

Felipe Carvalho Álvares da Silva

***Simulação e Análise Numérica de Equações
Diferenciais Estocásticas Unidimensionais: Uma
Abordagem Introdutória***

Belo Horizonte - MG, Brasil

14/03/2013

Felipe Carvalho Álvares da Silva

***Simulação e Análise Numérica de Equações
Diferenciais Estocásticas Unidimensionais: Uma
Abordagem Introdutória***

Dissertação apresentada para obtenção do Grau
de Mestre em Estatística pela Universidade Fe-
deral de Minas Gerais.

Orientador:

Luiz Henrique Duczmal

Co-orientador:

Denise Burgarelli Duczmal

DEPARTAMENTO DE ESTATÍSTICA
INSTITUTO DE CIÊNCIAS EXATAS
UNIVERSIDADE FEDERAL DE MINAS GERAIS

Belo Horizonte - MG, Brasil

14/03/2013

Dissertação de Mestrado sob o título "*Simulação e Análise Numérica de Equações Diferenciais Estocásticas Unidimensionais: Uma Abordagem Introdutória*", defendida por Felipe Carvalho Álvares da Silva e aprovada em 14/03/2013, em Belo Horizonte, Estado de Minas Gerais, pela banca examinadora constituída pelos professores:

Prof. Luiz H. Duczmal
Orientador

Prof.(a) Denise Buragerelli Duczmal
Co-Orientadora

Sokol Ndreca
Universidade Federal de Minas Gerais

Gregório Saravia Atuncar
Universidade Federal de Minas Gerais

Alberto Masayoshi Faria Ohashi
Insper Instituto de Ensino e Pesquisa

Resumo

Este trabalho apresenta uma discussão introdutória com respeito aos métodos numéricos para a simulação de equações diferenciais estocásticas. Começamos com a apresentação de conceitos matemáticos essenciais a respeito de probabilidade e processos estocásticos. Em seguida apresentamos os esquemas de discretização clássicos e fazemos uma análise teórica e experimental a respeito do comportamento destes métodos. Ao final tratamos de um algoritmo de simulação que explora o método da rejeição para gerar uma solução através da medida induzida pelas trajetórias correspondentes às soluções das equações.

Sumário

Lista de Figuras

Lista de Tabelas

1	Introdução	p. 9
2	Conceitos Matemáticos Preliminares	p. 11
2.1	Probabilidade e Processos Estocásticos	p. 11
2.2	Movimento Browniano	p. 16
2.3	A Integral de Itô	p. 22
2.4	Fórmula de Itô	p. 28
2.5	Transformação de Lamperti	p. 34
2.6	Transformação de Cameron-Martin-Girsanov	p. 37
3	Equações Diferenciais Estocásticas	p. 42
3.1	Soluções Fortes	p. 43
3.2	Soluções Fracas	p. 44
4	Métodos de Discretização	p. 47
4.1	Expansão de Itô-Taylor	p. 47
4.2	Métodos de Euler-Maruyama e Milstein	p. 50
4.3	Esquema Livre de Derivada	p. 55
4.4	Método θ Estocástico	p. 57
4.5	Método da Linearização Local	p. 59

4.6	Miscelânea	p. 62
5	Análise Numérica	p. 65
5.1	Convergência Forte	p. 65
5.2	Estabilidade	p. 69
5.2.1	Funções e Regiões de Estabilidade de Alguns Esquemas Numéricos .	p. 72
5.2.2	Método da Linearização Local	p. 76
5.3	Avaliação Experimental	p. 77
5.3.1	Convergência	p. 77
5.3.2	Estabilidade	p. 84
6	Simulação Exata	p. 93
6.1	O Algoritmo	p. 93
6.2	Avaliação Experimental	p. 100
7	Conclusões e Trabalhos Futuros	p. 105
	Referências Bibliográficas	p. 108

Lista de Figuras

2.1	Realizações do movimento Browniano discretizado.	p. 23
4.1	Exemplo de aproximação pelo método de Euler-Maruyama para o movimento Browniano Geométrico.	p. 53
4.2	Exemplo de aproximação pelo método de Milstein para o movimento Browniano Geométrico.	p. 54
4.3	Comparação entre os métodos de Milstein e livre de derivada.	p. 57
4.4	Euler-Maruyama $\times \theta$ Estocástico ($\theta = 1/4$).	p. 59
5.1	Esquema de Euler-Maruyama avaliado no intervalo $[0, 1]$	p. 79
5.2	Esquema de Euler-Maruyama avaliado nos intervalos $[0, 5]$ e $[0, 10]$	p. 80
5.3	Esquema Euler-trapezoidal avaliado nos intervalos $[0, 1]$, $[0, 5]$ e $[0, 10]$	p. 81
5.4	Esquema Euler-implícito avaliado nos intervalos $[0, 1]$, $[0, 5]$ e $[0, 10]$	p. 82
5.5	Esquema de Milstein avaliado nos intervalos $[0, 1]$ e $[0, 5]$	p. 83
5.6	Esquema livre de derivada avaliado nos intervalos $[0, 1]$ e $[0, 5]$	p. 84
5.7	Estabilidade do esquema de Euler-Maruyama.	p. 86
5.8	Estabilidade do método θ estocástico.	p. 86
5.9	Estabilidade do método de Milstein.	p. 87
5.10	Comparação de estabilidade entre os métodos de Milstein e de Euler-Maruyama.	p. 88
5.11	Método da linearização local aplicado ao problema (5.39)-(5.40).	p. 89
5.12	Método de Milstein aplicado ao problema (5.39)-(5.40).	p. 90
5.13	Método de Euler-Maruyama aplicado ao problema (5.39)-(5.40).	p. 90
5.14	Método da linearização local aplicado ao problema (5.43)-(5.44).	p. 91
5.15	Método de Euler-Maruyama aplicado ao problema (5.43)-(5.44).	p. 92

- 6.1 Densidade estimada da variável X_1 correspondente ao problema (6.17)-(6.18). p. 102
- 6.2 Comparação entre as densidades estimadas da variável X_1 - Euler-Maruyama \times EA1. p. 102
- 6.3 Comparação entre as densidades estimadas da variável X_1 - Linearização
Local \times EA1. p. 103

Lista de Tabelas

5.1	Esquema de Euler-Maruyama avaliado no intervalo $[0, 1]$	p. 78
5.2	Esquema de Euler-Maruyama avaliado nos intervalos $[0, 5]$ e $[0, 10]$	p. 80
5.3	Esquema Euler-trapezoidal avaliado nos intervalos $[0, 1]$, $[0, 5]$ e $[0, 10]$	p. 80
5.4	Ajustes lineares para avaliação experimental do esquema Euler-trapezoidal.	p. 80
5.5	Ajustes lineares para avaliação experimental do esquema Euler-implícito.	p. 81
5.6	Esquema Euler-implícito avaliado nos intervalos $[0, 1]$, $[0, 5]$ e $[0, 10]$	p. 82
5.7	Ajustes lineares para avaliação experimental do esquema de Milstein.	p. 83
5.8	Esquema de Milstein avaliado nos intervalos $[0, 1]$ e $[0, 5]$	p. 83
5.9	Ajustes lineares para avaliação experimental do esquema livre de derivada.	p. 83
5.10	Esquema livre de derivada avaliado nos intervalos $[0, 1]$ e $[0, 5]$	p. 84
5.11	Evolução do esquema trapezoidal no intervalo $[0, 150]$	p. 87
6.1	Esquema de Euler-Maruyama \times algoritmo EA1.	p. 103
6.2	Esquema de Milstein \times algoritmo EA1.	p. 104
6.3	Método da linearização local \times algoritmo EA1.	p. 104

1 *Introdução*

Equações diferenciais ordinárias são instrumentos matemáticos utilizados para explicar e modelar o comportamento de sistemas que evoluem continuamente no tempo. Estes modelos aparecem em diversas áreas do conhecimento. Podemos citar, por exemplo, as equações de Lotka-Volterra para descrever dinâmicas nos sistemas biológicos (modelo predador-presa), a segunda lei de Newton que relaciona a força resultante em uma partícula à variação de seu momento linear, dentre vários outros. À medida que os modelos se tornam mais complexos a dificuldade em desenvolver estudos puramente analíticos aumenta e, neste caso, utilizamos métodos numéricos como alternativa. A utilização de métodos numéricos nesta área já é bastante tradicional e tais métodos têm sido cada vez mais estudados devido ao avanço da computação.

Modelos envolvendo equações diferenciais ordinárias são indiscutivelmente úteis em teorias que lidam com o estudo do comportamento de sistemas ao longo do tempo. No entanto, a entrada de incertezas, aleatoriedade ou ruídos no sistema, induz o desenvolvimento de uma teoria que permita ampliar o espaço de fenômenos bem modelados por equações ordinárias. Considere, por exemplo, o simples modelo de crescimento populacional:

$$\frac{dN}{dt} = a(t)N(t), \quad (1.1)$$

$$N(0) = N_0, \quad (1.2)$$

onde $N(t)$ é o tamanho da população no instante t , $a(t)$ é taxa relativa de crescimento no instante t e N_0 é uma constante que indica a população inicial. É possível que o coeficiente $a(t)$ não seja completamente conhecido devido a ações aleatórias do meio. Desta forma, o coeficiente poderia ser expresso por:

$$a(t) = r(t) + \text{''ruído''},$$

onde o comportamento do termo de ruído não é conhecido exatamente, no entanto, podemos conhecer sua distribuição de probabilidade. Já a função $r(t)$ é considerada puramente determinística. Este tipo de abordagem nos leva naturalmente ao desenvolvimento das equações dife-

renciais estocásticas, que descrevem o comportamento físico através da teoria da probabilidade e dos processos estocásticos.

O tratamento analítico das equações diferenciais estocásticas é ainda mais complicado que o das equações diferenciais no caso determinístico. Neste contexto, a simples definição do conceito de solução da equação estocástica já não é tão simples. Desta forma, a necessidade do desenvolvimento de métodos numéricos é ainda maior. Alguns dos métodos numéricos são análogos ao caso determinístico: discretizamos o domínio de interesse e de alguma forma avaliamos aproximações do processo real nos pontos da malha discretizada. Outras linhas de métodos para a simulação de soluções de equações diferenciais estocásticas exploram o fato de estarmos lidando com elementos aleatórios. Nestes casos, algoritmos como o método da rejeição, amostragem por importância, MCMC, etc., podem ser empregados.

Tão importante como o desenvolvimento do método é o entendimento das condições sob as quais eles podem ser aplicados e o desenvolvimento de critérios para avaliar a qualidade destes algoritmos. Seguindo esta linha, este trabalho tem como objetivo apresentar alguns dos mais tradicionais métodos para a simulação de soluções de equações estocásticas e realizar uma breve análise destes métodos (tanto de maneira teórica quanto experimental).

2 *Conceitos Matemáticos Preliminares*

2.1 Probabilidade e Processos Estocásticos

Esta seção é dedicada à apresentação de notações e conceitos básicos necessários para o desenvolvimento dos próximos tópicos. Uma vez que este trabalho tem como foco equações unidimensionais, os conceitos a seguir levam em consideração apenas funções em \mathbb{R} . Estes mesmos conceitos podem, sem exceção, ser generalizados para vetores em \mathbb{R}^n , $n \in \mathbb{N}$.

Definição 2.1 *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) tal que $P(\Omega) = 1$. 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 2.2 *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*

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

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

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

A verificação de que P_X é de fato uma medida de probabilidade é imediata.

Definição 2.3 *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

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

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.

Definição 2.4 Seja X uma variável aleatória definida no espaço de probabilidade (Ω, \mathcal{F}, P) . Definimos o p -ésimo momento de X como sendo

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

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

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

e a variância

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

Definição 2.5 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 ($P - q.c.$) 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 garantidas 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 Shiryaev (1995).

Definição 2.6 *Um processo estocástico é uma família*

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

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

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

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

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

Definição 2.7 *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 $P(X_t = Y_t) = 1$.

(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

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

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

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

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

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

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

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

no entanto,

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

Definição 2.8 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

$$C(A) := \{\omega \in \mathbb{R}^{[0, \infty)}; (\omega(t_1), \dots, \omega(t_n)) \in A\}, \quad (2.1)$$

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

Se $A \in \mathcal{B}(\mathbb{R}^n)$, então $A = B_1 \times B_2 \times \dots \times B_n$, 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 2.7.

Definição 2.9 *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

$$Q_\tau(A) = P(C(A)) = P\left[\omega \in \mathbb{R}^{[0,\infty)}; (\omega(t_1), \dots, \omega(t_n)) \in A\right], \quad (2.2)$$

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*

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

- (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*

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

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

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

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 Karatzas & Shrev (1991).

Definição 2.10 *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,

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

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

Definição 2.11 *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 $t \in T$.*

Definição 2.12 *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 $\{\mathcal{F}_t\}_{t \geq 0}$.

2.2 Movimento Browniano

Matematicamente, um movimento Browniano unidimensional $B = \{B_t; 0 \leq t < +\infty\}$ em um espaço de probabilidade (Ω, \mathcal{F}, P) é completamente especificado pelo conjunto de hipóteses:

1. B é um processo Gaussiano em Ω ;
2. B tem trajetórias contínuas P -q.c.;

$$3. \mathbb{E}B_t = 0 \text{ e } \mathbb{E}[B_t B_s] \propto (t \wedge s), \forall t, s \geq 0.$$

É comum utilizarmos a expressão *movimento Browniano padrão* quando queremos enfatizar que $\mathbb{E}[B_t B_s] = (t \wedge s)$, $\forall t, s \geq 0$. Além disso, assumimos que o processo é sempre iniciado na origem, isto é, $P(B_0 = 0) = 1$. Daqui em diante, sempre que mencionarmos o termo movimento Browniano, estaremos nos referindo ao processo padrão iniciado na origem.

Esta descrição matemática foi proposta por Norbert Wiener com o objetivo de descrever o movimento de pequenas partículas sob a superfície da água causado pelo bombardeamento contínuo das moléculas do fluido. O modelo traduz bem a natureza do fenômeno físico representado pelo movimento Browniano em um intervalo de tempo suficientemente pequeno. Na prática, pode ser interessante também definir este processo em função de seus incrementos.

Definição 2.13 *Seja (Ω, \mathcal{F}, P) um espaço de probabilidade. Denotamos por movimento Browniano um processo estocástico $B = \{B_t, 0 \leq t < \infty\}$ tal que:*

1. $P(B_0 = 0) = 1$;
2. $(B_t - B_s) \sim N(0, t - s)$, $\forall 0 \leq s < t < +\infty$;
3. $(B_t - B_s)$ é independente da σ -álgebra $\mathcal{F}_t = \sigma(B_u, \forall 0 \leq u \leq s)$, $\forall 0 \leq s < t < +\infty$.

A existência de um processo estocástico com estas propriedades é garantida pelo teorema 2.1. Considere um processo estocástico $B = \{B_t; 0 \leq t < +\infty\}$ tal que, para quaisquer $0 = s_0 < s_1 < s_2 < \dots < s_n$, possui a seguinte função de distribuição acumulada para $(B_{s_1}, \dots, B_{s_n})$:

$$F_{(s_1, \dots, s_n)}(x_1, \dots, x_n) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \dots \int_{-\infty}^{x_n} p(s_1; 0, y_1) p(s_2 - s_1; y_1, y_2) \dots \dots p(s_n - s_{n-1}; y_{n-1}, y_n) dy_n \dots dy_2 dy_1, \quad (2.3)$$

onde $(x_1, \dots, x_n) \in \mathbb{R}^n$ e p é o núcleo Gaussiano:

$$p(t; x, y) := \frac{1}{\sqrt{2\pi t}} e^{-(x-y)^2/2t}, \quad t > 0; x, y \in \mathbb{R}. \quad (2.4)$$

Aplicando o teorema 2.1, podemos assegurar a existência de uma medida de probabilidade P no espaço $(\mathbb{R}^{[0, \infty)}, \mathcal{B}(\mathbb{R}^{[0, \infty)})$, sob a qual o processo estocástico B cumpre as três propriedades da definição anterior. Logo, podemos obter um espaço (Ω, \mathcal{F}, P) onde B é um movimento Browniano.

Teorema 2.2 (Kolmogorov & Cestov) *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

$$\mathbb{E}|X_t - X_s|^\alpha \leq C|t - s|^{1+\beta}, \quad 0 \leq s, t \leq T, \quad (2.5)$$

para algum trio de constantes positivas α, β e C . Então, existe uma modificação contínua $X' = \{X'; 0 \leq t \leq T\}$ de X .

Proposição 2.2 *Existe uma modificação contínua do movimento Browniano.*

Demonstração: Dado $T \geq 0$, já verificamos que é possível obter um processo estocástico $B = \{B_t; 0 \leq t \leq T\}$ que satisfaz as condições da definição 2.13. Como $(B_t - B_s) \sim N(0, t - s)$, para quaisquer $0 \leq s < t \leq T$, temos a seguinte função geradora de momentos:

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

Desta forma,

$$\begin{aligned} \mathbb{E}(|B_t - B_s|^4) &= \mathbb{E}((B_t - B_s)^4) \\ &= \frac{d^4 \phi}{d\lambda^4}(0) = 3(t - s)^2, \end{aligned}$$

o que mostra que B satisfaz o teorema de Kolmogorov-Cestov para as constantes $\alpha = 4$, $\beta = 1$ e $C = 3$.

□

Neste trabalho estamos particularmente interessados em processos com trajetórias contínuas. Assim, sempre que nos referirmos ao movimento Browniano, estaremos implicitamente considerando as modificações contínuas.

Teorema 2.3 *Todo movimento Browniano $B = \{B_t; 0 \leq t \leq T\}$ é uma martingale com relação à escolha natural de filtração $\mathcal{F}_t = \sigma(B_s; 0 \leq s \leq t)$.*

Demonstração: Note inicialmente que

$$\begin{aligned} \mathbb{E}^2|B_t| &\leq \mathbb{E}[|B_t|^2] \\ &= \text{Var}[B_t] = t < +\infty. \end{aligned}$$

Além disso, dado $s \geq t$, temos que $(B_s - B_t)$ é independente de \mathcal{F}_t . Logo,

$$\mathbb{E}[B_s - B_t | \mathcal{F}_t] = \mathbb{E}[B_s - B_t] = 0.$$

Como B_t é \mathcal{F}_t -mensurável, segue que

$$\mathbb{E}[B_t | \mathcal{F}_t] = B_t.$$

Assim,

$$\begin{aligned} \mathbb{E}[B_s | \mathcal{F}_t] &= \mathbb{E}[B_s - B_t + B_t | \mathcal{F}_t] \\ &= \mathbb{E}[B_s - B_t | \mathcal{F}_t] + \mathbb{E}[B_t | \mathcal{F}_t] = 0 + B_t = B_t. \end{aligned}$$

□

Teorema 2.4 *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

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

à medida que $n \rightarrow \infty$. Então

i. (Variação quadrática do movimento browniano)

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

ii. (Variação total do movimento browniano)

$$\sum_{k=0}^{m_n-1} |B_{t_{k+1}^n} - B_{t_k^n}| \rightarrow \infty;$$

em $L_2(\Omega, \mathcal{F}, P)$ à medida que $n \rightarrow \infty$.

Demonstração: (i) De modo a simplificar a notação, considere:

$$\Delta_k^n(B) := B_{t_{k+1}^n} - B_{t_k^n};$$

$$\Delta_k^n := t_{k+1}^n - t_k^n.$$

Seja $Q_n = \sum_{k=0}^{m_n-1} (\Delta_k^n(B))^2$, então

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

Elevando ao quadrado os dois lados da última igualdade:

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

Pela linearidade da esperança, tem-se

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

Vejamos agora que para todo $k \neq j$, o termo no somatório duplo acima é zero. Como $\Delta_k^n(B) \sim N(0, \Delta_k^n)$, segue que

$$\mathbb{E} (\Delta_k^n(B)^2) = \text{Var}(\Delta_k^n(B)) + \mathbb{E}(\Delta_k^n(B)) = \Delta_k^n.$$

Explorando novamente a linearidade da esperança:

$$\mathbb{E} [\Delta_k^n(B) - \Delta_k^n] = \Delta_k^n - \Delta_k^n = 0.$$

Para todo $k \neq j$, as variáveis $(\Delta_k^n(B) - \Delta_k^n)$ e $(\Delta_j^n(B) - \Delta_j^n)$ são independentes, logo

$$\mathbb{E} [(\Delta_k^n(B)^2 - \Delta_k^n) (\Delta_j^n(B)^2 - \Delta_j^n)] = \mathbb{E} [\Delta_k^n(B)^2 - \Delta_k^n] \mathbb{E} [\Delta_j^n(B)^2 - \Delta_j^n] = 0.$$

O somatório duplo se resume então apenas aos termos tais que $k = j$, ou seja, teremos agora apenas um somatório simples:

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

Note que a variável aleatória $\frac{\Delta_k^n(B)}{\sqrt{\Delta_k^n}}$ possui distribuição normal com média 0 e variância 1, independentemente de k . Neste caso, a esperança no último somatório pode ser tratada como uma constante, digamos C .

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

Como $|P^n| \rightarrow 0$, segue-se que $\mathbb{E}[(Q_n - (\beta - \alpha))^2] \rightarrow 0$ à medida que $n \rightarrow \infty$. Podemos então concluir que

$$Q_n \rightarrow \beta - \alpha.$$

$$\therefore \sum_{k=0}^{m_n-1} (B(t_{k+1}^n) - B(t_k^n))^2 \rightarrow \beta - \alpha,$$

à medida que $n \rightarrow \infty$.

(ii) Suponha, por absurdo, que B possua variação total limitada no intervalo $[\alpha, \beta]$. Nesse caso, teríamos

$$\sum_{k=0}^{m_n-1} (B_{t_{k+1}^n} - B_{t_k^n})^2 \leq \max_{0 \leq k \leq m_n-1} |B_{t_{k+1}^n} - B_{t_k^n}| \cdot \sum_{k=0}^{m_n-1} |B_{t_{k+1}^n} - B_{t_k^n}|.$$

Pela continuidade dos caminhos amostrais de B , segue-se que

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

à medida que $|P_n| \rightarrow 0$. Assim, concluiríamos que

$$\sum_{k=0}^{m_n-1} (B_{t_{k+1}^n} - B_{t_k^n})^2 \rightarrow 0$$

à medida que $n \rightarrow \infty$, o que contradiz o resultado verificado em (i).

□

Simulação

Para fins computacionais, é importante considerarmos o movimento Browniano discretizado. Ao utilizar o termo discretizado, fazemos referência a simulações do movimento Browniano (como um processo contínuo) avaliadas em quantidades finitas de pontos. Este tipo de simulação é trivial se explorarmos as propriedades de definição deste processo estocástico. Se $B = \{B_t; 0 \leq t \leq T\}$, $T > 0$ é um movimento Browniano, então temos que $P(B_0 = 0) = 1$. Desta forma, o processo discretizado partirá sempre da origem. Considerando que o intervalo de interesse foi particionado em N pontos igualmente espaçados, segue que os demais pontos da sequência podem ser gerados da relação:

$$B_{t_{n+1}} = B_{t_n} + \Delta Z,$$

onde $\Delta = \frac{T}{N-1}$ e Z é uma variável aleatória com distribuição normal padrão. Este algoritmo se justifica pela propriedade $B_t - B_s \sim N(0, t - s)$ para quaisquer $0 \leq s < t \leq T$. Este método pode

ainda ser facilmente estendido para partições com pontos espaçados de maneira arbitrária, para tal, basta ajustar o passo Δ a cada iteração. Perceba que o custo de simulação deste método é basicamente o custo de geração de $N - 1$ variáveis gaussianas. Um algoritmo eficiente para a geração destas variáveis é o de Box-Muler, o qual se baseia na geração na transformação de variáveis uniformes.

Box-Muler. Se u_1 e u_2 são variáveis aleatórias independentes e uniformemente distribuídas no intervalo $(0, 1)$, então

$$z_1 := \cos(2\pi u_2) \sqrt{-2\ln(u_1)}, \quad z_2 := \sin(2\pi u_2) \sqrt{-2\ln(u_1)},$$

são normalmente distribuídas com média $\mu = 0$ e variância $\sigma^2 = 1$. Para verificar o resultado basta isolar u_1 e u_2 ,

$$u_1 = e^{-(z_1^2 + z_2^2)/2}, \quad u_2 = \frac{1}{2\pi} \tan^{-1} \left(\frac{z_2}{z_1} \right),$$

e calcular o determinante jacobiano

$$\frac{\partial(u_1, u_2)}{\partial(z_1, z_2)} = \begin{vmatrix} \frac{\partial u_1}{\partial z_1} & \frac{\partial u_1}{\partial z_2} \\ \frac{\partial u_2}{\partial z_1} & \frac{\partial u_2}{\partial z_2} \end{vmatrix} = - \left(\frac{1}{2\pi} e^{-z_1^2/2} \right) \left(\frac{1}{2\pi} e^{-z_2^2/2} \right).$$

Desta forma,

$$f(z_1, z_2) = \frac{1}{2\pi} e^{-z_1^2/2} \frac{1}{2\pi} e^{-z_2^2/2},$$

o que mostra que $z_1, z_2 \sim N(0, 1)$.

Neste trabalho, todas as simulações envolvendo variáveis gaussianas, foram executadas a partir deste algoritmo.

2.3 A Integral de Itô

Apresentaremos agora o conceito de integração estocástica. Mais precisamente, trabalharemos a definição e algumas propriedades da integral

$$I(f) = \int_S^T f(t, \omega) dB_t(\omega), \quad (2.6)$$

onde $B = \{B_t; 0 \leq t < +\infty\}$ é um movimento Browniano e f é uma função estocástica pertencente a uma determinada classe que será definida adiante. Esta definição é não trivial pois as trajetórias $t \mapsto B_t(\omega)$ possuem variação total infinita, o que nos impede de construir a integral $I(f)$ segundo a teoria Lebesgue-Stieltjes. De uma maneira mais geral, integrais semelhantes podem ser definidas com relação a uma classe mais ampla de processos estocásticos mas tal

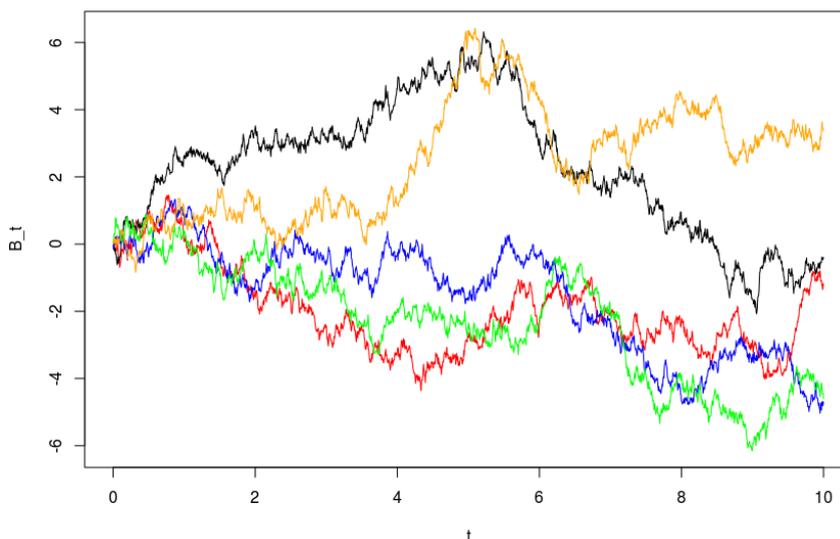


Figura 2.1: Realizações do movimento Browniano discretizado.

abordagem não será discutida neste trabalho.

Sejam $B = \{B_t; 0 \leq t < +\infty\}$ um movimento Browniano em um espaço de probabilidade (Ω, \mathcal{F}, P) e $\{\mathcal{F}_t\}_{t \geq 0}$ uma filtração deste espaço tal que para todo $t \geq 0$

- (a) $\mathcal{F}_t \subset \mathcal{F}$;
- (b) $\mathcal{F}(B_s, 0 \leq s \leq t) \subset \mathcal{F}_t$;
- (c) $\mathcal{F}(B_{t+\lambda} - B_t, \lambda \geq 0)$ é independente de \mathcal{F}_t .

Caso não seja explicitada uma outra filtração, consideraremos que \mathcal{F}_t é a σ -álgebra \mathcal{F}_t^B gerada pelo movimento Browniano até o instante t . Esta convenção é apenas de uma simplificação e pode 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.

Definição 2.14 *Suponha que $\{\mathcal{H}_t\}_{t \geq 0}$ é uma filtração associada ao espaço (Ω, \mathcal{F}, P) , com respeito a qual o movimento Browniano $B = \{B_t; 0 \leq t < +\infty\}$ é uma martingale. Considere as classes das funções*

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

tais que

- (i) $(t, \omega) \rightarrow f(t, \omega)$ é $\mathcal{B} \times \mathcal{F}$ -mensurável, onde \mathcal{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;

$$(iii) P \left[\int_s^r |f(t, \omega)|^p dt < +\infty \right] = 1;$$

$$(iii') \mathbb{E} \left[\int_s^r |f(t, \omega)|^p dt \right] < +\infty.$$

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

É imediato que $\mathbb{M}_{\mathcal{H}}^p \subseteq \mathbb{L}_{\mathcal{H}}^p$ pois toda variável aleatória integrável é finita quase certamente. No entanto, a recíproca não é verdadeira.

Durante a construção da integral de Itô (2.6) estaremos particularmente interessados em funções da classe $\mathbb{M}_{\mathcal{H}}^2$. A idéia é análoga à teoria de Lebesgue-Stieltjes: definimos $I(\phi)$ para uma classe elementar de funções $\phi \in \mathbb{M}_{\mathcal{H}}^2$, em seguida mostramos que cada $f \in \mathbb{M}_{\mathcal{H}}^2$ pode ser aproximada por uma sequência de funções elementares e, por fim, definimos $I(f)$ como o limite de $I(\phi)$ à medida que $\phi \rightarrow f$.

Definição 2.15 Um processo estocástico $\phi \in \mathbb{M}^2$ é dito elementar se puder ser representado por

$$\phi(t, \omega) = \sum_j^n e_j(\omega) \mathbf{I}_{[t_j, t_{j+1})}(t), \quad (2.7)$$

onde $n \in \mathbb{N}$, e_1, e_2, \dots, e_n são variáveis aleatórias $\mathcal{F}_{t_j}^B$ -mensuráveis, $j = 1, 2, \dots, n$ e $\cup_j [t_j, t_{j+1}) = [s, r)$.

Definição 2.16 A integral de Itô $I(\phi)$ de uma função elementar $\phi \in \mathbb{M}_{\omega}^2$ é dada pelo somatório

$$\int_s^r \phi(t, \omega) dB_t(\omega) = \sum_{j \geq 0} e_j(t_j, \omega) (B_{t_{j+1}} - B_{t_j})(\omega). \quad (2.8)$$

Teorema 2.5 (Isometria de Itô) Se $\phi(t, \omega) \in \mathbb{M}^2$ é elementar, então

$$\mathbb{E} \left[\left(\int_s^r \phi(t, \omega) dB_t(\omega) \right)^2 \right] = \mathbb{E} \left[\int_s^r \phi(t, \omega)^2 dt \right]. \quad (2.9)$$

Demonstração: Para simplificar a notação, considere

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

Para cada $i < j$, temos que os elementos aleatórios $\Delta_j(\omega)$ e $e_i(\omega)e_j(\omega)\Delta_i(\omega)$ são independentes. Neste caso,

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

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

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

Portanto,

$$\begin{aligned} \mathbb{E} \left[\left(\int_s^r \phi(t, \omega) dB_t(\omega) \right)^2 \right] &= \mathbb{E} \left[\sum_{i \geq 0} \sum_{j \geq 0} e_i(t_i, \omega)e_j(t_j, \omega)\Delta_i(\omega)\Delta_j(\omega) \right] \\ &= \sum_{j \geq 0} \mathbb{E} [e_j(t_j, \omega)^2] (t_{j+1} - t_j) \\ &= \mathbb{E} \left[\sum_{j \geq 0} e_j(t_j, \omega)^2 (t_{j+1} - t_j) \right] \\ &= \mathbb{E} \left[\int_s^r \phi(t, \omega)^2 dt \right]. \end{aligned}$$

□

A partir dos resultados obtidos para funções elementares podemos estender a definição de integral estocásticas para funções um pouco mais gerais. A aproximação de uma função $f \in \mathbb{M}^2$ será apresentada em três etapas, as quais são dadas pelos lemas abaixo.

Lema 2.1 *Seja $g \in \mathbb{M}^2(0, T)$ uma função limitada tal que $g(\cdot, \omega)$ é contínua para todo $\omega \in \Omega$. Então, existe uma sequência de funções elementares $\{\phi_n\}_{n \in \mathbb{N}} \subset \mathbb{M}^2(0, T)$ tal que*

$$\lim_{n \rightarrow +\infty} \mathbb{E} \left[\int_0^T (g - \phi_n)^2 dt \right] = 0. \quad (2.10)$$

Demonstração: Basta definir a sequência $\phi_n(t, \omega) := \sum_j g(t_j, \omega) \cdot \mathbf{1}_{[t_j, t_{j+1})}(t)$. Desta forma, ϕ_n é

elementar para todo $n \in \mathbb{N}$. Como g é contínua para cada $\omega \in \Omega$, segue que $(\phi_n - g)^2 \rightarrow 0$ para cada ω . Por fim, o resultado é obtido a partir do teorema da convergência dominada.

□

Lema 2.2 *Seja $h \in \mathbb{M}^2(0, T)$ uma função limitada. Então, existe uma sequência de funções limitadas $\{g_n\}_{n \in \mathbb{N}} \subset \mathbb{M}^2(0, T)$, tal que $g_n(\cdot, \omega)$ é contínua para quaisquer $\omega \in \Omega$ e $n \in \mathbb{N}$, e*

$$\lim_{n \rightarrow +\infty} \mathbb{E} \left[\int_0^T (h - g_n)^2 dt \right] = 0. \quad (2.11)$$

Demonstração: Como h é uma função limitada, podemos obter uma constante $M > 0$ tal que $|h(t, \omega)| \leq M$, para quaisquer $t \in [0, T]$ 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

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

Definindo

$$g_n(t, \omega) = \int_0^t \phi_n(s-t) h(s, \omega) ds,$$

temos que $g_n(\cdot, \omega)$ é uma função contínua para quaisquer $\omega \in \Omega$ e $n \in \mathbb{N}$, e $g_n(t, \cdot)$ é \mathcal{F}_t^B -mensurável pois $h(s, \cdot)$ não é avaliada em nenhum instante posterior a t . Além disso, g_n é limitada

$$\begin{aligned} \left| \int_0^t \phi_n(s-t) h(s, \omega) ds \right| &\leq \int_0^t \phi_n(s-t) |h(s, \omega)| ds \\ &\leq \int_0^t \phi_n(s-t) M ds \leq M \int_{\mathbb{R}} \phi_n(s-t) ds = M, \end{aligned}$$

e podemos verificar também (Hoffman (1962)) que

$$\lim_{n \rightarrow +\infty} \int_0^T |h(t, \omega) - g_n(t, \omega)|^2 dt = 0, \quad (2.12)$$

pois as funções ϕ_n definem aproximações da distribuição generalizada delta de Dirac. Por fim, o resultado segue do teorema da convergência dominada.

□

Lema 2.3 *Seja $f \in \mathbb{M}^2(0, T)$ uma função arbitrária. Então, existe uma sequência de funções limitadas $\{h_n\}_n \subset \mathbb{M}^2(0, T)$ tal que*

$$\lim_{n \rightarrow +\infty} \mathbb{E} \left[\int_0^T (f - h_n)^2 dt \right] = 0. \quad (2.13)$$

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

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

e aplicar o teorema da convergência dominada.

□

Os lemas anteriores nos permitem concluir que, dada $f \in \mathbb{M}^2(0, T)$, $T < \infty$, podemos obter uma sequência de funções elementares $\phi_n \in \mathbb{M}^2(0, T)$ tal que

$$\mathbb{E} \left[\int_0^t |f - \phi_n|^2 dt \right] \rightarrow 0. \quad (2.14)$$

Considerando o limite em $L^2(\Omega, \mathcal{F}, P)$, podemos então definir

$$I(f) = \int_0^T f(t, \omega) dB_t(\omega) := \lim_{n \rightarrow +\infty} \int_0^T \phi_n(t, \omega) dB_t(\omega). \quad (2.15)$$

O limite está bem definido já que a isometria de Itô garante que $\left\{ \int_0^T \phi_n(t, \omega) dB_t(\omega) \right\}_{n \in \mathbb{N}}$ é uma sequência de Cauchy no espaço de Hilbert $L^2(\Omega, \mathcal{F}, P)$

Definição 2.17 *Seja $f \in \mathbb{M}^2(0, T)$. A integral de Itô de f com respeito a um movimento Browniano $B = \{B_t, 0 \leq t \leq T\}$ é dada pelo limite em $L^2(\Omega, \mathcal{F}, P)$*

$$\int_0^T f(t, \omega) dB_t(\omega) = \lim_{n \rightarrow +\infty} \int_0^T \phi_n(t, \omega) dB_t(\omega), \quad (2.16)$$

onde $\{\phi_n\}_{n \in \mathbb{N}} \subset \mathbb{M}^2(0, T)$ é uma sequência de funções elementares tal que

$$\mathbb{E} \left[\int_0^T |f(t, \omega) - \phi_n(t, \omega)|^2 dt \right] \rightarrow 0. \quad (2.17)$$

Aproximar funções mais gerais por funções elementares é uma técnica que nos garante de imediato uma série de corolários devido à facilidade de se trabalhar com tais funções. Em particular, apresentaremos a seguir a extensão da isometria de Itô, verificaremos a linearidade da integral e obteremos sua esperança.

Corolário 2.1 *Sejam $f, g \in \mathbb{M}^2(0, T)$ e $\alpha \in \mathbb{R}$. Então*

$$(a) \quad \mathbb{E} \left[\left(\int_0^T f(t, \omega) dB_t(\omega) \right)^2 \right] = \mathbb{E} \left[\int_0^T f(t, \omega)^2 dt \right]; \quad (2.18)$$

$$(b) \quad \int_0^T (cf(t, \omega) + g(t, \omega)) dB_t(\omega) = c \cdot \int_0^T f(t, \omega) dB_t(\omega) + \int_0^T g(t, \omega) dB_t(\omega); \quad (2.19)$$

$$(c) \quad \mathbb{E} \left[\int_0^T f dB_t(\omega) \right] = 0. \quad (2.20)$$

Demonstração: Os dois primeiros itens são verificados de maneira imediata. Para demonstrar o item (c), consideramos inicialmente uma sequência de funções elementares $\{\phi_n\}_{n \in \mathbb{N}}$, tal que $\phi_n \rightarrow f$ em $L^2(\Omega, \mathcal{F}, P)$. Neste caso, ϕ é da forma:

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

onde e_1, e_2, \dots, e_m são variáveis aleatórias \mathcal{F}_{t_j} -mensuráveis, $j = 1, 2, \dots, m$. Utilizando as propriedades da esperança condicional, obtemos

$$\begin{aligned} \mathbb{E} \left[\int_0^T \phi_n(t, \omega) dB_t(\omega) \right] &= \mathbb{E} \left[\sum_{j=1}^m e_j(\omega) (B_{t_{j+1}} - B_{t_j})(\omega) \right] \\ &= \sum_{j=1}^m \mathbb{E} [e_j(\omega) (B_{t_{j+1}} - B_{t_j})(\omega)] \\ &= \sum_{j=1}^m \mathbb{E} \{ \mathbb{E} [e_j(\omega) (B_{t_{j+1}} - B_{t_j})(\omega) | \mathcal{F}_{t_j}] \} \\ &= \sum_{j=1}^m \mathbb{E} \{ e_j(\omega) \mathbb{E} [(B_{t_{j+1}} - B_{t_j})(\omega) | \mathcal{F}_{t_j}] \} = \sum_{j=1}^m \mathbb{E} [e_j(\omega) \cdot 0] = 0. \end{aligned}$$

O resultado segue tomando o limite em $L^2(\Omega, \mathcal{F}, P)$.

□

2.4 Fórmula de Itô

A fórmula de Itô descreve uma transformação utilizada para determinar o diferencial de funções de um processo de difusão. Basicamente, a fórmula funciona como uma versão estocástica da regra da cadeia utilizada de maneira ampla no cálculo clássico.

Definição 2.18 *Seja $X = \{X_t; 0 \leq t < +\infty\}$ um processo estocástico tal que, para quaisquer $0 \leq s < t \leq T$,*

$$X_t - X_s = \int_s^t f(t, X_t) dt + \int_s^t g(t, X_t) dB_t, \quad (2.21)$$

onde $f \in \mathbb{L}_\omega^1(0, T)$ e $g \in \mathbb{L}_\omega^2(0, T)$. Então dizemos que X é um processo de Itô com diferencial estocástico dX_t , em $[0, T]$, dado por

$$dX_t = f(t, X_t)dt + g(t, X_t)dB_t. \quad (2.22)$$

Antes de enunciar a fórmula de Itô, apresentamos um caso particular que será utilizado como resultado auxiliar na demonstração. Em seguida apresentamos uma versão estocástica da regra do produto.

Lema 2.4 Se $B = \{B_t; 0 \leq t < +\infty\}$ é um movimento Browniano padrão, então

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

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

Demonstração: Utilizando a definição da integral de Itô como um limite de funções elementares, mostra-se diretamente que

$$\int_s^r B_t dB_t = \frac{1}{2}(B_r^2 - B_s^2) - \frac{1}{2}(r - s) \quad \text{e} \quad \int_s^r t dB_t + \int_s^r B_t dt = rB_r - sB_s.$$

Em notação diferencial,

$$B_t dB_t = \frac{1}{2}d(B_t^2) - \frac{1}{2}dt \quad \text{e} \quad t dB_t + B_t dt = d(tB_t),$$

ou ainda,

$$d(B_t^2) = 2B_t dB_t + dt \quad \text{e} \quad d(tB_t) = B_t dt + t dB_t.$$

□

Teorema 2.6 (Regra do Produto) Suponha que $X = \{X_t; 0 \leq t < +\infty\}$ e $Y = \{Y_t; 0 \leq t < +\infty\}$ são processos de Itô cujos diferenciais estocásticos são dados por

$$\begin{cases} dX_t = f_1(t, X_t)dt + g_1(t, X_t)dB_t \\ dY_t = f_2(t, X_t^2)dt + g_2(t, Y_t)dB_t \end{cases},$$

$t \in [0, T]$, para qualquer $T > 0$. Então,

$$d(X_t Y_t) = X_t dY_t + Y_t dX_t + g_1(t, X_t)g_2(t, Y_t)dt. \quad (2.23)$$

Demonstração: Mostrar que

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

é equivalente a demonstrar que, para qualquer $[s, r] \subseteq [0, T]$, vale:

$$\begin{aligned} X_r Y_r &= X_s Y_s + \int_s^r (X_t f_2(t, Y_t) + Y_t f_1(t, X_t) + g_1(t, X_t) g_2(t, Y_t)) dt \\ &\quad + \int_s^r (X_t g_2(t, Y_t) + Y_t g_1(t, X_t)) dB_t. \end{aligned} \quad (2.24)$$

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

$$\begin{aligned} dX_t &= f_1 dt + g_1 dB_t \Rightarrow X_r = X_s + (r - s) f_1 + (B_r - B_s) g_1; \\ dY_t &= f_2 dt + g_2 dB_t \Rightarrow Y_r = Y_s + (r - s) f_2 + (B_r - B_s) g_2. \end{aligned}$$

Sem perda de generalidade, considere que $s = 0$. Neste caso,

$$X_r = X_s + r f_1 + B_r g_1; \quad (2.25)$$

$$Y_r = Y_s + r f_2 + B_r g_2. \quad (2.26)$$

Substituindo (2.25)-(2.26) na relação integral (2.24), obtemos:

$$\begin{aligned} \int_s^r [X_t f_2 + Y_t f_1 + g_1 g_2] dt + \int_s^r [X_t g_2 + Y_t g_1] dB_t &= \\ &= \int_s^r \{ [X_s + t f_1 + B_r g_1] f_2 + [Y_s + t f_2 + B_r g_2] f_1 + g_1 g_2 \} dt \\ &\quad + \int_s^r \{ [X_s + t f_1 + B_r g_1] g_2 + [Y_s + t f_2 + B_r g_2] g_1 \} dB_t \\ &= X_s (r f_2 + B_r g_2) + Y_s (r f_1 + B_r g_1) + r^2 f_1 f_2 + r g_1 g_2 \\ &\quad + (f_1 g_2 + f_2 g_1) \left(\int_s^r t dB_t + \int_s^r B_t dt \right) + 2 g_1 g_2 \int_s^r B_t dB_t. \end{aligned}$$

Aplicando o Lema (2.4),

$$\begin{aligned} \int_s^r [X_t f_2 + Y_t f_1 + g_1 g_2] dt + \int_s^r [X_t g_2 + Y_t g_1] dB_t &= \\ &= X_s f_2 r + X_s g_2 B_r + Y_s f_1 r + Y_s g_1 B_r + f_1 g_2 r B_r + g_1 f_2 r B_r + g_1 g_2 B_r^2 + f_1 f_2 r^2 \\ &= (X_s + r f_1 + B_r g_1) (Y_s + r f_2 + B_r g_2) - X_s Y_s \\ &= X_r Y_r - X_s Y_s, \end{aligned}$$

como queríamos mostrar. Considere 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 nos intervalos onde f_i e g_i são constantes de modo a obter,

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

$n = 0, 1, \dots, N-1$, onde $\tau = \{0 = t_0 < t_1 < \dots < t_N = T\}$ é uma partição do intervalo $[0, T]$, $T > 0$. Para o caso geral, utilizamos 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 considerar o limite em probabilidade. Da definição da integral de Itô, obtemos

$$\int_s^r [X_t f_2(t, Y_t) + Y_t f_1(t, X_t) + g_1(t, X_t) g_2(t, X_t)] dt + \int_s^r [X_t g_2(t, Y_t) + Y_t g_1(t, X_t)] dB_t = X_r Y_r - X_s Y_s,$$

como queríamos mostrar.

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

□

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:

$$\int_s^r X_t dY_t = X_t Y_t \Big|_s^r - \int_s^r Y_t dX_t - \int_s^r g_1(t, X_t) g_2(t, Y_t) dt.$$

O coeficiente $g_1 g_2 dt$ presente em (2.23) pode ser interpretado como um termo de correção devido aos efeitos aleatórios do movimento Browniano. Note que se $g_1 \equiv g_2 \equiv 0$, então voltaríamos ao método clássico de integração por partes:

$$\int_s^r X_t dY_t = X_t Y_t \Big|_s^r - \int_s^r Y_t dX_t,$$

onde X e Y são processos absolutamente contínuos.

A regra do produto pode ser adaptada naturalmente a potências X^n de um processo de Itô e, conseqüentemente, a polinômios.

Teorema 2.7 *Seja $u : \mathbb{R} \rightarrow \mathbb{R}$ uma função da forma $u(x) = x^n$, $n \in \mathbb{N}$. Se $X = \{X_t; 0 \leq t \leq T\}$, $T > 0$, é um processo de Itô satisfazendo*

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

então o diferencial estocástico do processo $u(X_t)$ é dado por

$$du_t = d(X_t^n) = nX_t^{n-1}dX_t + \frac{1}{2}n(n-1)X_t^{n-2}g^2(t, X_t)dt. \quad (2.27)$$

Demonstração: Seja $\mathcal{X} = \left\{ n \in \mathbb{N}; d(X_t^n) = nX_t^{n-1}dX + \frac{1}{2}n(n-1)X_t^{n-2}g^2(t, X_t)dt \right\}$. Verificaremos por indução em n , que $\mathcal{X} = \mathbb{N}$. Para $n = 1$, temos:

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

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

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

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

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

Aplicando a regra do produto e a hipótese de indução:

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

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

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

Concluimos assim que $\mathcal{X} = \mathbb{N}$, isto é, (2.23) é válida para potências $u(x) = x^n$, $n \in \mathbb{N}$.

□

Corolário 2.2 Seja $X = \{X_t; 0 \leq t \leq T\}$, $T > 0$, um processo de Itô conforme o teorema ante-

rior. Se Q é um polinômio qualquer, então o diferencial estocástico $d(Q(X_t))$ é dado por

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

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.

□

A fórmula de Itô é uma extensão dos resultados acima para funções $u \in C^{1,2}([0, +\infty) \times \mathbb{R})$ mais genéricas. Esta extensão é natural pois estamos interessados em processos com trajetórias contínuas, as quais podem ser aproximadas por polinômios segundo o teorema de Stone-Weirstrass.

Teorema 2.8 (Fórmula de Itô) *Seja $X = \{X_t; 0 \leq t < T\}$, $T > 0$, um processo de Itô com diferencial estocástico*

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

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

$$Y_t := u(t, X_t).$$

Neste caso, Y_t é um processo de Itô satisfazendo

$$\begin{aligned} dY &= \frac{\partial u}{\partial t}(t, X_t)dt + \frac{\partial u}{\partial x}(t, X_t)dX_t + \frac{1}{2} \frac{\partial^2 u}{\partial X^2}(t, X_t)(dX_t)^2 \\ &= \left(\frac{\partial u}{\partial t}(t, X_t) + \frac{\partial u}{\partial x}(t, X_t)f(t, X_t) + \frac{1}{2} \frac{\partial^2 u}{\partial x^2}(t, X_t)g^2(t, X_t) \right) dt \\ &\quad + \frac{\partial u}{\partial x}(t, X_t)g(t, X_t)dB_t. \end{aligned} \tag{2.28}$$

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

$$u(t, x) = v(x)w(t),$$

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:

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

Pelo corolário (2.2), podemos reescrever esta equação do seguinte modo:

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

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

$$u(x, t) = \sum_{i=1}^m v^i(x) w^i(t), \quad (2.29)$$

onde v^i e w^i são polinômios com respeito a x e t , respectivamente. Ou seja, 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

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

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

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

Passando ao limite, concluímos o resultado.

□

2.5 Transformação de Lamperti

A transformação de Lamperti é uma aplicação particular da fórmula de Itô que nos permite estabilizar o coeficiente g do diferencial estocástico

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

isto é, o método transforma o processo inicial X em um processo de Itô Y que satisfaz

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

Todos os resultados (simulação e inferência) obtidos a partir desta transformação serão dados em função do processo Y , logo, é necessário que a transformação seja inversível para que possamos adaptar as conclusões ao processo original.

Teorema 2.9 (Transformação de Lamperti) *Seja X um processo de Itô conforme o teorema (2.8) e assumamos que a transformação*

$$Y_t := S(t, X_t) = \int_{\xi}^x \frac{1}{g(t, u)} du \Big|_{x=X_t}, \quad \forall t \in [0, T], \quad (2.30)$$

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 é um processo de Itô cujo diferencial estocástico é dado por:

$$\begin{aligned} dY_t &= \left(\frac{f(t, X_t)}{g(t, X_t)} - \frac{1}{2} \frac{\partial g}{\partial x}(t, X_t) - \int_{\xi}^{X_t} \frac{1}{g^2(t, u)} \frac{\partial g}{\partial t}(t, u) du \right) dt + dB_t \\ &= \hat{f}(t, X_t)dt + dB_t = \hat{f}(t, S^{-1}(Y_t)) dt + dB_t. \end{aligned} \quad (2.31)$$

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

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

É imediato que a escolha

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

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

$$\frac{\partial S}{\partial t}(t, x) = \int_{\xi}^x \frac{\partial}{\partial t} \left(\frac{1}{g(t, u)} \right) du; \quad \frac{\partial S}{\partial x}(t, x) = \frac{1}{g(t, x)}; \quad \frac{\partial^2 S}{\partial x^2}(t, x) = -\frac{1}{g^2(t, x)} \frac{\partial g}{\partial x}(t, x).$$

□

As condições para a validade das hipóteses do teorema acima não são únicas. Um conjunto bem amplo de condições é dado por Luschy (2006):

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

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. Lembramos que a maior dificuldade no trato de um processo de Itô está relacionada ao cálculo da integral estocástica associada a este processo. Ao aplicar a transformação de Lamperti este problema fica resolvido pois a integral de Itô de uma constante pode ser resolvida facilmente. Quanto aos métodos numéricos, mostraremos mais adiante que a transformação de Lamperti não só melhora a taxa de convergência de certos métodos, mas também torna determinados algoritmos aplicáveis a classes mais abrangentes de processos de Itô.

Exemplo 2.1 Seja $X = \{X_t; 0 \leq t \leq T\}$, $T > 0$, um processo de Itô satisfazendo

$$dX_t = \lambda X_t dt + \mu X_t dB_t, \quad X_0 = 1, \quad 0 \leq t \leq T,$$

onde $\lambda, \mu \in \mathbb{R}$. Aplicando a transformação de Lamperti, obtemos

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

e

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

Assim, Y_t será um processo de Itô dado por

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

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

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

ou seja,

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

Podemos agora obter o processo original:

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

onde LN denota a distribuição log-normal.

2.6 Transformação de Cameron-Martin-Girsanov

As transformações apresentadas até aqui são aplicadas em um processo estocástico X e devolvem um segundo processo \hat{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.

O resultado apresentado nesta seção é um importante dispositivo que transforma uma medida P em uma segunda medida Q de modo a garantir que se B é um movimento Browniano com respeito a P , então uma determinada translação T aplicada a B é um movimento Browniano com respeito a Q . Como motivação para este resultado, suponha 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 e possui diversas aplicações. Em particular, dado um vetor $\underline{\mu} = (\mu_1, \mu_2, \dots, \mu_n) \in \mathbb{R}^n$, considere a seguinte função

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

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

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

é uma medida. Mais precisamente, vale a relação

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

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,

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

O estudo do movimento Browniano mostra que há uma relação forte entre este processo estocástico e as variáveis aleatórias Gaussianas. Desta forma, é natural procurarmos transformações similares que garantam a propriedade da invariância em se tratando do movimento Browniano.

Teorema 2.10 (Cameron-Martin-Girsanov) *Seja $u \in \mathbb{L}^2(0, T)$, $0 \leq T < +\infty$, e assumamos que*

$$M_t := \exp \left\{ - \int_0^t u(s, \omega) dB_s - \frac{1}{2} \int_0^t u^2(s, \omega) ds \right\} \quad (2.32)$$

é uma martingale. Defina o processo $\hat{B} = \{\hat{B}_t, 0 \leq t < +\infty\}$ por

$$\hat{B}_t := B_t - \int_0^t u(s, \omega) ds, \quad 0 \leq t < +\infty. \quad (2.33)$$

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

$$Q_T(A) = \mathbb{E} [I_A M_T], \quad \forall A \in \mathcal{F}_T. \quad (2.34)$$

A demonstração deste teorema utiliza a caracterização de Lévy do movimento Browniano, cuja descrição exige resultados matemáticos não abordados neste texto. Uma prova formal pode ser obtida em Karatzas e Shrev (1991) ou Øksendal (2010). Mesmo omitindo a demonstração é possível analisar o teorema e entender o papel das hipóteses e a importância do resultado. Note 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

$$\begin{aligned} \int_{\Omega} \mathbf{1}_A(\omega) M_T(\omega) P(d\omega) &= \mathbb{E} [\mathbf{1}_A M_T] \\ &= \mathbb{E} [\mathbb{E} [\mathbf{1}_A M_T | \mathcal{F}_t]] = \mathbb{E} [\mathbf{1}_A \mathbb{E} [M_T | \mathcal{F}_t]] \\ &= \mathbb{E} [\mathbf{1}_A M_t] = \int_{\Omega} \mathbf{1}_A(\omega) M_t(\omega) P(d\omega), \end{aligned} \quad (2.35)$$

o que garante que

$$M_T P(d\omega) = M_t P(d\omega), \quad \forall t \in [0, T], \quad (2.36)$$

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

$$Q_T(A) = Q_t(A); \quad \forall A \in \mathcal{F}_t, 0 \leq t \leq T. \quad (2.37)$$

Além disso,

$$\mathbb{E}[M_t] = \mathbb{E}[M_0] = \mathbb{E}[1] = 1, \quad 0 \leq t \leq T, \quad (2.38)$$

pois todo martingale possui esperança constante. Logo,

$$Q_T(\Omega) = \mathbb{E}_{Q_T}[\mathbf{1}_\Omega] = \mathbb{E}_P[M_T] = 1, \quad (2.39)$$

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 2.11 (Novikov (1972)) *Sejam $B = \{B_t; 0 \leq t < +\infty\}$ um movimento Browniano em (Ω, \mathcal{F}, P) e $u = \{u(t, \omega); 0 \leq t < +\infty\}$ um processo em $\mathbb{L}^2(0, T)$, para algum $T > 0$. Se*

$$\mathbb{E} \left[\exp \left(\frac{1}{2} \int_0^T u^2(t, \omega) dt \right) \right] < +\infty; \quad 0 \leq T < +\infty, \quad (2.40)$$

então M_t definido por (2.32) é um martingale.

Dentre as inúmeras aplicações do teorema (2.10), uma das mais importantes é a obtenção da densidade (derivada de Radon-Nykodim) das trajetórias de um processo de Itô com respeito as trajetórias de um segundo processo. Ao obter esta densidade estamos nos habilitando a realizar simulações a partir de métodos como amostragem por rejeição ou amostragem por importância, por exemplo, e\ou fazer inferência a partir dos métodos baseados em verossimilhança.

Teorema 2.12 *Sejam $u : [0, +\infty) \times \mathbb{R} \rightarrow \mathbb{R}$, M_t e \hat{B}_t conforme o teorema (2.10). Denote por P a medida induzida pelo processo de Itô*

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

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

$$g(t, X_t)u(t, X_t) = \hat{f}(t, X_t) - f(t, X_t), \quad (2.41)$$

então, com respeito à medida Q_T dada por

$$dQ_T(\omega) = M_T dP(\omega), \quad (2.42)$$

X possui o diferencial estocástico

$$dX_t = \hat{f}(t, X_t)dt + g(t, X_t)d\hat{B}_t, \quad 0 \leq t \leq T. \quad (2.43)$$

Demonstração: Pelo teorema (2.10), Q_T é uma medida de probabilidade em \mathcal{F}_T e \hat{B}_t é um movimento Browniano com respeito a Q_T . Desta forma, o processo de Itô (2.43) está bem

definido. Além disso,

$$\begin{aligned}
 dX_t &= f(t, X_t)dt + g(t, X_t)dB_t \\
 &= f(t, X_t)dt + g(t, X_t) (d\hat{B}_t + u(t, X_t)dt) \\
 &= (f(t, X_t) + g(t, X_t)u(t, X_t)) dt + g(t, X_t)d\hat{B}_t \\
 &= \hat{f}(t, X_t)dt + g(t, X_t)d\hat{B}_t.
 \end{aligned}$$

□

Exemplo 2.2 (Densidade com respeito à medida de Wiener) Sejam $X = \{X_t; 0 \leq t \leq T\}$ um processo de Itô com diferencial estocástico

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

e Q_T a medida de probabilidade induzida por X . Nosso objetivo é obter a derivada de Radon-Nykodim de Q_T com respeito à medida de Wiener \mathbb{W} induzida pelo movimento Browniano

$$dY_t = dB_t, \quad 0 \leq t \leq T.$$

Definindo

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

segue, do teorema (2.12), que

$$\frac{dQ_T}{d\mathbb{W}}(\omega) = M_T(B_t(\omega)) = \exp \left(\int_0^T f(t, B_t)dB_t - \frac{1}{2} \int_0^T f^2(t, B_t)dt \right). \quad (2.44)$$

▷

Exemplo 2.3 (Verossimilhança de uma trajetória) Suponha que $X = \{X_t; 0 \leq t \leq T\}$ é um processo de Itô que induz uma medida de probabilidade Q_T e cujo diferencial estocástico

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

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

$$dY_t = g(Y_t)dB_t, \quad 0 \leq t \leq T.$$

Definindo

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

segue, do teorema (2.12), que

$$\frac{dQ_T}{dP}(\omega) = M_T(Y_t(\omega)) = \left(\int_0^T \frac{f(Y_t, \theta)}{g^2(Y_t)} dY_t - \frac{1}{2} \int_0^T \frac{f^2(Y_t, \theta)}{g^2(Y_t)} dt\right). \quad (2.45)$$

Este procedimento é muito útil em se tratando de inferência em processos de difusão.

▷

3 *Equações Diferenciais Estocásticas*

Sejam $f : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ e $g : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ funções que satisfazem as condições (i) e (ii) da definição (2.14). Se $X = \{X_t; 0 \leq t \leq T\}$ é um processo estocástico tal que

$$dX_t = f(t, X_t)dt + g(t, X_t)dB_t, \quad t \in [0, T], \quad (3.1)$$

$$P(X_0 = \xi_0) = 1, \quad (3.2)$$

então dizemos que X satisfaz a *equação diferencial estocástica* (3.1) e a condição inicial (3.2). As funções f e g são denominadas coeficiente de deslocamento (ou deriva) e coeficiente de difusão (ou volatilidade), respectivamente. O problema (3.1)-(3.2) pode ainda ser expresso na forma integral

$$X_t = X_0 + \int_0^t f(s, X_s)ds + \int_0^t g(s, X_s)dB_s, \quad P(X_0 = \xi_0) = 1. \quad (3.3)$$

Note que se $g \equiv 0$, então nosso problema se reduz a uma equação diferencial aleatória

$$X_t = X_0 + \int_0^t f(t, X_t)dt, \quad (3.4)$$

cujas soluções são trajetórias contínuas que se iniciam no ponto inicial observado. Destacamos que não há nenhuma restrição ao fato de ξ_0 ser uma variável aleatória uma vez que, condicionado a uma observação desta variável, o instante inicial é puramente determinístico. Do estudo de equações diferenciais determinísticas, sabemos que, ao exigir que $x \mapsto f(\cdot, x)$ seja localmente Lipschitz contínua e que f seja limitada em subconjuntos compactos de $[0, T] \times \mathbb{R}$, estamos garantindo condições suficientes para que o problema (3.3) tenha solução e que tal solução seja única. A seguir, verificaremos que um resultado similar de existência e unicidade pode ser obtido para o problema (3.1)-(3.2), mas, primeiramente, devemos definir com um pouco mais de formalismo o significado de solução deste problema.

3.1 Soluções Fortes

Seja $B = \{B_t; 0 \leq t < +\infty\}$ um movimento Browniano padrão definido em (Ω, \mathcal{F}, P) . Associaremos a este espaço a filtração aumentada

$$\mathcal{F}_t := \sigma(\mathcal{G}_t \cup \mathcal{N}), \quad (3.5)$$

onde $\mathcal{G}_t = \sigma(\xi_0, B_s; 0 \leq s \leq t)$ e $\mathcal{N} = \{N \subseteq \Omega; \exists G \in \mathcal{G}_\infty \Rightarrow N \subseteq G, P(G) = 0\}$, e assumiremos que ele é rico o suficiente para construirmos uma variável aleatória ξ_0 independente de \mathcal{G}_t .

Definição 3.1 *Uma solução forte para o problema (3.1)-(3.2), no espaço (Ω, \mathcal{F}, P) e com respeito a uma realização (trajetória fixa) do movimento Browniano B e à condição inicial ξ_0 , é um processo estocástico $X = \{X_t; 0 \leq t < +\infty\}$, com 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

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

isto é, se quaisquer duas soluções forem indistinguíveis.

As duas últimas condições da definição acima são consideradas com o objetivo de trabalharmos com um problema bem posto. Mais precisamente, (ii) exige que a solução se comporte segundo a variável aleatória ξ_0 no instante inicial de observação e a condição (iii) garante o bom comportamento da integral de Riemann e da isometria de Itô. A hipótese mais forte é imposta por (i). Esta suposição nos permite entender o processo estocástico X como um sistema dinâmico descrito pelo par ordenado (f, g) , alimentado por B e ξ_0 . Neste caso, condicionado às observações de B e ξ_0 , faz sentido pensarmos em resultados de unicidade para X .

Teorema 3.1 *(Existência e Unicidade) Suponha que os coeficientes $f(t, x)$ e $g(t, x)$ satisfazem as condições*

$$|f(t, x) - f(t, y)| + |g(t, x) - g(t, y)| \leq K|x - y|, \quad (3.6)$$

$$|f(t, x)|^2 + |g(t, x)|^2 \leq K^2(1 + x^2), \quad (3.7)$$

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) , seja ξ_0 uma variável aleatória independente do movimento Browniano $B = \{B_t; 0 \leq t < +\infty\}$, e com segundo momento finito

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

Considere também que a filtração (3.5) é adotada. Neste caso, existe um processo estocástico

$$X = \{X_t; 0 \leq t < +\infty\}, \quad (3.8)$$

\mathcal{F}_t -adaptado, o qual representa a única solução forte do problema (3.1)-(3.2) relativo às observações de B e ξ_0 . Além disso,

$$\mathbb{E} \left[\int_0^T |X_t|^2 dt \right] < +\infty, \quad \forall T \in [0, +\infty). \quad (3.9)$$

A demonstração formal deste teorema é uma adaptação do resultado utilizado no caso determinístico. Basicamente, devemos construir uma sequência de aproximações $\{X^{(k)}\}_{k \in \mathbb{N}}$ tal que $X_t^{(0)} \equiv \xi_0$ e

$$X_t^{(k+1)} := \xi_0 + \int_0^t f(s, X_s^{(k)}) ds + \int_0^t g(s, X_s^{(k)}) dB_s.$$

O argumento é concluído provando que esta sequência de fato converge para a solução da equação. O detalhamento formal pode ser encontrado em Friedman (1975).

Em termos de soluções fortes, este trabalho tem como objetivo o estudo de métodos numéricos para a aproximação das trajetórias correspondentes a estas soluções. Alguns métodos para o cálculo analítico das soluções fortes podem ser obtidos em Kloeden & Platen (1999), porém, não serão abordados no texto.

3.2 Soluções Fracas

É possível obter uma segunda forma de solução mais fraca que a definida na seção anterior, porém, ainda muito útil tanto na teoria quanto nas aplicações. Esse tipo de solução é interessante pois sua existência está condicionada a hipóteses mais fracas com respeito à EDE de interesse.

Definição 3.2 Uma solução fraca para a equação (3.1) é uma tripla $(X, B), (\Omega, \mathcal{F}, P), \{\mathcal{H}_t\}_{t \geq 0}$, 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) $X = \{X_t; 0 \leq t < +\infty\}$ é um processo estocástico contínuo e $B = \{B_t; 0 \leq t < +\infty\}$ é um movimento Browniano, ambos \mathcal{H}_t -adaptados;

(c) X satisfaz

$$X_t = X_0 + \int_0^t f(s, X_s) ds + \int_0^t g(s, X_s) dB_s; \quad 0 \leq t < +\infty. \quad (3.10)$$

Diferentemente do caso das soluções fortes, a filtração na definição (3.2) não é necessariamente a filtração aumentada gerada pelo movimento Browniano em união com a condição inicial. Neste caso, dado um instante t , a variável aleatória $X_t(\omega)$ não é necessariamente dada por uma função mensurável do caminho Browniano e da condição inicial. Por outro lado, como B é adaptado à filtração $\{\mathcal{H}_t\}_{t \geq 0}$, a solução $X_t(\omega)$ no instante t não é suficiente para antecipar o futuro do movimento Browniano. Além disso, a informação utilizada para determinar $X_t(\omega)$ deve ser baseada no caminho amostral $\{B_s(\omega); 0 \leq s \leq t\}$ e da condição inicial $\xi_0(\omega)$, e qualquer informação adicional deve ser independente de $\{B_\theta(\omega) - B_t(\omega); t \leq \theta < +\infty\}$.

Definição 3.3 Dizemos que o problema (3.1)-(3.2) possui unicidade no sentido da distribuição de probabilidade se, para qualquer dupla de soluções fracas (X, B) , (Ω, \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, isto é,

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

os dois processos X, \hat{X} possuírem a mesma distribuição.

A existência de soluções fracas não implica a existência de soluções fortes, no entanto, a recíproca a verdadeira uma vez que uma solução forte em particular cumpre todas as propriedades da definição (3.2). Desta forma, todos os critérios de existência de soluções fortes são aplicáveis à existência de soluções fracas, porém, critérios mais fracos podem ser estabelecidos com o objetivo de garantirmos a existência e unicidade de soluções fracas. Em particular, estaremos interessados no teorema seguinte o qual nos fornecerá um ferramental importante para a simulação de soluções de uma classe de equações que será discutida mais adiante.

Teorema 3.2 Considere a equação diferencial estocástica

$$dX_t = f(t, X_t) dt + dB_t; \quad 0 \leq t \leq T, \quad (3.11)$$

onde T é uma constante positiva, $B = \{B_t; 0 \leq t \leq T\}$ é um movimento Browniano e $f(t, x)$ é uma função Borel-mensurável.

(a) Se f satisfaz

$$|f(t, x)| \leq K(1 + |x|); \quad 0 \leq t \leq T, x \in \mathbb{R} \quad (3.12)$$

para alguma constante K . Então, para qualquer medida μ em $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, a equação (3.11) 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 (3.11) com a mesma distribuição inicial. Se

$$P^{(i)} \left(\int_0^T |f(t, X_t)|^2 dt < +\infty \right) = 1; \quad i = 1, 2, \quad (3.13)$$

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

A demonstração deste teorema é baseada no teorema de Girsanov e pode ser obtida em Karatzas & Shreve (1991). Equações da forma (3.11) serão de interesse quando lidarmos com os métodos de simulação exata de soluções.

4 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,

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

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.

4.1 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 derivável, 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

$$f(x_0 + h) = f(x_0) + \sum_{i=1}^n \frac{1}{i!} f^{(i)}(x_0) \cdot h^i + o(h^n). \quad (4.1)$$

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

$$f(x_0 + h) \approx f(x_0) + \sum_{i=1}^n \frac{1}{i!} f^{(i)}(x_0) \cdot h^i. \quad (4.2)$$

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

$$dX_t = f(t, X_t)dt + g(t, X_t)dB_t, \quad X_0 = x_0, t \in [0, T]. \quad (4.3)$$

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

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

onde f e g são, respectivamente, os coeficientes de deslocamento e difusão da equação (4.3). Assim, dado um processo estocástico $Y_t = z(t, X_t)$, $z \in C^{1,2}([0, T] \times \mathbb{R})$, temos que

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

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

$$X_t = X_s + \int_s^t f(u, X_u)du + \int_s^t g(u, X_u)dB_u. \quad (4.4)$$

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

$$\begin{aligned}
X_t &= X_s + \int_s^t \left[f(s, X_s) + \int_s^u (\mathcal{L} \circ f)(s_1, X_{s_1}) ds_1 + \int_s^u (\mathcal{M} \circ f)(s_1, X_{s_1}) dB_{s_1} \right] du \\
&\quad + \int_s^t \left[g(s, X_s) + \int_s^u (\mathcal{L} \circ g)(s_1, X_{s_1}) ds_1 + \int_s^u (\mathcal{M} \circ g)(s_1, X_{s_1}) dB_{s_1} \right] dB_u \\
&= X_s + f(s, X_s) \cdot h + g(s, X_s)(B_t - B_s) \\
&\quad + \int_s^t \left[\int_s^u (\mathcal{L} \circ f)(s_1, X_{s_1}) ds_1 + \int_s^u (\mathcal{M} \circ f)(s_1, X_{s_1}) dB_{s_1} \right] du \\
&\quad + \int_s^t \left[\int_s^u (\mathcal{L} \circ g)(s_1, X_{s_1}) ds_1 + \int_s^u (\mathcal{M} \circ g)(s_1, X_{s_1}) dB_{s_1} \right] dB_u. \tag{4.5}
\end{aligned}$$

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

$$X_t = X_s + f(s, X_s) \cdot h + g(s, X_s) \cdot (\sqrt{h}Z) + R_1(t, X_t), \tag{4.6}$$

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

$$\int_s^t \int_s^u (\mathcal{M} \circ g)(s_1, X_{s_1}) dB_{s_1} dB_u \tag{4.7}$$

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 (4.6) 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 $(\mathcal{M} \circ g)$:

$$d[(\mathcal{M} \circ g)(t, X_t)] = (\mathcal{L} \circ \mathcal{M} \circ g)(t, X_t) dt + (\mathcal{M}^2 \circ g)(t, X_t) dB_t. \tag{4.8}$$

Assim,

$$\begin{aligned}
(\mathcal{M} \circ g)(t, X_t) &= (\mathcal{M} \circ g)(s, X_s) + \int_s^t (\mathcal{L} \circ \mathcal{M} \circ g)(u, X_u) du + \int_s^t (\mathcal{M}^2 \circ g)(u, X_u) dB_u \\
&= g(s, X_s) \frac{\partial g}{\partial x}(s, X_s) + \hat{R}_1(t, X_t). \tag{4.9}
\end{aligned}$$

Substituindo o resultado acima na integral (4.7):

$$\begin{aligned} \int_s^t \int_s^u (\mathcal{M} \circ g)(s_1, X_{s_1}) dB_{s_1} dB_u &= \int_s^t \int_s^u \left(g(s, X_s) \frac{\partial g}{\partial x}(s, X_s) + \hat{R}_1(s_1, X_{s_1}) \right) dB_{s_1} dB_u \\ &= g(s, X_s) \frac{\partial g}{\partial x}(s, X_s) \int_s^t \int_s^u dB_{s_1} dB_u + R_2(t, X_t). \end{aligned} \quad (4.10)$$

A integral estocástica dupla que aparece na relação acima pode ser tratada analiticamente através das fórmulas obtidas na seção 2.4:

$$\begin{aligned} \int_s^t \int_s^u dB_{s_1} dB_u &= \int_s^t (B_u - B_s) dB_u \\ &= \int_s^t B_u dB_u - B_s \int_s^t dB_s = \left[\frac{B_u^2 - u}{2} \right]_s^t - B_s(B_t - B_s) \\ &= \frac{1}{2} [B_t^2 - B_s^2 - (t - s)] - B_s(B_t - B_s) = \frac{1}{2} [(B_t - B_s)^2 - h]. \end{aligned} \quad (4.11)$$

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

$$X_t = X_s + f(s, X_s) \cdot h + g(s, X_s) \cdot (\sqrt{h}Z) + \frac{1}{2} g(s, X_s) \frac{\partial g}{\partial x}(s, X_s) (Z^2 - h) + R_2(t, X_t). \quad (4.12)$$

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 & Platen (1999).

4.2 Métodos de Euler-Maruyama e Milstein

A aproximação discreta mais simples para um processo de Itô é caracterizada pelo *método de Euler*, também denominado *método de Euler-Maruyama*. Considere um processo de Itô $X = \{X_t; 0 \leq t \leq T\}$ que satisfaz a seguinte equação diferencial estocástica

$$\begin{aligned} dX_t &= f(t, X_t)dt + g(t, X_t)dB_t, \quad t \in [0, T], \\ P(X_0 = \xi_0) &= 1. \end{aligned}$$

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 simples, isto é

$$f(t, \omega) \approx \sum_{i=0}^{N-1} f(t_i, \omega) \mathbf{1}_{[t_i, t_{i+1})}(t); \quad g(t, \omega) \approx \sum_{i=0}^{N-1} g(t_i, \omega) \mathbf{1}_{[t_i, t_{i+1})}(t). \quad (4.13)$$

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

$$\begin{aligned} X_{t_{n+1}}^N &= X_{t_n}^N + f(t_n, X_{t_n}^N) (t_{n+1} - t_n) + g(t_n, X_{t_n}^N) (B_{t_{n+1}} - B_{t_n}), \quad t_n \in \tau, \\ P(X_0^N = X_0) &= 1. \end{aligned}$$

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

$$X_{t_{n+1}}^N = X_{t_n}^N + f(t_n, X_{t_n}^N) \Delta + g(t_n, X_{t_n}^N) \Delta B_{t_n}, \quad t_n \in \tau, \quad (4.14)$$

$$P(X_0^N = X_0) = 1. \quad (4.15)$$

Este procedimento de discretização nos permite obter densidades de transição:

$$X_{t_{n+1}}^N \mid \{X_{t_n}^N = x\} \sim N(x + f(t_n, x)\Delta, g^2(t_n, x)\Delta), \quad (4.16)$$

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 4.1 *Considere a equação diferencial estocástica linear*

$$\begin{aligned} dX_t &= \lambda X_t dt + \mu X_t dB_t, \quad 0 \leq t \leq T, \\ X_0 &= 1; \end{aligned}$$

cuja solução exata, obtida no exemplo (2.1), é dada por

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

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

$$\begin{aligned} X_{t_{n+1}}^N &= \lambda X_{t_n}^N \Delta + \mu X_{t_n}^N \Delta B_{t_n}, \quad t_n \in \tau, \\ X_0^N &= 1, \end{aligned}$$

ou ainda,

$$\begin{aligned} X_{t_{n+1}}^N &= X_{t_n}^N [\lambda \Delta + \mu \Delta B_{t_n}], \quad t_n \in \tau, \\ X_0^N &= 1. \end{aligned}$$

Caminhos amostrais do processo discretizado podem ser obtidos facilmente através das densidades de transição (4.16). O custo da simulação é basicamente o custo de se gerar N variáveis aleatórias normalmente distribuídas.

As figuras (4.1a)-(4.1b) ilustram, a seguir, uma aplicação do método de Euler-Maruyama para a simulação de uma trajetória correspondente ao problema descrito no exemplo acima considerando os parâmetros $\lambda = 4$ e $\mu = 2$. Esta primeira simulação é um bom indicativo da qualidade que a aproximação pode atingir.

Uma análise da expansão de Itô-Taylor apresentada anteriormente nos mostra que o método de Euler-Maruyama pode ser obtido através desta expansão se negligenciarmos o termo R_1 da relação (4.6). Intuitivamente, se fizermos uma analogia com o caso determinístico, deveríamos esperar que aproximações melhores podem ser obtidas se levarmos em consideração mais coeficientes na expansão de Itô-Taylor. Em particular, se negligenciarmos o termo R_2 em (4.12), chegamos à seguinte aproximação:

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

Este procedimento sugere o seguinte esquema de discretização:

$$\begin{aligned} X_{t_{n+1}}^N &= X_{t_n}^N + f(t_n, X_{t_n}^N) \Delta + g(t_n, X_{t_n}^N) \Delta B_{t_n} \\ &\quad + \frac{1}{2}g(t_n, X_{t_n}^N) g_x(t_n, X_{t_n}^N) [(\Delta B_{t_n})^2 - \Delta], \quad t_n \in \tau, \end{aligned} \quad (4.17)$$

$$P(X_0^N = X_0) = 1, \quad (4.18)$$

o qual é denominado *método de Milstein*.

Ao se basear na aproximação (4.12), o método de Milstein obtém um processo discretizado que leva em consideração todos os coeficientes que dependem linearmente do espaçamento Δ na expansão de Itô-Taylor, o que não ocorre com o método de Euler. Neste caso, é razoável afirmarmos que o método de Milstein permite o desenvolvimento de simulações de trajetórias com um espaçamento maior entre os instantes de tempo. Essa característica é importante pois pode gerar uma boa economia de recursos computacionais. Voltaremos nesta questão quando abordarmos os conceitos de convergência e estabilidade. Por hora, os experimentos ilustrados abaixo fornecem uma evidência de que o método de Milstein oferece uma aproximação mais

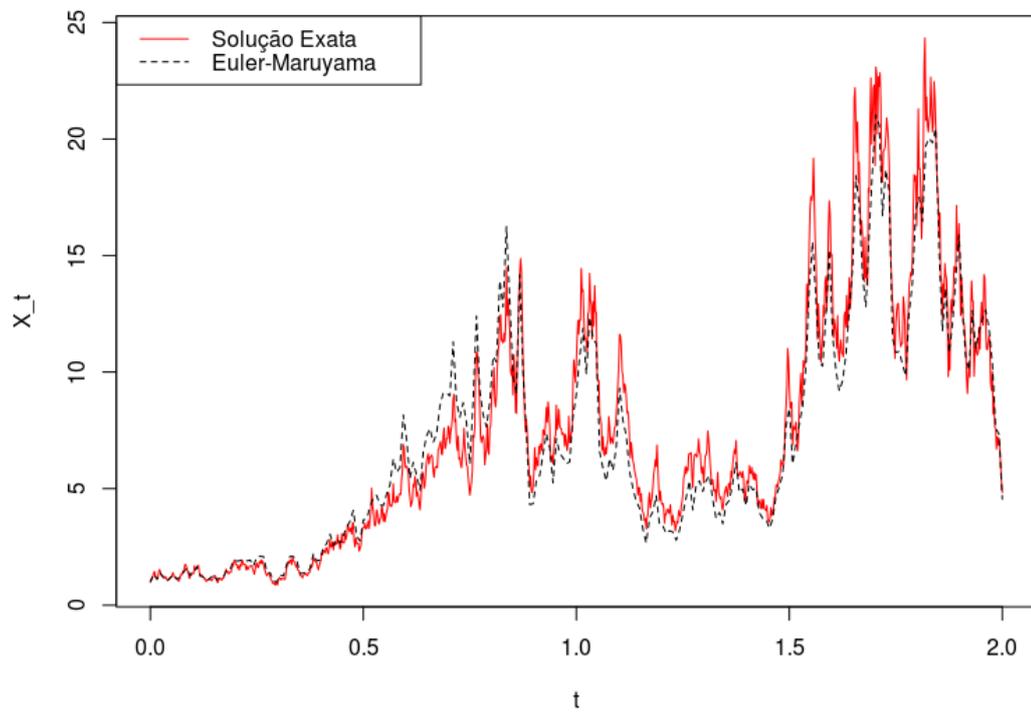
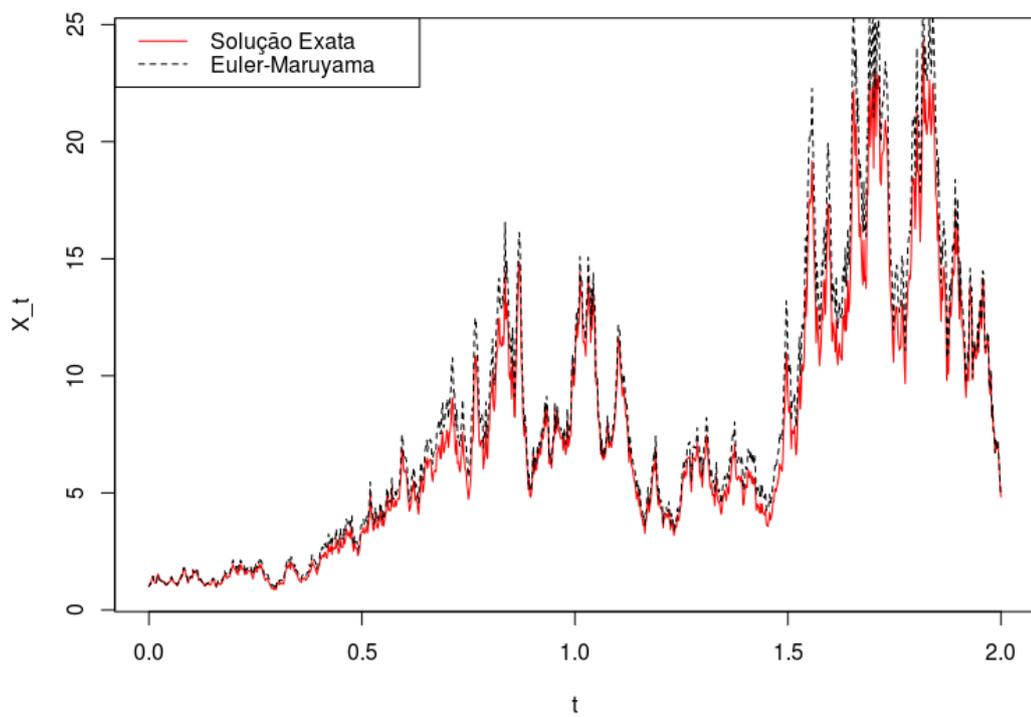
(a) Malha com $N = 257$ pontos(b) Malha com $N = 1025$ pontos

Figura 4.1: Exemplo de aproximação pelo método de Euler-Maruyama para o movimento Browniano Geométrico.

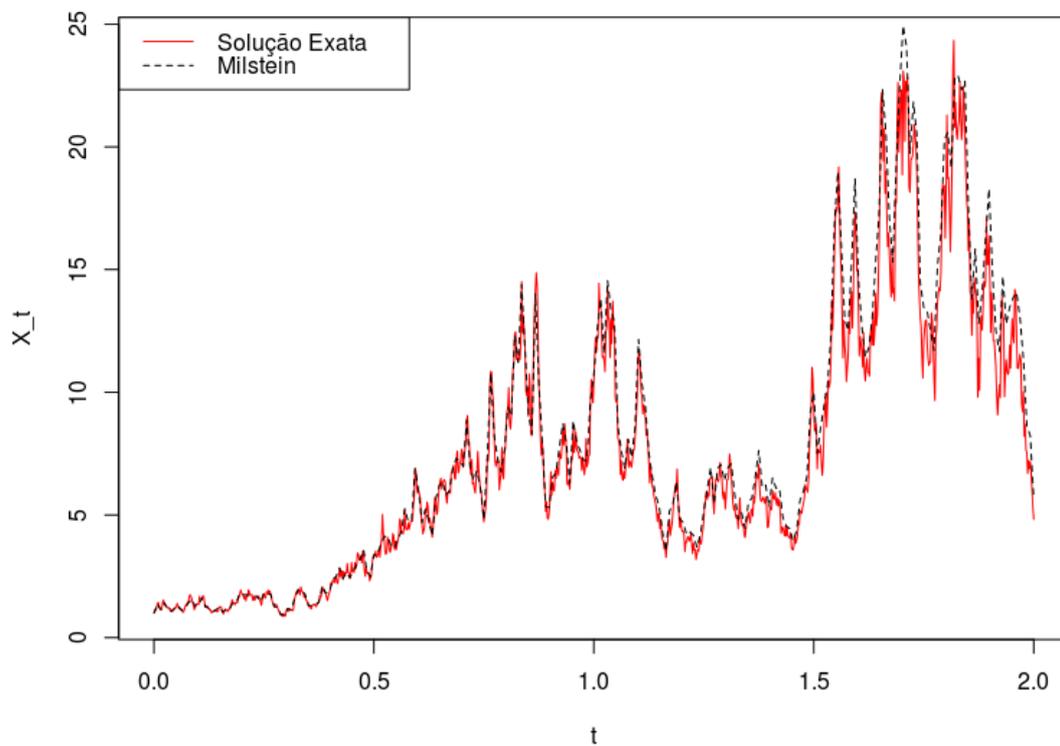
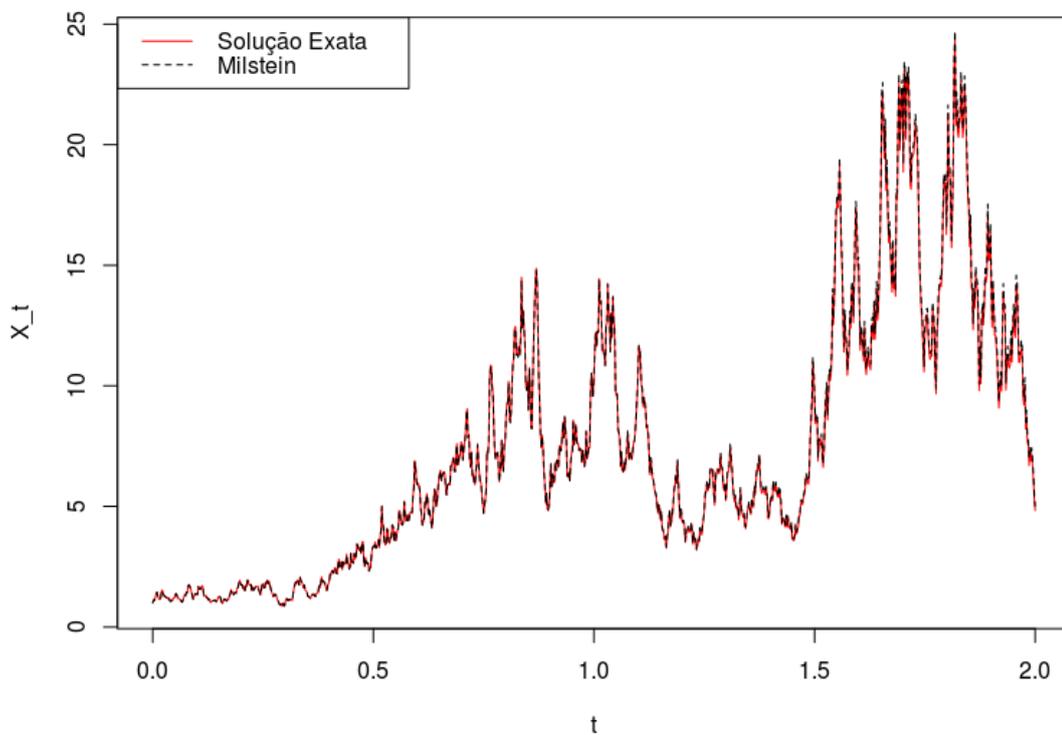
(a) Malha com $N = 257$ pontos(b) Malha com $N = 1025$ pontos

Figura 4.2: Exemplo de aproximação pelo método de Milstein para o movimento Browniano Geométrico.

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:

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

Note que se o coeficiente $g(\cdot, x)$ for constante, então $g_x(t, x) \equiv 0$ e os métodos de Euler e Milstein coincidem. À primeira vista esta condição é bastante restritiva, no entanto, ela pode ser obtida através da transformação de Lamperti. É importante destacar que se aplicarmos o método de Euler a uma processo de Itô obtido por esta 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.

4.3 Esquema Livre de Derivada

Dentre os dois métodos apresentados até aqui, temos que o esquema de Milstein é o mais preciso pois leva em consideração todos os termos lineares com respeito ao passo de discretização Δ na expansão de Itô-Taylor. Essa característica é importante porém não é a única que difere este método da aproximação mais simples proposta por Euler-Maruyama. O método de milstein faz uso também da derivada do coeficiente de deriva e tal procedimento pode ser custoso em determinadas aplicações. Uma abordagem trabalhada por Liske & Platen (1987) leva em conta uma aproximação da derivada com o objetivo de contornar este problema. Do cálculo clássico, temos que

$$g(t, x + dx) - g(t, x) = g_x(t, x)dx + O(dx^2). \quad (4.19)$$

Este resultado nos leva à seguinte aproximação:

$$g(t, x + dx) - g(t, x) \approx g_x(t, x)dx, \quad (4.20)$$

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

$$\begin{aligned} g(t, X_{t_n}^N + \Delta X_{t_n}^N) - g(t, X_{t_n}^N) &\approx g_x(t, X_{t_n}^N) \Delta X_{t_n}^N \\ &= g_x(t, X_{t_n}^N) [f(t, X_{t_n}^N) \Delta + g(t, X_{t_n}^N) \Delta B_t] \\ &= g_x(t, X_{t_n}^N) g(t, X_{t_n}^N) \Delta B_t + O(\Delta). \end{aligned} \quad (4.21)$$

Os autores do método utilizam também a aproximação $\Delta B_t \approx \sqrt{\Delta}$, a qual é razoável se pensarmos no teorema da variação quadrática apresentado no capítulo 1. Desta forma, obtemos

$$g_x(t, X_{t_n}^N) g(t, X_{t_n}^N) = \frac{g(t, X_{t_n}^N + \Delta X_{t_n}^N) - g(t, X_{t_n}^N)}{\sqrt{\Delta}} + O(\sqrt{\Delta}). \quad (4.22)$$

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

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

Assim,

$$\begin{aligned} g_x(t, X_{t_n}^N) g(t, X_{t_n}^N) &\approx \frac{g(t, X_{t_n}^N + f(t_n, X_{t_n}^N) \Delta + g(t_n, X_{t_n}^N) \Delta B_{t_n}) - g(t, X_{t_n}^N)}{\sqrt{\Delta}} \\ &\approx \frac{g(t, X_{t_n}^N + f(t_n, X_{t_n}^N) \Delta + g(t_n, X_{t_n}^N) \sqrt{\Delta}) - g(t, X_{t_n}^N)}{\sqrt{\Delta}}. \end{aligned} \quad (4.23)$$

Substituindo o resultado obtido no esquema de Milstein, chegamos ao seguinte método numérico:

$$\begin{aligned} X_{t_{n+1}}^N &= X_{t_n}^N + f(t_n, X_{t_n}^N) \Delta + g(t_n, X_{t_n}^N) \Delta B_{t_n} \\ &\quad + \left[g(t_n, X_{t_n}^N + f(t_n, X_{t_n}^N) \Delta + g(t_n, X_{t_n}^N) \sqrt{\Delta}) - g(t_n, X_{t_n}^N) \right] \frac{(\Delta B_{t_n})^2 - \Delta}{2\sqrt{\Delta}}. \end{aligned} \quad (4.24)$$

Exemplo 4.2 *Seja $X(\omega) = \{X_t(\omega); 0 \leq t \leq 5\}$ uma realização do movimento Browniano geométrico com $\lambda = 2, \mu = 2, x_0 = 1$. Neste caso, os esquemas de Milstein e livre de derivada serão dados por*

$$\begin{aligned} X_{t_{n+1}}^N &= X_{t_n}^N \left[1 + \lambda \Delta + \mu \Delta B_{t_n} + \mu^2 \frac{(\Delta B_{t_n})^2 - \Delta}{2} \right], \\ X_{t_0}^N &= 1, \end{aligned}$$

e,

$$\begin{aligned} X_{t_{n+1}}^N &= X_{t_n}^N \left[1 + \lambda \Delta + \mu \Delta B_{t_n} + \mu(\lambda \Delta + \mu \sqrt{\Delta}) \frac{(\Delta B_{t_n})^2 - \Delta}{2\sqrt{\Delta}} \right], \\ X_{t_0}^N &= 1, \end{aligned}$$

respectivamente. Ilustramos abaixo 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.

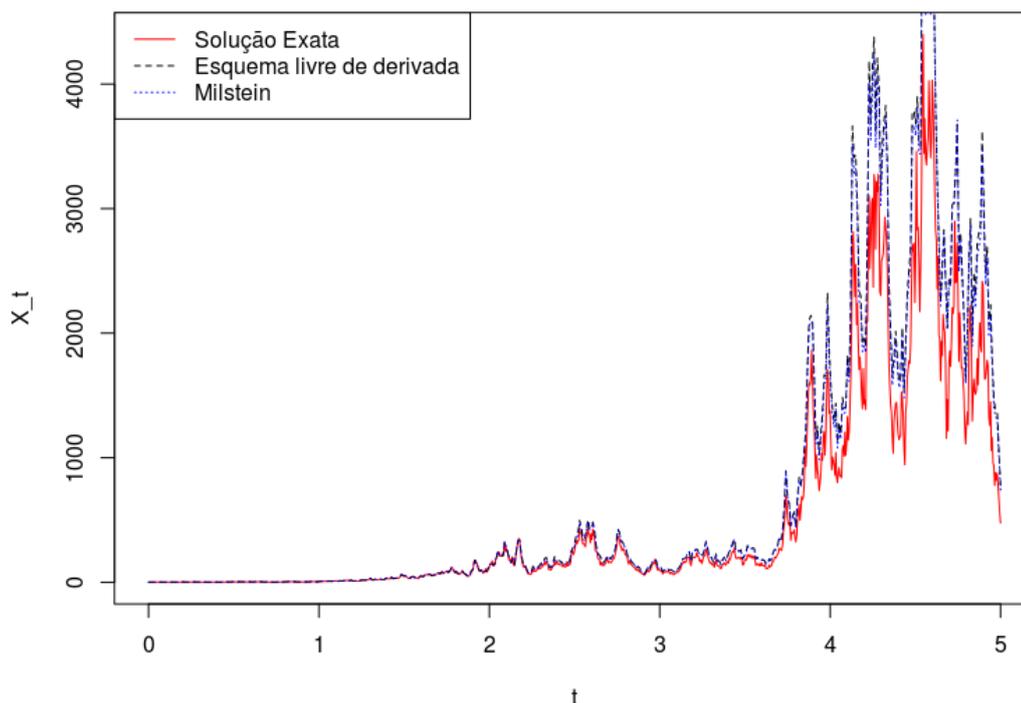


Figura 4.3: Comparação entre os métodos de Milstein e livre de derivada.

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 são encontrados na literatura, porém não serão detalhados aqui. Em particular, destacamos os métodos de Runge-Kutta abordados por Newton (1991), Saito & Mitsui (1992) e Burrage & Burrage (1996), os quais serão apresentados superficialmente na última seção.

4.4 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:

$$X_{t_{n+1}}^N = X_{t_n}^N + f(t_{n+1}, X_{t_{n+1}}^N) \Delta + g(t_n, X_{t_n}^N) \Delta B_{t_n}. \quad (4.25)$$

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 $\theta \in [0, 1]$:

$$\begin{aligned} X_{t_{n+1}}^N = & X_{t_n}^N + \theta \cdot f\left(t_{n+1}, X_{t_{n+1}}^N\right) \Delta + (1 - \theta) \cdot f\left(t_n, X_{t_n}^N\right) \Delta \\ & + g\left(t_n, X_{t_n}^N\right) \Delta B_{t_n}. \end{aligned} \quad (4.26)$$

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 a 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 de como o método θ pode ser utilizado para melhorar os resultados obtidos pelo método de Euler-Maruyama. A simulação abaixo mostra o comportamento 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:

$$\begin{aligned} X_{t_{n+1}}^N = & X_{t_n}^N + \left[\theta f\left(t_{n+1}, X_{t_{n+1}}^N\right) + (1 - \theta) f\left(t_n, X_{t_n}^N\right) \right] \Delta + g\left(t_n, X_{t_n}^N\right) \Delta B_{t_n} \\ & + \frac{1}{2} g_x\left(t_n, X_{t_n}^N\right) g\left(t_n, X_{t_n}^N\right) [(\Delta B_{t_n})^2 - \Delta], \end{aligned} \quad (4.27)$$

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

$$\begin{aligned} X_{t_{n+1}}^N = & X_{t_n}^N + \left[\theta f\left(t_n, X_{t_{n+1}}^N\right) + (1 - \theta) f\left(t_n, X_{t_n}^N\right) \right] \Delta + g\left(t_n, X_{t_n}^N\right) \Delta B_{t_n} \\ & + \left[g\left(t_n, X_{t_n}^N + f\left(t_n, X_{t_n}^N\right) \Delta + g\left(t_n, X_{t_n}^N\right) \sqrt{\Delta}\right) - g\left(t_n, X_{t_n}^N\right) \right] \frac{(\Delta B_{t_n})^2 - \Delta}{2\sqrt{\Delta}}, \end{aligned} \quad (4.28)$$

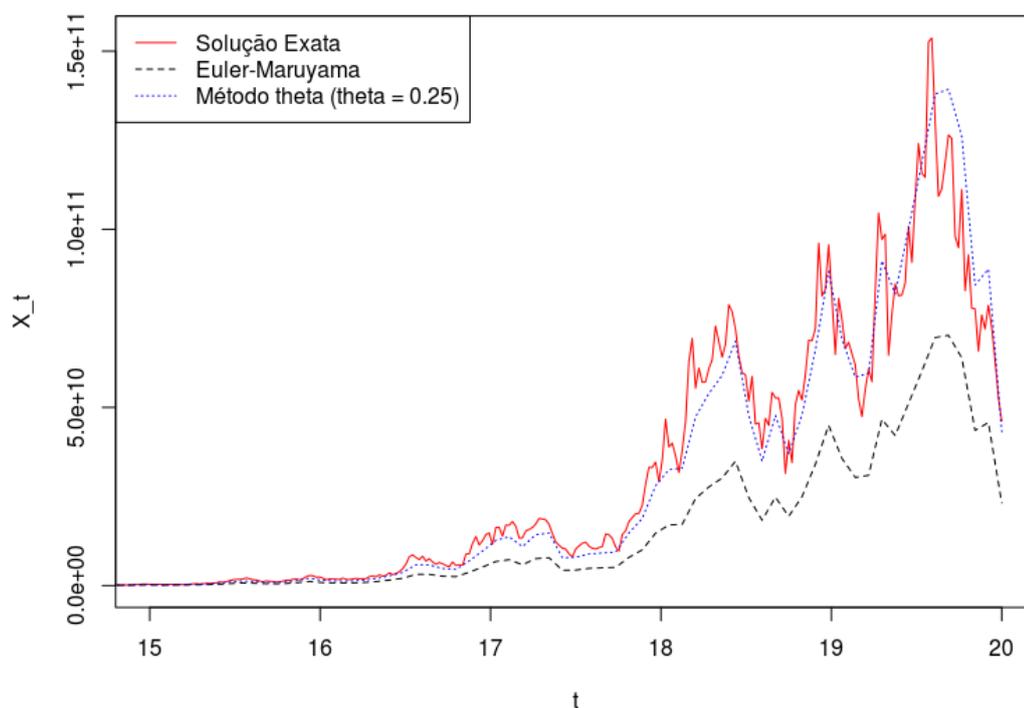


Figura 4.4: Euler-Maruyama $\times \theta$ Estocástico ($\theta = 1/4$).

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

4.5 Método da Linearização Local

O método da Linearização Local, introduzido por Shoji e Ozaki (1998), 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

$$dX_t = f(t, X_t)dt + g(X_t)dB_t, \quad 0 \leq t \leq T, \quad (4.29)$$

$$P(X_0 = \xi_0) = 1, \quad (4.30)$$

onde os coeficientes f e g satisfazem as condições do teorema (3.1) 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

$$dX_t = f(t, X_t)dt + \sigma dB_t, \quad 0 \leq t \leq T, \quad (4.31)$$

$$P(X_0 = \hat{\xi}_0) = 1, \quad (4.32)$$

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 (4.31) 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:

$$df = \left(\frac{\sigma^2}{2} f_{xx} + f_t \right) dt + f_x dx. \quad (4.33)$$

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:

$$\begin{aligned} df &\approx f(t, x_t) - f(s, x_s) \\ &= \left(\frac{\sigma^2}{2} f_{xx} + f_t \right) \cdot (t - s) + f_x \cdot (x_t - x_s). \end{aligned} \quad (4.34)$$

Definindo

$$\begin{aligned} L_s &:= f_x(s, X_s), & M_s &:= \frac{\sigma^2}{2} f_{xx}(s, X_s) + f_t(s, X_s), \\ N_s &:= f(s, X_s) - X_s f_x(s, X_s) - \left(\frac{\sigma^2}{2} f_{xx}(s, X_s) + f_t(s, X_s) \right) s, \end{aligned}$$

obtemos,

$$\hat{f}(t, x_t) := L_s x_t + t M_s + N_s, \quad t \geq s. \quad (4.35)$$

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

$$dX_t = (L_s X_t + t M_s + N_s) dt + \sigma dB_t = \hat{f}(t, X_t) + \sigma dB_t, \quad (4.36)$$

ao invés da equação (4.31), 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 (4.36) serão obtidas através de uma sequência de transformações que simplificarão o trabalho analítico. O primeiro passo é definir as funções

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

as quais cumprem a seguinte igualdade:

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

A relação acima nos permite aplicar a transformação de Girsanov de modo a obter uma equação

diferencial estocástica mais simples:

$$\begin{aligned} dX_t &= \alpha(t, X_t)dt + \sigma d\hat{B}_t \\ &= L_s X_t dt + \sigma d\hat{B}_t, \end{aligned} \quad (4.37)$$

onde

$$\hat{B}_t = B_t - \frac{1}{\sigma} \int_s^t (M_s u + N_s) du$$

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,

$$\begin{aligned} dY_t &= -L_s e^{-L_s t} X_t dt + e^{-L_s t} (L_s X_t dt + \sigma d\hat{B}_t) + 0 \cdot (dX_t)^2 \\ &= e^{-L_s t} \sigma d\hat{B}_t \\ &= e^{-L_s t} \sigma dB_t + e^{-L_s t} \sigma (M_s t + N_s) dt. \\ \therefore Y_t &= Y_s + \int_s^t (M_s u + N_s) e^{-L_s u} du + \sigma \int_s^t e^{-L_s u} dB_u. \end{aligned} \quad (4.38)$$

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

$$\begin{aligned} X_t &= e^{L_s t} \left(e^{-L_s s} X_s + \int_s^t (M_s u + N_s) e^{-L_s u} du + \sigma \int_s^t e^{-L_s u} dB_u \right) \\ &= e^{L_s(t-s)} X_s + \frac{M_s}{L_s} \left(s e^{L_s(t-s)} - t \right) + \frac{M_s}{L_s^2} \left(e^{L_s(t-s)} - 1 \right) + \frac{N_s}{L_s} \left(e^{L_s(t-s)} - 1 \right) \\ &\quad + \sigma \int_s^t e^{L_s(t-u)} dB_u \\ &= X_s + X_s \left(e^{L_s(t-s)} - 1 \right) + s \frac{M_s}{L_s} \left(e^{L_s(t-s)} - 1 \right) - \frac{M_s}{L_s} (t-s) \\ &\quad + \frac{M_s}{L_s^2} \left(e^{L_s(t-s)} - 1 \right) + \frac{N_s}{L_s} \left(e^{L_s(t-s)} - 1 \right) + \sigma \int_s^t e^{L_s(t-u)} dB_u \\ &= X_s + \frac{L_s X_s + M_s s + N_s}{L_s} \left(e^{L_s(t-s)} - 1 \right) + \frac{M_s}{L_s^2} \left[\left(e^{L_s(t-s)} - 1 \right) - L_s (t-s) \right] \\ &\quad + \sigma \int_s^t e^{L_s(t-u)} dB_u \\ &= X_s + \frac{\hat{f}(s, X_s)}{L_s} \left(e^{L_s(t-s)} - 1 \right) + \frac{M_s}{L_s^2} \left[\left(e^{L_s(t-s)} - 1 \right) - L_s (t-s) \right] \\ &\quad + \sigma \int_s^t e^{L_s(t-u)} dB_u. \end{aligned} \quad (4.39)$$

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

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

o qual é normalmente distribuído com média

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

e variância

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

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

$$X_t | \{X_s = x\} \sim N(A(x), B^2(x)), \quad (4.40)$$

com a notação:

$$A(x) := X_s + \frac{\hat{f}(s, X_s)}{L_s} \left(e^{L_s(t-s)} - 1 \right) + \frac{M_s}{L_s^2} \left[\left(e^{L_s(t-s)} - 1 \right) - L_s(t-s) \right],$$

$$B(x) := \sigma \sqrt{\frac{e^{2L_s(t-s)} - 1}{2L_s}}.$$

O processo discretizado obtido pelo método da linearização local é simples de ser implementando tanto para fins de simulação quanto para inferência.

4.6 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 da dissertação. Em conjunto com os esquemas de discretização, apresentamos também as respectivas ordens de convergência forte (ver Definição 5.1). Mesmo não entrando em detalhes, o conhecimento desses métodos é relevante para trabalhos futuros.

Método de Heun (McShane (1974))

$$X_{t_{n+1}}^N = X_{t_n}^N + \frac{1}{2}[F_1 + F_2]\Delta + \frac{1}{2}[G_1 + G_2]\Delta B_{t_n}, \quad (4.41)$$

onde

$$\begin{aligned} F(t, x) &= \left[f(t, x) - \frac{1}{2} g_x(t, x) g(t, x) \right], & F_1 &= F(X_{t_n}^N) \\ G_1 &= g(t_{n+1}, X_{t_{n+1}}^N), & F_2 &= F(t_n, X_{t_n}^N + F_1 \Delta + G_1 \Delta B_{t_n}), \\ G_2 &= g(t_n, X_{t_n}^N + F_1 \Delta + G_1 \Delta B_{t_n}). \end{aligned}$$

Ordem de convergência forte: 1.0

Runge-Kutta de Primeira Ordem (Newton (1991))

$$X_{t_{n+1}}^N = X_{t_n}^N + F_1 \Delta + G_2 \Delta B_{t_n} + [G_2 - G_1] \sqrt{\Delta}, \quad (4.42)$$

onde

$$F_1 = f(t_n, X_{t_n}^N), \quad G_1 = g(t_n, X_{t_n}^N), \quad G_2 = g\left(t_n, X_{t_n}^N + \frac{G_1(\Delta B_{t_n} - \sqrt{\Delta})}{2}\right). \quad (4.43)$$

Ordem de convergência forte: 1.0

Runge-Kutta de Três Estágios (Saito & Mitsui (1992))

$$\begin{aligned} X_{t_{n+1}}^N &= X_{t_n}^N + \frac{1}{4}[F_1 + 3F_3]\Delta + \frac{1}{4}[G_1 + 3G_3]\Delta B_{t_n} \\ &\quad + \frac{1}{2\sqrt{3}} [f_x(t_n, X_{t_n}^N) g(t_n, X_{t_n}^N) - g_x(t_n, X_{t_n}^N) f(t_n, X_{t_n}^N) \\ &\quad - \frac{1}{2} g_{xx}(t_n, X_{t_n}^N) g(t_n, X_{t_n}^N)^2] \Delta \cdot \Delta \hat{B}_{t_n}, \end{aligned} \quad (4.44)$$

onde ξ_1^n e ξ_2^n são variáveis aleatórias independentes com distribuição Gaussiana padrão,

$$\begin{aligned} \Delta B_{t_n} &= \xi_1^n \sqrt{\Delta}, & \Delta \hat{B}_{t_n} &= \xi_2^n \sqrt{\Delta}, & F(t, x) &= \left[f(t, x) - \frac{1}{2} g_x(t, x) g(t, x) \right], \\ F_1 &= F(t_n, X_{t_n}^N), & G_1 &= g(t_n, X_{t_n}^N), \\ F_2 &= F\left(t_n, X_{t_n}^N + \frac{1}{3} F_1 \Delta + \frac{1}{3} G_1 \Delta B_{t_n}\right), & G_2 &= g\left(t_n, X_{t_n}^N + \frac{1}{3} F_1 \Delta + \frac{1}{3} G_1 \Delta B_{t_n}\right), \\ F_3 &= F\left(t_n, X_{t_n}^N + \frac{2}{3} F_2 \Delta + \frac{2}{3} G_2 \Delta B_{t_n}\right), & G_3 &= g\left(t_n, X_{t_n}^N + \frac{2}{3} F_2 \Delta + \frac{2}{3} G_2 \Delta B_{t_n}\right). \end{aligned}$$

Ordem de convergência forte: 1.0

Esquema de Taylor (Milstein (1974))

$$\begin{aligned}
X_{t_{n+1}}^N &= X_{t_n}^N + f(t_n, X_{t_n}^N) \Delta + g(t_n, X_{t_n}^N) \Delta B_{t_n} \\
&+ \frac{1}{2} [g_x(t_n, X_{t_n}^N) g(t_n, X_{t_n}^N)] ((\Delta B_{t_n})^2 - \Delta) + [f_x(t_n, X_{t_n}^N) g(t_n, X_{t_n}^N)] \Delta Z_{t_n} \\
&+ \left[g_x(t_n, X_{t_n}^N) f(t_n, X_{t_n}^N) + \frac{1}{2} g_{xx}(t_n, X_{t_n}^N) g(t_n, X_{t_n}^N)^2 \right] \Delta \hat{Z}_{t_n} \\
&+ \frac{1}{6} \left[g_x^2(t_n, X_{t_n}^N) g(t_n, X_{t_n}^N) + g_{xx}(t_n, X_{t_n}^N) g(t_n, X_{t_n}^N)^2 ((\Delta B_{t_n})^3 - 3\Delta \cdot \Delta B_{t_n}) \right],
\end{aligned} \tag{4.45}$$

onde ξ_1^n e ξ_2^n são variáveis aleatórias independentes com distribuição Gaussiana padrão,

$$\Delta B_{t_n} = \xi_1^n \sqrt{\Delta}, \quad \Delta Z_{t_n} = \frac{1}{2} \left(\xi_1^n + \frac{\xi_2^n}{\sqrt{3}} \right) \Delta^{3/2}, \quad \Delta \hat{Z}_{t_n} = \frac{1}{2} \left(\xi_1^n - \frac{\xi_2^n}{\sqrt{3}} \right) \Delta^{3/2}.$$

Ordem de convergência forte: 1.5

Runge-Kutta de Quatro Estágios (Newton (1991))

$$\begin{aligned}
X_{t_{n+1}}^N &= X_{t_n}^N + \frac{1}{2} [F_1 + F_2] \Delta + \frac{1}{40} [37G_1 + 30G_3 - 27G_4] \Delta B_{t_n} \\
&+ \frac{1}{16} [8G_1 + G_2 - 9G_3] \sqrt{3\Delta},
\end{aligned} \tag{4.46}$$

onde

$$\begin{aligned}
F_1 &= f(t_n, X_{t_n}^N), \quad G_1 = g(t_n, X_{t_n}^N), \\
F_2 &= f\left(t_n, X_{t_n}^N + F_1 \Delta + g_1 \Delta B_{t_n}\right), \quad G_2 = g\left(t_n, X_{t_n}^N - \frac{2}{3} G_1 (\Delta B_{t_n} + \sqrt{2\Delta})\right), \\
G_3 &= g\left(t_n, X_{t_n}^N + \frac{2}{9} G_1 (3\Delta B_{t_n} + \sqrt{3\Delta})\right), \\
G_4 &= g\left(t_n, X_{t_n}^N - \frac{20}{27} F_1 \Delta + \frac{10}{27} (G_2 - G_1) \Delta B_{t_n} - \frac{10}{27} G_2 \sqrt{3\Delta}\right).
\end{aligned}$$

Ordem de convergência forte: 1.5

5 *Análise Numérica*

5.1 **Convergência Forte**

Ao utilizar um método numérico para obter aproximações das trajetórias associadas às soluções fortes de uma equação diferencial estocástica, 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:

$$\varepsilon := \mathbb{E} |X_T^N - X_T|. \quad (5.1)$$

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

$$\lim_{\delta \downarrow 0} \mathbb{E} |X_T^N - X_T| = 0. \quad (5.2)$$

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 5.1 *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*

$$\varepsilon(\delta) = \mathbb{E} |X^N - X_T| \leq M\delta^\alpha, \quad \forall \delta \in (0, \delta_0). \quad (5.3)$$

Perceba que a relação (5.3) 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 .

A definição anterior pode ser considerada superficial pois leva em consideração o comportamento do método apenas no último instante de observação. Esta questão é resolvida por um teorema de convergência (Kloden & Platen (1999)) que garante a consistência do método numérico ao longo de toda a trajetória caso algumas hipóteses adicionais sejam consideradas no conceito de convergência forte.

Teorema 5.1 *Seja $X^N = \{X_{t_n}^N; t_n \in \tau_\delta\}$ uma aproximação discreta, associada à partição τ_δ do intervalo $[0, T]$, $T > 0$, para o processo $X = \{X_t; 0 \leq t \leq T\}$, gerada por um esquema numérico de ordem de convergência forte $\gamma \in \{0.5, 1.0, 1.5, 2.0\}$. Se*

$$\mathbb{E}(|X_0|^2) < +\infty \quad (5.4)$$

e existir uma constante $K > 0$ tal que

$$\sqrt{\mathbb{E}(|X_0 - X_{t_0}^N|^2)} \leq K\delta^\gamma, \quad (5.5)$$

então

$$\mathbb{E} \left(\max_{0 \leq n \leq N-1} |X_{t_n}^N - X_{t_n}|^2 \right) \leq K\delta^{2\gamma}. \quad (5.6)$$

Dentre as aproximações discretas apresentadas, Shoji (2011) mostra que o método de linearização local possui a maior ordem de convergência forte: $\alpha = 2$. Esse resultado é expressivo pois esta é, segundo Rümelin (1982), a maior ordem de convergência que pode ser atingida para uma equação diferencial estocástica com estas características. Quanto aos métodos de Euler, θ estocástico e Milstein, podemos inspecionar os erros $R1$ e $R2$ da expansão de Itô-Taylor para inferir que estes métodos convergem fortemente com ordens 0.5, 0.5 e 1.0. Uma análise parecida à do método de Milstein é aplicável ao esquema livre de derivada pois este é apenas 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 experimentalmente 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 5.2 *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

$$\lim_{\delta \downarrow 0} c(\delta) = 0, \quad (5.7)$$

tal que

$$\mathbb{E} \left[\left| \mathbb{E} \left(\frac{X_{t_{n+1}}^N - X_{t_n}^N}{\Delta_n} \middle| \mathcal{F}_{t_n} \right) - f(t_n, X_{t_n}^N) \right|^2 \right] \leq c(\delta), \quad (5.8)$$

e

$$\mathbb{E} \left[\frac{1}{\Delta_n} \left| X_{t_{n+1}}^N - X_{t_n}^N - \mathbb{E} \left(X_{t_{n+1}}^N - X_{t_n}^N \middle| \mathcal{F}_{t_n} \right) - g(t_n, X_{t_n}^N) \Delta B_n \right|^2 \right] \leq c(\delta) \quad (5.9)$$

para quaisquer valores fixos $X_{t_n}^N = x$ e $n = 0, 1, 2, \dots, N-1$.

A condição (5.8) 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 (5.9) 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 5.1 *Seja X^N a aproximação discreta do método de Euler para a equação diferencial estocástica*

$$\begin{aligned} dX_t &= f(t, X_t)dt + g(t, X_t)dB_t, \quad t \in [0, T], \\ P(X_0 = \xi_0) &= 1. \end{aligned}$$

Perceba que

$$\begin{aligned} \mathbb{E} \left(\frac{X_{t_{n+1}}^N - X_{t_n}^N}{\Delta_N} \middle| \mathcal{F}_{t_n} \right) &= \mathbb{E} \left(\frac{f(t_n, X_{t_n}^N) \Delta_N + g(t_n, X_{t_n}^N) \Delta B_{t_n}}{\Delta_N} \middle| \mathcal{F}_{t_n} \right) \\ &= \mathbb{E} \left(\frac{g(t_n, X_{t_n}^N) \Delta B_{t_n}}{\Delta_N} \middle| \mathcal{F}_{t_n} \right) + \mathbb{E} (f(t_n, X_{t_n}^N) | \mathcal{F}_{t_n}) \\ &= \mathbb{E}(\Delta B_{t_n}) \mathbb{E} \left(\frac{g(t_n, X_{t_n}^N)}{\Delta_N} \middle| \mathcal{F}_{t_n} \right) + f(t_n, X_{t_n}^N) \\ &= f(t_n, X_{t_n}^N). \end{aligned}$$

Assim, para qualquer passo de discretização δ , podemos definir $c(\delta) \equiv 0$, de modo a obter

$$\mathbb{E} \left[\left| \mathbb{E} \left(\frac{X_{t_{n+1}}^N - X_{t_n}^N}{\Delta_N} \middle| \mathcal{F}_{t_n} \right) - f(t_n, X_{t_n}^N) \right|^2 \right] = \mathbb{E} \left[|f(t_n, X_{t_n}^N) - f(t_n, X_{t_n}^N)|^2 \right] = 0;$$

$$\begin{aligned}
& \mathbb{E} \left[\frac{1}{\Delta_n} \left| X_{t_{n+1}}^N - X_{t_n}^N - \mathbb{E} \left[X_{t_{n+1}}^N - X_{t_n}^N \middle| \mathcal{F}_{t_n} \right] - g(t_n, X_{t_n}^N) \Delta B_n \right|^2 \right] = \\
& = \mathbb{E} \left[\frac{1}{\Delta_N} \left| f(t_n, X_{t_n}^N) \Delta_N + g(t_n, X_{t_n}^N) \Delta B_{t_n} - f(t_n, X_{t_n}^N) \Delta_n - g(t_n, X_{t_n}^N) \Delta B_{t_n} \right|^2 \right] = \\
& = 0.
\end{aligned}$$

Portanto, o método de Euler é fortemente consistente. \triangleright

Exemplo 5.2 Considere novamente a equação diferencial estocástica do exemplo anterior e defina a aproximação discreta X^N a partir do método de Milstein. Perceba que

$$\begin{aligned}
\mathbb{E} \left(\frac{X_{t_{n+1}}^N - X_{t_n}^N}{\Delta_N} \middle| \mathcal{F}_{t_n} \right) &= \mathbb{E} \left(\frac{f(t_n, X_{t_n}) \Delta_N + g(t_n, X_{t_n}) \Delta B_{t_n}}{\Delta_N} \middle| \mathcal{F}_{t_n} \right) \\
&\quad + \mathbb{E} \left(\frac{\frac{1}{2} g(t_n, X_{t_n}) g_x(t_n, X_{t_n}) (\Delta B_{t_n}^2 - \Delta_N)}{\Delta_N} \middle| \mathcal{F}_{t_n} \right) \\
&= f(t_n, X_{t_n}^N) + \frac{1}{2\Delta_N} g(t_n, X_{t_n}) g_x(t_n, X_{t_n}) \mathbb{E} [\Delta B_{t_n}^2 - \Delta_N] \\
&= f(t_n, X_{t_n}^N).
\end{aligned}$$

Além disso, temos que $\Delta B_{t_n}^2 = \Delta_N Z^2$, onde $Z \sim N(0, 1)$. Neste caso, $Z^2 \sim \chi_{(1)}$ e

$$\text{Var}(\Delta B_{t_n}^2) = \Delta_N^2 \text{Var}(Z^2) = 2\Delta_N^2.$$

Se $\frac{1}{2} g g_x$ é uma aplicação limitada, então podemos encontrar um $M > 0$ tal que $\frac{1}{2} g g_x \leq M, \forall (t, x) \in [0, T] \times \mathbb{R}$. Desta forma, definindo $c(\delta) = 2M^2 \delta$, temos que $\lim_{\delta \downarrow 0} c(\delta) = 0$, e

$$\mathbb{E} \left[\left| \mathbb{E} \left(\frac{X_{t_{n+1}}^N - X_{t_n}^N}{\Delta_N} \middle| \mathcal{F}_{t_n} \right) - f(t_n, X_{t_n}^N) \right|^2 \right] = \mathbb{E} \left[|f(t_n, X_{t_n}^N) - f(t_n, X_{t_n}^N)|^2 \right] = 0;$$

$$\begin{aligned}
& \mathbb{E} \left[\frac{1}{\Delta_n} \left| X_{t_{n+1}}^N - X_{t_n}^N - \mathbb{E} \left[X_{t_{n+1}}^N - X_{t_n}^N \middle| \mathcal{F}_{t_n} \right] - g(t_n, X_{t_n}^N) \Delta B_n \right|^2 \right] = \\
& = \mathbb{E} \left[\frac{1}{\Delta_N} \left| \frac{1}{2} g(t_n, X_{t_n}) g_x(t_n, X_{t_n}) (\Delta B_{t_n}^2 - \Delta_N) \right|^2 \right] \\
& \leq \frac{M^2}{\Delta_N} \mathbb{E} [(\Delta B_{t_n}^2 - \Delta_N)] = \frac{M^2}{\Delta_N} \text{Var}(\Delta B_{t_n}^2) = 2M^2 \Delta_N \leq 2M^2 \delta = c(\delta) \rightarrow 0.
\end{aligned}$$

Portanto, o método de Milstein é fortemente consistente desde que a aplicação $\frac{1}{2} g g_x$ seja limitada. \triangleright

Dadas as propriedades dos conceitos envolvidos nas definições (5.1)-(5.2), é intuitivo pen-

sar que consistência implica convergência. De fato, Kloeden e Platen (1999) 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.

5.2 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

$$X' = \lambda X, \quad (5.10)$$

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

$$\begin{aligned} dX_t &= \lambda X dt + \mu X dB_t, \quad t \in [0, T], \\ X_0 &= x_0, \end{aligned} \quad (5.11)$$

onde x_0 é uma constante real fixa. Já verificamos anteriormente que este problema possui uma

solução explícita dada por

$$X_t = x_0 \exp \left\{ \left(\lambda - \frac{1}{2} \mu^2 \right) t + \mu B_t \right\}. \quad (5.12)$$

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 5.3 *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 $\varepsilon > 0$ existir um $\delta > 0$ tal que*

$$\mathbb{E}|X_t|^2 < \varepsilon, \quad \forall t \geq 0, |X_0| < \delta, \quad (5.13)$$

e, além disso, se existir um $\delta_0 > 0$ tal que

$$\lim_{t \rightarrow +\infty} \mathbb{E}|X_t|^2 = 0, \quad \forall |X_0| < \delta_0. \quad (5.14)$$

Obs.: Por simplicidade, daqui em diante utilizaremos a expressão *solução estável* para fazer referência a uma solução assintoticamente estável em média quadrática.

Proposição 5.1 *A solução (5.12) é estável se, e somente se,*

$$2\lambda + \mu^2 < 0. \quad (5.15)$$

Demonstração: Considere $Y_t := \mathbb{E}|X_t|^2$. Assim,

$$\begin{aligned} Y_t &= \mathbb{E} \left| x_0 \exp \left\{ \left(\lambda - \frac{1}{2} \mu^2 \right) t + \mu B_t \right\} \right|^2 \\ &= |x_0|^2 \exp \{ (2\lambda - \mu^2) t \} \mathbb{E} | \exp \{ 2\mu B_t \} | = |x_0|^2 \exp \{ (2\lambda - \mu^2) t \} \exp \{ 2t \mu^2 \} \\ &= |x_0|^2 \exp \{ (2\lambda + \mu^2) t \}. \end{aligned}$$

O resultado segue trivialmente da igualdade acima.

□

Entendido o comportamento analítico da solução (5.12), 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

$$\hat{Y}_n = \mathbb{E}|X_{t_n}^N|^2.$$

Quando aplicamos um esquema numérico à equação linear de teste (5.11), obtemos uma equação de diferenças da forma

$$\hat{Y}_{t_{n+1}} = R(\lambda, \mu, \Delta)\hat{Y}_{t_n}, \quad (5.16)$$

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

$$\hat{Y}_{t_{n+1}} = (R(\lambda, \mu, \Delta))^n \hat{Y}_{t_0}, \quad (5.17)$$

o que nos mostra que a solução numérica será estável se

$$|R(\lambda, \mu, \Delta)| < 1. \quad (5.18)$$

Definição 5.4 *Um esquema numérico é estável para qualquer tripla (λ, μ, Δ) que satisfaz a relação (5.18). A região \mathcal{R} dada por*

$$\mathcal{R} := \{(\lambda, \mu, \Delta); |R(\lambda, \mu, \Delta)| < 1\},$$

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.

Obs.: 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.

5.2.1 Funções e Regiões de Estabilidade de Alguns Esquemas Numéricos

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:

$$X_{t_{n+1}}^N = X_{t_n}^N [(1 + \lambda\Delta) + \mu\Delta B_{t_n}]. \quad (5.19)$$

Elevando ao quadrado os dois membros da igualdade e tomando a esperança, obtemos, segundo a notação apresentada anteriormente,

$$\begin{aligned} \hat{Y}_{t_{n+1}} &= \hat{Y}_{t_n} \mathbb{E} [(1 + \lambda\Delta)^2 + 2(1 + \lambda\Delta)\mu\Delta B_{t_n} + \mu^2(\Delta B_{t_n})^2] \\ &= \hat{Y}_{t_n} \cdot [(1 + \lambda\Delta)^2 + \mu^2\Delta] := R(\lambda, \mu, \Delta)\hat{Y}_{t_n}, \end{aligned} \quad (5.20)$$

onde $R(\lambda, \mu, \Delta) = (1 + \lambda\Delta)^2 + \mu^2\Delta$.

A região de estabilidade será dada pelas triplas (λ, μ, Δ) , tais que $|R(\lambda, \mu, \Delta)| < 1$. Utilizando a igualdade obtida acima, chegamos à seguinte região de estabilidade:

$$\mathcal{R}_{em} = \left\{ (\lambda, \mu, \Delta) \in \mathcal{R} \times \mathbb{R}^+; \Delta < -\frac{2\lambda + \mu^2}{\lambda^2} \right\}, \quad (5.21)$$

onde \mathcal{R} é a região delimitada pela proposição 5.1, 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:

$$X_{t_{n+1}}^N = X_{t_n}^N \left[(1 + \lambda\Delta) + \mu\Delta B_{t_n} + \frac{1}{2}\mu^2((\Delta B_{t_n})^2 - \Delta) \right]. \quad (5.22)$$

Elevando ao quadrado os dois membros da igualdade e tomando a esperança, obtemos,

$$\begin{aligned}
\hat{Y}_{t_{n+1}} &= \hat{Y}_{t_n} \mathbb{E} \left[(1 + \lambda \Delta)^2 + \mu^2 (\Delta B_{t_n})^2 + \frac{1}{4} \mu^4 ((\Delta B_{t_n})^4 - 2\Delta (\Delta B_{t_n})^2 + \Delta^2) + 2(1 + \lambda \Delta)(\mu \Delta B_{t_n}) \right. \\
&\quad \left. + 2 \cdot \frac{1}{2} \mu^2 (1 + \lambda \Delta) ((\Delta B_{t_n})^2 - \Delta) + 2 \cdot \frac{1}{2} \mu^3 \Delta B_{t_n} ((\Delta B_{t_n})^2 - \Delta) \right] \\
&= \hat{Y}_{t_n} \left[(1 + \lambda \Delta)^2 + \mu^2 \Delta + \frac{1}{4} \mu^4 (3\Delta^2 - 2\Delta^2 + \Delta^2) \right] \\
&= \hat{Y}_{t_n} \left[(1 + \lambda \Delta)^2 + \mu^2 \Delta + \frac{1}{2} \mu^4 \Delta^2 \right] := R(\lambda, \mu, \Delta) \hat{Y}_{t_n}, \tag{5.23}
\end{aligned}$$

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:

$$\mathcal{R}_{mils} = \left\{ (\lambda, \mu, \Delta) \in \mathcal{R} \times \mathbb{R}; \Delta < -\frac{2\lambda + \mu^2}{\lambda^2 + \frac{1}{2}\mu^4} \right\}. \tag{5.24}$$

Como $2\lambda + \mu^2 < 0$, segue que

$$-\frac{2\lambda + \mu^2}{\lambda^2 + \frac{1}{2}\mu^4} < -\frac{2\lambda + \mu^2}{\lambda^2}.$$

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 (4.2), temos que o esquema livre de derivada aplicado à equação linear de teste resulta no seguinte método recursivo:

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

Elevando ao quadrado os dois membros desta igualdade e tomando a esperança, obtemos,

$$\begin{aligned}
\hat{Y}_{t_{n+1}} &= \hat{Y}_{t_n} \mathbb{E} \left[(1 + \lambda \Delta + \mu \Delta B_{t_n})^2 + \left(\mu (\lambda \Delta + \mu \sqrt{\Delta}) \frac{(\Delta B_{t_n})^2 - \Delta}{2\sqrt{\Delta}} \right)^2 \right. \\
&\quad \left. + 2(1 + \lambda \Delta + \mu \Delta B_{t_n}) \left(\mu (\lambda \Delta + \mu \sqrt{\Delta}) \frac{(\Delta B_{t_n})^2 - \Delta}{2\sqrt{\Delta}} \right) \right] \\
&= \hat{Y}_{t_n} \left[(1 + \lambda \Delta)^2 + \mu^2 \Delta + \mu^2 (\lambda^2 \Delta^2 + 2\lambda \mu \Delta^{3/2} + \mu^2 \Delta) \frac{3\Delta^2 - 2\Delta \cdot \Delta + \Delta^2}{4\Delta} \right] \\
&= \hat{Y}_{t_n} \left[(1 + \lambda \Delta)^2 + \mu^2 \Delta + \frac{\mu^2 \Delta^2}{2} (\lambda^2 \Delta + 2\lambda \mu \Delta^{1/2} + \mu^2) \right] := R(\lambda, \mu, \Delta) \hat{Y}_{t_n}, \tag{5.25}
\end{aligned}$$

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

$$\left[g\left(t_n, X_{t_n}^N + f\left(t_n, X_{t_n}^N\right)\Delta + g\left(t_n, X_{t_n}^N\right)\sqrt{\Delta}\right) - g\left(t_n, X_{t_n}^N\right) \right]$$

por

$$\left[g\left(t_n, X_{t_n}^N + g\left(t_n, X_{t_n}^N\right)\sqrt{\Delta}\right) - g\left(t_n, X_{t_n}^N\right) \right].$$

Neste caso, chegaríamos de imediato à função de estabilidade

$$R(\lambda, \mu, \Delta)\hat{Y}_{t_n} = \left[(1 + \lambda\Delta)^2 + \mu^2\Delta + \frac{1}{2}\mu^4\Delta^2 \right],$$

cujos tratamentos são mais simples. Mais precisamente, chegaríamos à mesma função de estabilidade do método de Milstein, portanto

$$\mathcal{R}_{Id} = \left\{ (\lambda, \mu, \Delta) \in \mathcal{R} \times \mathbb{R}; \Delta < -\frac{2\lambda + \mu^2}{\lambda^2 + \frac{1}{2}\mu^4} \right\}. \quad (5.26)$$

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 experimental.

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:

$$\begin{aligned} X_{t_{n+1}}^N &= X_{t_n}^N + \theta\lambda X_{t_{n+1}}^N\Delta + (1 - \theta)\lambda X_{t_n}^N\Delta + \mu X_{t_n}^N\Delta B_{t_n} \\ &= X_{t_n}^N \cdot \frac{1 + (1 - \theta)\lambda\Delta + \mu\Delta B_{t_n}}{1 - \theta\lambda\Delta}. \end{aligned} \quad (5.27)$$

Elevando ao quadrado os dois membros da igualdade e tomando a esperança, obtemos,

$$\begin{aligned}\hat{Y}_{t_{n+1}} &= \hat{Y}_{t_n} \mathbb{E} \left\{ \frac{[1 + (1 - \theta)\lambda\Delta]^2 + 2\mu\Delta B_{t_n} [1 + (1 - \theta)\lambda\Delta] + \mu^2(\Delta B_{t_n})^2}{(1 - \theta\lambda\Delta)^2} \right\} \\ &= \hat{Y}_{t_n} \frac{[1 + (1 - \theta)\lambda\Delta]^2 + \mu^2\Delta}{(1 - \theta\lambda\Delta)^2} := R(\lambda, \mu, \Delta)\hat{Y}_{t_n},\end{aligned}\quad (5.28)$$

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 5.2 *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

$$R(\lambda, \mu, \Delta) = \frac{[1 + (1 - \theta)\lambda\Delta]^2 + \mu^2\Delta}{(1 - \theta\lambda\Delta)^2}.$$

Considerando primeiramente $\theta \in (\frac{1}{2}, 1]$, obtemos:

$$\begin{aligned}R(\lambda, \mu, \Delta) < 1 &\Leftrightarrow \frac{[1 + (1 - \theta)\lambda\Delta]^2 + \mu^2\Delta}{(1 - \theta\lambda\Delta)^2} < 1 \\ &\Leftrightarrow \lambda^2\Delta^2 + 2\lambda\Delta(1 - \theta\lambda\Delta) + (1 - \theta\lambda\Delta)^2 + \mu^2\Delta < (1 - \theta\lambda\Delta)^2 \\ &\Leftrightarrow \lambda^2\Delta + 2\lambda(1 - \theta\lambda\Delta) + \mu^2 < 0 \\ &\Leftrightarrow \Delta > \frac{\mu^2 + 2\lambda}{\lambda^2(2\theta - 1)}.\end{aligned}$$

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

$$\begin{aligned} R(\lambda, \mu, \Delta) < 1 &\Leftrightarrow \lambda^2\Delta + 2\lambda\left(1 - \frac{1}{2}\lambda\Delta\right) + \mu^2 < 0 \\ &\Leftrightarrow \mu^2 + 2\lambda < 0, \end{aligned}$$

o que é garantido pela hipótese inicial. □

Obs.: Um esquema numérico incondicionalmente estável é também denominado *esquema A-estável*.

5.2.2 Método da Linearização Local

A análise de estabilidade apresentada aqui leva em consideração a equação linear de teste (5.11). 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

$$Z = S(t, X) = \frac{\log X}{\mu}$$

à equação linear de teste, chegamos à nova equação diferencial estocástica:

$$dZ_t = \left(\frac{\lambda}{\mu} - \frac{1}{2}\mu \right) dt + dB_t. \quad (5.29)$$

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:

$$Z_t = e^{L_s t} \left(e^{-L_s s} Z_s + \int_s^t (M_s u + N_s) e^{-L_s u} du + \sigma \int_s^t e^{-L_s u} dB_u \right), \quad (5.30)$$

onde

$$L_s \equiv 0, \quad M_s \equiv 0, \quad N_s = f(s, Z_s) = \left(\frac{\lambda}{\mu} - \frac{1}{2}\mu \right), \quad \sigma = 1. \quad (5.31)$$

Perceba que esta abordagem resultará no seguinte esquema:

$$Z_{t_{n+1}} = Z_{t_n} + \left(\frac{\lambda}{\mu} - \frac{1}{2}\mu \right) \Delta + \Delta B_{t_n}. \quad (5.32)$$

Aplicando a transformação inversa de Lamperti, segue que o esquema de discretização com respeito ao processo original é expresso por

$$\begin{aligned} X_{t_{n+1}}^N &= e^{\mu Z_{t_{n+1}}} \\ &= \exp \left\{ \mu Z_{t_n} + \left(\lambda - \frac{1}{2}\mu^2 \right) \Delta + \mu \Delta B_{t_n} \right\}, \end{aligned} \quad (5.33)$$

e pode ser resolvido como uma relação de recorrência

$$\begin{aligned} X_{t_{n+1}}^N &= x_0 \exp \left\{ \left(\lambda - \frac{1}{2}\mu^2 \right) (n+1)\Delta + \mu \sum_{j=0}^n \Delta B_{t_j} \right\} \\ &= x_0 \exp \left\{ \left(\lambda - \frac{1}{2}\mu^2 \right) t_{n+1} + \mu B_{t_{n+1}} \right\}. \end{aligned} \quad (5.34)$$

A relação acima pode ser estudada de maneira idêntica à proposição 5.1, 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 experimentalmente que este método é de fato mais estável que os anteriores.

5.3 Avaliação Experimental

Esta seção é reservada à análise experimental 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. Os resultados foram obtidos de implementações em C++ as quais foram executadas em um processador intel i7-2670QM CPU 2.20GHz \times 8, 7.7Gb de memória RAM e sistema operacional Ubuntu 12.04 64-bits. Procedimentos mais simples como montagem de gráficos e análise de dados, que demandam menos recursos computacionais, foram realizados com o auxílio do software R versão 2.14.1 64-bits.

5.3.1 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 o movi-

mento Browniano geométrico

$$\begin{aligned} dX_t &= \lambda X_t dt + \mu X_t dB_t, \quad 0 \leq t \leq T, \\ X_0 &= x_0, \end{aligned}$$

com parâmetros $\lambda = 4$, $\mu = 2$ e $x_0 = 1$, como caso de teste. Segundo o conceito apresentado, um esquema numérico converge fortemente até o instante T , com ordem α , se

$$\varepsilon(\Delta) = \mathbb{E} |X^N - X_T| \leq M\Delta^\alpha,$$

onde $M > 0$ é uma constante e Δ é o passo de discretização $(\frac{T}{N-1})$. Se a desigualdade acima fosse aproximadamente uma igualdade, poderíamos tomar o logaritmo dos dois lados da equação de modo a chegarmos a uma relação linear:

$$\log(\varepsilon(\Delta)) \approx \log M + \alpha \log \Delta. \quad (5.35)$$

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 (5.35) for razoável, ao plotarmos $\log(\varepsilon(\Delta)) \times \Delta$, deveríamos esperar uma reta cuja inclinação indica a ordem de convergência forte do algoritmo.

Passo de discretização		Erro esperado	
Δ	$\log \Delta$	$\varepsilon(\Delta)$	$\log(\varepsilon(\Delta))$
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 5.1: Esquema de Euler-Maruyama avaliado no intervalo $[0, 1]$.

Comparando os resultados em escala real e logarítmica, percebemos que a relação (5.35) é de fato razoável. A partir da técnica dos mínimos quadrados, chegamos à seguinte relação:

$$\log(\varepsilon(\Delta)) = 0.53 \log \Delta + 4.96, \quad (5.36)$$

o que indica uma ordem de convergência forte $\alpha \approx 0.53$, confirmando assim nossas expectativas

quanto ao método de Euler-Maruyama.

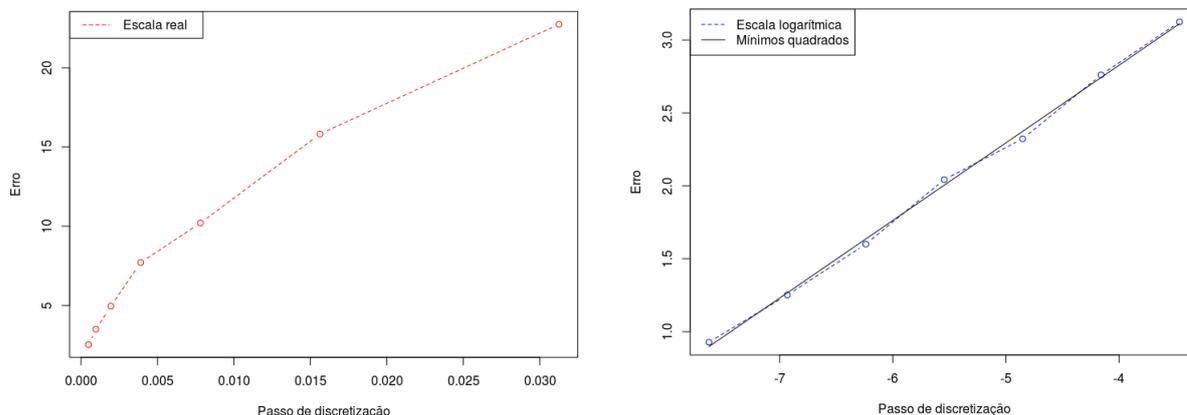


Figura 5.1: Esquema de Euler-Maruyama avaliado no intervalo $[0, 1]$.

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 da constante M . A tabela abaixo mostra que a relação linear já não é tão razoável, porém, o ajuste de mínimos quadrados continua evidenciando uma ordem de convergência forte $\alpha \approx 0.5$. Os ajustes lineares obtidos para $T = 5$ e $T = 10$ foram:

$$\log(\varepsilon(\Delta)) = 0.55 \log \Delta + 39.43, \quad (5.37)$$

e

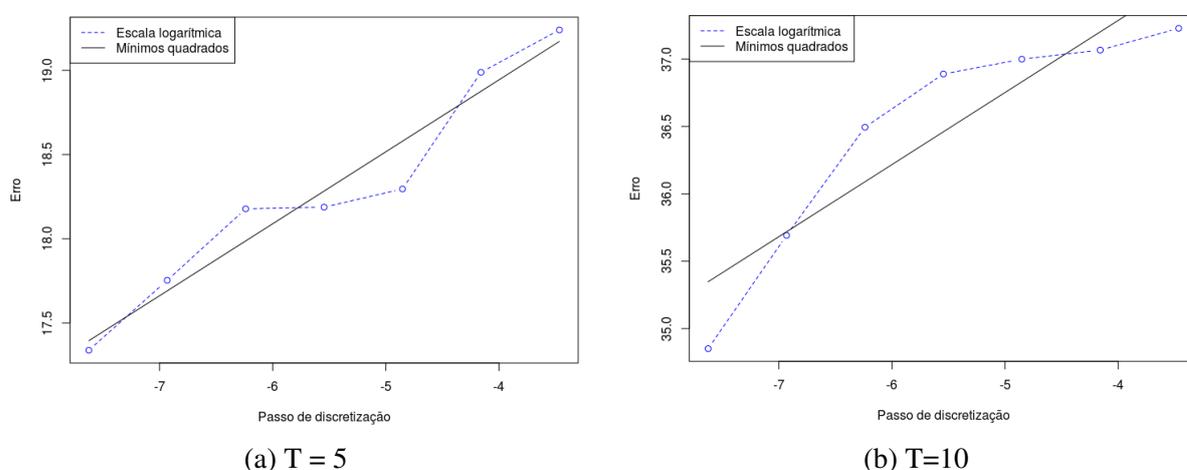
$$\log(\varepsilon(\Delta)) = 0.54 \log \Delta + 42.02, \quad (5.38)$$

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 coeficientes angulares correspondentes aos ajustes de mínimos quadrados foram um pouco superiores, porém, a diferença não foi significativa.

O resultado se repete quando consideramos $\theta = 1$ (Euler-implícito). Apresentamos inicialmente os resultados obtidos pelo ajuste de mínimos quadrados, os quais evidenciam uma ordem

Passo de discretização		Erro esperado	
Δ	$\log \Delta$	$\log(\varepsilon(\Delta)) (T = 5)$	$\log(\varepsilon(\Delta)) (T = 10)$
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 5.2: Esquema de Euler-Maruyama avaliado nos intervalos $[0, 5]$ e $[0, 10]$.Figura 5.2: Esquema de Euler-Maruyama avaliado nos intervalos $[0, 5]$ e $[0, 10]$.

Passo de discretização		Erro esperado		
Δ	$\log \Delta$	$\log(\varepsilon(\Delta)) (T = 1)$	$\log(\varepsilon(\Delta)) (T = 5)$	$\log(\varepsilon(\Delta)) (T = 10)$
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 5.3: Esquema Euler-trapezoidal avaliado nos intervalos $[0, 1]$, $[0, 5]$ e $[0, 10]$.

	$T = 1$	$T = 5$	$T = 10$
$\log M$	5.41	22.97	39.96
α	0.60	0.74	0.69

Tabela 5.4: Ajustes lineares para avaliação experimental do esquema Euler-trapezoidal.

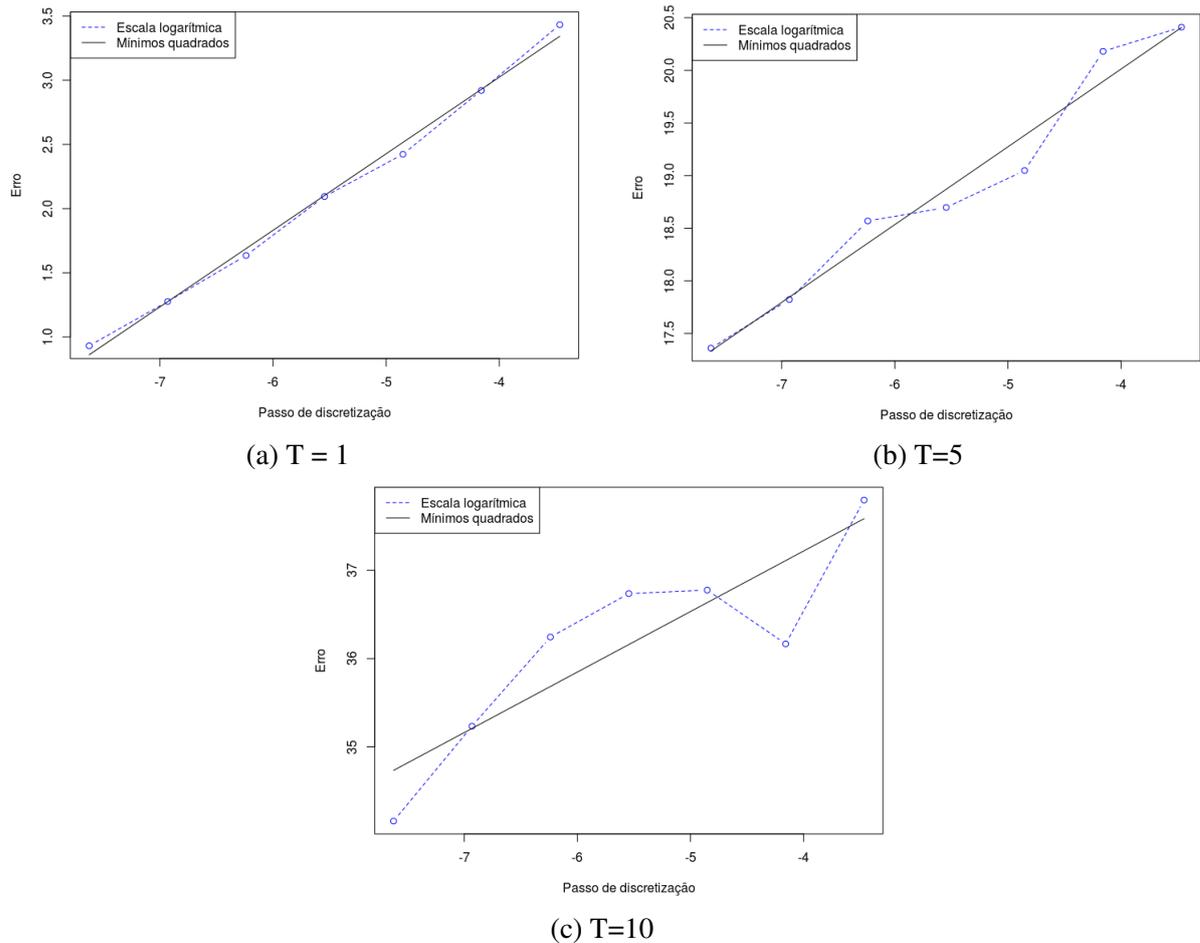


Figura 5.3: Esquema Euler-trapezoidal avaliado nos intervalos $[0, 1]$, $[0, 5]$ e $[0, 10]$.

de convergência forte similar às dos dois esquemas anteriores.

	$T = 1$	$T = 5$	$T = 10$
$\log M$	5.03	21.08	39.60
α	0.54	0.48	0.59

Tabela 5.5: Ajustes lineares para avaliação experimental 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.

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 e livre de derivada, uma vez que o segundo é uma modificação simples do primeiro.

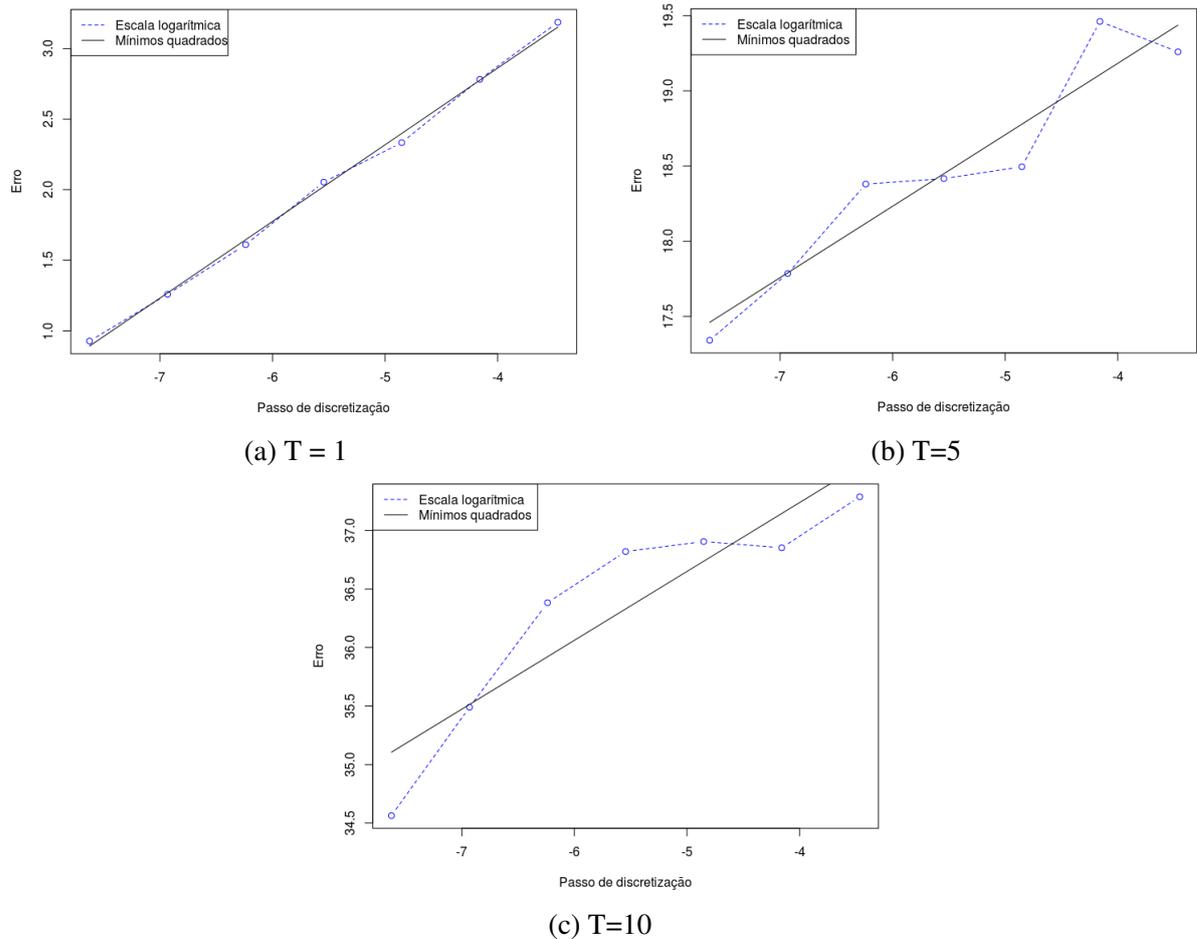


Figura 5.4: Esquema Euler-implícito avaliado nos intervalos $[0, 1]$, $[0, 5]$ e $[0, 10]$.

Passo de discretização		Erro esperado		
Δ	$\log \Delta$	$\log(\varepsilon(\Delta)) (T = 1)$	$\log(\varepsilon(\Delta)) (T = 5)$	$\log(\varepsilon(\Delta)) (T = 10)$
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 5.6: Esquema Euler-implícito avaliado nos intervalos $[0, 1]$, $[0, 5]$ e $[0, 10]$.

Para não tornar a análise repetitiva, apresentamos apenas os resultados obtidos para os intervalos $[0, 1]$ e $[0, 15]$.

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 se-

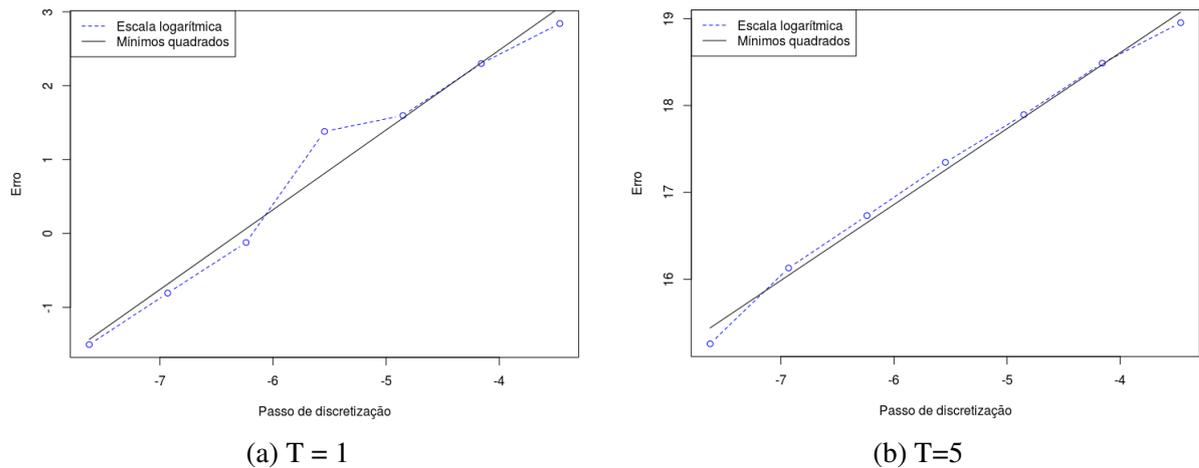


Figura 5.5: Esquema de Milstein avaliado nos intervalos $[0, 1]$ e $[0, 5]$.

	$T = 1$	$T = 5$
$\log M$	6.80	22.10
α	1.08	0.88

Tabela 5.7: Ajustes lineares para avaliação experimental do esquema de Milstein.

Passo de discretização		Erro esperado	
Δ	$\log \Delta$	$\log(\varepsilon(\Delta)) (T = 1)$	$\log(\varepsilon(\Delta)) (T = 5)$
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 5.8: Esquema de Milstein avaliado nos intervalos $[0, 1]$ e $[0, 5]$.

	$T = 1$	$T = 5$
$\log M$	6.85	22.26
α	1.07	0.88

Tabela 5.9: Ajustes lineares para avaliação experimental do esquema livre de derivada.

rão todos nulos. A avaliação experimental 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).

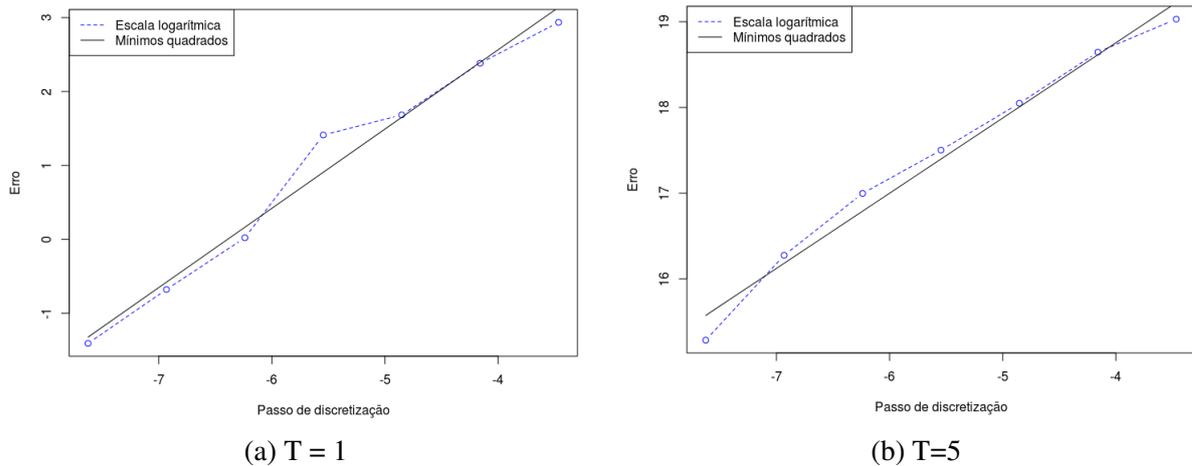


Figura 5.6: Esquema livre de derivada avaliado nos intervalos $[0, 1]$ e $[0, 5]$.

Passo de discretização		Erro esperado	
Δ	$\log \Delta$	$\log(\varepsilon(\Delta)) (T = 1)$	$\log(\varepsilon(\Delta)) (T = 5)$
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 5.10: Esquema livre de derivada avaliado nos intervalos $[0, 1]$ e $[0, 5]$.

5.3.2 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 1000000 de simulações;

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

$$2\lambda + \mu^2 = 2(-3) + 3 = -3 < 0,$$

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

$$\mathcal{R}_{em} = \left\{ \Delta \in \mathbb{R}^+; \Delta < \frac{1}{3} \right\}.$$

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.

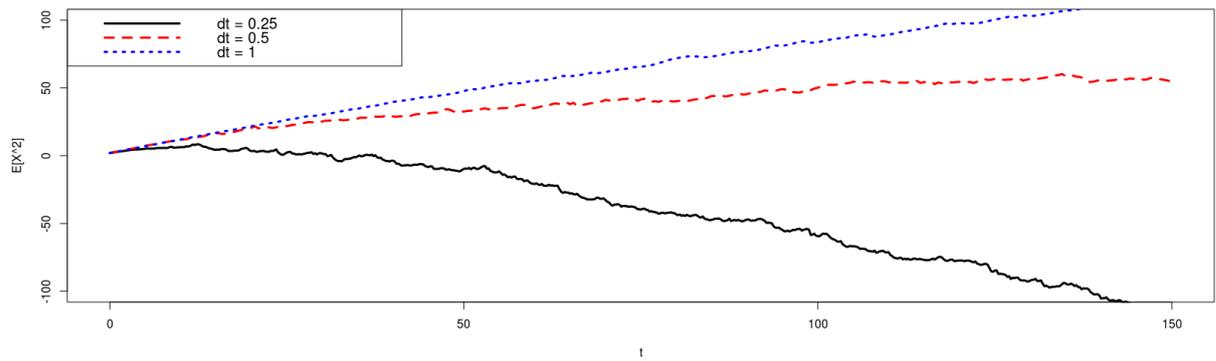
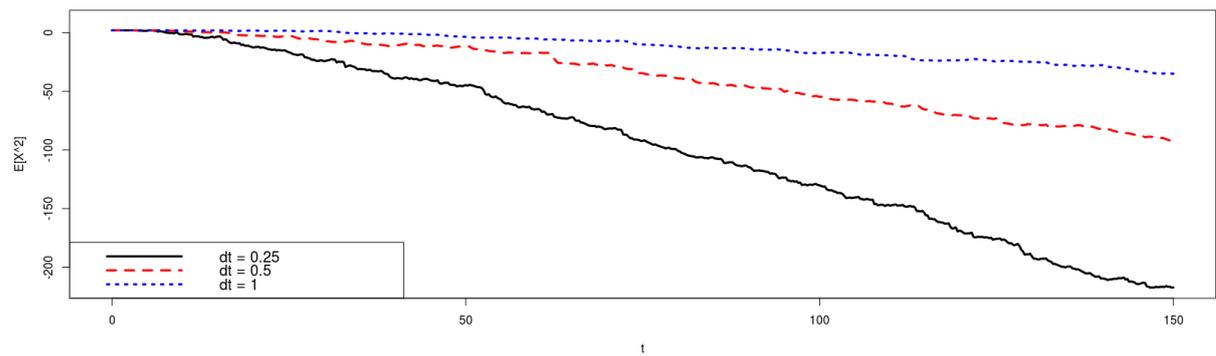
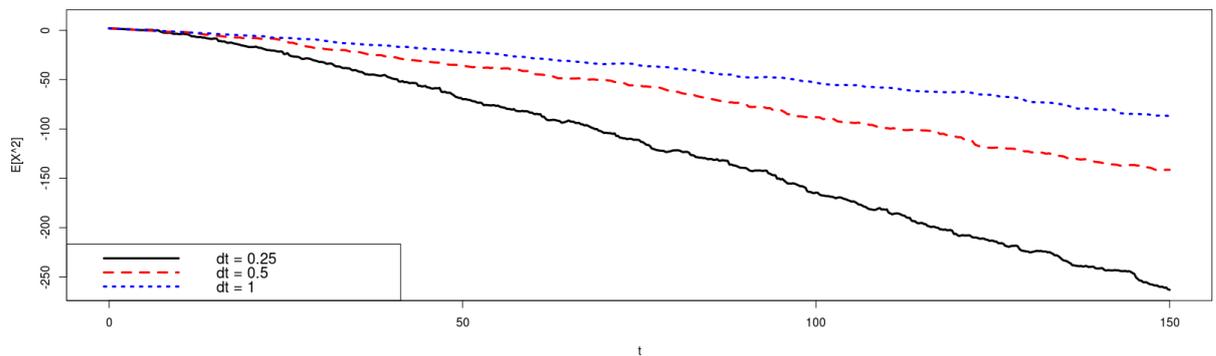


Figura 5.7: Estabilidade do esquema de Euler-Maruyama.

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, devemos esperar que as três curvas geradas por este experimento tendam a $-\infty$.



(a) Euler-trapezoidal ($\theta = 1/2$)



(b) Euler-implícito ($\theta = 1$)

Figura 5.8: Estabilidade do método θ estocástico.

A princípio, a figura 5.8(a) pode nos induzir a pensar 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 5.11: Evolução do esquema trapezoidal no intervalo $[0, 150]$.

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

$$\mathcal{R}_{ld} = \mathcal{R}_{mils} = \left\{ \Delta \in \mathbb{R}^+; \Delta < \frac{2}{9} \approx 0.22 \right\}.$$

Neste caso, ao utilizar os passos de discretização das simulações anteriores, devemos esperar que nenhum das três trajetórias apresente o comportamento decrescente. No entanto, este resultado não foi reproduzido corretamente. Como $0.25 > 0.22$, a trajetória em cor preta na figura

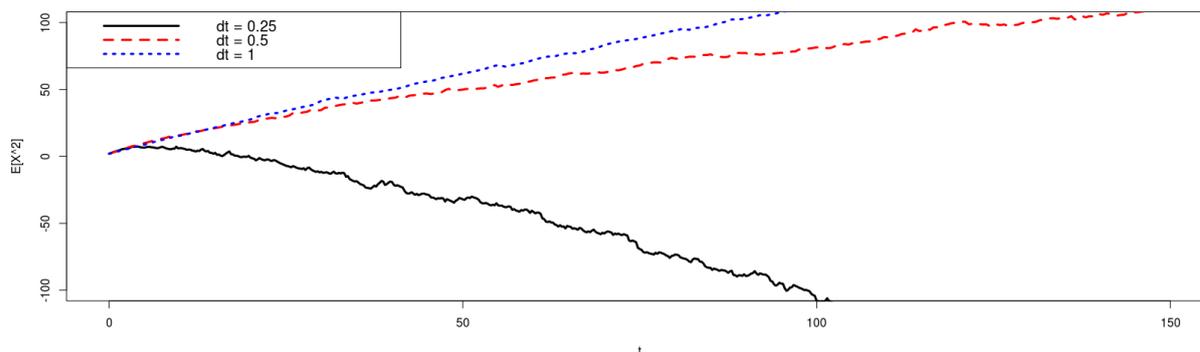


Figura 5.9: Estabilidade do método de Milstein.

acima não deveria 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 $\Delta = 0.3$ 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 $\Delta = \frac{1}{3}$. 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 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 $|R(\lambda, \mu, \Delta)| \approx 1$.

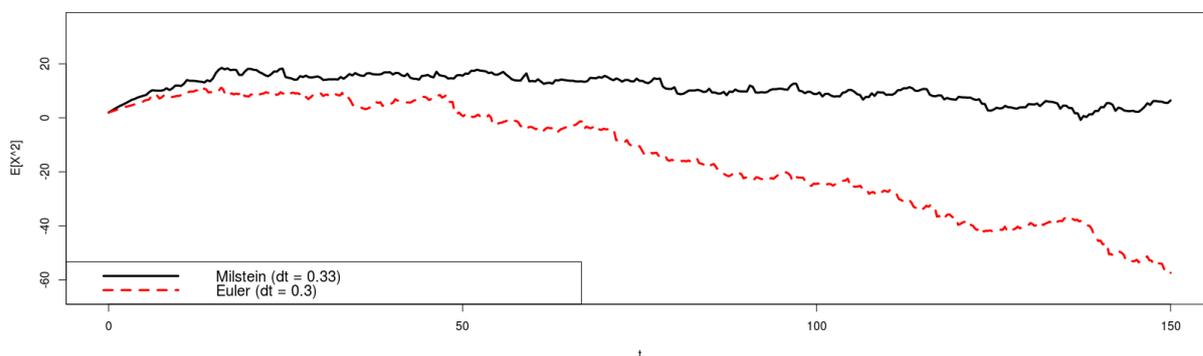


Figura 5.10: Comparação de estabilidade entre os métodos de Milstein e de Euler-Maruyama.

Os resultados obtidos para o esquema livre de derivada são similares aos do método de Milstein portanto serão omitidos.

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:

$$dX_t = X_t dt + \sqrt{X_t} dB_t, \quad t \in [0, 2], \quad (5.39)$$

$$X_0 = 2. \quad (5.40)$$

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:

$$dY_t = \frac{1}{2} \left(Y_t dt - \frac{1}{Y_t} \right) + dB_t, \quad t \in [0, 2], \quad (5.41)$$

$$Y_0 = 2\sqrt{2}. \quad (5.42)$$

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: $\Delta \in \{\frac{1}{64}, \frac{1}{16}, \frac{1}{4}\}$. Ilustramos primeiramente os resultados obtidos pelo método da linearização local. Perceba que a variação do passo de discretização

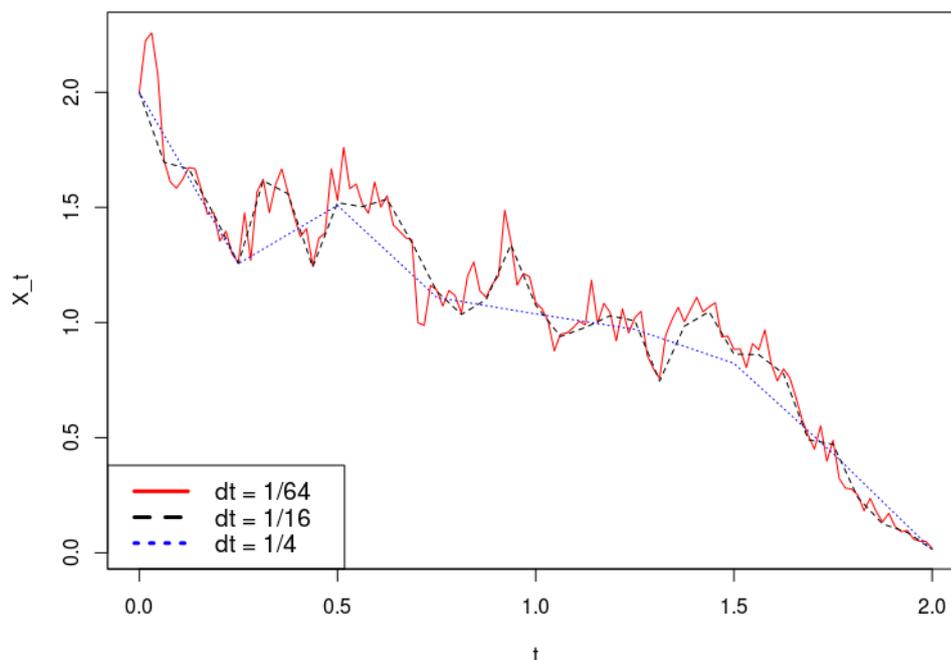


Figura 5.11: Método da linearização local aplicado ao problema (5.39)-(5.40).

não é refletida de maneira significativa 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:

$$\mathcal{R}_{mils} \subset \mathcal{R}_{em}.$$

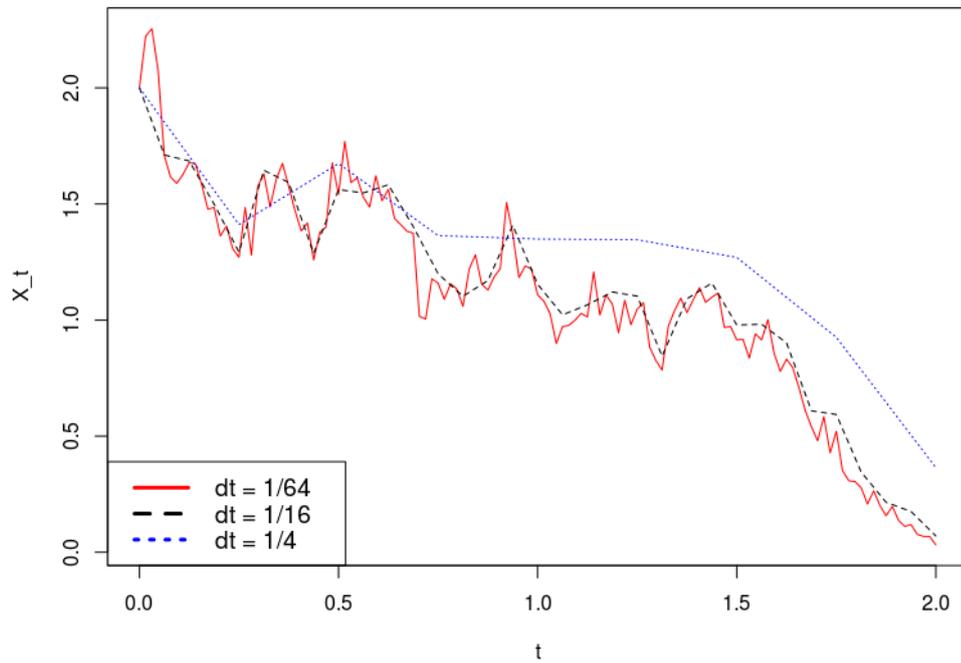


Figura 5.12: Método de Milstein aplicado ao problema (5.39)-(5.40).

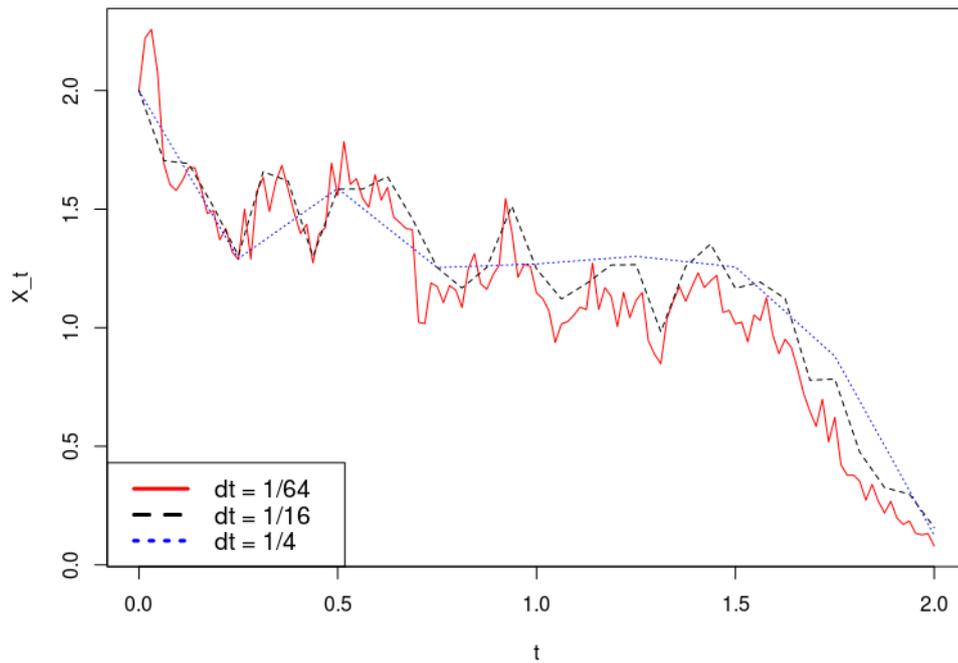


Figura 5.13: Método de Euler-Maruyama aplicado ao problema (5.39)-(5.40).

Os resultados acima 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:

$$dX_t = (5 - 11X_t + 6X_t^2 - X_t^3) dt + dB_t, \quad t \in [0, 10], \quad (5.43)$$

$$X_0 = 2. \quad (5.44)$$

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: $\Delta \in \{\frac{1}{16}, \frac{1}{4}, 1\}$. 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 um bom

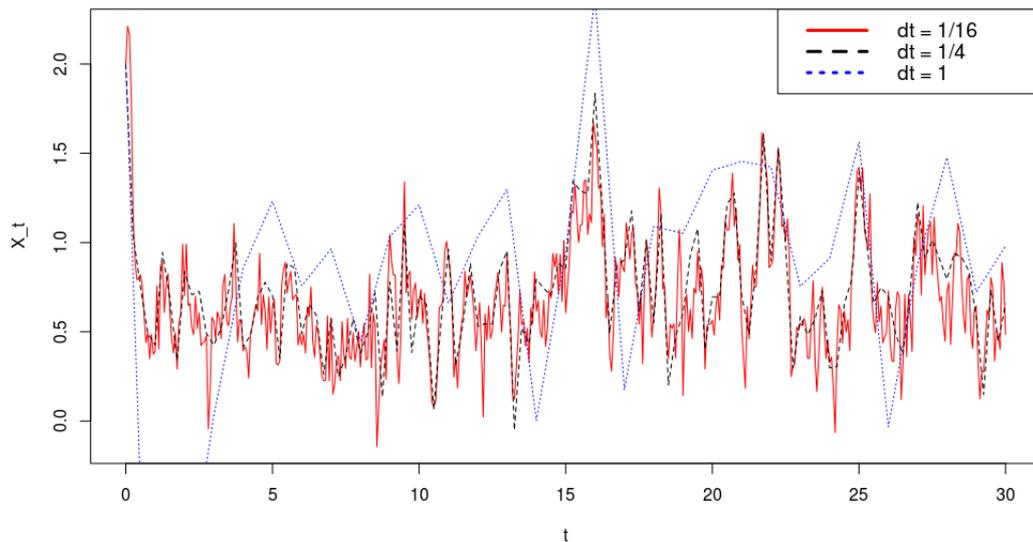


Figura 5.14: Método da linearização local aplicado ao problema (5.43)-(5.44).

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 no gráfico abaixo, onde a trajetória em azul pode até mesmo ser desprezada.

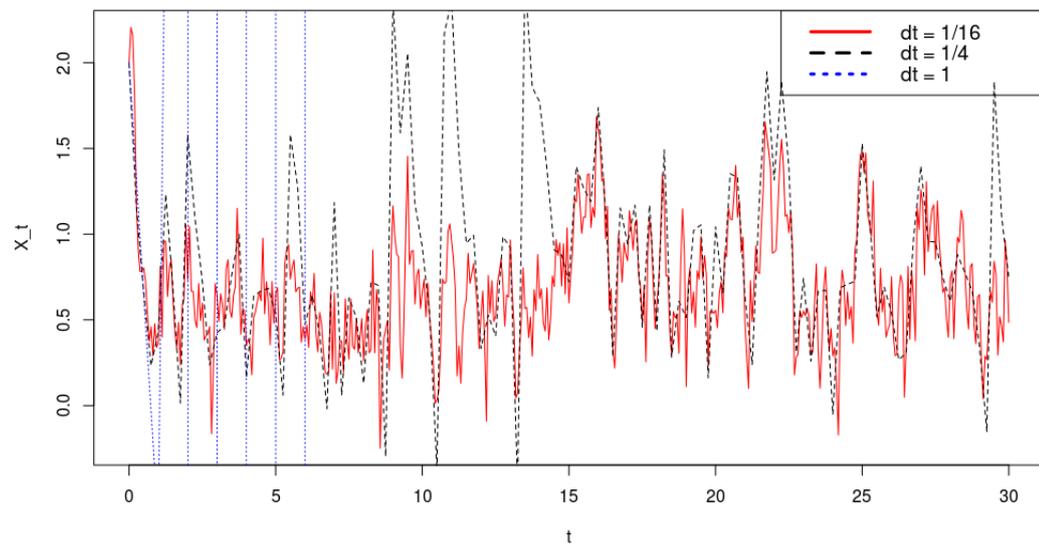


Figura 5.15: Método de Euler-Maruyama aplicado ao problema (5.43)-(5.44).

6 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.

6.1 O 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 6.1 (*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 $\varepsilon > 0$ tal que $f := \varepsilon \frac{d\mu}{d\lambda} \leq 1$, 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*

$$\tau := \min\{i \in \mathbb{N}; I_i = 1\},$$

temos que

$$P[Y_\tau \in B] = \mu(B), \forall B \in \mathcal{B}. \quad (6.1)$$

Demonstração: Dado $i \in \mathbb{N}$, temos que

$$\begin{aligned} P[I_i = 1] &= \int_{\Omega} P[I_i = 1 | Y = y] d\lambda(y) \\ &= \int_{\Omega} f(y) d\lambda(y) = \int_{\Omega} \varepsilon d\mu(y) = \varepsilon. \end{aligned}$$

Além disso, dado $B \in \mathcal{B}$, podemos utilizar a independência entre Y_τ e I_1 de modo a obter

$$\begin{aligned} P[Y_\tau \in B] &= P[Y_\tau \in B, I_1 = 1] + P[Y_\tau \in B, I_1 = 0] \\ &= P[Y_\tau \in B, I_1 = 1] + P[Y_\tau \in B | I_1 = 1] P[I_1 = 0] \\ &= P[Y_\tau \in B, I_1 = 1] + P[Y_\tau \in B] P[I_1 = 0] \\ &= P[Y_\tau \in B, I_1 = 1] + P[Y_\tau \in B](1 - \varepsilon). \end{aligned}$$

A probabilidade conjunta utilizada acima é equivalente a

$$\begin{aligned} P[Y_\tau \in B, I_1 = 1] &= \int_B P[I_1 = 1 | Y_\tau = y] d\lambda(y) \\ &= \int_B f(y) d\lambda(y) = \int_B \varepsilon d\mu(y) = \varepsilon \mu(B). \end{aligned}$$

Assim,

$$\begin{aligned} P[Y_\tau \in B] &= \varepsilon \mu(B) + P[Y_\tau \in B](1 - \varepsilon) \\ \therefore P[Y_\tau \in B] &= \mu(B). \end{aligned}$$

□

O algoritmo que será desenvolvido para a simulação exata de trajetórias é aplicado a processos de Itô $X = \{X_t; 0 \leq t \leq T\}$ que satisfazem

$$dX_t = f(X_t) + dB_t, \quad 0 \leq t \leq T, \quad (6.2)$$

$$X_0 = x_0. \quad (6.3)$$

Assumimos aqui que a equação 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 geral. De uma maneira em 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:

$$B_t(\omega) := \omega(t), t \in [0, T]; \quad \mathcal{C} := \sigma(\{B_t; 0 \leq t \leq T\}). \quad (6.4)$$

Introduzimos a seguinte notação para distinguir um movimento Browniano padrão das funções coordenadas B_t :

$$W := \{W_t; 0 \leq t \leq T\}. \quad (6.5)$$

Desta forma, se \mathbb{W} é a medida de Wiener em (C, \mathcal{C}) , então $B = \{B_t; 0 \leq t \leq T\}$ é um movimento Browniano.

Seja \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 (6.2)-(6.3). Denote por \mathbb{W} a medida de Wiener e suponha que o coeficiente f satisfaz a condição de Novikov:

$$\mathbb{E}_{\mathbb{W}} \left[\exp \left\{ \frac{1}{2} \int_0^T f^2(B_t) dt \right\} \right] < +\infty.$$

Neste caso, a derivada de Radon-Nykodim \mathbb{Q}/\mathbb{W} pode ser obtida a partir da transformação de Cameron-Martin-Girsanov:

$$\frac{d\mathbb{Q}}{d\mathbb{W}}(\omega) = \exp \left\{ \int_0^T f(B_t) dB_t - \frac{1}{2} \int_0^T f^2(B_t) dt \right\} =: G(B). \quad (6.6)$$

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

$$F(B_t) := \int_0^b f(x) dx \Big|_{b=B_t}$$

e aplicando a fórmula de Itô:

$$\begin{aligned} dF_t &= \frac{\partial}{\partial t} \int_0^b f(x) dx \Big|_{b=B_t} + \left(\frac{\partial}{\partial x} \int_0^b f(x) dx \Big|_{b=B_t} \right) dB_t \\ &\quad + \frac{1}{2} \left(\frac{\partial^2}{\partial x^2} \int_0^b f(x) dx \Big|_{b=B_t} \right) (dB_t)^2 \\ &= 0 + f(B_t) dB_t + \frac{1}{2} f'(B_t) dt. \end{aligned}$$

Integrando a expressão acima em $[0, T]$, obtemos

$$F(B_t) - F(B_0) = \int_0^T f(B_t) dB_t + \frac{1}{2} \int_0^T f'(B_t) dt.$$

Assim,

$$\int_0^T f(B_t) dB_t + F(B_t) - F(B_0) - \frac{1}{2} \int_0^T f'(B_t) dt, \quad (6.7)$$

e a densidade G será então expressa por

$$G(B) = \exp \left\{ F(B_t) - F(B_0) - \frac{1}{2} \int_0^T (f^2(B_t) + f'(B_t)) dt \right\}. \quad (6.8)$$

O método da rejeição exige que $G(\cdot)$ seja limitada quase certamente e esta restrição está ligada à limitação de $F(\cdot)$. Esse problema pode ser muito restritivo e, portanto, precisa ser contornado. Para tal, será introduzida uma terceira medida de probabilidade, a qual será utilizada para gerar os candidatos a amostra para o método da rejeição. Considere a ponte Browniana $PB = \{PB_t; 0 \leq t \leq T\}$, heurísticamente definida por $W | W_T = \rho$, com ρ distribuída de acordo com alguma função densidade de probabilidade $h : \mathbb{R} \rightarrow [0, +\infty)$ com respeito à medida de Lebesgue.

Teorema 6.2 *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

$$\chi_h = \{x \in \mathbb{R}; h(x) > 0\} = \mathbb{R},$$

então \mathbb{Z} é equivalente a \mathbb{W} e

$$\frac{d\mathbb{Z}}{d\mathbb{W}}(\omega) = \frac{h(B_T)}{(1/\sqrt{2\pi T}) \exp\{-(B_T - x_0)^2/(2T)\}}. \quad (6.9)$$

Demonstração: Defina

$$f(\omega) := h(B_T) \sqrt{2\pi T} / \exp\{-(B_T - x_0)^2/(2T)\}, \quad \omega \in C,$$

de modo que f seja $\sigma(B_T)$ -mensurável. Dado $A \in \mathcal{C}$, segue, da definição do processo PB , que

$$\mathbb{Z}[A | \sigma(B_T)] = \mathbb{W}[A | \sigma(B_T)] =: g(B_T), \quad \mathbb{W} - q.c.,$$

para alguma função $g : \mathbb{R} \rightarrow \mathbb{R}$, Borel-mensurável. Uma vez que $B_T \sim N(x_0, T)$, com respeito a \mathbb{W} , temos que

$$\begin{aligned} \mathbb{E}_{\mathbb{W}} &= \mathbb{E}_{\mathbb{W}} [\mathbb{E}_{\mathbb{W}} [\mathbf{1}_A f | \sigma(B_T)]] \\ &= \mathbb{E}_{\mathbb{W}} [f \cdot \mathbb{W}[A | \sigma(B_T)]] = \mathbb{E}_{\mathbb{W}} [fg] = \int_{\mathbb{R}} f(u)g(u)\mathbb{W}(du) \\ &= \int_{\mathbb{R}} \frac{h(u)\sqrt{2\pi T}}{\exp\{-(u-x_0)^2/(2T)\}} \cdot \frac{\exp\{-(u-x_0)^2/(2T)\}}{\sqrt{2\pi T}} g(u)du = \int_{\mathbb{R}} h(u)g(u)du. \end{aligned}$$

Concluimos assim que

$$f = d\mathbb{Z}/d\mathbb{W}.$$

Além disso, como o suporte de h é a reta, temos que $f > 0$. Desta forma, \mathbb{W} é equivalente a \mathbb{Z} e

$$1/f = d\mathbb{W}/d\mathbb{Z}.$$

□

A partir dos resultados apresentados acima, fica claro que

$$\begin{aligned} \frac{d\mathbb{Q}}{d\mathbb{Z}}(\omega) &= \frac{d\mathbb{Q}}{d\mathbb{W}}(\omega) \frac{d\mathbb{W}}{d\mathbb{Z}}(\omega) \\ &= \frac{\exp\left\{F(B_T) - F(B_0) - \frac{1}{2} \int_0^T (f^2(B_t) + f'(B_t)) dt\right\}}{h(B_T) / (1/\sqrt{2\pi T}) \exp\{-(B_T - x_0)^2/(2T)\}} \\ &= \sqrt{2\pi T} \cdot \exp\{-F(B_0)\} \cdot \exp\left\{F(B_T) - \frac{(B_T - x_0)^2}{2T} - \frac{1}{2} \int_0^T (f^2(B_t) + f'(B_t)) dt\right\} / h(B_T) \\ &\propto \exp\left\{F(B_T) - \frac{(B_T - x_0)^2}{2T} - \frac{1}{2} \int_0^T (f^2(B_t) + f'(B_t)) dt\right\} / h(B_T). \end{aligned} \quad (6.10)$$

Assumindo que

$$\int_{\mathbb{R}} \exp\{F(u) - (u - x_0)^2/2T\} du := c < +\infty, \quad (6.11)$$

podemos definir (convenientemente) a função densidade de probabilidade

$$h(u) := \exp\{F(u) - (u - x_0)^2/2T\}/c, \quad (6.12)$$

de modo a obter

$$\frac{d\mathbb{Q}}{d\mathbb{Z}}(\omega) \propto \exp\left\{-\frac{1}{2} \int_0^T (f^2(B_t) + f'(B_t)) dt\right\}. \quad (6.13)$$

Assumindo agora que o funcional em (6.13) 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 (6.13) como

$$\frac{d\mathbb{Q}}{d\mathbb{Z}} \propto \exp\left\{-\int_0^T \phi(B_t) dt\right\} \leq 1, \quad \mathbb{Z} - q.c., \quad (6.14)$$

onde $\phi > 0$ é dada por

$$\phi(u) = \frac{f^2(u) + f'(u)}{2} - K, \quad \forall u \in \mathbb{R}. \quad (6.15)$$

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. Apresentamos abaixo um resumo das hipóteses utilizadas para a construção da densidade (6.14):

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.

Apresentaremos agora a versão do algoritmo correspondente ao caso em que ϕ é também limitada superiormente. A versão final deste algoritmo, desenvolvida por Beskos et al. (2006), foi denominada 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.

Em uma situação ideal onde caminhos amostrais $\omega \sim \mathbb{Z}$ em $[0, T]$ são disponíveis e a integral (6.14) pode ser obtida de maneira analítica, temos que o método da rejeição pode ser aplicado de maneira simples. Na prática, a informação $\omega \sim \mathbb{Z}$ é disponível apenas em uma quantidade discreta de pontos e, neste caso, o método da rejeição precisa ser adaptado. O teorema a seguir resolve o problema gerado por estas duas restrições, porém, sua utilização está condicionada à existência de uma cota superior para ϕ .

Teorema 6.3 *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 $\{(t, \phi(B_t)); t \in [0, T]\}$, então*

$$P[N = 0 | \omega] \propto \exp \left\{ - \int_0^T \phi(B_t) dt \right\}. \quad (6.16)$$

Demonstração: Dada uma área $A \in [0, T] \times [0, M(\omega)]$, temos que

$$P(N = K | A) = \frac{e^{-T \cdot M \cdot P(A)} [T \cdot M \cdot P(A)]^K}{K!},$$

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 $\{(t, \phi(B_t)); t \in [0, T]\}$, então

$$P(A) = \frac{1}{T \cdot M} \int_0^T \phi(B_t) dt,$$

e

$$\begin{aligned} P(N = 0 | \omega) &= \frac{\exp\left\{-\int_0^T \phi(B_t) dt\right\} \left(\int_0^T \phi(B_t) dt\right)^0}{0!} \\ &= \exp\left\{-\int_0^T \phi(B_t) dt\right\}. \end{aligned}$$

□

O teorema acima nos permite aplicar a método da rejeição baseado em observações discretas das trajetórias de X .

Algoritmo EA1

1. Simule uma realização $\{(x_{i,1}, x_{i,2})\}_{i=1}^{\tau}$ 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 5, senão, volte para 1;
 4. Retorne o esqueleto atual $S(\omega)$ de ω .
-

Um dos pontos curiosos do algoritmo EA1 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.

6.2 Avaliação Experimental

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:

$$dX_t = \text{sen}(X_t) + dB_t, \quad 0 \leq t \leq T, \quad (6.17)$$

$$X_0 = 0. \quad (6.18)$$

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:

$$F(u) = \int_0^u \text{sen}(x)dx = 1 - \text{cos}(u).$$

Como $1 - \cos(x) \geq 0, \forall x \in \mathbb{R}$, segue que

$$\begin{aligned} \exp \left\{ F(u) - \frac{(x_0 - u)^2}{2T} \right\} &= \exp \left\{ 1 - \cos(u) - \frac{u^2}{2T} \right\} \\ &\leq e \cdot \exp \left\{ \frac{u^2}{2T} \right\}. \end{aligned}$$

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

$$\frac{1}{2}(f^2(x) + f'(x)) = \frac{1}{2}(\text{sen}^2(x) + \cos(x)) \geq -\frac{1}{2}.$$

Verificadas as três condições acima, podemos utilizar o problema (6.17)-(6.18) como caso de teste.

O procedimento realizado é descrito pelos seguintes passos:

- Passo 1: simulação de 1000000 de observações da variável aleatória X_T através do algoritmo de simulação exata;
- Passo 2: simulação de 1000000 de 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.

As simulações foram executadas na mesma máquina descrita na seção 5.3. Por uma questão de eficiência, adotamos o intervalo de tempo $[0, 1]$, desta forma, estaremos interessados na simulação da variável aleatória X_1 . A simulação das observações através do algoritmo EA1 foi impressionantemente rápida; o tempo de execução foi de apenas 4.08s. É importante ressaltar que este tempo de execução só foi obtido pois este método não carece da construção de toda uma trajetória até o instante imediatamente anterior. Utilizando o método do núcleo estimador, chegamos à densidade estimada ilustrada pela figura abaixo.

Lembramos que este método reproduz exatamente a distribuição da variável X_1 ; 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,

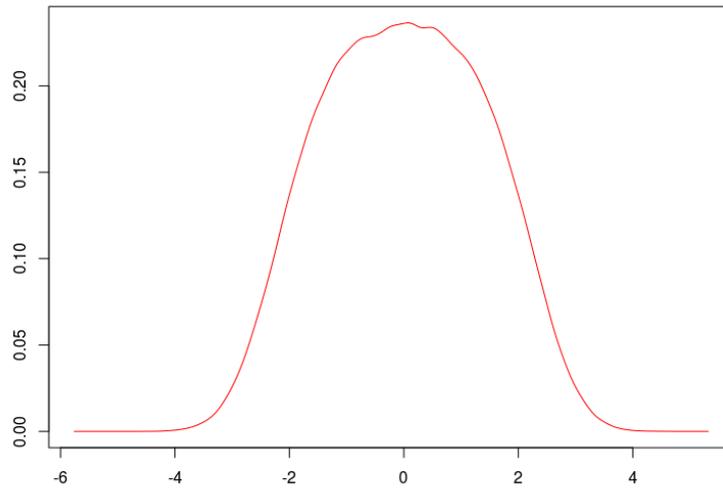


Figura 6.1: Densidade estimada da variável X_1 correspondente ao problema (6.17)-(6.18).

devemos obter um comportamento parecido ao simular X_1 a partir destas aproximações. Este fato fica evidenciado pela figura abaixo, 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. Um aproximação ainda mais

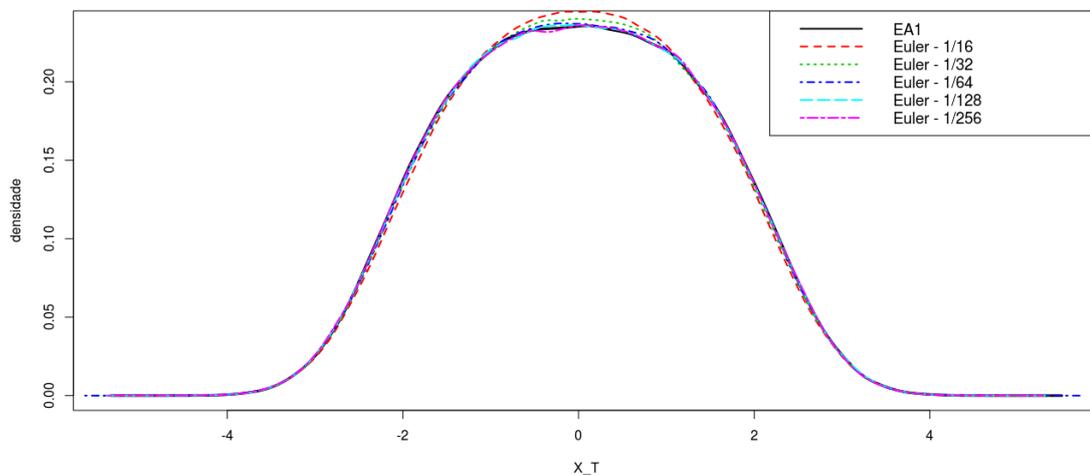


Figura 6.2: Comparação entre as densidades estimadas da variável X_1 - Euler-Maruyama \times EA1.

interessante é observada quando utilizamos o método da linearização local. Mesmo utilizando espaçamentos razoavelmente grandes, o núcleo estimado para a variável X_1 a partir deste esquema coincide com o obtido pelo algoritmo EA1.

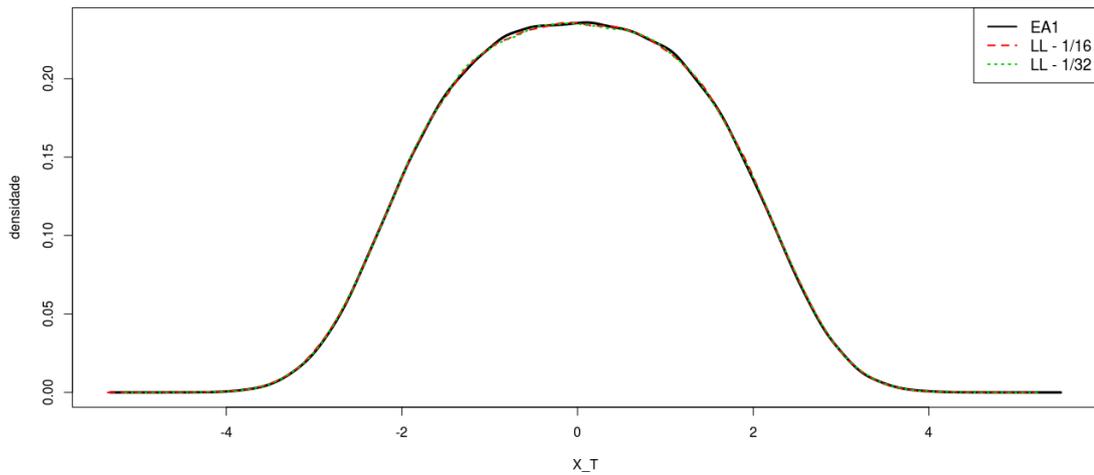


Figura 6.3: Comparação entre as densidades estimadas da variável X_1 - Linearização Local \times EA1.

Apresentamos agora os resultados obtidos pelo método de Euler-Maruyama. A tabela a seguir ilustra o tempo de execução de cada simulação, o passo de discretização considerado e o p -valor do teste de Kolmogorov-Smirnov para a comparação entre as amostras.

Tempo de execução	Δ	p -valor
1.93s	2^{-4}	8.02e-35
3.75s	2^{-5}	2.26e-12
7.65s	2^{-6}	2.98e-04
14.80s	2^{-7}	0.13
29.55s	2^{-8}	0.13
58.13s	2^{-9}	0.73

Tabela 6.1: Esquema de Euler-Maruyama \times algoritmo EA1.

O teste é significativo apenas a partir de $\Delta = 2^{-7}$. Este resultado é razoável devido à baixa ordem de convergência forte do esquema. Perceba ainda que o método é computacionalmente eficiente. Resultados similares foram obtidos pelo método de Milstein. A tabela abaixo 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.

Tempo de execução	Δ	p -valor
2.25s	2^{-4}	7.84e-40
4.27s	2^{-5}	4.73e-13
8.58s	2^{-6}	1.52e-06
16.81s	2^{-7}	0.08
33.32s	2^{-8}	0.06
65.48s	2^{-9}	0.70

Tabela 6.2: Esquema de Milstein \times algoritmo EA1.

Os resultados mais expressivos foram novamente obtidos pelo método da linearização local. Perceba que todos os passos de discretização levaram à aceitação no teste de Kolmogorov-Smirnov. 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 supere em muito os demais devido à necessidade do cálculo de exponenciais a cada iteração do algoritmo.

Tempo de execução	Δ	p -valor
3.47s	2^{-4}	0.52
6.62s	2^{-5}	0.06
13.30s	2^{-6}	0.19
26.39s	2^{-7}	0.31
52.64s	2^{-8}	0.66
103.82s	2^{-9}	0.52

Tabela 6.3: Método da linearização local \times algoritmo EA1.

7 *Conclusões e Trabalhos Futuros*

Apresentamos aqui alguns dos métodos numéricos mais importantes no estudo de equações diferenciais estocásticas, explorando tanto os algoritmos mais tradicionais quanto os mais contemporâneos. Ficamos restritos ao caso unidimensional pois este tópico já é rico o suficiente a ponto de nos permitir abordar muitos dos conceitos cruciais na área de simulação e análise numérica de equações diferenciais estocásticas. Os algoritmos discutidos aqui, com exceção ao método da linearização local e o algoritmo EA1, podem ser facilmente estendidos para o caso n -dimensional. Estes dois casos em particular, apresentam uma dificuldade com relação à estabilização do coeficiente de difusão. A aplicação da transformada de Lamperti a um sistema n -dimensional não é trivial e restrita a determinadas condições de regularidade que podem não ser razoáveis do ponto de vista da modelagem. Ainda assim, é possível obter uma versão n -dimensional destes algoritmos.

Um ponto importante a ser destacado é o método de Euler-Maruyama. Os resultados apresentados neste trabalho evidenciam o quanto este método é interessante sobretudo em se tratando de estabilidade. Além disso, sua maior virtude é a simplicidade: apresenta poucas restrições de aplicação, é extremamente eficiente do ponto de vista computacional e é de fácil implementação. Mesmo nos casos onde o passo de discretização deve ser pequeno, este fato pode ser compensado pela eficiência da simulação. O método de Milstein mostrou ter um desempenho similar ao de Euler-Maruyama. Casos onde ambos são estáveis, pode ser interessante aplicarmos o método de Milstein para explorar sua ordem de convergência superior. Lembramos ainda que a equação diferencial estocástica pode, em vários casos, ser simplificada pela transformada de Lamperti. Ao aplicar esta transformação, temos que o método de Euler-Maruyama e Milstein nos levarão ao mesmo resultado. O esquema livre de derivada, conforme os experimentos evidenciaram, pode ser entendido como uma versão alternativa do método de Milstein a ser utilizada quando a derivada do coeficiente de difusão é difícil de ser obtida ou tratada computacionalmente. Com respeito a este tipo de estratégia, existe toda uma família de esquemas numéricos denominados métodos de Runge-Kutta. Tais métodos merecem muita atenção pois o tratamento de derivadas pode complicar muito o desenvolvimento de vários mo-

delos. Não entramos em maiores detalhes a respeito desta família de métodos pois a riqueza de detalhes de tais métodos já seria suficiente para o desenvolvimento de um segundo trabalho.

Deixamos a desejar um pouco na análise teórica do método da linearização local porém os resultados evidenciados pelos experimentos computacionais foram muito satisfatórios. Eles deixaram bem claro que este método supera por muito os demais esquemas baseados na série de Taylor estocástica. Conforme citado anteriormente, sua maior carência reside na necessidade de um coeficiente de difusão constante. Tal restrição só não é considerada excessiva devido à possibilidade de aplicação da transformada de Lamperti. A qualidade deste esquema é superada apenas pelo algoritmo de simulação exata, o qual possui um preço alto correspondente ao grande número de restrições imposta pelo desenvolvimento do método. Uma versão mais interessante deste algoritmo (EA2) baseia-se em processos de Bessel ao invés de pontes Brownianas. Com esta abordagem, os autores conseguiram afrouxar as restrições com respeito à limitação da função ϕ , tornando assim o algoritmo mais geral.

O método θ estocástico foi o único esquema implícito abordado pois o trabalho não tinha como objetivo explorar a fundo a teoria de estabilidade assintótica. Este tipo de problema torna-se mais interessante quando passamos para modelos multidimensionais, onde surgem os problemas *stiff*. Tais problemas são caracterizados pela discrepância entre as velocidades de convergência entre as várias componentes do sistema, o que causa grande instabilidade aos métodos. Outros conceitos de estabilidade também poderiam ser levadas em consideração construindo assim uma discussão mais ampla sob as condições e casos específicos que levam um critério de estabilidade a ser mais interessante que os demais. Outras famílias de métodos numéricos, assim como o método θ , são ideias para este tipo de tratamento; por exemplo: método *preditor-corretor*, esquemas implícitos de passos múltiplos, etc.

Faltou abordarmos como aplicação de todos estes métodos, tópicos relacionados a inferência estatística. Os processos discretizados abordados aqui são interessantes neste contexto pois suas densidades de transição são obtidas de maneira mais simples, o que nos permite explorar com mais naturalidade conceitos relacionados a inferência paramétrica. Ainda com respeito a este tema, devemos ressaltar a importância da transformação de Camero-Martin-Girsanov. Este resultado nos permite associar densidades aos processos, tornando possível a aplicação do método da máxima verossimilhança. A curto prazo, o estudo dessa área será mais aprofundado como continuidade do trabalho.

Por fim, apontamos o desenvolvimento de métodos numéricos para equações diferenciais estocásticas parciais como o principal foco de trabalho futuro. Outras linhas de métodos serão necessárias nesta área devido à maior complexidade dos problemas. O ferramental adquirido

com a compreensão dos conceitos e métodos utilizados no caso das equações ordinárias será de extrema importância para esta transição.

Referências Bibliográficas

- [1] A. Beskos, O. Papaspiliopoulos, and G. O. Roberts. Retrospective exact simulation of diffusion sample paths with applications. *Bernoulli*, 12(6):1077–1098, 2006.
- [2] A. Beskos and G. O. Roberts. Exact simulation of diffusions. *Annals of Applied Probability*, 15:2422–2444, 2005.
- [3] J. Bishwal. *Parameter Estimation in Stochastic Differential Equations*. Lecture Notes in Mathematics. Springer, 2007.
- [4] K. Burrage, P. Burrage, and T. Mitsui. Numerical solutions of stochastic differential equations - implementation and stability issues. *Journal of Computational and Applied Mathematics*, 125:171–182, 2000.
- [5] K. Burrage and P. M. Burrage. High strong order explicit Runge-Kutta methods for stochastic ordinary differential equations. *Applied Numerical Mathematics*, 22:81–101, 1996.
- [6] K. Burrage and P. M. Burrage. General order conditions for stochastic Runge-Kutta methods for both commuting and non-commuting stochastic ordinary differential equation systems. *Applied Numerical Mathematics*, 28:161–177, 1998.
- [7] L. Chen and F. Wu. Choice of θ and its effects on stability in the stochastic θ -method of stochastic delay differential equations. *International Journal of Computer Mathematics*, 89(15):2106 – 2122, 2012.
- [8] A. Friedman. *Stochastic Differential Equations and Applications: Two Volumes Bound as One*. Dover Books on Mathematics Series. Dover Publications, 1975.
- [9] J. G. Gaines and T. J. Lyons. Variable Step Size Control in the Numerical Solution of Stochastic Differential Equations. *Siam Journal on Applied Mathematics*, 57, 1997.
- [10] D. J. Higham. Mean-square and asymptotic stability of the stochastic theta method. *SIAM J. Numer. Anal.*, 38:753–769, 2000.
- [11] D. J. Higham. An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM Review*, 43(3):525–546, 2001.
- [12] S. Iacus. *Simulation and Inference for Stochastic Differential Equations: With R Examples*. Springer Series in Statistics. Springer-Verlag New York, 2008.
- [13] I. Karatzas and S. Shreve. *Brownian Motion and Stochastic Calculus*. Graduate Texts in Mathematics Series. Springer-Verlag GmbH, 1991.
- [14] P. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Applications of Mathematics Series. Springer, 1999.

- [15] P. E. Kloeden and E. Platen. A survey of numerical methods for stochastic differential equations. *J. Stoch. Hydrol. Hydraulics*, 3:155–178, 1987.
- [16] H. Liske and E. Platen. Simulation studies on time discrete diffusion approximations. *Math. Comput. Simulation*, 29:253–260, 1987.
- [17] H. Luschgy and G. Pagès. Functional quantization of a class of brownian diffusions: A constructive approach. *Stochastic Processes and their Applications*, 116(2):310–336, 2006.
- [18] E. Meshane. *Stochastic Calculus and Stochastic Models*. Probability and Mathematical Statistics. Academic Press, 1974.
- [19] G. Milstein. *Numerical Integration of Stochastic Differential Equations*. Mathematics and Its Applications. Springer, 1994.
- [20] N. J. Newton. Asymptotically efficient runge-kutta methods for a class of ito and stratonovich equations. *J. Appl. Math.*, 51:542–567, 1991.
- [21] A. Novikov. On an identity for stochastic integrals. *Theory Probab. Appl.*, 17(3):717–720, 1972.
- [22] B. Øksendal. *Stochastic Differential Equations: An Introduction with Applications*. Universitext (1979). Springer, 2010.
- [23] S. Resnick. *A Probability Path*. Birkhäuser Boston, 1998.
- [24] G. O. Roberts and O. Stramer. On inference for partially observed nonlinear diffusion models using the metropolis-hastings algorithm. *Biometrika*, 88(3):603–621, 2001.
- [25] W. Rümelin. Numerical Treatment of Stochastic Differential Equations. *SIAM Journal on Numerical Analysis*, 19(3), 1982.
- [26] Y. Saito and T. Mitsui. Stability analysis of numerical schemes for stochastic differential equations. *SIAM Journal on Numerical Analysis*, 33(6):pp. 2254–2267, 1996.
- [27] A. Shiryaev and S. Wilson. *Probability*. Graduate Texts in Mathematics. Springer, 1995.
- [28] I. Shoji. A note on convergence rate of a linearization method for the discretization of stochastic differential equations. *Communications in Nonlinear Science and Numerical Simulation*, 16(7):2667–2671, 2011.
- [29] I. Shoji and T. Ozaki. Estimation for nonlinear stochastic differential equations by a local linearization method. *Stochastic Analysis and Applications*, 16(4):733–752, 1998.