de Euler*, também denominado *método de Euler-Maruyama*. Consideremos, por exemplo, o processo X solução da equação diferencial estocástica

$$dX_t = f(t, X_t)dt + g(t, X_t)dW_t, \quad t \in [0, T], \quad (5.2.15)$$

$$P(X_0 = \xi_0) = 1. \quad (5.2.16)$$

Dada uma partição $\tau = \{0 = t_0 < t_1 < t_2 < \dots < t_{N-1} < t_N = T\}$, é razoável supor que se os espaçamentos $(t_{n+1} - t_n)$, $i = 0, 1, 2, \dots, N-1$ forem suficientemente pequenos, então os processos $f(t, X_t)$ e $g(t, X_t)$ sobre a partição τ são aproximadamente elementares, isto é

$$f(t, \omega) \approx \sum_{i=0}^{N-1} f(t_i, \omega) \mathbf{1}_{[t_i, t_{i+1})}(t); \quad g(t, \omega) \approx \sum_{i=0}^{N-1} g(t_i, \omega) \mathbf{1}_{[t_i, t_{i+1})}(t). \quad (5.2.17)$$

Esta intuição nos leva ao processo discretizado $X^N = \{X_{t_n}^N; n = 0, 1, 2, \dots, N\}$, o qual satisfaz a equação discretizada

$$X_{t_{n+1}}^N = X_{t_n}^N + f(t_n, X_{t_n}^N) (t_{n+1} - t_n) + g(t_n, X_{t_n}^N) (W_{t_{n+1}} - W_{t_n}), \quad t_n \in \tau, \quad (5.2.18)$$

$$P(X_0^N = X_0) = 1. \quad (5.2.19)$$

Para simplificar a notação, consideramos $\Delta W_{t_n} := (W_{t_{n+1}} - W_{t_n})$. Assim, chegamos ao seguinte processo discretizado

$$X_{t_{n+1}}^N = X_{t_n}^N + f(t_n, X_{t_n}^N) \Delta + g(t_n, X_{t_n}^N) \Delta W_{t_n}, \quad t_n \in \tau, \quad (5.2.20)$$

$$P(X_0^N = X_0) = 1. \quad (5.2.21)$$

Este procedimento de discretização nos permite obter densidades de transição:

$$X_{t_{n+1}}^N \mid \{X_{t_n}^N = x\} \sim N(x + f(t_n, x)\Delta, g^2(t_n, x)\Delta), \quad (5.2.22)$$

as quais podem ser utilizadas para a simulação de trajetórias do processo real X ou para o estudo de inferência estatística (caso X dependa de um parâmetro desconhecido).

Exemplo 5.1. Considere a equação diferencial estocástica linear

$$dX_t = \lambda X_t dt + \mu X_t dW_t, \quad 0 \leq t \leq T, \quad (5.2.23)$$

$$P(X_0 = 1) = 1. \quad (5.2.24)$$

Já verificamos, no exemplo 4.5, que uma solução X deste problema cumpre a seguinte dinâmica

$$X_t = \exp \left\{ \left(\lambda - \frac{1}{2} \mu^2 \right) t + \mu W_t \right\}, \quad (5.2.25)$$

onde $\lambda, \mu \in \mathbb{R}$ e $T \in [0, +\infty)$. A equação discretizada sob a partição

$$\tau = \{0 = t_0 < t_1 < \dots < t_{N-1} < t_N = T\}$$

obtida pelo método de Euler-Maruyama será dada por

$$X_{t_{n+1}}^N = \lambda X_{t_n}^N \Delta + \mu X_{t_n}^N \Delta W_{t_n}, \quad t_n \in \tau, \quad (5.2.26)$$

$$P(X_0^N = 1) = 1, \quad (5.2.27)$$

ou ainda,

$$X_{t_{n+1}}^N = X_{t_n}^N [\lambda \Delta + \mu \Delta W_{t_n}], \quad t_n \in \tau, \quad (5.2.28)$$

$$P(X_0^N = 1) = 1. \quad (5.2.29)$$

Caminhos amostrais do processo discretizado podem ser obtidos facilmente através das densidades de transição 5.2.22 conforme ilustrado na Figura 5.1 a seguir. O custo da simulação é basicamente o custo de se gerar N variáveis aleatórias normalmente distribuídas.

Analisando a expansão de Itô-Taylor, é verificado prontamente que o método de Euler-Maruyama pode ser obtido através da exclusão do termo R_1 em 5.1.8. Em analogia com o caso determinístico, é intuitivo esperar que aproximações melhores possam ser obtidas se levarmos em consideração mais coeficientes na expansão de Itô-Taylor. Em particular, o negligenciamento do termo R_2 em 5.1.14 nos leva à seguinte aproximação:

$$X_t \approx X_s + f(s, X_s) \cdot h + g(s, X_s) \cdot (\sqrt{h}Z) + \frac{1}{2} g(s, X_s) \frac{\partial g}{\partial x}(s, X_s) (Z^2 - h). \quad (5.2.30)$$

Este procedimento sugere o seguinte esquema de discretização:

$$X_{t_{n+1}}^N = X_{t_n}^N + f(t_n, X_{t_n}^N) \Delta + g(t_n, X_{t_n}^N) \Delta W_{t_n} + \frac{1}{2} g(t_n, X_{t_n}^N) g_x(t_n, X_{t_n}^N) [(\Delta W_{t_n})^2 - \Delta], \quad t_n \in \tau, \quad (5.2.31)$$

$$P(X_0^N = X_0) = 1, \quad (5.2.32)$$

o qual é denominado *método de Milstein*.

O método de Milstein concebe um modelo de discretização baseado na totalidade dos coeficientes que dependem linearmente do espaçamento Δ na expansão de Itô-Taylor, o que não ocorre com o método de Euler. Desta forma, é razoável afirmarmos que o método de Milstein gera aproximações mais precisas e, a princípio, permite a simulação de trajetórias com um espaçamento maior entre os instantes de tempo. Este quesito é importante sob a perspectiva da economia de recursos

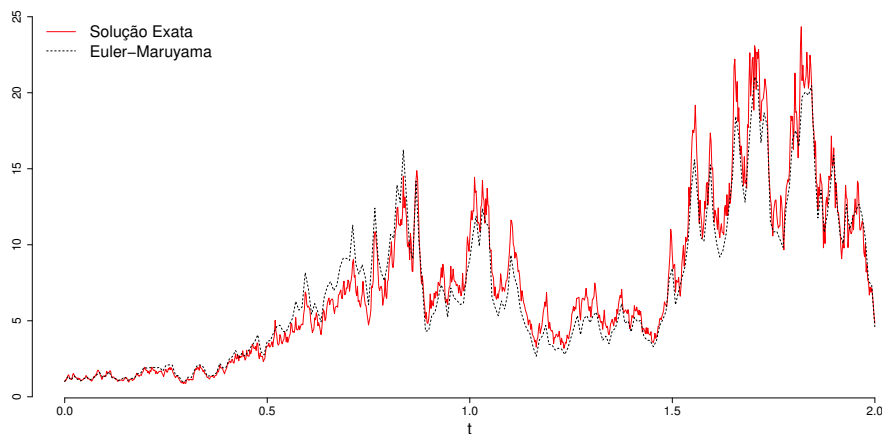
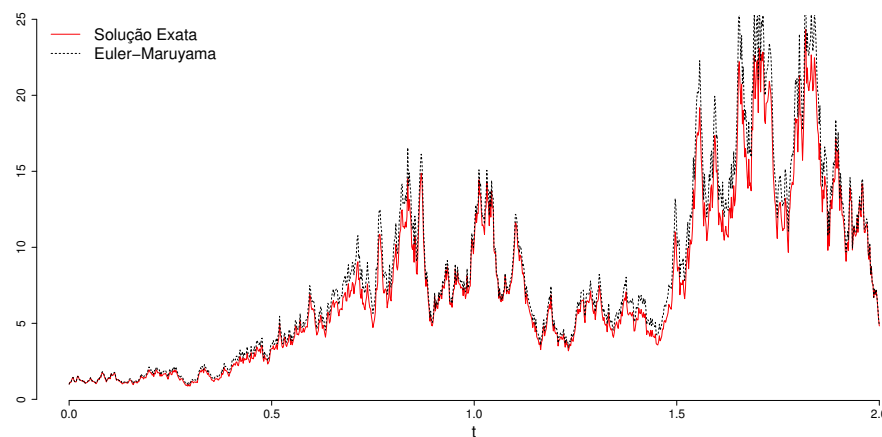
(a) Malha com $N = 257$ pontos(b) Malha com $N = 1025$ pontos

Figura 5.1: Exemplo de aproximação pelo método de Euler-Maruyama para o movimento Browniano Geométrico com $\lambda = 4$ e $\mu = 2$.

computacionais: maiores espaçamentos acarretam na utilização de menos pontos. Voltaremos a questão quando abordarmos os conceitos de convergência e estabilidade. Por hora, as simulações apresentadas na Figura 5.2 fornecem uma evidência de que o método de Milstein oferece uma aproximação mais precisa; perceba que as aproximações são melhores que as oferecidas pelo método de Euler-Maruyama considerando as mesmas malhas de discretização.

É possível obter casos em que o método de Euler coincide com o método de Milstein. Considere o diferencial estocástico padrão:

$$dX_t = f(t, X_t)dt + g(t, X_t)dW_t.$$

Notemos que se o coeficiente $g(\cdot, x)$ for constante, então $g_x(t, x) \equiv 0$ e os métodos de

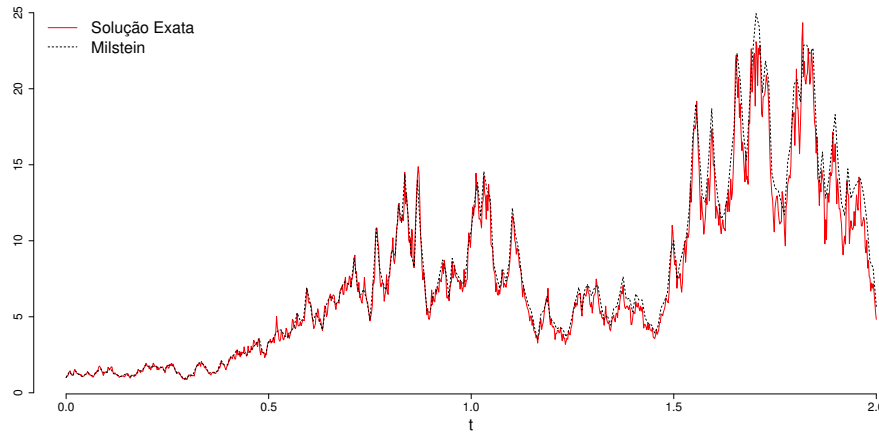
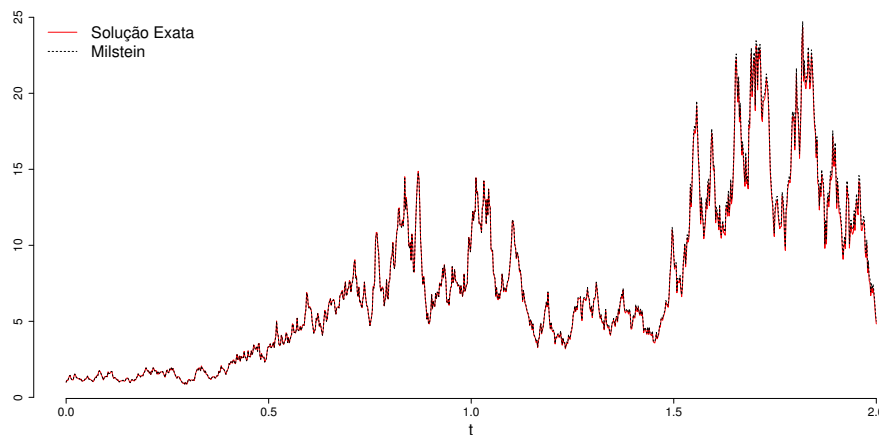
(a) Malha com $N = 257$ pontos(b) Malha com $N = 1025$ pontos

Figura 5.2: Exemplo de aproximação pelo método de Milstein para o movimento Browniano Geométrico com $\lambda = 4$ e $\mu = 2$.

Euler e Milstein coincidem. À primeira vista esta condição pode parecer bastante restritiva, no entanto ela pode ser obtida através da transformação de Lamperti. É importante destacar que se aplicarmos o método de Euler a EDEs obtidas a partir desta transformação, não estaremos obtendo aproximações para o processo original. Após a realização dos métodos de simulação ou inferência, devemos aplicar a transformação inversa para adaptar os dados obtidos ao problema inicial.

5.3 Esquema Livre de Derivada

Dentre os dois métodos apresentados até aqui, temos que o esquema de Milstein fornece a aproximação mais precisa por levar em consideração todos os termos

lineares com respeito ao passo de discretização Δ na expansão de Itô-Taylor. Entretanto, devemos observar que o método de milstein faz uso também da derivada do coeficiente de difusão, um procedimento que pode ser custoso em determinadas aplicações. A fim de contornar este problema, Liske e Platen [21] propuseram uma abordagem apoiada em aproximações de primeira ordem da derivada. O método baseia-se fundamentalmente na estimativa:

$$g(t, x + dx) - g(t, x) = g_x(t, x)dx + O(dx^2). \quad (5.3.33)$$

O que nos leva à seguinte aproximação:

$$g(t, x + dx) - g(t, x) \approx g_x(t, x)dx, \quad (5.3.34)$$

a qual é denominada *esquema de diferenças avançadas*. Adaptando esta aproximação ao nosso problema, obtemos:

$$\begin{aligned} g(t, X_{t_n}^N + \Delta X_{t_n}^N) - g(t, X_{t_n}^N) &\approx g_x(t, X_{t_n}^N) \Delta X_{t_n}^N \\ &= g_x(t, X_{t_n}^N) [f(t, X_{t_n}^N) \Delta + g(t, X_{t_n}^N) \Delta W_t] \\ &= g_x(t, X_{t_n}^N) g(t, X_{t_n}^N) \Delta W_t + O(\Delta). \end{aligned} \quad (5.3.35)$$

Os autores do método utilizam também a aproximação $\Delta W_t \approx \sqrt{\Delta}$, a qual é razoável se pensarmos na estimativa 3.1.11:

$$dW_t = \sqrt{dt}. \quad (5.3.36)$$

Desta forma, obtemos

$$g_x(t, X_{t_n}^N) g(t, X_{t_n}^N) = \frac{g(t, X_{t_n}^N + \Delta X_{t_n}^N) - g(t, X_{t_n}^N)}{\sqrt{\Delta}} + O(\sqrt{\Delta}). \quad (5.3.37)$$

Utilizando agora o método de Euler-Maruyama, podemos inferir a relação:

$$\begin{aligned} X_{t_n}^N + \Delta X_{t_n}^N &= X_{t_n}^N + X_{t_{n+1}}^N - X_{t_n}^N \\ &= X_{t_{n+1}}^N \\ &= X_{t_n}^N + f(t_n, X_{t_n}^N) \Delta + g(t_n, X_{t_n}^N) \Delta W_{t_n}. \end{aligned} \quad (5.3.38)$$

Assim,

$$\begin{aligned} g_x(t, X_{t_n}^N) g(t, X_{t_n}^N) &\approx \frac{g(t, X_{t_n}^N + f(t_n, X_{t_n}^N) \Delta + g(t_n, X_{t_n}^N) \Delta W_{t_n}) - g(t, X_{t_n}^N)}{\sqrt{\Delta}} \\ &\approx \frac{g(t, X_{t_n}^N + f(t_n, X_{t_n}^N) \Delta + g(t_n, X_{t_n}^N) \sqrt{\Delta}) - g(t, X_{t_n}^N)}{\sqrt{\Delta}}. \end{aligned} \quad (5.3.39)$$

Substituindo o resultado obtido no esquema de Milstein, chegamos ao seguinte *esquema livre de derivadas* :

$$X_{t_{n+1}}^N = X_{t_n}^N + f(t_n, X_{t_n}^N) \Delta + g(t_n, X_{t_n}^N) \Delta W_{t_n} + G_{t_n}^N, \quad (5.3.40)$$

onde

$$G_{t_n}^N := \frac{(\Delta W_{t_n})^2 - \Delta}{2\sqrt{\Delta}} \left[g(t_n, X_{t_n}^N + f(t_n, X_{t_n}^N) \Delta + g(t_n, X_{t_n}^N) \sqrt{\Delta}) - g(t_n, X_{t_n}^N) \right]. \quad (5.3.41)$$

Exemplo 5.2. Seja $X(\omega) = \{X_t(\omega); 0 \leq t \leq 5\}$ uma realização do movimento Browniano geométrico com $\lambda = 4, \mu = 2, x_0 = 1$. Neste caso, os esquemas de Milstein e livre de derivada serão dados por

$$X_{t_{n+1}}^N = X_{t_n} \left[1 + \lambda\Delta + \mu\Delta W_{t_n} + \mu^2 \frac{(\Delta W_{t_n})^2 - \Delta}{2} \right], \quad (5.3.42)$$

$$X_{t_0}^N = 1, \quad (5.3.43)$$

e,

$$X_{t_{n+1}} = X_{t_n} \left[1 + \lambda\Delta + \mu\Delta W_{t_n} + \mu(\lambda\Delta + \mu\sqrt{\Delta}) \frac{(\Delta W_{t_n})^2 - \Delta}{2\sqrt{\Delta}} \right], \quad (5.3.44)$$

$$X_{t_0}^N = 1, \quad (5.3.45)$$

respectivamente. A Figura 5.3 abaixo apresenta uma simulação com $N = 513$ pontos, envolvendo a solução exata e as aproximações fornecidas pelos dois esquemas acima. Observe que as trajetórias dos dois esquemas são praticamente indistinguíveis.

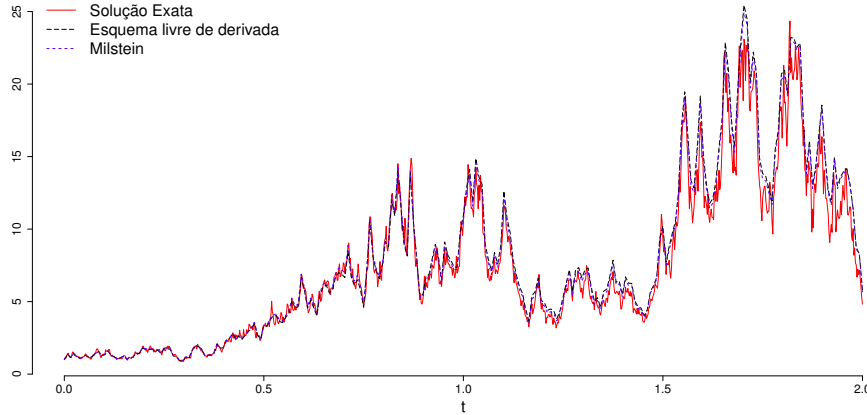


Figura 5.3: Comparação entre os métodos de Milstein e livre de derivada na simulação de trajetórias do movimento Browniano Geométrico com $\lambda = 4$ e $\mu = 2$.

A semelhança obtida nas trajetórias do exemplo anterior não é uma mera coincidência. Lembramos que o esquema livre de derivada é uma versão do método de Milstein com uma aproximação para a derivada g_x . Portanto, desde que o passo de discretização seja suficientemente pequeno, devemos esperar que os métodos tenham a mesma precisão.

Vários outros algoritmos que dispensam o uso de derivada podem ser encontrados na literatura, porém não serão descritos em detalhes aqui. Destacamos, em particular, os métodos de Runge-Kutta abordados por Newton [26], Saito e Mitsui [32] e Burrage e Burrage [9], os quais serão apresentados superficialmente na última seção.

5.4 Método θ Estocástico

Os métodos de discretização apresentados até aqui avaliam o processo discretizado X^N em um determinado instante t_{n+1} baseados apenas nas informações já conhecidas até o instante imediatamente anterior. A qualidade da aproximação pode ser melhorada se considerarmos outro tipo de informação além da disponível até o instante t_n . Poderíamos, por exemplo, tentar inserir o valor mais atual $X_{t_{n+1}}^N$ (ainda desconhecido) no processo de discretização:

$$X_{t_{n+1}}^N = X_{t_n}^N + f(t_{n+1}, X_{t_{n+1}}^N) \Delta + g(t_n, X_{t_n}^N) \Delta W_{t_n}. \quad (5.4.46)$$

Esquemas desta natureza são denominados *métodos implícitos* pois a quantidade desconhecida $X_{t_{n+1}}^N$ aparece nos dois lados da igualdade correspondente ao passo de atualização e, em geral, não pode ser isolada algebricamente. Esse problema pode ser contornado através de algoritmos numéricos padrão, tais como o método de Newton-Raphson.

Uma família mais geral de métodos implícitos pode ser construída a partir do método de Euler através da inserção da informação ainda desconhecida e de um parâmetro de controle $\theta \in [0, 1]$:

$$\begin{aligned} X_{t_{n+1}}^N &= X_{t_n}^N + \theta \cdot f(t_{n+1}, X_{t_{n+1}}^N) \Delta + (1 - \theta) \cdot f(t_n, X_{t_n}^N) \Delta \\ &\quad + g(t_n, X_{t_n}^N) \Delta W_{t_n}. \end{aligned} \quad (5.4.47)$$

Este esquema é denominado método θ -estocástico em alusão ao método θ utilizado para obter aproximações de equações diferenciais no contexto determinístico ($g \equiv 0$). Perceba que se $\theta = 0$, então voltaremos ao método de Euler-Maruyama que foi discutido na seção anterior. Utilizando agora a nomenclatura do caso determinístico, faremos referência aos casos $\theta = \frac{1}{2}$ e $\theta = 1$ como *método de Euler trapezoidal* e *método de Euler implícito*, respectivamente.

Métodos implícitos são importantes quando lidamos com problemas "instáveis". Voltaremos à esta discussão quando definirmos formalmente o conceito de estabilidade e em seguida verificaremos que esta abordagem de fato melhora os resultados obtidos pela aproximação de Euler-Maruyama. A princípio, ilustramos apenas um exemplo (ver Figura 5.4) de como o método θ pode ser utilizado para melhorar os resultados obtidos pela versão explícita. O experimento abaixo faz um comparativo entre os comportamentos dos dois métodos a longo prazo, aplicados ao movimento Browniano geométrico com parâmetros $\lambda = 2, \mu = 1, x_0 = 0, 2$, em uma malha com $N = 257$ pontos.

Esta abordagem pode ser estendida aos outros métodos já discutidos. No caso do esquema de Milstein, teríamos:

$$\begin{aligned} X_{t_{n+1}}^N &= X_{t_n}^N + \left[\theta f(t_{n+1}, X_{t_{n+1}}^N) + (1 - \theta) f(t_n, X_{t_n}^N) \right] \Delta + g(t_n, X_{t_n}^N) \Delta W_{t_n} \\ &\quad + \frac{1}{2} g_x(t_n, X_{t_n}^N) g(t_n, X_{t_n}^N) [(\Delta W_{t_n})^2 - \Delta], \end{aligned} \quad (5.4.48)$$

onde $\theta \in [0, 1]$. Já para o esquema livre de derivada:

$$\begin{aligned} X_{t_{n+1}}^N &= X_{t_n}^N + \left[\theta f(t_n, X_{t_{n+1}}^N) + (1 - \theta) f(t_n, X_{t_n}^N) \right] \Delta + g(t_n, X_{t_n}^N) \Delta W_{t_n} \\ &\quad + \left[g(t_n, X_{t_n}^N + f(t_n, X_{t_n}^N) \Delta) + g(t_n, X_{t_n}^N) \sqrt{\Delta} \right] \\ &\quad \times \frac{(\Delta W_{t_n})^2 - \Delta}{2\sqrt{\Delta}}, \end{aligned} \quad (5.4.49)$$

onde, $\theta \in [0, 1]$.

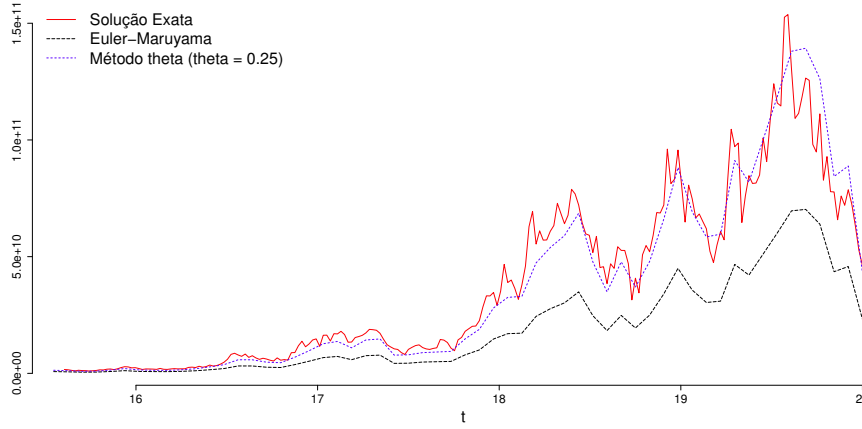


Figura 5.4: Euler-Maruyama $\times \theta$ Estocástico ($\theta = 1/4$).

5.5 Método da Linearização Local

O método da Linearização Local, introduzido por Shoji e Ozaki [35], consiste em aproximar localmente o coeficiente de deslocamento da equação diferencial estocástica por uma função linear. Basicamente, o método explora o fato de que uma aproximação linear é mais razoável que a aproximação por uma constante utilizada pelo método de Euler.

Considere inicialmente a seguinte equação diferencial estocástica

$$dX_t = f(t, X_t)dt + g(X_t)dW_t, \quad 0 \leq t \leq T, \quad (5.5.50)$$

$$P(X_0 = \xi_0) = 1, \quad (5.5.51)$$

onde os coeficientes f e g satisfazem as condições do Teorema 4.1 para garantir a existência de soluções fortes. A partir da transformação de Lamperti podemos simplificar a equação acima de modo a concentrarmos apenas em equações diferenciais estocásticas da forma

$$dX_t = f(t, X_t)dt + \sigma dW_t, \quad 0 \leq t \leq T, \quad (5.5.52)$$

$$P(X_0 = \hat{\xi}_0) = 1, \quad (5.5.53)$$

onde $f \in C^{2,1}(\mathbb{R}, [0, +\infty))$ e $\sigma \in \mathbb{R}$ é uma constante não nula.

O comportamento local da função em 5.5.52 pode ser entendido a partir da observação de sua dinâmica em um intervalo infinitesimal. Em outras palavras, estamos interessados no diferencial df , o qual pode ser obtido pela fórmula de Itô:

$$df = \left(\frac{\sigma^2}{2} f_{xx} + f_t \right) dt + f_x dx. \quad (5.5.54)$$

A linearização de f com respeito a x_t e t é obtida assumindo que f_{xx} , f_x , f_t são constantes, o que é razoável se estivermos trabalhando com um intervalo $[s, s + \Delta s)$ suficientemente pequeno:

$$\begin{aligned} df &\approx f(t, x_t) - f(s, x_s) \\ &= \left(\frac{\sigma^2}{2} f_{xx} + f_t \right) \cdot (t - s) + f_x \cdot (x_t - x_s). \end{aligned} \quad (5.5.55)$$

Definindo

$$L_s := f_x(s, X_s), \quad M_s := \frac{\sigma^2}{2} f_{xx}(s, X_s) + f_t(s, X_s), \quad (5.5.56)$$

$$N_s := f(s, X_s) - X_s f_x(s, X_s) - \left(\frac{\sigma^2}{2} f_{xx}(s, X_s) + f_t(s, X_s) \right) s, \quad (5.5.57)$$

obtemos,

$$\hat{f}(t, x_t) := L_s x_t + t M_s + N_s, \quad t \geq s. \quad (5.5.58)$$

Esta aproximação nos permite focar em resolver a equação diferencial estocástica linear

$$dX_t = (L_s X_t + t M_s + N_s) dt + \sigma dW_t = \hat{f}(t, X_t) + \sigma dW_t, \quad (5.5.59)$$

ao invés da equação (5.5.52), desde que t pertença ao intervalo $[s, s + \Delta s)$.

As probabilidades de transição $P(X_t|X_s)$ do processo X_t solução da equação (5.5.59) serão obtidas através de uma sequência de transformações que simplificarão o trabalho analítico. O primeiro passo é definir as funções

$$z(t, X_t) := -\frac{1}{\sigma} (M_s t + N_s), \quad \alpha(t, X_t) := L_s X_t, \quad (5.5.60)$$

as quais cumprem a seguinte igualdade:

$$\sigma \cdot z(t, X_t) = \hat{f}(t, X_t) - \alpha(t, X_t). \quad (5.5.61)$$

A relação acima nos permite aplicar a transformação de Girsanov de modo a obter uma equação diferencial estocástica mais simples:

$$\begin{aligned} dX_t &= \alpha(t, X_t) dt + \sigma d\widehat{W}_t \\ &= L_s X_t dt + \sigma d\widehat{W}_t, \end{aligned} \quad (5.5.62)$$

onde

$$\widehat{W}_t = W_t - \frac{1}{\sigma} \int_s^t (M_s u + N_s) du \quad (5.5.63)$$

é um movimento Browniano. O coeficiente de deslocamento da última equação pode ser eliminado a partir da transformação $Y_t = e^{-L_s t} X_t$. Aplicando a fórmula de Itô,

$$\begin{aligned} dY_t &= -L_s e^{-L_s t} X_t dt + e^{-L_s t} \left(L_s X_t dt + \sigma d\widehat{W}_t \right) + 0 \cdot (dX_t)^2 \\ &= e^{-L_s t} \sigma d\widehat{W}_t \\ &= e^{-L_s t} \sigma dW_t + e^{-L_s t} \sigma (M_s t + N_s) dt. \end{aligned} \quad (5.5.64)$$

$$\therefore Y_t = Y_s + \int_s^t (M_s u + N_s) e^{-L_s u} du + \sigma \int_s^t e^{-L_s u} dW_u. \quad (5.5.65)$$

Como $Y_t = e^{-L_s t} X_t$, podemos retornar ao processo X_t de maneira direta:

$$\begin{aligned}
X_t &= e^{L_s t} \left(e^{-L_s s} X_s + \int_s^t (M_s u + N_s) e^{-L_s u} du + \sigma \int_s^t e^{-L_s u} dW_u \right) \\
&= e^{L_s(t-s)} X_s + \frac{M_s}{L_s} \left(s e^{L_s(t-s)} - t \right) + \frac{M_s}{L_s^2} \left(e^{L_s(t-s)} - 1 \right) + \frac{N_s}{L_s} \left(e^{L_s(t-s)} - 1 \right) \\
&\quad + \sigma \int_s^t e^{L_s(t-u)} dW_u \\
&= X_s + X_s \left(e^{L_s(t-s)} - 1 \right) + s \frac{M_s}{L_s} \left(e^{L_s(t-s)} - 1 \right) - \frac{M_s}{L_s} (t-s) \\
&\quad + \frac{M_s}{L_s^2} \left(e^{L_s(t-s)} - 1 \right) + \frac{N_s}{L_s} \left(e^{L_s(t-s)} - 1 \right) \sigma \int_s^t e^{L_s(t-u)} dW_u \\
&= X_s + \frac{L_s X_s + M_s s + N_s}{L_s} \left(e^{L_s(t-s)} - 1 \right) + \frac{M_s}{L_s^2} \left[\left(e^{L_s(t-s)} - 1 \right) - L_s(t-s) \right] \\
&\quad + \sigma \int_s^t e^{L_s(t-u)} dW_u \\
&= X_s + \frac{\hat{f}(s, X_s)}{L_s} \left(e^{L_s(t-s)} - 1 \right) + \frac{M_s}{L_s^2} \left[\left(e^{L_s(t-s)} - 1 \right) - L_s(t-s) \right] \\
&\quad + \sigma \int_s^t e^{L_s(t-u)} dW_u. \tag{5.5.66}
\end{aligned}$$

Perceba que, ao condicionar o processo à observação $X_s = x$, resta-nos apenas um elemento aleatório à direita da equação acima:

$$\sigma \int_s^t e^{L_s(t-u)} dW_u,$$

o qual é normalmente distribuído com média

$$\mathbb{E} \left[\sigma \int_s^t e^{L_s(t-u)} dW_u \right] = 0, \tag{5.5.67}$$

e variância

$$\mathbb{E} \left[\sigma \int_s^t e^{L_s(t-u)} dW_u \right]^2 = \mathbb{E} \left[\sigma^2 \int_s^t e^{2L_s(t-u)} dt \right] = \sigma^2 \frac{e^{2L_s(t-s)} - 1}{2L_s}. \tag{5.5.68}$$

Desta forma, a dinâmica do processo X_t pode ser descrita por

$$X_t | \{X_s = x\} \sim N(A(x), B^2(x)), \tag{5.5.69}$$

com a notação:

$$A(x) := X_s + \frac{\hat{f}(s, X_s)}{L_s} \left(e^{L_s(t-s)} - 1 \right) + \frac{M_s}{L_s^2} \left[\left(e^{L_s(t-s)} - 1 \right) - L_s(t-s) \right], \tag{5.5.70}$$

$$B(x) := \sigma \sqrt{\frac{e^{2L_s(t-s)} - 1}{2L_s}}. \tag{5.5.71}$$

O processo discretizado obtido pelo método da linearização local é simples de ser implementando tanto para fins de simulação quanto para inferência.

5.6 Miscelânea

Esta seção é dedicada à apresentação de diversos outros métodos numéricos encontrados na literatura. Seria interessante poder discutir em mais detalhes todas essas abordagens porém isso resultaria em um trabalho complexo e fugiria do objetivo introdutório do texto. Mesmo não entrando em detalhes, o conhecimento desses métodos é relevante para trabalhos futuros.

Método de Heun (McShane [23])

$$X_{t_{n+1}}^N = X_{t_n}^N + \frac{1}{2}[F_1 + F_2]\Delta + \frac{1}{2}[G_1 + G_2]\Delta W_{t_n}, \quad (5.6.72)$$

onde

$$\begin{aligned} F(t, x) &= \left[f(t, x) - \frac{1}{2}g_x(t, x)g(t, x) \right], & F_1 &= F(X_{t_n}^N) \\ G_1 &= g(t_{n+1}, X_{t_{n+1}}^N), & F_2 &= F(t_n, X_{t_n}^N + F_1\Delta + G_1\Delta W_{t_n}), \\ G_2 &= g(t_n, X_{t_n}^N + F_1\Delta + G_1\Delta W_{t_n}). \end{aligned}$$

Runge-Kutta de Primeira Ordem (Newton [26])

$$X_{t_{n+1}}^N = X_{t_n}^N + F_1\Delta + G_2\Delta W_{t_n} + [G_2 - G_1]\sqrt{\Delta}, \quad (5.6.73)$$

onde

$$F_1 = f(t_n, X_{t_n}^N), \quad G_1 = g(t_n, X_{t_n}^N), \quad G_2 = g\left(t_n, X_{t_n}^N + \frac{G_1(\Delta W_{t_n} - \sqrt{\Delta})}{2}\right). \quad (5.6.74)$$

Runge-Kutta de Três Estágios (Saito e Mitsui [32])

$$\begin{aligned} X_{t_{n+1}}^N &= X_{t_n}^N + \frac{1}{4}[F_1 + 3F_3]\Delta + \frac{1}{4}[G_1 + 3G_3]\Delta W_{t_n} \\ &\quad + \frac{1}{2\sqrt{3}} [f_x(t_n, X_{t_n}^N)g(t_n, X_{t_n}^N) - g_x(t_n, X_{t_n}^N)f(t_n, X_{t_n}^N) \\ &\quad - \frac{1}{2}g_{xx}(t_n, X_{t_n}^N)g(t_n, X_{t_n}^N)^2] \Delta \cdot \Delta \widehat{W}_{t_n}, \end{aligned} \quad (5.6.75)$$

onde ξ_1^n e ξ_2^n são variáveis aleatórias independentes com distribuição Gaussiana padrão,

$$\begin{aligned} \Delta W_{t_n} &= \xi_1^n \sqrt{\Delta}, & \Delta \widehat{W}_{t_n} &= \xi_2^n \sqrt{\Delta}, & F(t, x) &= \left[f(t, x) - \frac{1}{2}g_x(t, x)g(t, x) \right], \\ F_1 &= F(t_n, X_{t_n}^N), & G_1 &= g(t_n, X_{t_n}^N), \\ F_2 &= F\left(t_n, X_{t_n}^N + \frac{1}{3}F_1\Delta + \frac{1}{3}G_1\Delta W_{t_n}\right), & G_2 &= g\left(t_n, X_{t_n}^N + \frac{1}{3}F_1\Delta + \frac{1}{3}G_1\Delta W_{t_n}\right), \\ F_3 &= F\left(t_n, X_{t_n}^N + \frac{2}{3}F_2\Delta + \frac{2}{3}G_2\Delta W_{t_n}\right), & G_3 &= g\left(t_n, X_{t_n}^N + \frac{2}{3}F_2\Delta + \frac{2}{3}G_2\Delta W_{t_n}\right). \end{aligned}$$

Esquema de Taylor (Milstein [24])

$$\begin{aligned}
X_{t_{n+1}}^N &= X_{t_n}^N + f(t_n, X_{t_n}^N) \Delta + g(t_n, X_{t_n}^N) \Delta W_{t_n} & (5.6.76) \\
&+ \frac{1}{2} [g_x(t_n, X_{t_n}^N) g(t_n, X_{t_n}^N)] ((\Delta W_{t_n})^2 - \Delta) + [f_x(t_n, X_{t_n}^N) g(t_n, X_{t_n}^N)] \Delta Z_{t_n} \\
&+ \left[g_x(t_n, X_{t_n}^N) f(t_n, X_{t_n}^N) + \frac{1}{2} g_{xx}(t_n, X_{t_n}^N) g(t_n, X_{t_n}^N)^2 \right] \Delta \hat{Z}_{t_n} \\
&+ \frac{1}{6} \left[g_x^2(t_n, X_{t_n}^N) g(t_n, X_{t_n}^N) + g_{xx}(t_n, X_{t_n}^N) g(t_n, X_{t_n}^N)^2 ((\Delta W_{t_n})^3 - 3\Delta \cdot \Delta W_{t_n}) \right],
\end{aligned}$$

onde ξ_1^n e ξ_2^n são variáveis aleatórias independentes com distribuição Gaussiana padrão,

$$\Delta W_{t_n} = \xi_1^n \sqrt{\Delta}, \quad \Delta Z_{t_n} = \frac{1}{2} \left(\xi_1^n + \frac{\xi_2^n}{\sqrt{3}} \right) \Delta^{3/2}, \quad \Delta \hat{Z}_{t_n} = \frac{1}{2} \left(\xi_1^n - \frac{\xi_2^n}{\sqrt{3}} \right) \Delta^{3/2}.$$

Runge-Kutta de Quatro Estágios (Newton [26])

$$\begin{aligned}
X_{t_{n+1}}^N &= X_{t_n}^N + \frac{1}{2} [F_1 + F_2] \Delta + \frac{1}{40} [37G_1 + 30G_3 - 27G_4] \Delta W_{t_n} \\
&+ \frac{1}{16} [8G_1 + G_2 - 9G_3] \sqrt{3\Delta}, & (5.6.77)
\end{aligned}$$

onde

$$\begin{aligned}
F_1 &= f(t_n, X_{t_n}^N), \quad G_1 = g(t_n, X_{t_n}^N), \\
F_2 &= f\left(t_n, X_{t_n}^N + F_1 \Delta + g_1 \Delta W_{t_n}\right), \quad G_2 = g\left(t_n, X_{t_n}^N - \frac{2}{3} G_1 (\Delta W_{t_n} + \sqrt{2\Delta})\right), \\
G_3 &= g\left(t_n, X_{t_n}^N + \frac{2}{9} G_1 (3\Delta W_{t_n} + \sqrt{3\Delta})\right), \\
G_4 &= g\left(t_n, X_{t_n}^N - \frac{20}{27} F_1 \Delta + \frac{10}{27} (G_2 - G_1) \Delta W_{t_n} - \frac{10}{27} G_2 \sqrt{3\Delta}\right).
\end{aligned}$$

5.7 Exercícios

1. Implemente, em qualquer linguagem de programação, os esquemas de Euler-Maruyama e Milstein para a construção de aproximações para o movimento Browniano geométrico :

$$dX_t = \lambda X_t dt + \mu X_t dW_t, \quad P(X_0 = x_0) = 1.$$

2. Utilizando os programas construídos no exercício anterior, construa soluções para a EDE

$$dX_t = -3X_t dt + \sqrt{3}X_t dW_t, \quad 0 \leq t \leq 10, \\ P(X_0 = 1) = 1.$$

Utilize os métodos de Euler-Maruyama e Milstein considerando os seguintes espaçamentos $\Delta = 2^{-n}$, $n = 1, 2, 3$. É possível observar alguma diferença significativa nos experimentos? Comente os resultados.

3. Ainda com respeito à equação do exercício anterior, implemente o método da linearização local para a construção de aproximações discretas. Utilize este programa para estimar $\mathbb{E}X_1$ e compare as estimativas com os resultados analíticos obtidos no Capítulo 4.
4. Todos os métodos abordados aqui lidaram com um passo de discretização Δ constante. Sugira uma estratégia adaptativa para a escolha dinâmica de espaçamentos e comente sua escolha. Implemente a sugestão para o método de Euler-Maruyama aplicado ao problema proposto no Exercício 2.
5. Considere a EDE

$$dX_t = \text{sen}(X_t)dt + dW_t, \quad 0 \leq t < 1, \\ P(X_0 = 2) = 1.$$

Utilizando o método de discretização de sua preferência, construa um programa para obter aproximações da função densidade de probabilidade referentes às soluções desta EDE com respeito à medida de Wiener.

Capítulo 6

Análise Numérica

A análise formal de procedimentos numéricos utilizados na construção de soluções aproximadas para equações diferenciais, sejam determinísticas ou estocásticas, é parte essencial do processo de aplicação desses métodos em um cenário prático. Três aspectos fundamentais devem ser levados em consideração durante a construção de um boa aproximação numérica:

convergência: trata da premissa básica de qualquer método numérico: garantir que o erro global de aproximação (a distância entre solução real e aproximada) tenda a zero à medida que tomamos partições cada vez mais refinadas;

consistência: refere-se à uma medida quantitativa do quanto a solução exata é condizente com o problema discretizado;

estabilidade: quantifica o controle do problema discretizado com respeito ao crescimento do erro de aproximação em um horizonte infinito de tempo.

Devido à natureza aleatória das variáveis envolvidas em uma EDE, os procedimentos clássicos de análise para métodos determinísticos não se aplicam diretamente. Conforme será apresentado a seguir, algumas adaptações são cruciais de modo a simultaneamente manter a precisão dos conceitos de convergência, consistência e estabilidade, e acomodar a estrutura estocástica do problema.

6.1 Convergência

Ao utilizar um método numérico para obter trajetórias associadas a uma EDE, devemos assegurar que o erro de aproximação seja controlado. Este erro pode ser calculado através do critério do erro absoluto, o qual toma o valor esperado da diferença entre o processo discretizado $X^N = \{X_{t_n}^N; t_n \in \tau_\delta\}$, associado à partição $\tau_\delta = \{0 = t_0 < t_1 < \dots < t_N = T\}$, e o processo real $X = \{X_t; 0 \leq t \leq T\}$, avaliada no instante final do intervalo de observação:

$$\epsilon := \mathbb{E} [|X_T^N - X_T|]. \quad (6.1.1)$$

Este critério nos dá uma boa estimativa da proximidade das trajetórias real e aproximada ao final das observações.

Dizemos que uma aproximação discreta X^N , com espaçamento máximo δ , converge fortemente para X no instante T se

$$\lim_{\delta \downarrow 0} \mathbb{E} [|X_T^N - X_T|] = 0. \quad (6.1.2)$$

Uma vez definido o conceito de convergência forte, é natural tentarmos desenvolver um critério para a comparação das aproximações obtidas por dois ou mais procedimentos distintos.

Definição 6.1. *Um processo discretizado X^N converge fortemente com ordem $\alpha > 0$ no instante T se existir uma constante M , independente de δ , e um $\delta_0 > 0$ tais que*

$$\epsilon(\delta) = \mathbb{E} [|X^N - X_T|] \leq M\delta^\alpha, \quad \forall \delta \in (0, \delta_0). \quad (6.1.3)$$

Perceba que a relação (6.1.3) nos dá uma ideia da "velocidade" da convergência do modelo discretizado. Analisando a fórmula acima podemos concluir que quanto maior o valor do coeficiente α , menor será a cota $M\delta^\alpha$, desde que δ seja suficientemente pequeno. Em outras palavras, este resultado garante que quanto maior a ordem de convergência forte α , maior poderá ser o passo δ .

Dentre as aproximações discretas apresentadas, Shoji [34] mostra que o método de linearização local possui a maior ordem de convergência forte: $\alpha = 2$. Esse resultado é bastante expressivo já que esta é, segundo Rümelin [31], a maior ordem de convergência acessível a uma aproximação desta natureza. Quanto aos métodos de Euler, θ -estocástico e Milstein, podemos inspecionar os erros R_1 e R_2 na expansão de Itô-Taylor para inferir as respectivas ordens de convergência forte: 0.5, 0.5 e 1.0. Uma análise semelhante à do método de Milstein pode ser aplicada ao esquema livre de derivada, lembrando que este é basicamente uma "simplificação" do primeiro. Caso o passo de discretização seja suficientemente pequeno, a aproximação utilizada para a derivada do coeficiente de difusão será boa o suficiente para não propagar erros adicionais significativos ao esquema. Desta forma, também chegamos a uma ordem de convergência forte $\alpha = 1.0$ para o esquema livre de derivada. Estes resultados serão verificados numericamente mais adiante.

Assim como na teoria das equações diferenciais determinísticas, o conceito de convergência forte está ligado ao de consistência, o qual pode ser mais fácil de se verificar em determinados casos.

Definição 6.2. *Uma aproximação discreta $X^N = \{X_{t_n}^N; t_n \in \tau_\delta\}$, associada à partição τ_δ do intervalo $[0, T]$, $T > 0$, é fortemente consistente se existir uma função não negativa $c = c(\delta)$ com*

$$\lim_{\delta \downarrow 0} c(\delta) = 0, \quad (6.1.4)$$

tal que

$$\mathbb{E} \left[\left| \mathbb{E} \left(\frac{X_{t_{n+1}}^N - X_{t_n}^N}{\Delta_N} \middle| \mathcal{F}_{t_n} \right) - f(t_n, X_{t_n}^N) \right|^2 \right] \leq c(\delta), \quad (6.1.5)$$

e

$$\mathbb{E} \left[\frac{1}{\Delta_n} \left| X_{t_{n+1}}^N - X_{t_n}^N - \mathbb{E} \left[X_{t_{n+1}}^N - X_{t_n}^N \middle| \mathcal{F}_{t_n} \right] - g(t_n, X_{t_n}^N) \Delta W_n \right|^2 \right] \leq c(\delta) \quad (6.1.6)$$

para quaisquer valores fixos $X_{t_n}^N = x$ e $n = 0, 1, 2, \dots, N - 1$.

A condição (6.1.5) garante que o valor esperado dos incrementos da aproximação converge para o coeficiente de deslocamento f em todos os instantes observados. Por outro lado, a condição (6.1.6) diz que a variância da diferença entre as partes aleatórias de X^N e X também converge para zero em cada instante. Em resumo, este critério nos dá indícios de que os caminhos amostrais de X^N se aproximam dos

de X , diferentemente da ideia de convergência forte que garante esta proximidade apenas no instante final. Os exemplos a seguir ilustram casos em que a consistência forte do esquema numérico pode ser verificada de maneira simples.

Exemplo 6.1. *Seja X^N a aproximação discreta do método de Euler para a equação diferencial estocástica*

$$\begin{aligned} dX_t &= f(t, X_t)dt + g(t, X_t)dW_t, \quad t \in [0, T], \\ P(X_0 = \xi_0) &= 1. \end{aligned}$$

Perceba que

$$\begin{aligned} \mathbb{E} \left[\frac{X_{t_{n+1}}^N - X_{t_n}^N}{\Delta_N} \middle| \mathcal{F}_{t_n} \right] &= \mathbb{E} \left[\frac{f(t_n, X_{t_n}) \Delta_N + g(t_n, X_{t_n}) \Delta W_{t_n}}{\Delta_N} \middle| \mathcal{F}_{t_n} \right] \\ &= \mathbb{E} \left[\frac{g(t_n, X_{t_n}^N) \Delta W_{t_n}}{\Delta_N} \middle| \mathcal{F}_{t_n} \right] + \mathbb{E} [f(t_n, X_{t_n}^N) | \mathcal{F}_{t_n}] \\ &= \mathbb{E}[\Delta W_{t_n}] \mathbb{E} \left[\frac{g(t_n, X_{t_n}^N)}{\Delta_N} \middle| \mathcal{F}_{t_n} \right] + f(t_n, X_{t_n}^N) \\ &= f(t_n, X_{t_n}^N). \end{aligned}$$

Assim, para qualquer passo de discretização δ , podemos definir $c(\delta) \equiv 0$, de modo a obter

$$\mathbb{E} \left[\left| \mathbb{E} \left[\frac{X_{t_{n+1}}^N - X_{t_n}^N}{\Delta_N} \middle| \mathcal{F}_{t_n} \right] - f(t_n, X_{t_n}^N) \right|^2 \right] = \mathbb{E} [|f(t_n, X_{t_n}^N) - f(t_n, X_{t_n}^N)|^2] = 0;$$

$$\begin{aligned} \mathbb{E} \left[\frac{1}{\Delta_n} \left| X_{t_{n+1}}^N - X_{t_n}^N - \mathbb{E} [X_{t_{n+1}}^N - X_{t_n}^N | \mathcal{F}_{t_n}] - g(t_n, X_{t_n}^N) \Delta W_{t_n} \right|^2 \right] &= \\ = \mathbb{E} \left[\frac{1}{\Delta_N} \left| f(t_n, X_{t_n}^N) \Delta_N + g(t_n, X_{t_n}^N) \Delta W_{t_n} - f(t_n, X_{t_n}^N) \Delta_n - g(t_n, X_{t_n}^N) \Delta W_{t_n} \right|^2 \right] &= \\ = 0. \end{aligned}$$

Portanto, o método de Euler é fortemente consistente.

Exemplo 6.2. *Considere novamente a equação diferencial estocástica do exemplo anterior e defina a aproximação discreta X^N a partir do método de Milstein. Perceba que*

$$\begin{aligned} \mathbb{E} \left[\frac{X_{t_{n+1}}^N - X_{t_n}^N}{\Delta_N} \middle| \mathcal{F}_{t_n} \right] &= \mathbb{E} \left[\frac{f(t_n, X_{t_n}) \Delta_N + g(t_n, X_{t_n}) \Delta W_{t_n}}{\Delta_N} \middle| \mathcal{F}_{t_n} \right] \\ &\quad + \mathbb{E} \left[\frac{\frac{1}{2} g(t_n, X_{t_n}) g_x(t_n, X_{t_n}) (\Delta W_{t_n}^2 - \Delta_N)}{\Delta_N} \middle| \mathcal{F}_{t_n} \right] \\ &= f(t_n, X_{t_n}^N) + \frac{1}{2\Delta_N} g(t_n, X_{t_n}) g_x(t_n, X_{t_n}) \mathbb{E} [\Delta W_{t_n}^2 - \Delta_N] \\ &= f(t_n, X_{t_n}^N). \end{aligned}$$

Além disso, temos que $\Delta W_{t_n}^2 = \Delta_N Z^2$, onde $Z \sim N(0, 1)$. Neste caso, $Z^2 \sim \chi_{(1)}$ e

$$\text{Var}(\Delta W_{t_n}^2) = \Delta_N^2 \text{Var}(Z^2) = 2\Delta_N^2. \quad (6.1.7)$$

Se $\frac{1}{2}gg_x$ é uma aplicação limitada, então podemos encontrar um $M > 0$ tal que $\frac{1}{2}gg_x \leq M, \forall (t, x) \in [0, T] \times \mathbb{R}$. Desta forma, definindo $c(\delta) = 2M^2\delta$, temos que $\lim_{\delta \downarrow 0} c(\delta) = 0$, e

$$\begin{aligned} \mathbb{E} \left[\left| \mathbb{E} \left[\frac{X_{t_{n+1}}^N - X_{t_n}^N}{\Delta_N} \middle| \mathcal{F}_{t_n} \right] - f(t_n, X_{t_n}^N) \right|^2 \right] &= \mathbb{E} \left[|f(t_n, X_{t_n}^N) - f(t_n, X_{t_n}^N)|^2 \right] = 0; \\ \mathbb{E} \left[\frac{1}{\Delta_n} \left| X_{t_{n+1}}^N - X_{t_n}^N - \mathbb{E} \left[X_{t_{n+1}}^N - X_{t_n}^N \middle| \mathcal{F}_{t_n} \right] - g(t_n, X_{t_n}^N) \Delta W_n \right|^2 \right] &= \\ &= \mathbb{E} \left[\frac{1}{\Delta_N} \left| \frac{1}{2}g(t_n, X_{t_n})g_x(t_n, X_{t_n}) (\Delta W_{t_n}^2 - \Delta_N) \right|^2 \right] \\ &\leq \frac{M^2}{\Delta_N} \mathbb{E} [\Delta W_{t_n}^2 - \Delta_N] \\ &= \frac{M^2}{\Delta_N} \text{Var}(\Delta W_{t_n}^2) = 2M^2\Delta_N \leq 2M^2\delta = c(\delta) \longrightarrow 0. \end{aligned} \quad (6.1.8)$$

Portanto, o método de Milstein é consistente desde que a aplicação $\frac{1}{2}gg_x$ seja limitada.

Dadas as propriedades dos conceitos envolvidos nas definições (6.1)-(6.2), é intuitivo pensar que consistência implica convergência. De fato, Kloeden e Platen [19] mostram que se a equação diferencial estocástica de interesse satisfaz as condições de regularidade para a existência e unicidade de soluções fortes, então uma aproximação discreta consistente é também convergente. Apesar desta relação entre os dois conceitos, consistência forte ainda não é um tópico amplamente explorado no contexto das equações diferenciais estocásticas ordinárias. A maioria dos trabalhos concentra-se nas idéias de convergência e estabilidade.

6.2 Estabilidade

Convergência de um processo discretizado não é uma condição suficiente para garantir que a aproximação obtida seja aplicável, uma vez que este conceito leva em consideração apenas o comportamento dos métodos em um intervalo finito $[0, T]$. Na prática, precisamos assegurar também que a propagação de erros iniciais e erros de arredondamento seja mantida sob controle à medida que $T \rightarrow +\infty$. O estudo deste comportamento a longo prazo é dado pelo conceito de estabilidade. De uma maneira mais simples, podemos dizer que um método numérico é estável se as soluções construídas próximas à solução real tendem a se manter próximas a longo prazo e, dizemos que o método é instável se essas soluções tendem a se afastar. Visualmente, em um problema estável os gráficos das soluções estão próximos, enquanto em um problema instável eles se separam em instantes de tempo suficientemente grandes.

No contexto determinístico das equações diferenciais ordinárias, uma vasta teoria de estabilidade já foi desenvolvida. Basicamente, neste tipo de estudo um método numérico é aplicado a uma classe de problemas cujo comportamento assintótico é conhecido e a habilidade do método em reproduzir este comportamento é avaliada. Dentre todas as classes de problemas encontradas na literatura, a mais simples (e talvez a mais interessante) envolve a equação linear de teste

$$X' = \lambda X, \quad (6.2.9)$$

onde $\lambda \in \mathbb{C}$ é um parâmetro fixo. É possível desenvolver uma teoria similar no caso estocástico utilizando como equação de teste o movimento Browniano geométrico

$$\begin{aligned} dX_t &= \lambda X_t dt + \mu X_t dW_t, \quad t \in [0, T], \\ X_0 &= x_0, \end{aligned} \quad (6.2.10)$$

onde x_0 é uma constante real fixa. Já verificamos anteriormente que este problema possui uma solução explícita dada por

$$X_t = x_0 \exp \left\{ \left(\lambda - \frac{1}{2} \mu^2 \right) t + \mu W_t \right\}. \quad (6.2.11)$$

Perceba que se $x_0 = 0$, então temos uma solução trivial $X_t \equiv 0, \forall t \in [0, T]$, a qual será denominada solução de equilíbrio. Nosso objetivo, primeiramente, é determinar formalmente o conceito de estabilidade e verificar sob quais condições a solução analítica acima tende à solução de equilíbrio.

Definição 6.3. *Seja $X = \{X_t; 0 \leq t < +\infty\}$ um processo estocástico. Dizemos que a posição de equilíbrio $X_t \equiv 0$ é assintoticamente estável em média quadrática se, para todo $\epsilon > 0$ existir um $\delta > 0$ tal que*

$$\mathbb{E} [|X_t|^2] < \epsilon, \quad \forall t \geq 0, |X_0| < \delta, \quad (6.2.12)$$

e, além disso, se existir um $\delta_0 > 0$ tal que

$$\lim_{t \rightarrow +\infty} \mathbb{E} [|X_t|^2] = 0, \quad \forall |X_0| < \delta_0. \quad (6.2.13)$$

Observação: por simplicidade, daqui em diante utilizaremos a expressão *solução estável* para fazer referência a uma solução assintoticamente estável em média quadrática.

Proposição 6.1. *A solução (6.2.11) é estável se, e somente se,*

$$2\lambda + \mu^2 < 0. \quad (6.2.14)$$

Demonstração: : Considere $Y_t := \mathbb{E} [|X_t|^2]$. Assim,

$$\begin{aligned} Y_t &= \mathbb{E} \left[\left| x_0 \exp \left\{ \left(\lambda - \frac{1}{2} \mu^2 \right) t + \mu W_t \right\} \right|^2 \right] \\ &= |x_0|^2 \exp \{ (2\lambda - \mu^2) t \} \mathbb{E} [|\exp \{ 2\mu W_t \}|] = |x_0|^2 \exp \{ (2\lambda - \mu^2) t \} \exp \{ 2t\mu^2 \} \\ &= |x_0|^2 \exp \{ (2\lambda + \mu^2) t \}. \end{aligned} \quad (6.2.15)$$

O resultado segue trivialmente da igualdade acima. ■

Uma vez entendido o comportamento analítico da solução (6.2.11), o próximo passo é determinar as condições sob as quais a solução $X_{t_n}^N$ gerada por um esquema numérico é estável, ou seja, determinar as condições para que os métodos numéricos reproduzam o comportamento a longo prazo da solução analítica. Assim como na demonstração do teorema anterior, considere

$$\hat{Y}_n = \mathbb{E} [|X_{t_n}^N|^2]. \quad (6.2.16)$$

Quando aplicamos um esquema numérico à equação linear de teste (6.2.10), obtemos uma equação de diferenças da forma

$$\hat{Y}_{t_{n+1}} = R(\lambda, \mu, \Delta)\hat{Y}_{t_n}, \quad (6.2.17)$$

onde R é uma função real denominada função de estabilidade, λ e μ são os parâmetros do problema, e Δ representa o tamanho do passo de discretização. Analisando a relação acima como uma equação de recorrência, chegamos ao resultado

$$\hat{Y}_{t_{n+1}} = (R(\lambda, \mu, \Delta))^n \hat{Y}_{t_0}, \quad (6.2.18)$$

o que nos mostra que a solução numérica será estável se

$$|R(\lambda, \mu, \Delta)| < 1. \quad (6.2.19)$$

Definição 6.4. *Um esquema numérico é estável para qualquer tripla (λ, μ, Δ) que satisfaz a relação (6.2.19). A região \mathcal{R} dada por*

$$\mathcal{R} := \{(\lambda, \mu, \Delta); |R(\lambda, \mu, \Delta)| < 1\}, \quad (6.2.20)$$

é denominada região de estabilidade numérica.

Cada um dos esquemas apresentados até aqui possui uma região de estabilidade distinta. Quanto mais próxima esta região estiver do domínio de estabilidade da solução analítica, mais interessante será o método.

É importante destacarmos que o conceito de estabilidade em média quadrática não é o único encontrado na literatura. Outros conceitos comuns são os de estabilidade assintótica e estabilidade-T. O primeiro difere dos conceitos abordados aqui pelo fato de explorar convergência quase certa ao invés de convergência em média quadrática. Já o segundo destaca a convergência da única trajetória proveniente da filtração natural imposta pelo movimento Browniano corrente. Cada uma destas abordagens pode se destacar com relação às outras em função do contexto de trabalho. Em particular, estabilidade em média quadrática supre bem a análise proposta neste trabalho por expandir a noção usual de estabilidade absoluta na análise de métodos para EDOs; uma análise dos demais conceitos pode ser obtida em Burrage et al. [8].

6.2.1 Funções e Regiões de Estabilidade de Alguns Esquemas

Apresentamos agora o cálculo explícito das funções e regiões de estabilidade dos métodos discutidos neste trabalho: Euler-Maruyama, Milstein e θ estocástico. Quanto ao método da linearização local, a análise será feita em separado devido às características do problema. A estabilidade deste método será então comentada em separado mais adiante.

Euler-Maruyama

Aplicando o método de Euler-Maruyama à equação linear de teste, obtemos:

$$X_{t_{n+1}}^N = X_{t_n}^N [(1 + \lambda\Delta) + \mu\Delta W_{t_n}]. \quad (6.2.21)$$

Elevando ao quadrado os dois membros da igualdade e tomando a esperança, obtemos, segundo a notação apresentada anteriormente,

$$\begin{aligned} \hat{Y}_{t_{n+1}} &= \hat{Y}_{t_n} \mathbb{E} [(1 + \lambda\Delta)^2 + 2(1 + \lambda\Delta)\mu\Delta W_{t_n} + \mu^2(\Delta W_{t_n})^2] \\ &= \hat{Y}_{t_n} \cdot [(1 + \lambda\Delta)^2 + \mu^2\Delta] =: R(\lambda, \mu, \Delta)\hat{Y}_{t_n}, \end{aligned} \quad (6.2.22)$$

onde $R(\lambda, \mu, \Delta) = (1 + \lambda\Delta)^2 + \mu^2\Delta$.

A região de estabilidade pode ser expressa em função do conjunto das triplas ordenadas (λ, μ, Δ) tais que

$$|R(\lambda, \mu, \Delta)| < 1. \quad (6.2.23)$$

Utilizando a igualdade obtida acima, chegamos à seguinte região de estabilidade:

$$\mathcal{R}_{em} = \left\{ (\lambda, \mu, \Delta) \in \mathcal{R} \times \mathbb{R}^+; \Delta < -\frac{2\lambda + \mu}{\lambda^2} \right\}, \quad (6.2.24)$$

onde \mathcal{R} é a região delimitada pela proposição 6.1, isto é,

$$\{(\lambda, \mu) \in \mathbb{R}^2; 2\lambda + \mu^2 < 0\}.$$

Milstein

Aplicando o método de Milstein à equação linear de teste, obtemos:

$$X_{t_{n+1}}^N = X_{t_n}^N \left[(1 + \lambda\Delta) + \mu\Delta W_{t_n} + \frac{1}{2}\mu^2((\Delta W_{t_n})^2 - \Delta) \right]. \quad (6.2.25)$$

Elevando ao quadrado os dois membros da igualdade e tomando a esperança, obtemos,

$$\begin{aligned} \hat{Y}_{t_{n+1}} &= \hat{Y}_{t_n} \mathbb{E} \left[(1 + \lambda\Delta)^2 + \mu^2(\Delta W_{t_n})^2 + \frac{1}{4}\mu^4((\Delta W_{t_n})^4 - 2\Delta(\Delta W_{t_n})^2 + \Delta^2) \right. \\ &\quad \left. + 2(1 + \lambda\Delta)(\mu\Delta W_{t_n}) + 2 \cdot \frac{1}{2}\mu^2(1 + \lambda\Delta)((\Delta W_{t_n})^2 - \Delta) \right. \\ &\quad \left. + 2 \cdot \frac{1}{2}\mu^3\Delta W_{t_n}((\Delta W_{t_n})^2 - \Delta) \right] \\ &= \hat{Y}_{t_n} \left[(1 + \lambda\Delta)^2 + \mu^2\Delta + \frac{1}{4}\mu^4(3\Delta^2 - 2\Delta^2 + \Delta^2) \right] \\ &= \hat{Y}_{t_n} \left[(1 + \lambda\Delta)^2 + \mu^2\Delta + \frac{1}{2}\mu^4\Delta^2 \right] \\ &=: R(\lambda, \mu, \Delta)\hat{Y}_{t_n}, \end{aligned} \quad (6.2.26)$$

onde $R(\lambda, \mu, \Delta) = \left[(1 + \lambda\Delta)^2 + \mu^2\Delta + \frac{1}{2}\mu^4\Delta^2 \right]$. Assim, chegamos à seguinte região de estabilidade:

$$\mathcal{R}_{mils} = \left\{ (\lambda, \mu, \Delta) \in \mathcal{R} \times \mathbb{R}; \Delta < -\frac{2\lambda + \mu^2}{\lambda^2 + \frac{1}{2}\mu^4} \right\}. \quad (6.2.27)$$

Como $2\lambda + \mu^2 < 0$, segue que

$$-\frac{2\lambda + \mu^2}{\lambda^2 + \frac{1}{2}\mu^4} < -\frac{2\lambda + \mu^2}{\lambda^2}. \quad (6.2.28)$$

Desta forma, o domínio de estabilidade do método de Euler-Maruyama é mais abrangente que o do método de Milstein:

$$\mathcal{R}_{mils} \subset \mathcal{R}_{em}.$$

Esquema livre de derivada

Do exemplo (5.2), temos que o esquema livre de derivada aplicado à equação linear de teste resulta no seguinte método recursivo:

$$X_{t_{n+1}} = X_{t_n} \left[1 + \lambda\Delta + \mu\Delta W_{t_n} + \mu(\lambda\Delta + \mu\sqrt{\Delta}) \frac{(\Delta W_{t_n})^2 - \Delta}{2\sqrt{\Delta}} \right]. \quad (6.2.29)$$

Elevando ao quadrado os dois membros desta igualdade e tomando a esperança, obtemos,

$$\begin{aligned} \hat{Y}_{t_{n+1}} &= \hat{Y}_{t_n} \mathbb{E} \left[(1 + \lambda\Delta + \mu\Delta W_{t_n})^2 + \left(\mu(\lambda\Delta + \mu\sqrt{\Delta}) \frac{(\Delta W_{t_n})^2 - \Delta}{2\sqrt{\Delta}} \right)^2 \right. \\ &\quad \left. + 2(1 + \lambda\Delta + \mu\Delta W_{t_n}) \left(\mu(\lambda\Delta + \mu\sqrt{\Delta}) \frac{(\Delta W_{t_n})^2 - \Delta}{2\sqrt{\Delta}} \right) \right] \\ &= \hat{Y}_{t_n} \left[(1 + \lambda\Delta)^2 + \mu^2\Delta + \mu^2(\lambda^2\Delta^2 + 2\lambda\mu\Delta^{3/2} + \mu^2\Delta) \frac{3\Delta^2 - 2\Delta \cdot \Delta + \Delta^2}{4\Delta} \right] \\ &= \hat{Y}_{t_n} \left[(1 + \lambda\Delta)^2 + \mu^2\Delta + \frac{\mu^2\Delta^2}{2}(\lambda^2\Delta + 2\lambda\mu\Delta^{1/2} + \mu^2) \right] \\ &=: R(\lambda, \mu, \Delta) \hat{Y}_{t_n}, \end{aligned} \quad (6.2.30)$$

onde $R(\lambda, \mu, \Delta) = \left[(1 + \lambda\Delta)^2 + \mu^2\Delta + \frac{\mu^2\Delta^2}{2}(\lambda^2\Delta + 2\lambda\mu\Delta^{1/2} + \mu^2) \right]$. A partir da função acima definimos a região de estabilidade mas, diferentemente dos casos anteriores, este resultado não pode ser expresso de maneira trivial isolando Δ das demais variáveis. Alguns autores (Liske & Platen (1987), Kloeden & Platen (1989), Saito & Mitsui (1996)) consideram uma simplificação deste esquema numérico que resulta em uma função de estabilidade mais simples. Tal simplificação consiste em substituir o termo

$$\left[g(t_n, X_{t_n}^N) + f(t_n, X_{t_n}^N) \Delta + g(t_n, X_{t_n}^N) \sqrt{\Delta} \right] - g(t_n, X_{t_n}^N) \quad (6.2.31)$$

por

$$\left[g(t_n, X_{t_n}^N) + g(t_n, X_{t_n}^N) \sqrt{\Delta} \right] - g(t_n, X_{t_n}^N). \quad (6.2.32)$$

Neste caso, chegaríamos de imediato à função de estabilidade

$$R(\lambda, \mu, \Delta) \hat{Y}_{t_n} = \left[(1 + \lambda\Delta)^2 + \mu^2\Delta + \frac{1}{2}\mu^4\Delta^2 \right], \quad (6.2.33)$$

cujo tratamento é mais simples. Mais precisamente, chegaríamos à mesma função de estabilidade do método de Milstein, portanto

$$\mathcal{R}_{ld} = \left\{ (\lambda, \mu, \Delta) \in \mathcal{R} \times \mathbb{R}; \Delta < -\frac{2\lambda + \mu^2}{\lambda^2 + \frac{1}{2}\mu^4} \right\}. \quad (6.2.34)$$

Ao lidarmos com um domínio de estabilidade mais "tratável", as comparações entre os métodos são realizadas de maneira mais natural. Neste caso, utilizaremos a segunda abordagem durante a avaliação numérica.

θ Estocástico

A equação linear de teste é um caso especial no qual o método θ Estocástico pode ser aplicado diretamente sem a utilização de algum algoritmo numérico auxiliar. Aplicando o método à equação linear de teste, obtemos:

$$\begin{aligned} X_{t_{n+1}}^N &= X_{t_n}^N + \theta \lambda X_{t_{n+1}}^N \Delta + (1 - \theta) \lambda X_{t_n}^N \Delta + \mu X_{t_n}^N \Delta W_{t_n} \\ &= X_{t_n}^N \cdot \frac{1 + (1 - \theta) \lambda \Delta + \mu \Delta W_{t_n}}{1 - \theta \lambda \Delta}. \end{aligned} \quad (6.2.35)$$

Elevando ao quadrado os dois membros da igualdade e tomando a esperança, obtemos,

$$\begin{aligned} \hat{Y}_{t_{n+1}} &= \hat{Y}_{t_n} \mathbb{E} \left\{ \frac{[1 + (1 - \theta) \lambda \Delta]^2 + 2\mu \Delta W_{t_n} [1 + (1 - \theta) \lambda \Delta] + \mu^2 (\Delta W_{t_n})^2}{(1 - \theta \lambda \Delta)^2} \right\} \\ &= \hat{Y}_{t_n} \frac{[1 + (1 - \theta) \lambda \Delta]^2 + \mu^2 \Delta}{(1 - \theta \lambda \Delta)^2} \\ &=: R(\lambda, \mu, \Delta) \hat{Y}_{t_n}, \end{aligned} \quad (6.2.36)$$

onde $R(\lambda, \mu, \Delta) = \frac{[1 + (1 - \theta) \lambda \Delta]^2 + \mu^2 \Delta}{(1 - \theta \lambda \Delta)^2}$. Em particular, obtemos para os métodos de Euler trapezoidal e implícito, as funções de estabilidade

$$R_T(\lambda, \mu, \Delta) = \frac{[1 + \frac{1}{2} \lambda \Delta]^2 + \mu^2 \Delta}{(1 - \frac{1}{2} \lambda \Delta)^2} \quad \text{e} \quad R_I(\lambda, \mu, \Delta) = \frac{1 + \mu^2 \Delta}{(1 - \lambda \Delta)^2},$$

respectivamente.

Uma primeira análise das regiões de estabilidade nos mostra que o método θ Estocástico é mais estável que os anteriores. De fato, podemos verificar formalmente que para determinadas combinações da dupla (λ, μ) , a estabilidade deste esquema numérico torna-se independente do passo de discretização Δ .

Proposição 6.2. *Se $\mu^2 + 2\lambda < 0$, então o método θ estocástico é incondicionalmente estável em média quadrática para $\frac{1}{2} \leq \theta \leq 1$.*

Demonstração: Lembramos inicialmente que a desigualdade $\mu^2 + 2\lambda < 0$ é exatamente a condição exigida para que a solução explícita da equação linear de teste seja assintoticamente estável em média quadrática. Ou seja, esta condição é o mínimo que devemos exigir para fazer sentido pensar em estabilidade.

Já verificamos que a função de estabilidade deste esquema é dada por

$$R(\lambda, \mu, \Delta) = \frac{[1 + (1 - \theta) \lambda \Delta]^2 + \mu^2 \Delta}{(1 - \theta \lambda \Delta)^2}. \quad (6.2.37)$$

Considerando primeiramente $\theta \in (\frac{1}{2}, 1]$, obtemos:

$$\begin{aligned} R(\lambda, \mu, \Delta) < 1 &\Leftrightarrow \frac{[1 + (1 - \theta) \lambda \Delta]^2 + \mu^2 \Delta}{(1 - \theta \lambda \Delta)^2} < 1 \\ &\Leftrightarrow \lambda^2 \Delta^2 + 2\lambda \Delta (1 - \theta \lambda \Delta) + (1 - \theta \lambda \Delta)^2 + \mu^2 \Delta < (1 - \theta \lambda \Delta)^2 \\ &\Leftrightarrow \lambda^2 \Delta + 2\lambda (1 - \theta \lambda \Delta) + \mu^2 < 0 \\ &\Leftrightarrow \Delta > \frac{\mu^2 + 2\lambda}{\lambda^2 (2\theta - 1)}. \end{aligned} \quad (6.2.38)$$

Como Δ deve ser um valor estritamente positivo e as restrições $\frac{1}{\lambda^2(2\theta-1)} > 0$ e $\mu^2 + 2\lambda < 0$ são verificadas, concluímos que a relação acima é satisfeita para qualquer Δ . Em outras palavras, o esquema é incondicionalmente estável.

Considerando agora $\theta = \frac{1}{2}$ e aproveitando os cálculos acima, obtemos

$$\begin{aligned} R(\lambda, \mu, \Delta) < 1 &\Leftrightarrow \lambda^2\Delta + 2\lambda\left(1 - \frac{1}{2}\lambda\Delta\right) + \mu^2 < 0 \\ &\Leftrightarrow \mu^2 + 2\lambda < 0, \end{aligned} \quad (6.2.39)$$

o que é garantido pela hipótese inicial. \blacksquare

Um esquema numérico incondicionalmente estável é também denominado *esquema A-estável*.

6.2.2 Método da Linearização Local

A análise de estabilidade apresentada aqui leva em consideração a equação linear de teste (6.2.10). Em se tratando do método da linearização local, este tipo de análise é um tanto quanto trivial pois estaríamos aplicando um método de linearização a um problema originalmente linear. Devemos esperar então que a estabilidade do esquema obtido não dependa do passo de discretização, mas sim da equação original. Em outras palavras, devemos esperar que o método seja incondicionalmente estável.

Aplicando a transformação de Lamperti

$$Z = S(t, X) = \frac{\log X}{\mu} \quad (6.2.40)$$

à equação linear de teste, chegamos à nova equação diferencial estocástica:

$$dZ_t = \left(\frac{\lambda}{\mu} - \frac{1}{2}\mu \right) dt + dW_t. \quad (6.2.41)$$

O método da linearização local é baseado no tratamento analítico simples desta equação. Utilizando a notação empregada na descrição do método, obtemos a seguinte relação:

$$Z_t = e^{L_s t} \left(e^{-L_s s} Z_s + \int_s^t (M_s u + N_s) e^{-L_s u} du + \sigma \int_s^t e^{-L_s u} dW_u \right), \quad (6.2.42)$$

onde

$$L_s \equiv 0, \quad M_s \equiv 0, \quad N_s = f(s, Z_s) = \left(\frac{\lambda}{\mu} - \frac{1}{2}\mu \right), \quad \sigma = 1. \quad (6.2.43)$$

Perceba que esta abordagem resultará no seguinte esquema:

$$Z_{t_{n+1}} = Z_{t_n} + \left(\frac{\lambda}{\mu} - \frac{1}{2}\mu \right) \Delta + \Delta W_{t_n}. \quad (6.2.44)$$

Aplicando a transformação inversa de Lamperti, segue que o esquema de discretização com respeito ao processo original é expresso por

$$\begin{aligned} X_{t_{n+1}}^N &= e^{\mu Z_{t_{n+1}}} \\ &= \exp \left\{ \mu Z_{t_n} + \left(\lambda - \frac{1}{2}\mu^2 \right) \Delta + \mu \Delta W_{t_n} \right\}, \end{aligned} \quad (6.2.45)$$

e pode ser resolvido como uma relação de recorrência

$$\begin{aligned} X_{t_{n+1}}^N &= x_0 \exp \left\{ \left(\lambda - \frac{1}{2} \mu^2 \right) (n+1) \Delta + \mu \sum_{j=0}^n \Delta W_{t_j} \right\} \\ &= x_0 \exp \left\{ \left(\lambda - \frac{1}{2} \mu^2 \right) t_{n+1} + \mu W_{t_{n+1}} \right\}. \end{aligned} \quad (6.2.46)$$

A relação acima pode ser estudada de maneira idêntica à proposição 6.1, o que confirma as expectativas iniciais. Voltaremos a estudar a estabilidade do método da linearização local através de problemas mais "relevantes" na próxima seção, onde confirmaremos numericamente que este método é de fato mais estável que os anteriores.

6.3 Avaliação Numérica

Esta seção é reservada à análise numérica de convergência e estabilidade dos esquemas numéricos discutidos anteriormente, através de simulações de Monte Carlo. Os experimentos trabalhados aqui são apenas ilustrativos e têm como objetivo dar evidências empíricas a respeito dos conceitos abordados, desta forma, não podemos tomá-los como demonstrações formais.

6.3.1 Convergência

Este conjunto de testes avalia a ordem de convergência forte dos esquemas: Euler-Maruyama, Milstein, livre de derivada, Euler trapezoidal e Euler implícito. Novamente, utilizamos como base de avaliação o modelo do movimento Browniano geométrico

$$dX_t = \lambda X_t dt + \mu X_t dW_t, \quad 0 \leq t \leq T, \quad (6.3.47)$$

$$X_0 = x_0, \quad (6.3.48)$$

com parâmetros $\lambda = 4$, $\mu = 2$ e $x_0 = 1$. Segundo o conceito apresentado na seção anterior, um esquema numérico converge fortemente no instante T , com ordem α , se

$$\epsilon(\Delta) = \mathbb{E} [|X^N - X_T|] \leq M \Delta^\alpha, \quad (6.3.49)$$

onde $M > 0$ é uma constante e Δ é o passo de discretização $\left(\frac{T}{N-1}\right)$. Se a desigualdade acima fosse aproximadamente uma igualdade, um resultado plausível para espaçamentos de tempo suficientemente pequenos, poderíamos tomar o logaritmo dos dois lados da equação de modo a chegarmos a uma relação linear:

$$\log(\epsilon(\Delta)) \approx \log M + \alpha \log \Delta. \quad (6.3.50)$$

Os experimentos foram baseados em 50000 trajetórias do movimento Browniano e, para cada uma delas, consideramos 7 passos de discretização distintos. Nestes 7 casos, consideramos a média aritmética da diferença em módulo entre o valor obtido pelos esquemas numéricos no instante final e a solução explícita do problema ao longo das 50000 simulações. Se a aproximação (6.3.50) for razoável, ao plotarmos $\log(\epsilon(\Delta)) \times \Delta$, deveríamos esperar uma reta cuja inclinação indica a ordem de convergência forte do algoritmo.

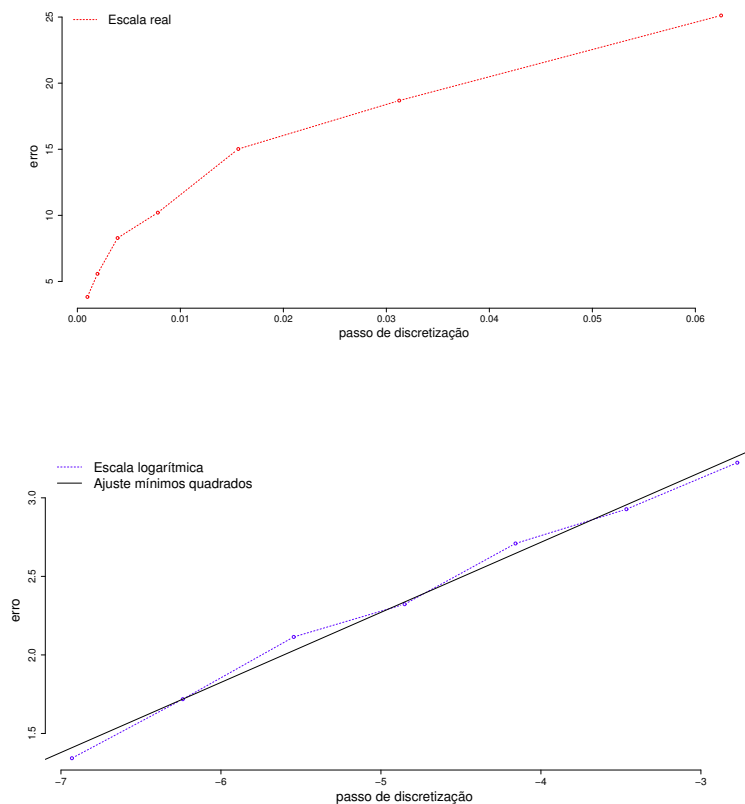
Passo de discretização		Erro esperado	
Δ	$\log \Delta$	$\epsilon(\Delta)$	$\log(\epsilon(\Delta))$
2^{-11}	-7.62	2.53	0.93
2^{-10}	-6.93	3.49	1.25
2^{-9}	-6.23	4.95	1.60
2^{-8}	-5.55	7.71	2.04
2^{-7}	-4.85	10.20	2.32
2^{-6}	-4.15	15.82	2.76
2^{-5}	-3.47	22.74	3.12

Tabela 6.1: Esquema de Euler-Maruyama avaliado no intervalo $[0, 1]$.

Comparando os resultados em escala real e logarítmica, percebemos que a relação (6.3.50) é de fato razoável. A partir da técnica dos mínimos quadrados, chegamos à seguinte relação:

$$\log(\epsilon(\Delta)) = 0.53 \log \Delta + 4.96, \quad (6.3.51)$$

o que indica uma ordem de convergência forte $\alpha \approx 0.53$, confirmando assim nossas expectativas quanto ao método de Euler-Maruyama. A reta correspondente à relação anterior é ilustrada na Figura 6.1.

Figura 6.1: Esquema de Euler-Maruyama avaliado no intervalo $[0, 1]$.

Resultados similares são obtidos quando consideramos intervalos de tempo mais longos, no entanto, este tipo de simulação é mais custosa pois precisamos considerar uma quantidade N grande de pontos para manter os mesmos espaçamentos do experimento anterior. O ponto mais importante dos testes com intervalos mais longos é a verificação de que a constante M cresce bastante junto com o intervalo. Esta observação deve ser levada em consideração pois deixa a análise de convergência forte sem significado à medida que $T \rightarrow +\infty$. Perceba ainda que os erros crescem arbitrariamente mas acabam sendo compensados pelo crescimento ilimitado da constante M . A tabela abaixo mostra que a relação linear não apresentou um ajuste

Passo de discretização		Erro esperado - $\log(\epsilon(\Delta))$	
Δ	$\log \Delta$	$T = 5$	$T = 10$
2^{-11}	-7.62	17.34	34.85
2^{-10}	-6.93	17.75	35.69
2^{-9}	-6.23	18.18	36.49
2^{-8}	-5.55	18.19	36.89
2^{-7}	-4.85	18.30	36.99
2^{-6}	-4.15	18.99	37.07
2^{-5}	-3.47	19.24	37.23

Tabela 6.2: Esquema de Euler-Maruyama avaliado nos intervalos $[0, 5]$ e $[0, 10]$.

satisfatório, no entanto o modelo de mínimos quadrados continua evidenciando uma ordem de convergência forte da ordem de $\alpha \approx 0.5$. Para $T = 5$ e $T = 10$ os ajustes lineares obtidos foram:

$$\log(\epsilon(\Delta)) = 0.55 \log \Delta + 39.43, \quad (6.3.52)$$

e

$$\log(\epsilon(\Delta)) = 0.54 \log \Delta + 42.02, \quad (6.3.53)$$

respectivamente.

Aplicando o mesmo procedimento ao método θ estocástico, devemos observar um comportamento similar ao esquema de Euler-Maruyama já que ambos possuem ordem de convergência forte $\alpha = 0.5$. De fato, considerando $\theta = \frac{1}{2}$ (Euler-trapezoidal), obtivemos resultados praticamente semelhantes aos do experimento anterior. Os co-

Passo de discretização		Erro esperado - $\log(\epsilon(\Delta))$		
Δ	$\log \Delta$	$T = 1$	$T = 5$	$T = 10$
2^{-11}	-7.62	0.93	17.36	34.16
2^{-10}	-6.93	1.28	17.82	35.23
2^{-9}	-6.23	1.64	18.57	36.24
2^{-8}	-5.55	2.09	18.70	36.73
2^{-7}	-4.85	2.42	19.05	36.78
2^{-6}	-4.15	2.92	20.18	36.17
2^{-5}	-3.47	3.43	20.41	37.80

Tabela 6.3: Esquema Euler-trapezoidal avaliado nos intervalos $[0, 1]$, $[0, 5]$ e $[0, 10]$.

eficientes angulares correspondentes aos ajustes de mínimos quadrados foram um pouco superiores, porém, a diferença não foi significativa conforme ilustrado na Figura 6.2.

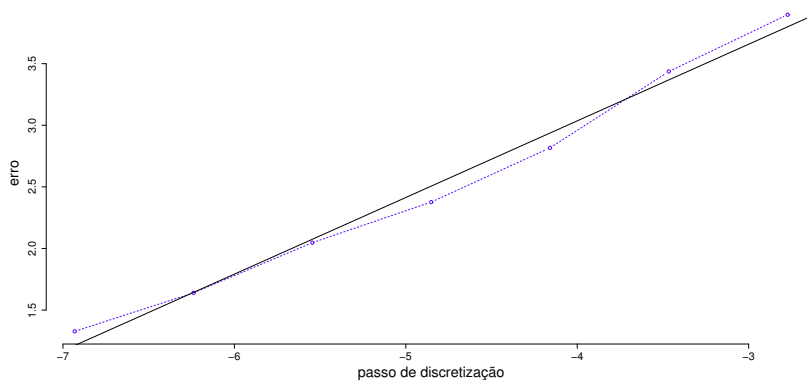


Figura 6.2: Esquema Euler-trapezoidal avaliado no intervalo $[0, 1]$.

	$T = 1$	$T = 5$	$T = 10$
$\log M$	5.41	22.97	39.96
α	0.60	0.74	0.69

Tabela 6.4: Ajustes lineares para avaliação numérica do esquema Euler-trapezoidal.

O resultado se repete quando consideramos $\theta = 1$ (Euler-implícito). Apresentamos inicialmente os resultados obtidos pelo ajuste de mínimos quadrados (ver Figura 6.3), os quais evidenciam uma ordem de convergência forte similar às dos dois esquemas anteriores.

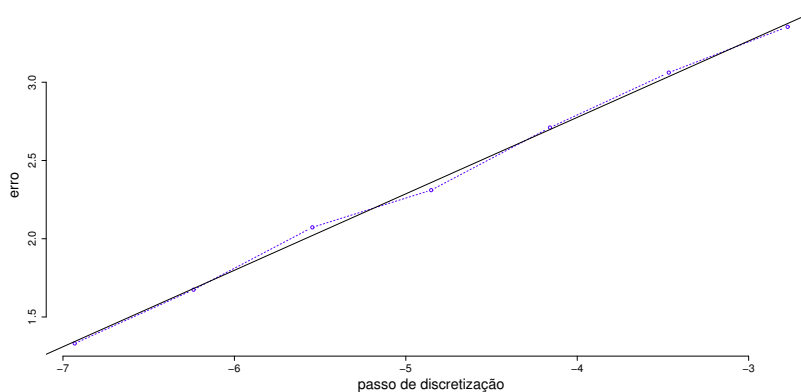


Figura 6.3: Esquema Euler-implícito avaliado no intervalo $[0, 1]$.

	$T = 1$	$T = 5$	$T = 10$
$\log M$	5.03	21.08	39.60
α	0.54	0.48	0.59

Tabela 6.5: Ajustes lineares para avaliação numérica do esquema Euler-implícito.

Conforme podemos acompanhar pela tabela abaixo, a semelhança entre os desempenhos dos métodos é ainda mais evidente quando analisamos mais minuciosamente os erros médios obtidos para cada passo de discretização.

Passo de discretização		Erro esperado $-\log(\epsilon(\Delta))$		
Δ	$\log \Delta$	$T = 1$	$T = 5$	$T = 10$
2^{-11}	-7.62	0.93	17.34	34.56
2^{-10}	-6.93	1.26	17.78	35.49
2^{-9}	-6.23	1.61	18.38	36.38
2^{-8}	-5.55	2.05	18.42	36.82
2^{-7}	-4.85	2.33	18.50	36.91
2^{-6}	-4.15	2.78	19.46	36.85
2^{-5}	-3.47	3.19	19.26	37.28

Tabela 6.6: Esquema Euler-implícito avaliado nos intervalos $[0, 1]$, $[0, 5]$ e $[0, 10]$.

Os próximos resultados correspondem aos algoritmos de Milstein e livre de derivada. Assim como no conjunto de testes anterior, devemos esperar um desempenho similar dos esquemas de Milstein (Figura 6.4) e livre de derivada (Figura 6.5), uma vez que o segundo é uma modificação simples do primeiro. Para não tornar a análise repetitiva, apresentamos apenas os resultados obtidos para os intervalos $[0, 1]$ e $[0, 5]$.

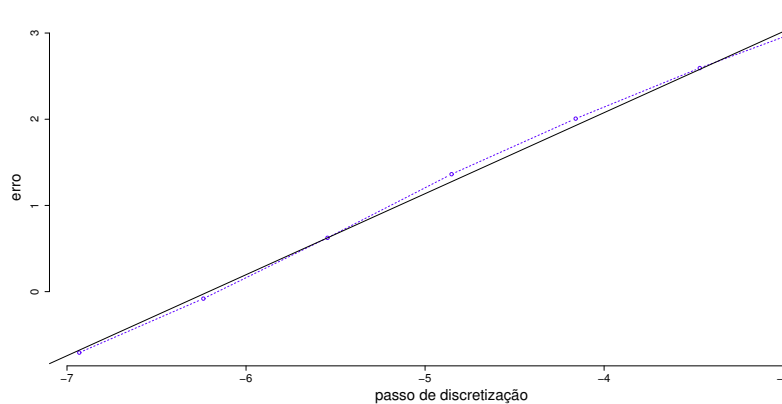
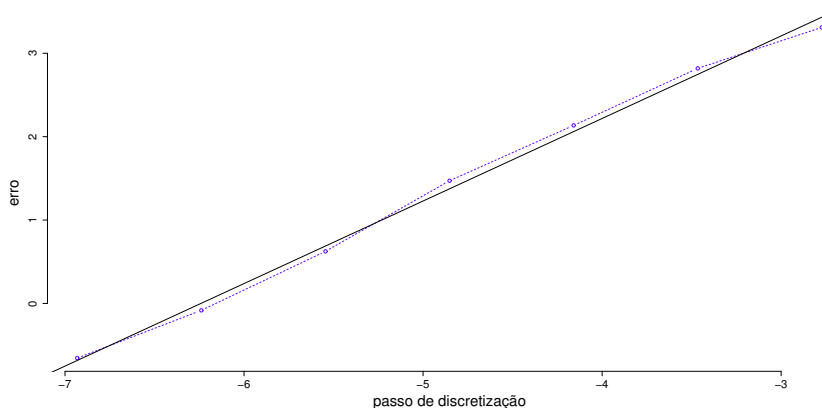


Figura 6.4: Esquema de Milstein avaliado no intervalo $[0, 1]$.

	$T = 1$	$T = 5$
$\log M$	6.80	22.10
α	1.08	0.88

Tabela 6.7: Ajustes lineares para avaliação numérica do esquema de Milstein.

Passo de discretização		Erro esperado $-\log(\epsilon(\Delta))$	
Δ	$\log \Delta$	$T = 1$	$T = 5$
2^{-11}	-7.62	-1.50	15.26
2^{-10}	-6.93	-0.81	16.13
2^{-9}	-6.23	-0.12	16.73
2^{-8}	-5.55	1.38	17.35
2^{-7}	-4.85	1.60	17.90
2^{-6}	-4.15	2.30	18.48
2^{-5}	-3.47	2.84	18.95

Tabela 6.8: Esquema de Milstein avaliado nos intervalos $[0, 1]$ e $[0, 5]$.Figura 6.5: Esquema livre de derivada avaliado no intervalo $[0, 1]$.

	$T = 1$	$T = 5$
$\log M$	6.85	22.26
α	1.07	0.88

Tabela 6.9: Ajustes lineares para avaliação numérica do esquema livre de derivada.

Passo de discretização		Erro esperado $-\log(\epsilon(\Delta))$	
Δ	$\log \Delta$	$T = 1$	$T = 5$
2^{-11}	-7.62	-1.40	15.29
2^{-10}	-6.93	-0.68	16.27
2^{-9}	-6.23	0.02	16.99
2^{-8}	-5.55	1.41	17.50
2^{-7}	-4.85	1.68	18.05
2^{-6}	-4.15	2.38	18.65
2^{-5}	-3.47	2.93	19.03

Tabela 6.10: Esquema livre de derivada avaliado nos intervalos $[0, 1]$ e $[0, 5]$.

Assim como na análise teórica da estabilidade dos métodos numéricos, a aplicação

do método da linearização local a este problema não faz sentido pois a simulação do método coincide exatamente com a simulação da solução real. Neste caso, os erros médios de aproximação serão todos nulos. A avaliação numérica da convergência forte deste método é um pouco mais delicada pois precisamos conhecer a solução explícita do problema de teste para efetuarmos as comparações de erro e grande parte dos problemas cuja solução explícita é conhecida se reduzem a casos lineares triviais após a aplicação da transformada de Lamperti (ponto crucial para a aplicação do método da linearização local).

6.3.2 Estabilidade

O próximo conjunto de testes evidencia empiricamente as regiões de estabilidade obtidas para cada um dos esquemas numéricos. O procedimento é simples:

passo 1: determinar uma tripla (λ, μ, Δ) ;

passo 2: analisar o comportamento médio do esquema ao longo de 10^5 repetições;

passo 3: estudar os resultados; se a tripla pertence à região de estabilidade, devemos esperar uma trajetória convergindo para zero.

Consideramos inicialmente os parâmetros $\lambda = -3$, $\mu = \sqrt{3}$ com a condição inicial $x_0 = 10$. Note que

$$2\lambda + \mu^2 = 2(-3) + 3 = -3 < 0, \quad (6.3.54)$$

o que mostra que estamos na região de estabilidade do problema inicial. Começando pelo método de Euler-Maruyama, temos que a região de estabilidade será dada por

$$\mathcal{R}_{em} = \left\{ \Delta \in \mathbb{R}^+; \Delta < \frac{1}{3} \right\}. \quad (6.3.55)$$

Assim, testando os passos de discretização $\Delta = 1, \Delta = 0.5, \Delta = 0.25$, devemos esperar que apenas no último caso a solução convirja em média quadrática para zero. Este comportamento foi de fato reproduzido nos experimentos e é ilustrado a seguir. Para melhorar a visualização dos resultados, plotamos o eixo das ordenadas em escala \log_{10} , portanto, devemos esperar que a curva vá para $-\infty$ e não para zero. Ainda com relação a este mesmo caso de teste, é interessante verificarmos os comportamentos dos métodos Euler-trapezoidal e Euler-implícito. Como vimos anteriormente, estes esquemas são incondicionalmente estáveis, logo, podemos esperar que as três curvas geradas por este experimento tendam a $-\infty$ conforme de fato verificado nas Figuras 6.6(b)-(c).

A princípio, a Figura 6.6(a) pode nos conduzir equivocadamente à conclusão de que a evolução do esquema trapezoidal permanece estagnada no caso $\Delta = 1$. Na verdade, a convergência neste caso é simplesmente mais lenta que as demais e esta diferença fica acentuada devido à escala do gráfico. Analisando mais detalhadamente os resultados obtidos por este método, a convergência fica mais visível, como mostra a tabela a seguir.

	$t = 0$	$t = 25$	$t = 50$	$t = 75$	$t = 100$	$t = 125$	$t = 150$
$\mathbb{E}X_t^2$	100	58.94	2.92e-4	8.31e-11	5.20e-18	1.99e-25	1.13e-35

Tabela 6.11: Evolução do esquema trapezoidal no intervalo $[0, 150]$.

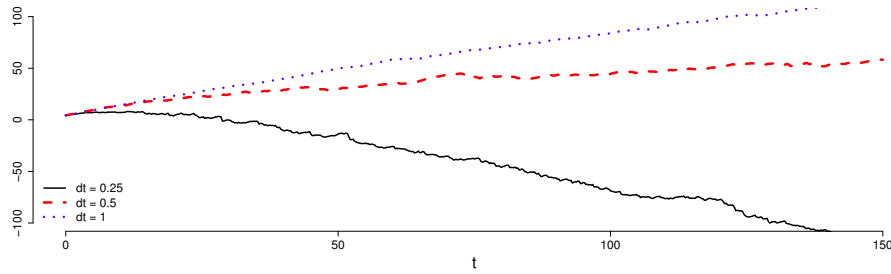
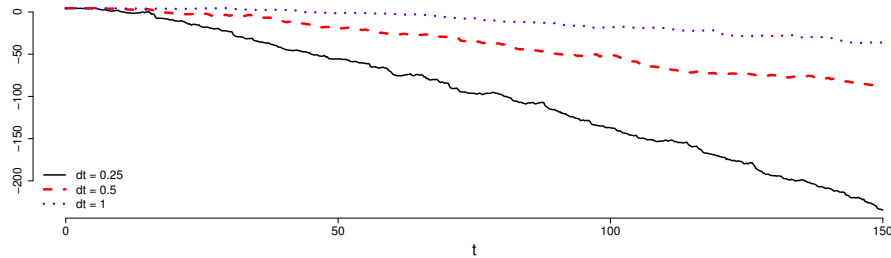
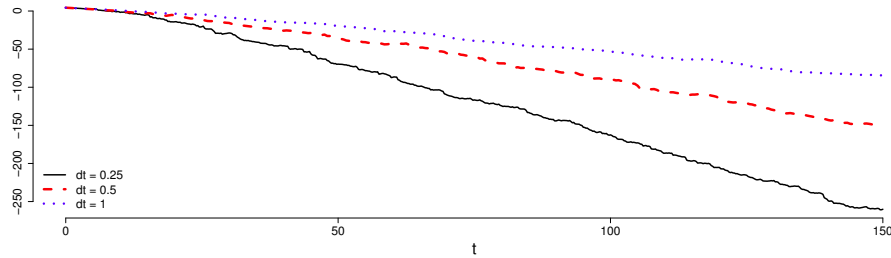
(a) Estabilidade do esquema de Euler-Maruyama ($\theta = 0$).(b) Estabilidade do esquema Euler-trapezoidal ($\theta = 1/2$).(c) Estabilidade do esquema Euler-implícito ($\theta = 1$).

Figura 6.6: Estabilidade do método θ estocástico para diferentes escolhas do parâmetro θ .

As funções de estabilidade obtidas para os esquemas de Milstein e livre de derivada nos indicaram que seus domínios de estabilidade estão contidos no domínio do método de Euler-Maruyama. Mais precisamente, para este conjunto de testes, segue que

$$\mathcal{R}_{ld} = \mathcal{R}_{mils} = \left\{ \Delta \in \mathbb{R}^+; \Delta < \frac{2}{9} \approx 0.22 \right\}. \quad (6.3.56)$$

Neste caso, ao utilizar os passos de discretização das simulações anteriores, devemos esperar que nenhum das três trajetórias apresente o comportamento decrescente. No

entanto, a Figura 6.7 indica que este resultado não foi reproduzido corretamente. Como $0.25 > 0.22$, a trajetória em cor preta na Figura 6.7 acima não deveria

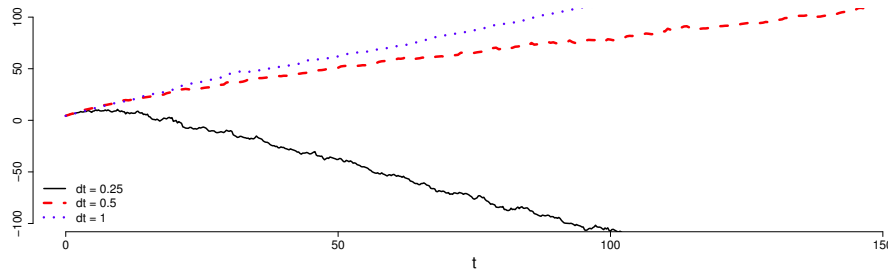


Figura 6.7: Estabilidade do método de Milstein.

apresentar esse comportamento decrescente. No entanto, devemos ressaltar que o passo de discretização $\Delta = 0.25$ está quase no limiar da região de estabilidade, de modo que comportamentos imprevistos sejam aceitáveis. É provável que erros de precisão computacional provenientes de truncamentos ou da imperfeição da geração de números pseudo-aleatórios, influenciem significativamente os métodos em casos limite como este.

Para tentar contornar o problema do último experimento, realizamos uma simulação com $\Delta = \frac{1}{3}$ de modo afastar o passo de discretização um pouco mais dos limiares da região de estabilidade. Em seguida, comparamos o resultado com o método de Euler-Maruyama executado com $\Delta = 0.3$. Neste caso, o comportamento dos métodos foi mais próximo do previsto e a relação $\mathcal{R}_{mils} \subset \mathcal{R}_{em}$ foi evidenciada. Perceba que a curva gerada pelo método de Milstein (Figura 6.8) ainda não apresenta o comportamento de crescimento caracterizado por $|R(\lambda, \mu, \Delta)| > 1$. O mais provável é que devido a erros de aproximação, o algoritmo tenha chegado à relação $|R(\lambda, \mu, \Delta)| \approx 1$. Os resultados obtidos para o esquema livre de derivada são similares aos do método de Milstein portanto serão omitidos.

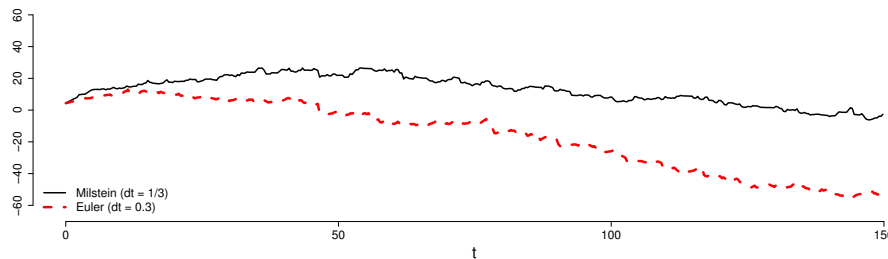


Figura 6.8: Comparação de estabilidade entre os métodos de Milstein e de Euler-Maruyama.

O restante da seção é dedicada a testes envolvendo o método da linearização local. Os testes serão baseados em equações diferenciais estocásticas com coeficientes de deriva linear e polinomial, respectivamente. O procedimento será o seguinte:

passo 1: simulação de um movimento Browniano a ser utilizado como base para as simulações (filtração natural);

passo 2: avaliar três trajetórias dos esquemas numéricos considerando passos de discretização distintos;

passo 3: comparar as trajetórias; quanto mais estável o método, menos sensível ele será quanto a mudança do passo de discretização.

O primeiro caso é dado pela seguinte equação:

$$dX_t = X_t dt + \sqrt{X_t} dW_t, \quad t \in [0, 2], \quad (6.3.57)$$

$$X_0 = 2. \quad (6.3.58)$$

Para a utilização do método da linearização local, devemos primeiramente aplicar a transformada de Lamperti $y = 2\sqrt{x}$, o que nos leva à seguinte equação:

$$dY_t = \frac{1}{2} \left(Y_t dt - \frac{1}{Y_t} \right) + dW_t, \quad t \in [0, 2], \quad (6.3.59)$$

$$Y_0 = 2\sqrt{2}. \quad (6.3.60)$$

Ao final da simulação, o resultado obtido em função do processo Y deve ser transformado de volta para o processo original através da inversão $x = \frac{1}{4}y^2$.

Por se tratar de um problema com coeficiente de deriva linear, devemos esperar que o desempenho dos diferentes algoritmos seja similar. Este experimento leva em consideração os seguintes espaçamentos entre pontos: $\Delta \in \{\frac{1}{64}, \frac{1}{16}, \frac{1}{4}\}$. A Figura 6.9 ilustra primeiramente os resultados obtidos pelo método da linearização local. Perceba que a variação do passo de discretização não é refletida de maneira signifi-

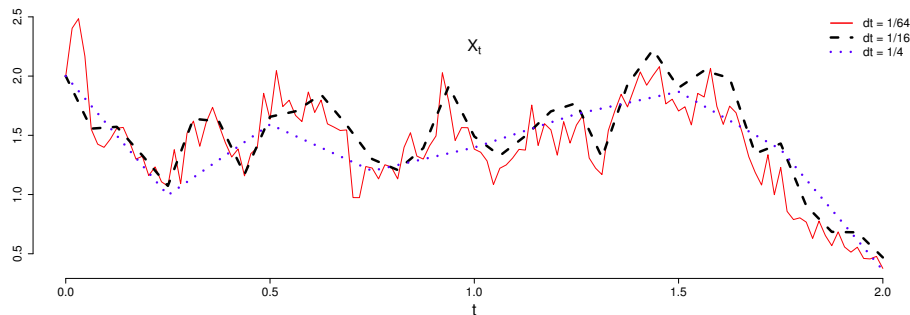


Figura 6.9: Método da linearização local aplicado ao problema (6.3.57)-(6.3.58).

cativa pelas trajetórias. A grande diferença entre as trajetórias se dá simplesmente pela discrepância entre o número de pontos considerados no eixo do tempo em cada uma das simulações. Este é um forte indício da estabilidade do método: variações nas condições iniciais não são propagadas adiante.

Os métodos de Milstein e Euler-Maruyama se comportaram relativamente bem, porém, diferentemente do caso anterior, o método de Milstein demonstrou uma certa

sensibilidade quanto à mudança de $\Delta = \frac{1}{16}$ para $\Delta = \frac{1}{4}$. Já o esquema de Euler-Maruyama superou o de Milstein e não acusou esta sensibilidade significativa quanto à variação dos espaçamentos. Este resultado não é surpreendente pois, conforme o estudo realizado com a equação linear de teste, o método de Euler-Maruyama possui um domínio de estabilidade mais abrangente:

$$\mathcal{R}_{mils} \subset \mathcal{R}_{em}.$$

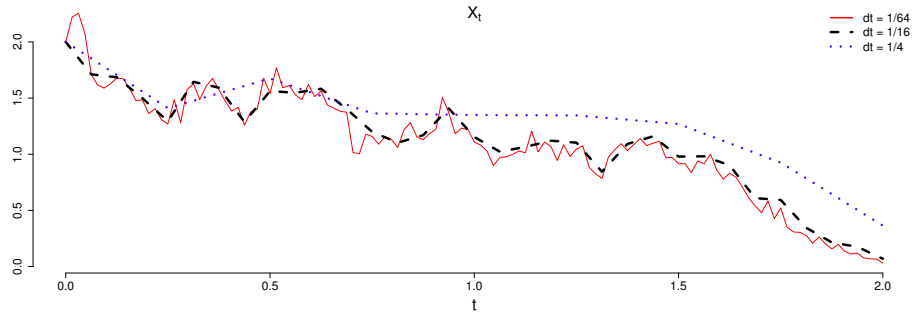


Figura 6.10: Método de Milstein aplicado ao problema (6.3.57)-(6.3.58).

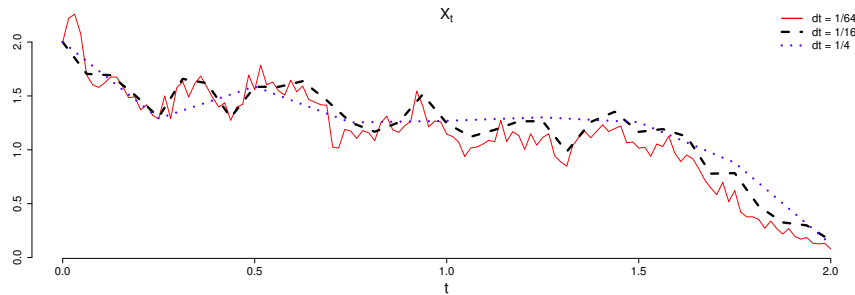


Figura 6.11: Método de Euler-Maruyama aplicado ao problema (6.3.57)-(6.3.58).

Os resultados observados na Figura 6.11 eram esperados já que o método da linearização local não oferece vantagens tão consideráveis quando o coeficiente de deriva é linear. O próximo conjunto de teste tem como objetivo mudar esta característica de modo a evidenciar o ganho de estabilidade do método. Para tal, considere o seguinte problema:

$$dX_t = (5 - 11X_t + 6X_t^2 - X_t^3) dt + dW_t, \quad t \in [0, 10], \quad (6.3.61)$$

$$X_0 = 2. \quad (6.3.62)$$

Como o coeficiente de difusão é constante, não há necessidade de aplicarmos a transformada de Lamperti. Este fator também deixa o sistema naturalmente mais estável, o que nos permitirá lidar com passos de discretização ainda maiores. Por fim, como este coeficiente não depende do processo X , segue que os métodos de Euler-Maruyama e Milstein coincidem, logo, basta implementarmos um deles (o primeiro por simplicidade).

Começamos novamente pelo método da linearização local. Desta vez utilizamos o seguinte conjunto de passos de discretização: $\Delta \in \{\frac{1}{16}, \frac{1}{4}, 1\}$. A Figura 6.12 indica que o ótimo desempenho apresentado no primeiro conjunto de testes não foi repetido, o que já era esperado pelas características do novo problema. No entanto, o comportamento ainda é bastante satisfatório considerando que os passos de discretização são quatro vezes maiores que no caso anterior. Temos assim mais

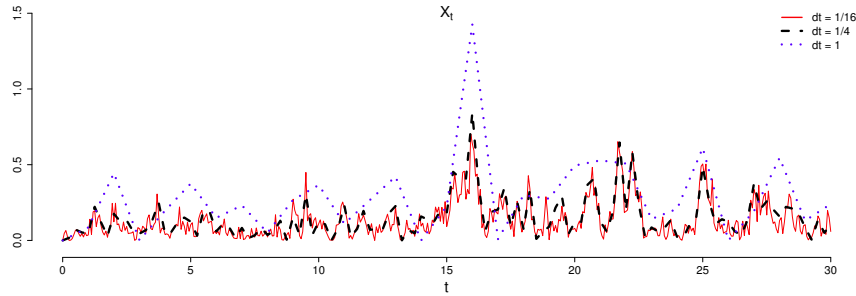


Figura 6.12: Método da linearização local aplicado ao problema (6.3.61)-(6.3.62).

um bom indício do comportamento estável deste esquema. Por outro lado, o método de Euler-Maruyama se mostrou de Euler-Maruyama acusou bastante a mudança do tamanho dos espaçamentos de $\Delta = \frac{1}{16}$ para $\Delta = \frac{1}{4}$ e se mostrou completamente instável com a mudança para $\Delta = 1$. Este é comportamento pode ser perfeitamente observado na Figura 6.13, onde a trajetória em azul pode até mesmo ser desprezada.

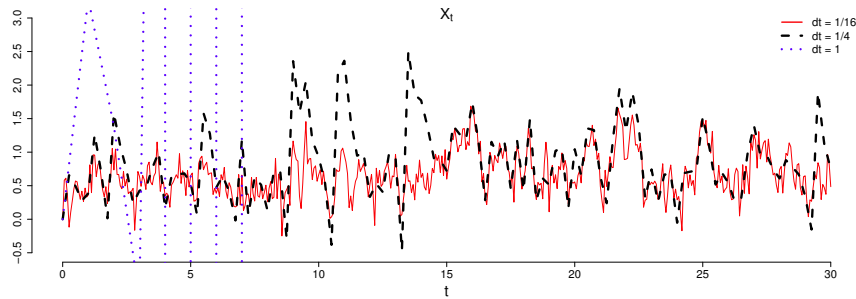


Figura 6.13: Método de Euler-Maruyama aplicado ao problema (6.3.61)-(6.3.62).

6.4 Exercícios

1. Considere o seguinte problema de valor inicial

$$\begin{aligned} dX_t &= X_t(1 - X_t)dt + 2X_t dW_t, & 0 \leq t < 10, \\ X_0 &= 2, \end{aligned}$$

Implemente os métodos de Euler-Maruyama, Milstein e Linearização Local para a construção de aproximações para este problema.

- Simule $N = 100$ trajetórias para cada uma das três metodologias considerando $\Delta = 2^{-4}$.
 - Utilize os resultados do item anterior para estimar intervalos de 95% de confiança para os erros absolutos de cada um dos métodos. Os resultados coincidem com as expectativas construídas a partir da exposição teórica dos métodos.
 - Repita os procedimentos anteriores considerando agora os espaçamentos $\Delta = 2^{-n}$, $n = 5, 6, 7$. Os resultados obtidos apresentam diferenças significativas?
 - Considerando os tempos computacionais para a execução dos experimentos, assim como os intervalos de confiança obtidos, defina qual seria o método mais apropriado para o tratamento deste problema. Comente sua resposta.
2. Considere o seguinte problema de valor inicial

$$\begin{aligned} dX_t &= 4X_t dt + 2X_t dW_t, & 0 \leq t < 1, \\ X_0 &= 1, \end{aligned}$$

o qual foi utilizado para a avaliação numérica dos métodos de discretização. Considere ainda a seguinte quantidade

$$\epsilon^* := |\mathbb{E}X_1^N - \mathbb{E}X_1|. \quad (6.4.63)$$

- Utilizando as propriedades do movimento Browniano geométrico, determine analiticamente a quantidade $\mathbb{E}X_1$.
- Implemente os métodos de Euler-Maruyama, Milstein e Linearização Local para a construção de aproximações para este problema.
- Repita a avaliação numérica de convergência dos três métodos considerando agora o critério ϵ^* . Os resultados apresentam alguma diferença?

(A análise de convergência segundo o critério ϵ^* define o conceito de convergência fraca no contexto numérico das EDEs. Este critério é interessante quando estamos interessados apenas em propriedades a respeito dos valores esperados de funções das soluções, o que nos permite construir soluções até mesmo mais "relaxadas". Um tratamento mais aprofundado de aproximações fracas pode ser obtido em [20].)

Capítulo 7

Simulação Exata

Os métodos de simulação discutidos até aqui eram baseados no conceito de solução forte, isto é, dadas as observações de um movimento Browniano e da condição inicial, trajetórias eram construídas baseadas nestas observações. As aproximações obtidas podiam ser comparadas com uma solução exata condicionada às mesmas observações. Este último procedimento fazia sentido devido à unicidade das trajetórias. Uma outra abordagem seria a simulação de trajetórias baseadas no conceito de solução fraca. Tal metodologia é possível se tomarmos como base de geração de dados a medida induzida pelo processo de interesse ao invés de tomarmos como base trajetórias de um movimento Browniano. Neste caso, não faz sentido compararmos trajetórias de soluções diferentes uma vez que se X e Y são duas soluções fracas de uma mesma equação diferencial estocástica, então estes processos possuem as mesmas distribuições finito-dimensionais, porém, seus caminhos amostrais podem ter características distintas.

Os tópicos apresentados aqui constituem basicamente uma discussão simplificada das inspiradoras obras de Beskos e Roberts [3] e Beskos et al. [2]. Uma discussão mais profunda do tema pode ser obtida nos textos originais.

7.1 Algoritmo

O primeiro passo para a construção do algoritmo de simulação exata é o estudo do método da rejeição, o qual nos permite gerar amostras provenientes de uma variável aleatória com distribuição μ a partir de uma segunda distribuição λ , desde que estas medidas sejam relacionadas por certas condições de regularidade. Tais condições são detalhadas no próximo teorema.

Teorema 7.1. (*Método da Rejeição*) *Seja (Ω, \mathcal{B}) um espaço mensurável e considere duas medidas de probabilidade λ e μ neste espaço, tais que $\mu \ll \lambda$ ($B \in \mathcal{B}, \lambda(B) = 0 \Rightarrow \mu(B) = 0$). Assuma a existência de uma constante $\epsilon > 0$ tal que $f := \epsilon \frac{d\mu}{d\lambda} \leq 1$, λ -q.c., e que é fácil amostrar de λ . Se $(Y_n, I_n)_{n \in \mathbb{N}}$ é uma sequência de variáveis aleatórias independentes e identicamente distribuídas tomando valores em $\Omega \times \{0, 1\}$, $Y_1 \sim \lambda$ e $P[I_1 = 1 | Y_1 = y] = f(y)$, $\forall y \in \Omega$, então, definindo*

$$\tau := \min\{i \in \mathbb{N}; I_i = 1\}, \quad (7.1.1)$$

temos que

$$P[Y_\tau \in B] = \mu(B), \forall B \in \mathcal{B}. \quad (7.1.2)$$

Demonstração: Dado $i \in \mathbb{N}$, temos que

$$\begin{aligned} P[I_i = 1] &= \int_{\Omega} P[I_i = 1 | Y_i = y] d\lambda(y) \\ &= \int_{\Omega} f(y) d\lambda(y) = \int_{\Omega} \epsilon d\mu(y) = \epsilon. \end{aligned} \quad (7.1.3)$$

Além disso, dado $B \in \mathcal{B}$, temos que

$$\begin{aligned} P[Y_{\tau} \in B] &= P[Y_{\tau} \in B, I_1 = 1] + P[Y_{\tau} \in B, I_1 = 0] \\ &= P[Y_{\tau} \in B, I_1 = 1] + P[Y_{\tau} \in B | I_1 = 0] P[I_1 = 0] \\ &= P[Y_{\tau} \in B, I_1 = 1] + P[Y_{\tau} \in B] P[I_1 = 0] \\ &= P[Y_{\tau} \in B, I_1 = 1] + P[Y_{\tau} \in B](1 - \epsilon). \end{aligned} \quad (7.1.4)$$

Na relação acima temos que $P[Y_{\tau} \in B | I_1 = 0] = P[Y_{\tau} \in B]$ devido à independência entre os elementos da família $(Y_n, I_n)_{n \in \mathbb{N}}$. A probabilidade conjunta utilizada acima é equivalente a

$$\begin{aligned} P[Y_{\tau} \in B, I_1 = 1] &= \int_B P[I_1 = 1 | Y_1 = y] d\lambda(y) \\ &= \int_B f(y) d\lambda(y) = \int_B \epsilon d\mu(y) = \epsilon \mu(B). \end{aligned} \quad (7.1.5)$$

Assim,

$$\begin{aligned} P[Y_{\tau} \in B] &= \epsilon \mu(B) + P[Y_{\tau} \in B](1 - \epsilon) \\ \therefore P[Y_{\tau} \in B] &= \mu(B). \end{aligned}$$

■

A representação geral do método da rejeição conforme especificado pelo teorema acima nos permite elaborar um algoritmo capaz de simular soluções exatas para EDEs da forma

$$dX_t = f(X_t) + dW_t, \quad 0 \leq t \leq T, \quad (7.1.6)$$

$$X_0 = x_0. \quad (7.1.7)$$

Neste contexto, solução exata indica que as trajetórias simuladas por tal algoritmo induzem (no espaço das possíveis trajetórias) a mesma medida de probabilidade que uma curva solução da EDE correspondente. Em linhas gerais, nosso objetivo é gerar trajetórias de um determinado processo estocástico cuja simulação é viável, e aceitar algumas dessas trajetórias como solução do problema (7.1.6-7.1.7) a partir do método da rejeição. Ao final temos que o conjunto de trajetórias aceitas é distribuído segundo a medida de probabilidade que governa as soluções do problema.

Buscar um algoritmo desta natureza é plausível uma vez que a ideia de solução fraca nos permite pensar em unicidade de medidas de probabilidade com respeito às soluções de uma EDE. Desta forma, devemos assumir de antemão que o problema acima satisfaz as condições necessárias para a existência e unicidade de uma solução fraca. Um modelo com estas características não fica limitado a modelos com coeficiente de difusão constante pois esta propriedade pode ser obtida através da transformação de Lamperti aplicada em um modelo mais abrangente. De uma forma geral, exigir que o modelo seja governado por um ruído aditivo é uma restrição muito severa.

Para trabalharmos com o método da rejeição, é conveniente pensar nos processos estocásticos envolvidos como medidas induzidas no espaço $C^1([0, T])$. Se ω é um elemento qualquer deste espaço, estaremos interessados em suas funções coordenadas e suas respectivas σ -álgebras mínimas:

$$W_t(\omega) := \omega(t), \quad t \in [0, T]; \quad \mathcal{C} := \sigma(\{W_t; 0 \leq t \leq T\}). \quad (7.1.8)$$

Assim, denotemos por \mathbb{Q} a medida de probabilidade induzida em (C, \mathcal{C}) pelo processo $X = \{X_t; 0 \leq t \leq T\}$ o qual satisfaz o problema (7.1.6)-(7.1.7). Consideremos ainda que \mathbb{W} é a medida de Wiener e suponha que o coeficiente f satisfaz a condição de Novikov:

$$\mathbb{E}_{\mathbb{W}} \left[\exp \left\{ \frac{1}{2} \int_0^T f^2(W_t) dt \right\} \right] < +\infty. \quad (7.1.9)$$

Neste caso, a derivada de Radon-Nykodim $\frac{d\mathbb{Q}}{d\mathbb{W}}$ pode ser obtida a partir da transformação de Cameron-Martin-Girsanov:

$$\frac{d\mathbb{Q}}{d\mathbb{W}}(\omega) = \exp \left\{ \int_0^T f(W_t) dW_t - \frac{1}{2} \int_0^T f^2(W_t) dt \right\} =: G(W). \quad (7.1.10)$$

Devido à dificuldade de obter uma expressão fechada para a solução da integral estocástica acima, aplicamos a fórmula de Itô de modo a simplificar $G(B)$. Para tal, basta assumir que o coeficiente $f(\cdot, \cdot)$ é diferenciável. Neste caso, definindo

$$F(W_t) := \int_0^b f(x) dx \Big|_{b=W_t} \quad (7.1.11)$$

e aplicando a fórmula de Itô:

$$\begin{aligned} dF_t &= \frac{\partial}{\partial t} \int_0^b f(x) dx \Big|_{b=W_t} + \left(\frac{\partial}{\partial x} \int_0^b f(x) dx \Big|_{b=W_t} \right) dW_t \\ &\quad + \frac{1}{2} \left(\frac{\partial^2}{\partial x^2} \int_0^b f(x) dx \Big|_{b=W_t} \right) (dW_t)^2 \\ &= 0 + f(W_t) dW_t + \frac{1}{2} f'(W_t) dt. \end{aligned} \quad (7.1.12)$$

Integrando a expressão acima em $[0, T]$, obtemos

$$F(W_t) - F(W_0) = \int_0^T f(W_t) dW_t + \frac{1}{2} \int_0^T f'(W_t) dt. \quad (7.1.13)$$

Assim,

$$\int_0^T f(W_t) dW_t + F(W_t) - F(W_0) - \frac{1}{2} \int_0^T f'(W_t) dt, \quad (7.1.14)$$

e a densidade G será então expressa por

$$G(W) = \exp \left\{ F(W_t) - F(W_0) - \frac{1}{2} \int_0^T (f^2(W_t) + f'(W_t)) dt \right\}. \quad (7.1.15)$$

O método da rejeição exige que $G(\cdot)$ seja limitada quase certamente e esta restrição está ligada à limitação de $F(\cdot)$. Este problema pode ser muito restritivo em

diversos casos, o que nos leva à busca por alternativas mais viáveis. Uma abordagem mais flexível pode ser construída a partir da introdução de uma terceira medida de probabilidade, a qual será utilizada para gerar os candidatos a amostra para o método da rejeição. Esta nova medida diz respeito a um importante processo estocástico conhecido como *ponte Browniana*.

Uma ponte Browniana pode ser definida informalmente como um caso particular do processo de Wiener padrão "forçado" a passar por um ponto específico ρ em seu instante final de observação T . Se $W = \{W_t; 0 \leq t \leq T\}$ é um processo de Wiener padrão iniciado na origem, então uma ponte Browniana passando pelo ponto ρ no instante T pode ser definida a partir de W como

$$PB_t(\omega) = W_t(\omega) - \frac{t}{T} \cdot (W_T(\omega) - \rho), \quad t \in [0, T]. \quad (7.1.16)$$

A Figura 7.1 evidencia as diferenças esperadas com respeito às trajetórias de um movimento Browniano e de uma ponte Browniana.

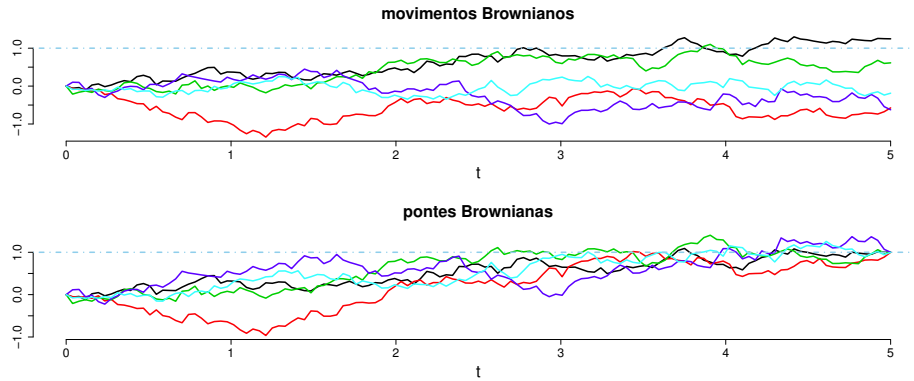


Figura 7.1: Cinco realizações de movimentos Brownianos no intervalo $[0, 5]$ e suas respectivas pontes com $\rho = 1$.

Para a formulação geral do algoritmo de simulação exata consideraremos pontes Brownianas $PB = \{PB_t; 0 \leq t \leq T\}$ heurísticamente definidas por $W | W_T = \rho$, onde ρ é uma variável aleatória distribuída de acordo com alguma função densidade de probabilidade $h : \mathbb{R} \rightarrow [0, +\infty)$ com respeito à medida de Lebesgue.

Teorema 7.2. *Sejam \mathbb{Z} a medida de probabilidade induzida pela ponte Browniana*

$$PB = \{PB_t; 0 \leq t \leq T\}$$

em (C, \mathbb{C}) , $X_0 = x_0$ a condição inicial do processo X e h a densidade associada à variável ρ . Se

$$\chi_h = \{x \in \mathbb{R}; h(x) > 0\} = \mathbb{R}, \quad (7.1.17)$$

então \mathbb{Z} é equivalente a \mathbb{W} e

$$\frac{d\mathbb{Z}}{d\mathbb{W}}(\omega) = \frac{h(W_T)}{\left(1/\sqrt{2\pi T}\right) \exp\{-(W_T - x_0)^2/(2T)\}}. \quad (7.1.18)$$

Demonstração: Defina

$$f(\omega) := h(W_T) \sqrt{2\pi T} / \exp\{-(W_T - x_0)^2/(2T)\}, \quad \omega \in C, \quad (7.1.19)$$

de modo que f seja $\sigma(W_T)$ -mensurável. Dado $A \in \mathcal{C}$, segue, da definição do processo PB , que

$$\mathbb{Z}[A|\sigma(W_T)] = \mathbb{W}[A|\sigma(W_T)] =: g(W_T), \quad \mathbb{W} - q.c., \quad (7.1.20)$$

para alguma função $g : \mathbb{R} \rightarrow \mathbb{R}$, Borel-mensurável. Uma vez que $W_T \sim N(x_0, T)$, com respeito a \mathbb{W} , temos que

$$\begin{aligned} \mathbb{E}_{\mathbb{W}} &= \mathbb{E}_{\mathbb{W}} [\mathbb{E}_{\mathbb{W}} [\mathbf{1}_A f | \sigma(W_T)]] \\ &= \mathbb{E}_{\mathbb{W}} [f \cdot \mathbb{W}[A|\sigma(W_T)]] = \mathbb{E}_{\mathbb{W}}[fg] = \int_{\mathbb{R}} f(u)g(u)\mathbb{W}(du) \\ &= \int_{\mathbb{R}} \frac{h(u)\sqrt{2\pi T}}{\exp\{-(u-x_0)^2/(2T)\}} \cdot \frac{\exp\{-(u-x_0)^2/(2T)\}}{\sqrt{2\pi T}} g(u)du = \int_{\mathbb{R}} h(u)g(u)du. \end{aligned} \quad (7.1.21)$$

Concluimos assim que

$$f = d\mathbb{Z}/d\mathbb{W}. \quad (7.1.22)$$

Além disso, como o suporte de h é a reta, temos que $f > 0$. Desta forma, \mathbb{W} é equivalente a \mathbb{Z} e

$$1/f = d\mathbb{W}/d\mathbb{Z}. \quad (7.1.23)$$

■

A partir dos resultados apresentados acima, fica claro que

$$\begin{aligned} \frac{d\mathbb{Q}}{d\mathbb{Z}}(\omega) &= \frac{d\mathbb{Q}}{d\mathbb{W}}(\omega) \frac{d\mathbb{W}}{d\mathbb{Z}}(\omega) \\ &= \frac{\exp\left\{F(W_T) - F(W_0) - \frac{1}{2} \int_0^T (f^2(W_t) + f'(W_t)) dt\right\}}{h(W_T)\exp\{-(W_T - x_0)^2/(2T)\}} \Big/ \left(1/\sqrt{2\pi T}\right) \\ &= \frac{\exp\left\{F(W_T) - F(W_0) - \frac{(W_T - x_0)^2}{2T} - \frac{1}{2} \int_0^T (f^2(W_t) + f'(W_t)) dt\right\}}{h(W_T)/\sqrt{2\pi T}}. \end{aligned} \quad (7.1.24)$$

Assim, desconsiderando todos os termos que independem da realização ω , obtemos

$$\frac{d\mathbb{Q}}{d\mathbb{Z}}(\omega) \propto \frac{\exp\left\{F(W_T) - \frac{(W_T - x_0)^2}{2T} - \frac{1}{2} \int_0^T (f^2(W_t) + f'(W_t)) dt\right\}}{h(W_T)}. \quad (7.1.25)$$

Assumindo que

$$\int_{\mathbb{R}} \exp\{F(u) - (u - x_0)^2/2T\} du := c < +\infty, \quad (7.1.26)$$

podemos definir (convenientemente) a função densidade de probabilidade

$$h(u) := \exp\{F(u) - (u - x_0)^2/2T\}/c, \quad (7.1.27)$$

de modo a obter

$$\frac{d\mathbb{Q}}{d\mathbb{Z}}(\omega) \propto \exp\left\{-\frac{1}{2} \int_0^T (f^2(W_t) + f'(W_t)) dt\right\}. \quad (7.1.28)$$

Assumindo agora que o funcional em (7.1.28) é limitado inferiormente, isto é, supondo a existência de uma constante $K \in \mathbb{R}$ tal que $K \leq \frac{f^2(u) + f'(u)}{2}$, $\forall u \in \mathbb{R}$, podemos reescrever (7.1.28) como

$$\frac{d\mathbb{Q}}{d\mathbb{Z}} \propto \exp \left\{ - \int_0^T \phi(W_t) dt \right\} \leq 1, \quad \mathbb{Z} - q.c., \quad (7.1.29)$$

onde $\phi > 0$ é dada por

$$\phi(u) = \frac{f^2(u) + f'(u)}{2} - K, \quad \forall u \in \mathbb{R}. \quad (7.1.30)$$

Segue abaixo um resumo das hipóteses utilizadas para a construção da densidade (7.1.29):

1. o coeficiente f é limitado;
2. a função $\exp\{F(u) - (u - x_0)^2/(2T)\}$, $u \in \mathbb{R}$, para $F(u) = \int_0^u a(x)dx$, é integrável;
3. a função $(f^2 + f')/2$ é limitada inferiormente.

A derivada de Radon-Nykodim de \mathbb{Q} com respeito à medida \mathbb{Z} induzida pela ponte Browniana em conjunto com o método da rejeição, nos fornecem as bases para a construção do algoritmo de simulação exata. Basicamente, o procedimento consistirá em gerar trajetórias de pontes Brownianas cujas observações finais são distribuídas de acordo com a densidade h . Notemos que a realização desta etapa demanda um amostrador de variáveis aleatórias governadas por h , o qual pode ser obtido através de versões usuais do método da rejeição (Exercícios 1 e 2). Para concluir o algoritmo, falta apenas estabelecer um dispositivo que nos permita aceitar tais trajetórias segundo a distribuição de probabilidade estabelecida em (7.1.29).

A fim de facilitar o desenvolvimento de tal dispositivo, Beskos et al. [2] acrescentaram às propriedades acima a seguinte restrição:

4. ϕ é limitada superiormente;

O algoritmo de simulação exata desenvolvido a partir do acréscimo desta hipótese foi denominado EA1. No mesmo trabalho, o autor desenvolve uma versão mais abrangente onde a hipótese de limitação de ϕ é substituída por

$$\limsup_{u \rightarrow +\infty} \phi(u) < +\infty \quad \text{ou} \quad \limsup_{u \rightarrow -\infty} \phi(u) < +\infty.$$

Esta generalização foi denominada algoritmo EA2 e não será abordada neste trabalho.

Teorema 7.3. *Sejam $\omega \in C([0, T])$ e $M = M(\omega)$ uma cota superior para a função $t \mapsto \phi(\omega_t)$, $t \in [0, T]$. Se Φ é um processo de Poisson homogêneo de intensidade $T \cdot M$ em $[0, T] \times [0, M(\omega)]$ e N é número de pontos de Φ localizados abaixo do gráfico $\{(t, \phi(W_t)); t \in [0, T]\}$, então*

$$P[N = 0 | \omega] \propto \exp \left\{ - \int_0^T \phi(W_t) dt \right\}. \quad (7.1.31)$$

Demonstração: Dada uma área $A \in [0, T] \times [0, M(\omega)]$, temos que

$$P(N = K|A) = \frac{e^{-T \cdot M \cdot P(A)} [T \cdot M \cdot P(A)]^K}{K!}, \quad (7.1.32)$$

onde N é o número de pontos de Φ que residem em A , pois Φ é um processo de Poisson homogêneo com intensidade $\lambda = T \cdot M$. Em particular, se A é a região delimitada pelo gráfico $\{(t, \phi(W_t)); t \in [0, T]\}$, então

$$P(A) = \frac{1}{T \cdot M} \int_0^T \phi(W_t) dt, \quad (7.1.33)$$

e

$$\begin{aligned} P(N = 0|\omega) &= \frac{\exp\left\{-\int_0^T \phi(W_t) dt\right\} \left(\int_0^T \phi(W_t) dt\right)^0}{0!} \\ &= \exp\left\{-\int_0^T \phi(W_t) dt\right\}. \end{aligned} \quad (7.1.34)$$

■

O teorema acima nos permite aplicar a método da rejeição baseado em observações discretas das trajetórias de X . Dada uma trajetória $\omega \sim \mathbb{Z}$ e uma realização p do processo de Poisson homogêneo Φ em $A \in [0, T] \times [0, M(\omega)]$, basta contarmos o número de pontos de p observados abaixo do gráfico de ω . Se este número for positivo, rejeitamos a trajetória; caso contrário, aceitamos. Formalmente, obtemos o algoritmo abaixo.

Algoritmo de Simulação Exata

1. Simule uma realização

$$\{(x_{i,1}, x_{i,2})\}_{i=1}^T$$

do processo de Poisson Φ em $[0, T] \times [0, M]$;

2. Simule um esqueleto $\omega \sim \mathbb{Z}$ nos instantes

$$\{x_{1,1}, x_{2,1}, \dots, x_{\tau,1}\}$$

e avalie N ;

3. Se $N = 0$, vá para 4, senão, volte para 1;
4. Retorne o esqueleto atual $S(\omega)$ de ω .

Um dos pontos curiosos do algoritmo de simulação exata é o fato de não determinarmos de antemão os instantes de tempo onde esqueleto da trajetória será simulado. Uma alternativa para contornar esta situação seria uma abordagem do tipo "força bruta", onde o domínio de interesse seria refinado em pequenos intervalos e algoritmo seria aplicado diversas vezes. Nesta abordagem, cada ponto simulado para o extremo direito do intervalo seria utilizado como condição inicial para a próxima simulação, e assim por diante. Não é necessário uma análise profunda para concluirmos que este tipo de procedimento nos levaria a um algoritmo

ineficiente. A solução proposta pelos autores do algoritmo é mais simples, no entanto, não garante que os pontos fora do esqueleto sejam exatamente distribuídos com a medida de interesse. Tal solução consiste em preencher com pontes Brownianas os espaços do esqueleto obtido com uma iteração. Apesar da simplicidade, esta alternativa é capaz de gerar boas aproximações.

7.2 Avaliação Numérica

O algoritmo de simulação exata tem várias aplicações, sobretudo na área de inferência estatística. Neste trabalho, utilizaremos o algoritmo EA1 para avaliar, sob um ponto de vista distinto, a eficácia dos procedimentos de discretização discutidos nos capítulos anteriores. Outro ponto importante será a avaliação da eficiência deste algoritmo. Como veremos nos testes a seguir, o tempo de execução é surpreendentemente superior ao dos esquemas discutidos até agora.

O tipo de comparação a ser realizada aqui é diferente das anteriores pois não temos uma observação prévia do movimento Browniano de modo a gerar trajetórias adaptadas a uma única filtração. Em outras palavras, não faz sentido compararmos trajetórias. Nosso objetivo será então comparar as distribuições de probabilidade das soluções obtidas através dos métodos de discretização, em instantes de tempo específico. Este tipo de comparação é interessante pois quanto melhor for a aproximação oferecida por um esquema, mais próximas as distribuições em pontos específicos da trajetória discretizada devem estar das distribuições exatas da solução nestes instantes.

Nosso problema de teste será a seguinte equação:

$$dX_t = \text{sen}(X_t) + dW_t, \quad 0 \leq t \leq T, \quad (7.2.35)$$

$$X_0 = 0. \quad (7.2.36)$$

Antes de aplicar o algoritmo devemos checar se o coeficiente de deriva $f(x) = \text{sen}(x)$ satisfaz as condições de regularidade necessárias.

1. f é limitada e diferenciável;

Não há o que mostrar. Sabemos que $f(x) = \text{sen}(x) \leq 1$ e f é diferenciável com $f'(x) = \text{cos}(x)$.

2. $\exp\left\{F(u) - \frac{(x_0 - u)^2}{2T}\right\}$, $u \in \mathbb{R}$ é integrável, onde $F(u) := \int_0^u f(x)dx$;

Da definição do problema:

$$F(u) = \int_0^u \text{sen}(x)dx = 1 - \text{cos}(u). \quad (7.2.37)$$

Como $1 - \text{cos}(x) \geq 0$, $\forall x \in \mathbb{R}$, segue que

$$\begin{aligned} \exp\left\{F(u) - \frac{(x_0 - u)^2}{2T}\right\} &= \exp\left\{1 - \text{cos}(u) - \frac{u^2}{2T}\right\} \\ &\leq e \cdot \exp\left\{\frac{u^2}{2T}\right\}. \end{aligned} \quad (7.2.38)$$

Podemos concluir que esta função é integrável pois trata-se de uma constante multiplicada pelo núcleo Gaussiano $N(0, T)$.

3. a função $\frac{1}{2}(f^2 + f')$ é limitada inferiormente;
Basta observar que

$$\frac{1}{2}(f^2(x) + f'(x)) = \frac{1}{2}(\text{sen}^2(x) + \cos(x)) \geq -\frac{1}{2}. \quad (7.2.39)$$

Verificadas as três condições acima, podemos utilizar o problema (7.2.35)-(7.2.36) como caso de teste.

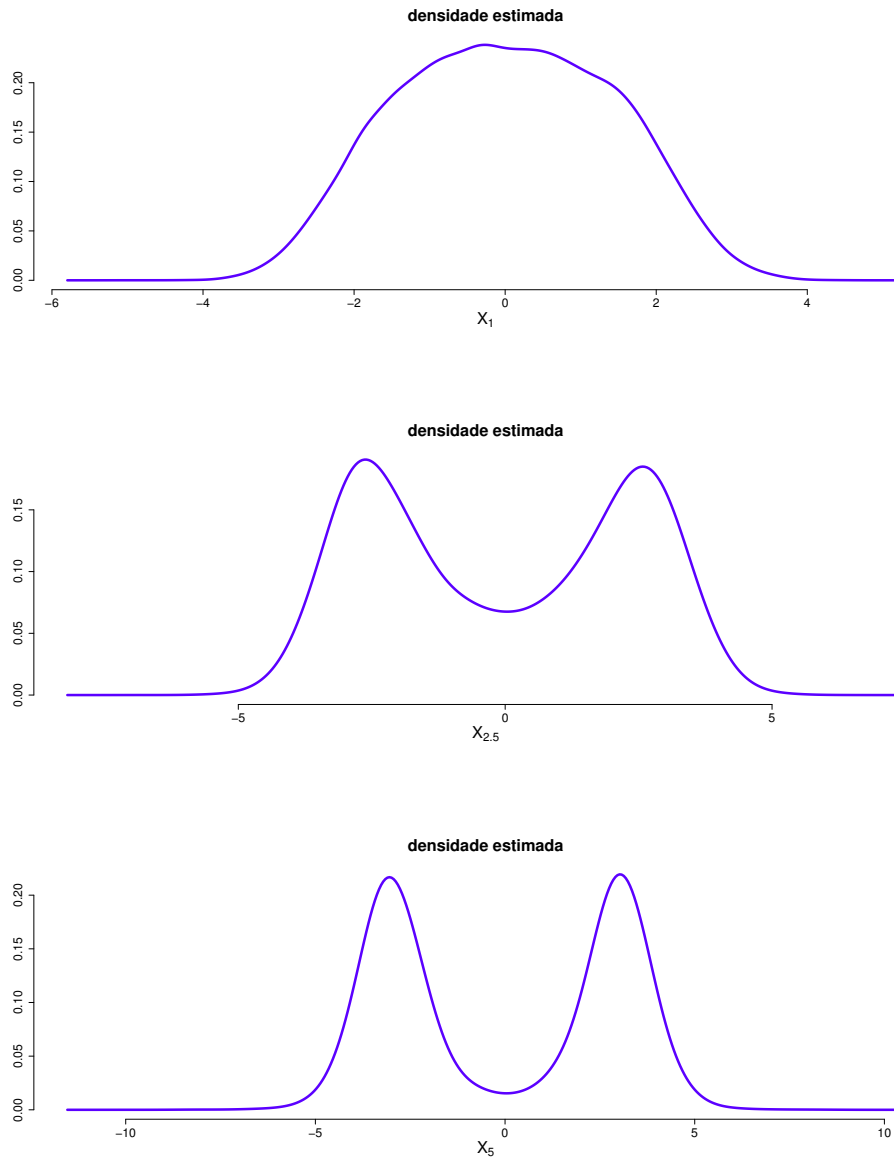


Figura 7.2: Densidades estimadas das variáveis X_1 , $X_{2,5}$ e X_5 correspondentes ao problema (7.2.35)-(7.2.36), segundo o algoritmo de simulação exata.

O procedimento realizado é descrito pelos seguintes passos:

passo 1: simulação de 5×10^4 observações da variável aleatória X_T através do algoritmo de simulação exata;

passo 2: simulação de 5×10^4 observações da variável aleatória X_T através dos esquemas de Euler-Maruyama, Milstein e linearização local;

passo 3: comparar as distribuições das variáveis obtidas nos passos 1 e 2 através do teste de Kolmogorov-Smirnov; concluiremos que os métodos induzem a mesma distribuição caso o teste acuse um p -valor superior a 0.05.

Os resultados foram construídos considerando apenas os instantes de tempo $T = 1$, $T = 2, 5$ e $T = 5$, isto é, avaliamos as distribuições das variáveis aleatórias X_1 , $X_{2,5}$ e X_5 conforme apresentado na Figura 7.2. Lembramos que este método reproduz exatamente a distribuição das variáveis X_1 , $X_{2,5}$ e X_5 ; isto é, o algoritmo é formalmente exato, devemos levar em consideração apenas desvios computacionais naturais de qualquer simulação. Logo, se as aproximações discretas forem de fato razoáveis, devemos obter um comportamento parecido ao simular tais variáveis a partir destas aproximações. Este fato fica evidenciado pela Figura 7.3, onde os núcleos estimados para o método de Euler-Maruyama são comparados à densidade estimada para a amostra do algoritmo EA1. Note como a semelhança aumenta à medida que os espaçamentos diminuem.

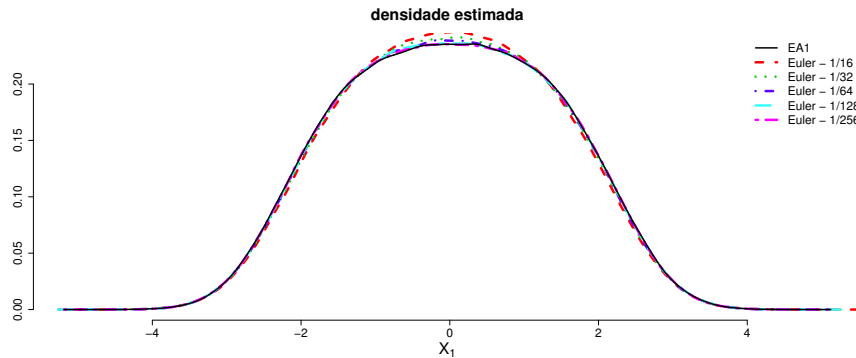


Figura 7.3: Comparação entre as densidades estimadas da variável X_1 - Euler-Maruyama \times EA1.

Δ	p -valor
2^{-4}	8.02e-35
2^{-5}	2.26e-12
2^{-6}	2.98e-04
2^{-7}	0.13
2^{-8}	0.13
2^{-9}	0.73

Tabela 7.1: Esquema de Euler-Maruyama \times algoritmo EA1.

A tabela acima ilustra os resultados do método de Euler-Maruyama. Para não tornar o texto repetitivo, apresentamos apenas as estatísticas do teste de

Kolmogorov-Smirnoff para a variável X_1 ; os demais resultados foram rigorosamente idênticos, portanto redundantes. Notemos que o teste é significativo apenas a partir de $\Delta = 2^{-7}$. Este resultado é razoável devido à baixa ordem de convergência forte do esquema.

Resultados similares foram obtidos pelo método de Milstein (Figura 7.4). A tabela 7.2 mostra que este esquema também é significativo apenas a partir do passo $\Delta = 2^{-7}$. Seu desempenho computacional também não difere muito do apresentado pelo método anterior.

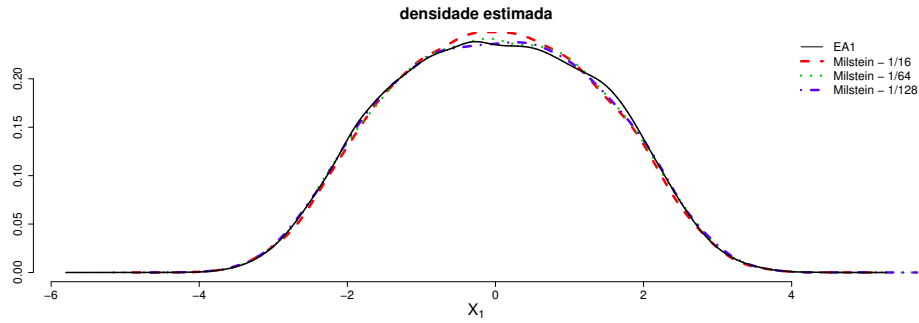


Figura 7.4: Comparação entre as densidades estimadas da variável X_1 - Milstein \times EA1.

Δ	p -valor
2^{-4}	7.84e-40
2^{-5}	4.73e-13
2^{-6}	1.52e-06
2^{-7}	0.08
2^{-8}	0.06
2^{-9}	0.70

Tabela 7.2: Esquema de Milstein \times algoritmo EA1.

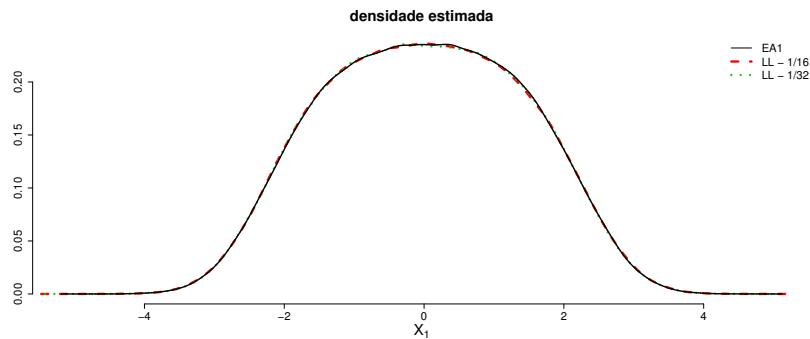
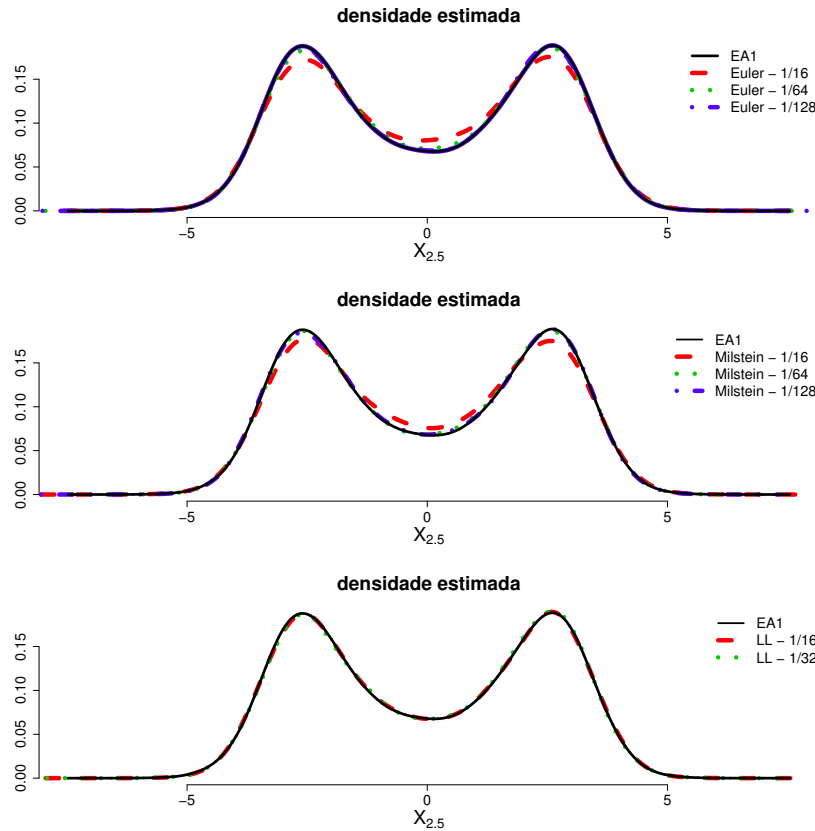


Figura 7.5: Comparação entre as densidades estimadas da variável X_1 - Linearização Local \times EA1.

Δ	p -valor
2^{-4}	0.52
2^{-5}	0.06
2^{-6}	0.19
2^{-7}	0.31
2^{-8}	0.66
2^{-9}	0.52

Tabela 7.3: Método da linearização local \times algoritmo EA1.

Os resultados mais expressivos foram novamente obtidos pelo método da linearização local (Figura 7.5). Perceba que todos os passos de discretização levaram à aceitação no teste de Kolmogorov-Smirnov. Resultados análogos são observados na construção de aproximações para as densidades de $X_{2,5}$ e X_5 , conforme ilustrado nas Figuras 7.6 e 7.7. Este é talvez um dos melhores indicativos de que este método supera os demais em convergência e estabilidade. Sua desvantagem é o tempo de execução que ultrapassa em muito os demais devido à necessidade do cálculo de exponenciais a cada iteração do algoritmo.

Figura 7.6: Comparação entre as densidades estimadas da variável $X_{2,5}$.

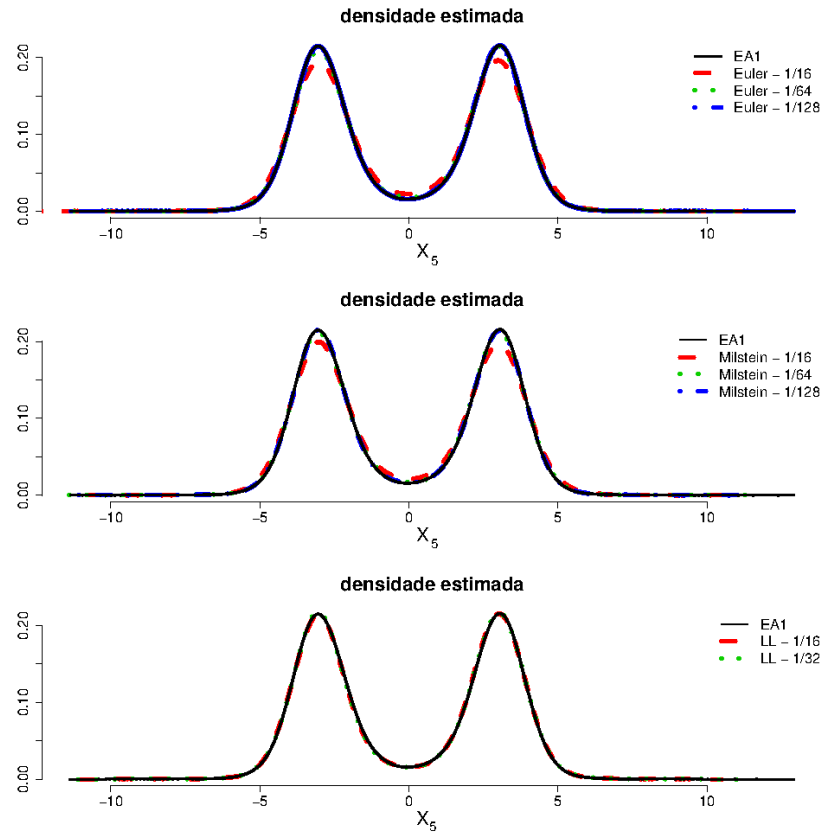


Figura 7.7: Comparação entre as densidades estimadas da variável $X_{2,5}$.

7.3 Exercícios

1. (Método da Rejeição para variáveis aleatórias em \mathbb{R}^n) Suponha que f e g são duas funções de probabilidade com respeito à medida de Lebesgue, tais que

$$f(x) \leq \epsilon g(x), \quad \forall x,$$

para algum $\epsilon > 0$. Utilize o Teorema 7.1 para mostrar que o seguinte algoritmo

- (a) amostre uma variável aleatória y distribuída segundo g ;
- (b) amostre uma variável u uniformemente distribuída em $(0, 1)$;
- (c) se $u < \epsilon f(y)/g(y)$, retorne y ; caso contrário, volte para o passo 1;

gera variáveis distribuídas segundo f a partir da geração de variáveis distribuídas segundo g .

2. Utilize o exercício anterior para desenvolver um algoritmo para a geração de variáveis aleatórias associadas a uma densidade h da forma (7.1.27).

3. Considere o seguinte problema de valor inicial:

$$dX_t = -3X_t dt + 2\sqrt{1 + X_t^2} dW_t, \quad 0 \leq t \leq 10$$

$$P(X_0 = 0) = 1.$$

- (a) A partir do problema acima, utilize a transformação de Lamperti para obter uma EDE com coeficiente de difusão constante.
- (b) Mostre que os coeficientes da nova EDE satisfazem as condições para a aplicação do método de simulação exata.
- (c) Siga os passos da avaliação experimental apresentada anteriormente e implemente o método de simulação exata para a construção de soluções deste problema, comparando os resultados obtidos com os demais métodos de discretização.

Bibliografia

- [1] BACHELIER, L. *Théorie de la speculation*. Tese (Doutorado) — Annales Scientifiques de l' École Normale Supérieure, 1900.
- [2] BESKOS, A.; PAPASPILIOPOULOS, O.; ROBERTS, G. O. Retrospective exact simulation of diffusion sample paths with applications. *Bernoulli*, International Statistical Institute (ISI), Voorburg; Bernoulli Society for Mathematical Statistics and Probability, Voorburg, v. 12, n. 6, p. 1077–1098, 2006. ISSN 1350-7265.
- [3] BESKOS, A.; ROBERTS, G. O. Exact simulation of diffusions. *Annals of Applied Probability*, v. 15, p. 2422–2444, 2005.
- [4] BILLINGSLEY, P. *Probability and measure*. 3. ed. ed. New York: Wiley, 1995. (A Wiley-Interscience publication). ISBN 0471007102.
- [5] BISHWAL, J. *Parameter Estimation in Stochastic Differential Equations*. [S.l.]: Springer, 2007. (Lecture Notes in Mathematics). ISBN 9783540744474.
- [6] BREIMAN, L. *Probability*. [S.l.]: Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1968. (Classics in Applied Mathematics). ISBN 9781611971286.
- [7] BROWN, R. Robert brown on microscopical observations of active molecules (brownian motion). 1827. Disponível em: <<http://sciweb.nybg.org/science2/pdfs/dws/Brownian.pdf>>.
- [8] BURRAGE, K.; BURRAGE, P.; MITSUI, T. Numerical solutions of stochastic differential equations - implementation and stability issues. *Journal of Computational and Applied Mathematics*, v. 125, p. 171–182, 2000.
- [9] BURRAGE, K.; BURRAGE, P. M. High strong order explicit Runge-Kutta methods for stochastic ordinary differential equations. *Applied Numerical Mathematics*, v. 22, p. 81–101, 1996.
- [10] EINSTEIN, A. On the motion of small particles suspended in liquids at rest required by the molecular-kinetic theory of heat. *Annalen der Physik*, v. 17, p. 549–560, 1905.
- [11] EVANS, L. C. *An Introduction to Stochastic Differential Equations*. [S.l.]: American Mathematical Society, 2012. ISBN 9781470410544.
- [12] FRIEDMAN, A. *Stochastic Differential Equations and Applications: Two Volumes Bound as One*. [S.l.]: Dover Publications, 1975. (Dover Books on Mathematics Series). ISBN 9780486453590.

- [13] HOFFMAN, K. *Banach Spaces of Analytic Functions*. [S.l.]: Prentice Hall, 1962.
- [14] IKEDA, N.; WATANABE, S. *Stochastic differential equations and diffusion processes*. Amsterdam, New York: North-Holland Pub. Co. Tokyo, 1989. (North-Holland mathematical library).
- [15] ITÔ, K. *On stochastic differential equations*. New York: American Mathematical Society, 1951. (Memories of the American Mathematical Society; No. 4).
- [16] JACOD, J.; MEMIN, J. Weak and strong solutions of stochastic differential equations: Existence and stability. In: WILLIAMS, D. (Ed.). *Stochastic Integrals*. Springer Berlin Heidelberg, 1981, (Lecture Notes in Mathematics, v. 851). p. 169–212. ISBN 978-3-540-10690-6. Disponível em: <<http://dx.doi.org/10.1007/BFb0088728>>.
- [17] KARATZAS, I.; SHREVE, S. *Brownian Motion and Stochastic Calculus*. [S.l.]: Springer-Verlag GmbH, 1991. (Graduate Texts in Mathematics Series). ISBN 9780387976556.
- [18] KARLIN, S.; TAYLOR, H. M. *A First Course in Stochastic Processes*. [S.l.]: Academic Press, 1975. ISBN 9780123985521.
- [19] KLOEDEN, P.; PLATEN, E. *Numerical Solution of Stochastic Differential Equations*. [S.l.]: Springer, 1999. (Applications of Mathematics Series). ISBN 9783540540625.
- [20] KLOEDEN, P. E.; PLATEN, E. A survey of numerical methods for stochastic differential equations. *J. Stoch. Hydrol. Hydraulics*, v. 3, p. 155–178, 1987.
- [21] LISKE, H.; PLATEN, E. Simulation studies on time discrete diffusion approximations. *Math. Comput. Simulation*, v. 29, p. 253–260, 1987.
- [22] LUSCHGY, H.; PAGÈS, G. Functional quantization of a class of brownian diffusions: A constructive approach. *Stochastic Processes and their Applications*, v. 116, n. 2, p. 310–336, 2006.
- [23] MCSHANE, E. *Stochastic Calculus and Stochastic Models*. [S.l.]: Academic Press, 1974. (Probability and Mathematical Statistics).
- [24] MILSTEIN, G. *Numerical Integration of Stochastic Differential Equations*. [S.l.]: Springer, 1994. (Mathematics and Its Applications). ISBN 9780792332138.
- [25] NELSON, E. *Dynamical Theories of Brownian Motion*. Princeton, N.J.: Princeton University Press, 1967.
- [26] NEWTON, N. J. Asymptotically efficient runge-kutta methods for a class of ito and stratonovich equations. *J. Appl. Math.*, SIAM, v. 51, p. 542–567, 1991.
- [27] NOVIKOV, A. On an identity for stochastic integrals. *Theory Probab. Appl.*, Society for Industrial and Applied Mathematics, v. 17, n. 3, p. 717–720, 1972.
- [28] ØKSENDAL, B. *Stochastic Differential Equations: An Introduction with Applications*. [S.l.]: Springer, 2010. (Universitext (1979)). ISBN 9783540047582.
- [29] PRESS, W. H. et al. *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. 3. ed. New York, NY, USA: Cambridge University Press, 2007.

- [30] PROTTER, P. E. *Stochastic Integration and Differential Equations*. [S.l.]: Springer-Verlag Berlin Heidelberg, 2005. (Stochastic Modelling and Applied Probability, v. 21).
- [31] RÜMELIN, W. Numerical Treatment of Stochastic Differential Equations. *SIAM Journal on Numerical Analysis*, Society for Industrial and Applied Mathematics, v. 19, n. 3, 1982.
- [32] SAITO, Y.; MITSUI, T. Stability analysis of numerical schemes for stochastic differential equations. *SIAM Journal on Numerical Analysis*, Society for Industrial and Applied Mathematics, v. 33, n. 6, p. pp. 2254–2267, 1996. ISSN 00361429.
- [33] SHIRYAEV, A.; WILSON, S. *Probability*. [S.l.]: Springer, 1995. (Graduate Texts in Mathematics). ISBN 9780387945491.
- [34] SHOJI, I. A note on convergence rate of a linearization method for the discretization of stochastic differential equations. *Communications in Nonlinear Science and Numerical Simulation*, Elsevier Science, Amsterdam, v. 16, n. 7, p. 2667–2671, 2011. ISSN 1007-5704.
- [35] SHOJI, I.; OZAKI, T. Estimation for nonlinear stochastic differential equations by a local linearization method. *Stochastic Analysis and Applications*, Taylor & Francis, v. 16, n. 4, p. 733–752, 1998. ISSN 0736-2994.
- [36] STROOCK, D. W.; VARADHAN, S. R. S. *Multidimensional Diffusion Processes*. New York: Spinger-Verlag, 1979.
- [37] WIENER, N. Differential space. *Journal of Mathematical Physics*, v. 2, p. 131–174, 1923.
- [38] WIENER, N. Un problème de probabilités énombrables. *Bull. Soc. Math. de France*, v. 52, p. 569–578, 1924.

Índice

- p -ésimo momento, 4
- cilindro n -dimensional, 6
- coeficiente de difusão, 34
- coeficiente de tendência, 34
- condição de Novikov, 52
- convergência forte, 71
- critério de convergência de Kolmogorov, 14
- densidade de transição, 8
- distribuições finito-dimensionais, 7
- equação de Tanaka, 40
- equação diferencial estocástica, 35
- espaço de probabilidade, 4
- esperança condicional, 5
- esquema A-estável, 80
- esquema de diferenças avançadas, 62
- esquema de Taylor, 69
- esquema livre de derivadas, 62
- estabilidade, 74
- existência e unicidade de soluções fortes, 37
- fórmula de Itô, 42, 44
- filtração, 7
- função geradora de momentos, 5
- independência, 4
- integração por partes, 45
- integral de Itô, 20, 22, 31
- integral de Stratonovich, 31
- integral estocástica, 33
- Lei do logaritmo iterado, 16
- método θ -estocástico, 64
- método da Linearização Local, 65
- método de Euler implícito, 64
- método de Euler trapezoidal, 64
- método de Euler-Maruyama, 58
- método de Heun, 68
- método de Milstein, 59
- método de Runge-Kutta de primeira ordem, 68
- método de Runge-Kutta de quatro estágios, 69
- método de Runge-Kutta de três estágios, 68
- métodos implícitos, 64
- martingale, 9, 52
- modificação, 6
- movimento Browniano, 11
- movimento Browniano Geométrico, 70
- movimento Browniano geométrico, 49, 93
- ordem de convergência forte, 72
- ponte Browniana, 54
- ponte Browniana, 98
- processo de difusão, 33
- processo de Ornstein-Uhlenbeck, 35
- processo de Wiener, 15
- processo estocástico, 5
- processo Markoviano, 8
- processos indistinguíveis, 6
- propriedade de Markov, 8
- regra do produto, 44
- ruído branco Gaussiano, 34
- solução forte, 36
- solução fraca, 39
- submartingale, 9
- supermartingale, 9
- Teorema de Cameron-Martin-Girsanov, 51
- Teorema de Girsanov, 52
- Teoremas de extensão de Kolmogorov, 7
- transformação de Lamperti, 48
- truncamento, 27
- variável aleatória, 4
- variação quadrática, 17
- variação total, 17