



Uma Introdução à Resolução de Equações Diferenciais Estocásticas: Soluções Exatas e Aproximações Numéricas

Dalila Gil Franco Alves

Dissertação para a obtenção do Grau de

Mestre em Métodos Quantitativos em Finanças

Júri

Presidente: Paulo Eduardo Aragão Aleixo e Neves de Oliveira

Orientador: Ana Cristina Martins Rosa

Vogal: José Augusto Mendes Ferreira

Data: Setembro de 2012

Resumo

A teoria das equações diferenciais estocásticas é de extrema importância em muitas áreas científicas, em particular, na economia. De facto, em finanças, estas equações são instrumentos essenciais para modelar a evolução dos preços dos activos, assim como taxas de juro e seus derivados. Dado que o tempo e a incerteza são os elementos centrais que influenciam o comportamento dos agentes económicos, as equações diferenciais estocásticas constituem a ferramenta matemática privilegiada de análise da teoria económica e financeira avançada. Uma das razões da sua crescente utilização deve-se às potencialidades do cálculo estocástico de Itô. Nesta dissertação, começamos por efetuar uma introdução ao problema da resolução das equações diferenciais estocásticas, nomeadamente aquelas em que é possível encontrar a forma explícita da solução forte. No entanto, para a grande maioria, tal não é possível, sendo então necessário utilizar técnicas numéricas para encontrar uma aproximação da solução. Iremos, assim, introduzir métodos numéricos que permitem construir uma aproximação da solução forte de uma equação diferencial estocástica. Abordam-se os métodos de Euler-Maruyama e de Milstein, casos particulares dos designados métodos de Taylor estocásticos, que se obtêm a partir da expansão de Itô-Taylor. O desempenho dos métodos será analisado no sentido da convergência forte. Por último, apresentam-se algumas aplicações destes dois métodos a exemplos clássicos do contexto das finanças.

Palavras Chave:

Equações Diferenciais Estocásticas, solução forte, métodos numéricos, convergência forte, método de Euler-Maruyama, método de Milstein.

Abstract

The theory of stochastic differential equations is of most importance for many scientific fields, notably economics. In fact, these equations are essential instruments, in finance, to model asset prices, interest rates and their derivatives. Since time and uncertainty are key elements that influence economic agents' behavior, stochastic differential equations have been the main tool for analyzing the advanced economic and financial theory. One of the reasons for the growing use of such equations is due to the potential of Itô stochastic calculus. In this thesis, we begin with an introduction to the resolution of stochastic differential equations, in particular, those that can be solved exactly in the strong sense. However, for the great majority, this is not possible and numerical techniques are needed in order to approximate the solution. Thus, we introduce numerical methods that allow us to construct an approximation of the solution of a stochastic differential equation. We address Euler-Maruyama and Milstein methods, which are special cases of the so-called Taylor methods, obtained from Itô-Taylor expansion. Their performance will be evaluated by means of strong convergence. Finally, we present some applications of Euler-Maruyama and Milstein methods to classical examples in the finance framework.

Keywords:

Stochastic Differential Equations, strong solution, numerical methods, strong convergence, Euler-Maruyama method, Milstein method.

Agradecimentos

À Professora Doutora Ana Cristina Rosa, minha orientadora, pelo apoio, ensinamentos e disponibilidade reveladas ao longo deste trabalho, assim como pelas críticas, correções e sugestões feitas durante a orientação.

À minha mãe pelo seu amor e compreensão e pelo constante incentivo e apoio.

A toda a minha família, especialmente à minha avó Zé (in memoriam), que sempre fizeram o possível para as realizações da minha vida.

Ao Hugo por toda a dedicação e apoio que me deu e por me incentivar a perseguir os meus sonhos.

Ao meu colega Rui Brito pela amizade e pelos bons momentos de trabalho e companheirismo ao longo do mestrado.

Conteúdo

1	Introdução	1
2	Equações Diferenciais Estocásticas	5
2.1	Considerações e resultados preliminares	5
2.2	Motivação e definição geral	8
2.3	Solução exata de uma equação diferencial estocástica	12
2.3.1	Equações diferenciais estocásticas lineares	12
2.3.2	Equações diferenciais estocásticas redutíveis	19
3	Métodos Numéricos para resolver Equações Diferenciais Estocásticas	25
3.1	Introdução	25
3.2	Expansão de Itô-Taylor	29
3.3	Método de Euler-Maruyama	30
3.4	Método de Milstein	35
3.5	Aplicações	36
3.5.1	Aplicação de métodos numéricos na modelação de taxas de juro	42
4	Conclusão	45
A	Algoritmos utilizados na implementação dos métodos numéricos	51
A.1	Algoritmos dos exemplos (3.5.1) e (3.5.2)	51
A.2	Algoritmos para as aproximações do Modelo de Vasicek	56

Lista de Tabelas

3.1	Erros no método de Euler-Maruyama	40
3.2	Erros no método de Milstein	40

Lista de Figuras

3.1	<i>Aproximação de Euler e solução exata com $\delta = 2^{-2}$</i>	38
3.2	<i>Aproximação de Milstein e solução exata com $\delta = 2^{-2}$</i>	38
3.3	<i>Aproximação de Euler e solução exata com $\delta = 2^{-4}$</i>	39
3.4	<i>Aproximação de Milstein e solução exata com $\delta = 2^{-4}$</i>	39
3.5	<i>Aproximação de Euler e solução exata da equação redutível</i>	41
3.6	<i>Aproximação de Milstein e solução exata da equação redutível</i>	41
3.7	<i>Aproximação de Euler-Maruyama do Modelo de Vasicek</i>	43
3.8	<i>Simulação de 100 aproximações de Euler-Maruyama do Modelo de Vasicek</i>	44

Capítulo 1

Introdução

É sobejamente conhecida a utilidade das equações diferenciais ordinárias para modelar e analisar o comportamento de diversos fenômenos que evoluem com o tempo, bem como a sua aplicação às mais variadas áreas da ciência.

No entanto, quando o fenômeno em estudo é influenciado por fatores incontroláveis de natureza aleatória, os modelos clássicos devem ser convenientemente adaptados por forma a incorporar as flutuações casuais subjacentes. As Equações Diferenciais Estocásticas, aqui designadas abreviadamente por EDE's, constituem instrumentos matemáticos eficazes para responder a esta exigência.

As EDE's são fundamentalmente equações diferenciais com um termo estocástico adicional que descreve o "ruído", ou seja, as referidas perturbações aleatórias.

Tanto quanto sabemos, a primeira EDE surgiu em 1908 e foi proposta por Langevin no âmbito dos estudos sobre o famoso movimento browniano que muito surpreendeu o botânico Brown, em 1827, nas suas observações ao microscópico.

De entre os muitos matemáticos e físicos que, durante o primeiro quartel do século XX, contribuíram para a construção da teoria das EDE's, destacam-se Bachelier, Einstein, Smoluchowsky, Ornstein, Uhlenbeck, Wiener e Lévy. Contudo, foram indubitavelmente os trabalhos de Itô e Stratonovich que, ao desenvolverem o cálculo estocástico, a impulsionaram de forma decisiva, conferindo-lhe uma base matemática mais sólida.

Itô pretendia estudar os processos de difusão através de um método probabilístico, em alternativa ao método analítico usado na altura. O método analítico passava por analisar as soluções de duas equações com derivadas parciais, as equações de Kolmogorov, relativas às funções densidade de transição do modelo de Markov que descreve tais processos. Mas as técnicas da teoria das equações com derivadas parciais eram muito restritivas na medida em que exigiam condições muito fortes para assegurar a existência de solução dessas equações.

Em contrapartida, o método probabilístico, sugerido por Lévy para construir difu-

sões e estudar as suas propriedades, conduziu Itô à criação da teoria das equações diferenciais estocásticas. Itô demonstrou que as soluções destas equações são processos de Markov e desenvolveu o cálculo estocástico para, deste modo, poder analisar as propriedades dos processos estocásticos obtidos. Efetivamente, o cálculo de Itô fornece ferramentas poderosas para estabelecer propriedades importantes dos processos de difusão, tais como a regularidade das trajetórias, estimativas dos momentos e dependência em relação a parâmetros ou condições iniciais.

Desde então, o cálculo estocástico foi alvo de um progresso extraordinário, motivado pelo facto de, em muitas aplicações, ser necessário considerar EDE's construídas à custa de processos integradores mais gerais do que o processo de Wiener, nomeadamente com trajetórias descontínuas. Neste sentido, surgiu a integração estocástica relativamente a semimartingalas, com base nos trabalhos pioneiros de Doléans (1976) e Protter (1976). De facto, este tipo de processos revela-se mais flexível quando se pretende modelar dados empíricos ou experimentais nos quais são detetados "saltos", que surgem, por exemplo, em mercados de taxas de câmbio com fortes oscilações ou em situações relacionadas com "crashes" nos mercados financeiros.

Mas o benefício da teoria das EDE's não se limita ao campo financeiro, encontrando atualmente inúmeras aplicações em áreas científicas muito diversas, que vão desde a Química, a Física e a Engenharia até à Biologia ou às Ciências Sociais. A título de exemplo, podem referir-se as seguintes aplicações.

Nas ciências biológicas, quando se pretende modelar a dinâmica de uma população, são usadas equações deste tipo, onde o "ruído" aleatório representa, em geral, a incerteza ambiental (c.f. [5]). De igual modo, as EDE's foram propostas em vários modelos estocásticos que descrevem a atividade de descarga de um neurónio (c.f. [5]).

Nas ciências sociais, têm sido utilizadas em modelos que descrevem o processo de polarização política (c.f. [6]).

No contexto da avaliação das opções, que nos é mais familiar, estipula-se, em geral, uma EDE linear, relativamente simples, para o comportamento do ativo subjacente: a equação proposta por Black e Sholes, em 1973. Para este tipo de EDE's, é possível obter a solução explícita e, deste modo, exibir a expressão das variáveis do processo estocástico que descreve a evolução temporal do preço das opções. No entanto, as séries financeiras apresentam, quase sempre, fortes efeitos não lineares (sendo a volatilidade o efeito mais importante) que apenas podem ser representados à custa de

EDE's não lineares. Ora, a resolução de equações deste tipo reveste-se de dificuldades consideráveis (c.f. [1], por exemplo), não sendo mesmo possível, na grande maioria dos casos, encontrar a forma explícita da solução.

A impossibilidade de, na grande maioria dos casos, determinar explicitamente a expressão da solução de uma EDE, tornou o estudo dos métodos numéricos um campo particularmente profícuo da análise estocástica, razão pela qual conheceu um desenvolvimento considerável nas últimas décadas.

Métodos numéricos para aproximar a solução de uma EDE podem ser obtidos generalizando os métodos numéricos deterministas tais como o de Euler e de Runge-Kutta. Contudo, tais métodos não são muito precisos. Para obter métodos numéricos apropriados deve ser feita uma análise detalhada da ordem de convergência bem como da estabilidade numérica do método e do comportamento dos erros.

O método de Euler-Maruyama para EDE's é o método numérico mais simples e obtém-se por generalização do método de Euler determinista. Este método não sendo muito preciso é um ponto de partida para métodos numéricos mais avançados.

Esta dissertação tem por objetivo apresentar uma introdução às abordagens elementares para o problema da resolução de uma EDE, quer no sentido da determinação da sua solução (forte) exata, quer da aplicação de métodos numéricos para construir aproximações dessa solução.

O trabalho está dividido em duas partes. Na primeira, elaborada com base em [8], [11], [18], [19] e [22], começa-se por introduzir as classes de processos estocásticos subjacentes à teoria das EDE's, o Lema de Itô, ferramenta matemática indispensável para o seu manuseamento, seguidos da definição de EDE e do teorema relativo à existência e unicidade da sua solução, no sentido forte; a última secção do capítulo, dedicada aos métodos de resolução propriamente ditos, contempla essencialmente a determinação da solução exata de uma EDE linear geral e da classe de equações não lineares que se podem reduzir à forma linear através de uma transformação adequada. A segunda parte do trabalho é dedicada à construção de aproximações numéricas para a solução de uma EDE, quando ela existe e é única. Tendo [19] e [13] como referência principal, mas também apoiada em [4], [20], [28], [29] e [30], são inicialmente apresentados os critérios mais usuais para avaliar o desempenho de um método numérico neste contexto, assim como a expansão de Itô-Taylor, enquanto ponto de partida da construção dos métodos de Euler-Maruyama e de Milstein, que

serão abordados nas duas secções seguintes; será dado particular destaque às propriedades de convergência forte do primeiro, que, devido à sua simplicidade, é um dos mais referidos na literatura. Este capítulo termina com a aplicação destes métodos a algumas equações particularmente populares no contexto financeiro, para as quais se conhece a forma explícita da solução exata, a fim de comparar as aproximações construídas com essa solução.

Os algoritmos dos programas que permitiram construir as trajetórias pretendidas encontram-se em apêndice; esses programas foram implementados recorrendo ao software MatLab, versão 7.6.

Capítulo 2

Equações Diferenciais Estocásticas

2.1. Considerações e resultados preliminares

Ao longo do presente texto, denotaremos por (Ω, \mathcal{A}, P) um espaço de probabilidade completo sobre o qual estão definidos todas as variáveis aleatórias e todos os processos estocásticos intervenientes. Consideraremos apenas processos estocásticos reais, representados genericamente por $(X_t, t \in \mathbf{T})$, cujo espaço dos tempos, \mathbf{T} , será \mathbb{R}_0^+ ou um intervalo limitado da forma $[t_0, T]$, com $t_0, T \in \mathbb{R}_0^+$, $T > t_0$. Como é usual, admitiremos que o conjunto dos números reais, bem como o conjunto \mathbf{T} , estão munidos das correspondentes tribos de Borel, denotadas por \mathcal{B} e $\mathcal{B}_{\mathbf{T}}$, respetivamente. A notação $(W_t, t \in \mathbb{R}_0^+)$ será usada exclusivamente para designar o processo de Wiener standard.

Designaremos ainda por $(\mathcal{F}_t)_{t \in \mathbb{R}_0^+}$ a filtração completa de (Ω, \mathcal{A}, P) gerada por este processo estocástico, isto é, para cada $t \in \mathbb{R}_0^+$, \mathcal{F}_t é a tribo gerada pela família de acontecimentos $\mathcal{N} \cup \sigma(W_s, 0 \leq s \leq t)$, com $\mathcal{N} = \{A \in \mathcal{A} : P(A) = 0\}$ ¹.

Trabalharemos essencialmente com a classe dos processos de Itô. A sua definição requer a introdução de duas outras classes de processos, que representaremos por $L^i(\Omega, [t_0, T])$, $i = 1, 2$, constituídas pelos processos estocásticos $(X_t, t \in [t_0, T])$ satisfazendo as seguintes condições:

- (a) $(X_t, t \in [t_0, T])$ é adaptado a $(\mathcal{F}_t)_{t \in [t_0, T]}$;
- (b) $(X_t, t \in [t_0, T])$ é mensurável, ou seja, a aplicação

$$\begin{aligned} [t_0, T] \times \Omega &\rightarrow \mathbb{R} \\ (t, \omega) &\rightsquigarrow X_t(\omega) \end{aligned}$$

é mensurável relativamente à tribo produto $\mathcal{B}_{[t_0, T]} \otimes \mathcal{A}$;

- (c) $\int_{t_0}^T |X_t|^i dt < +\infty$ *q.c.*, $i = 1, 2$.

¹Note-se que a filtração natural do processo de Wiener, assim construída, constitui apenas um caso particular de filtrações mais gerais consideradas habitualmente na literatura.

Definição 2.1.1. O processo estocástico $(X_t, t \in [t_0, T])$ diz-se um **processo de Itô** se as suas trajectórias forem q.c. contínuas e, para cada $t \in [t_0, T]$, X_t puder ser escrita na forma

$$X_t = X_{t_0} + \int_{t_0}^t B_s ds + \int_{t_0}^t C_s dW_s \quad q.c., \quad (2.1)$$

com X_{t_0} uma variável aleatória real mensurável relativamente a \mathcal{F}_{t_0} e $(B_t, t \in [t_0, T])$ e $(C_t, t \in [t_0, T])$ processos estocásticos pertencentes a $L^1(\Omega, [t_0, T])$ e $L^2(\Omega, [t_0, T])$, respetivamente (c.f. [22]).²

Observação 2.1.1. Na igualdade anterior, o primeiro integral é um integral de Lebesgue, enquanto que o segundo é um integral de Itô. A fim de simplificar a notação, é habitual representar simbolicamente (2.1) na forma diferencial

$$dX_t = B_t dt + C_t dW_t, \quad t \in [t_0, T]$$

e dizer que dX_t é o diferencial estocástico do processo $(X_t, t \in [t_0, T])$.

Assim, no que se segue, todas as expressões envolvendo diferenciais estocásticos deverão ser forçosamente interpretadas neste sentido.

Apresentamos agora uma versão do lema de Itô cuja demonstração podemos encontrar em [19], por exemplo, que será útil nas próximas secções. Como é sabido, este lema constitui uma ferramenta matemática crucial do cálculo estocástico, sendo, em particular, imprescindível na manipulação e na resolução de equações diferenciais estocásticas.

Lema de Itô: Sejam $(X_t^{(k)}, t \in [t_0, T])$, $k = 1, \dots, n$, n processos de Itô com diferenciais estocásticos

$$dX_t^{(k)} = B_t^{(k)} dt + C_t^{(k)} dW_t, \quad t \in [t_0, T], \quad k = 1, \dots, n$$

²Na verdade, a definição de processo de Itô requer que os processos estocásticos $(B_t, t \in [t_0, T])$ e $(C_t, t \in [t_0, T])$ sejam progressivamente mensuráveis em relação a $(\mathcal{F}_t)_{t \in [t_0, T]}$. Apesar desta condição ser mais forte do que as hipóteses ((a) e (b)) formuladas, é possível mostrar (c.f. [17]) que, sob (a) e (b), existe uma versão de $(X_t, t \in [t_0, T])$ que é progressivamente mensurável em relação a $(\mathcal{F}_t)_{t \in [t_0, T]}$, o que nos leva a considerar, implicitamente, uma tal versão.

Diz-se que o processo estocástico $(X_t, t \in [t_0, T])$ é progressivamente mensurável em relação a $(\mathcal{F}_t)_{t \in [t_0, T]}$ se, para todo o $t \in [t_0, T]$, a aplicação

$$\begin{aligned} [t_0, t] \times \Omega &\rightarrow \mathbb{R} \\ (s, \omega) &\rightsquigarrow X_s(\omega) \end{aligned}$$

é $\mathcal{B}_{[t_0, t]} \otimes \mathcal{F}_t$ mensurável.

e $h(t, x_1, \dots, x_n)$, $(t, x_1, \dots, x_n) \in [t_0, T] \times \mathbb{R}^n$, uma função real, com derivadas parciais

$$h_t = \frac{\partial h}{\partial t}, \quad h_i = \frac{\partial h}{\partial x_i}, \quad h_{ij} = \frac{\partial^2 h}{\partial x_i \partial x_j}, \quad i, j = 1, \dots, n$$

contínuas.

Então, o processo estocástico $(Z_t, t \in [t_0, T])$ definido por $Z_t = h(t, X_t^{(1)}, \dots, X_t^{(n)})$, $t \in [t_0, T]$, é um processo de Itô e verifica

$$Z_t = Z_{t_0} + \int_{t_0}^t \tilde{B}_s ds + \int_{t_0}^t \tilde{C}_s dW_s \quad q.c., \quad t \in [t_0, T], \quad (2.2)$$

com

$$\tilde{C}_t = \sum_{i=1}^n h_i(t, X_t^{(1)}, \dots, X_t^{(n)}) C_t^{(i)}$$

e

$$\begin{aligned} \tilde{B}_t &= h_t(t, X_t^{(1)}, \dots, X_t^{(n)}) + \sum_{i=1}^n h_i(t, X_t^{(1)}, \dots, X_t^{(n)}) B_t^{(i)} + \\ &+ \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n h_{ij}(t, X_t^{(1)}, \dots, X_t^{(n)}) C_t^{(i)} C_t^{(j)}, \quad t \in [t_0, T]. \end{aligned}$$

Observação 2.1.2.

(i) Em alternativa à notação diferencial equivalente a (2.2), escrevemos ainda

$$\begin{aligned} dZ_t &= [h_t(t, X_t^{(1)}, \dots, X_t^{(n)}) + \\ &+ \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n h_{ij}(t, X_t^{(1)}, \dots, X_t^{(n)}) C_t^{(i)} C_t^{(j)}] dt + \\ &+ \sum_{i=1}^n h_i(t, X_t^{(1)}, \dots, X_t^{(n)}) dX_t^{(i)}, \quad t \in [t_0, T], \end{aligned}$$

agrupando os termos $\sum_{i=1}^n h_i(t, X_t^{(1)}, \dots, X_t^{(n)}) B_t^{(i)} dt$ e $\sum_{i=1}^n h_i(t, X_t^{(1)}, \dots, X_t^{(n)}) C_t^{(i)} dW_t$

e substituindo, formalmente, $B_t^{(i)} dt + C_t^{(i)} dW_t$ por $dX_t^{(i)}$, para $i = 1, \dots, n$.

(ii) Tomando $n = 1$, obtemos apenas

$$\begin{aligned} dh(t, X_t) &= \left(h_t(t, X_t) + h_x(t, X_t) B_t + \frac{1}{2} h_{xx}(t, X_t) C_t^2 \right) dt + \\ &+ h_x(t, X_t) C_t dW_t, \quad t \in [t_0, T], \end{aligned} \quad (2.3)$$

onde $h_t = \frac{\partial h}{\partial t}$, $h_x = \frac{\partial h}{\partial x}$ e $h_{xx} = \frac{\partial^2 h}{\partial x^2}$.³

Ou ainda, mais abreviadamente,

$$dh(t, X_t) = \left(h_t(t, X_t) + \frac{1}{2} h_{xx}(t, X_t) C_t^2 \right) dt + h_x(t, X_t) dX_t, \quad t \in [t_0, T].$$

³No caso $n = 1$ omitimos o índice superior.

Se, em particular, $B_t = 0$ e $C_t = 1$, $t \in [t_0, T]$, então $(X_t, t \in [t_0, T])$ é o próprio processo de Wiener, pelo que, neste caso, tem-se

$$dh(t, W_t) = \left(h_t(t, W_t) + \frac{1}{2} h_{xx}(t, W_t) \right) dt + h_x(t, W_t) dW_t, \quad t \in [t_0, T].$$

(iii) Considerando $n = 2$ e

$$h(t, x_1, x_2) = x_1 x_2, \quad (t, x_1, x_2) \in [t_0, T] \times \mathbb{R}^2,$$

somos conduzidos à denominada “fórmula” da integração por partes:

$$d(X_t^{(1)} X_t^{(2)}) = X_t^{(2)} dX_t^{(1)} + X_t^{(1)} dX_t^{(2)} + C_t^{(1)} C_t^{(2)} dt. \quad (2.4)$$

Por fim, chamamos à atenção de que, no texto, representaremos as funções deterministas por letras minúsculas, sendo as letras maiúsculas reservadas para designar os processos estocásticos, de acordo com a notação introduzida anteriormente.

2.2. Motivação e definição geral

Suponhamos que pretendemos modelar a evolução do valor de um ativo financeiro ao longo do tempo, a partir de determinado instante inicial, $t_0 \in \mathbb{R}_0^+$. Uma abordagem usual para construir um tal modelo consiste em analisar o comportamento da variação relativa do valor do ativo num intervalo de tempo de amplitude suficientemente pequena, $\Delta t > 0$, ou seja

$$\frac{x(t + \Delta t) - x(t)}{x(t)},$$

onde $x(t)$ denota o preço do ativo no instante t .

Se o ativo se comportasse, durante o intervalo de tempo em questão, como um ativo financeiro sem risco, então poderíamos considerar

$$\frac{x(t + \Delta t) - x(t)}{x(t)} = \mu \Delta t, \quad (2.5)$$

onde μ é uma constante positiva que designa a taxa média de crescimento do ativo (por exemplo, $\mu = r$, com r uma taxa de juro com capitalização contínua).

Então, escrevendo

$$\frac{x(t + \Delta t) - x(t)}{\Delta t} = \mu x(t)$$

e tomando o limite quando Δt tende para zero, obtemos (admitindo que x é uma função diferenciável) uma equação diferencial ordinária (EDO) linear e de primeira ordem

$$\frac{dx}{dt}(t) = \mu x(t), \quad t \in [t_0, T],$$

cuja solução, sob a condição inicial $x(t_0) = x_0$, $x_0 \in \mathbb{R}^+$, é trivialmente dada por

$$x(t) = x_0 \exp [\mu(t - t_0)], \quad t \in [t_0, T].$$

Contudo, a relação (2.5) não traduz adequadamente a realidade, já que o retorno $[x(t + \Delta t) - x(t)] / x(t)$ está sujeito a múltiplos fatores externos e casuais cujo efeito não é desprezável, devendo pois ser incorporado no modelo.

A introdução desse efeito passa, em geral, por adicionar ao segundo membro de (2.5) um termo proporcional ao acréscimo de um processo estocástico $(Y_t, t \in [t_0, T])$, no intervalo considerado:

$$\frac{X_{t+\Delta t} - X_t}{X_t} = \mu\Delta t + \sigma(Y_{t+\Delta t} - Y_t), \quad (2.6)$$

onde σ é um número real positivo. O termo $\sigma(Y_{t+\Delta t} - Y_t)$ representa, assim, um impulso aleatório que influencia a dinâmica do preço do ativo.

Conseqüentemente, se as trajetórias dos processos estocásticos envolvidos forem diferenciáveis, um raciocínio idêntico ao anterior conduz-nos novamente a uma equação diferencial de primeira ordem

$$\frac{dX_t(\omega)}{dt} = \mu X_t(\omega) + \sigma X_t(\omega) \frac{dY_t(\omega)}{dt}, \quad t \in [t_0, T], \omega \in \Omega. \quad (2.7)$$

As equações deste tipo são designadas por **Equações Diferenciais Aleatórias** e resolvem-se, para cada trajetória, como equações diferenciais deterministas, não necessitando, portanto, do recurso aos métodos do cálculo estocástico. Assim, sob condições de regularidade análogas às do caso determinista, a solução da equação (2.7), para uma dada condição inicial $X_0(\omega)$ em t_0 , é dada por

$$X_t(\omega) = X_0(\omega) \exp [\mu(t - t_0) + \sigma (Y_t(\omega) - Y_{t_0}(\omega))], \quad t \in [t_0, T].$$

Ora, a escolha do processo estocástico $(Y_t, t \in [t_0, T])$ deve procurar refletir o que se observa empiricamente no comportamento temporal do preço de um ativo. Nesse sentido, a literatura financeira propõe que a fonte de aleatoriedade seja representada pelo processo de Wiener standard, o que nos conduz à equação

$$dX_t = \mu X_t dt + \sigma X_t dW_t, \quad t \in [t_0, T], \quad X_{t_0} = X_0 \text{ q.c..}$$

Surge, deste modo, o conceito de Equação Diferencial Estocástica, cuja resolução requer técnicas específicas uma vez que a não diferenciabilidade (q.c.) das trajetórias do processo de Wiener impede-nos de utilizar a abordagem clássica.

Mais geralmente, dada uma variável aleatória real X_0 , mensurável em relação a \mathcal{F}_{t_0} , e duas funções reais, f e g , definidas sobre $[t_0, T] \times \mathbb{R}$, mensuráveis em relação a $\mathcal{B}_{[t_0, T]} \otimes \mathcal{B}$, uma **Equação Diferencial Estocástica (EDE)**, com condição inicial X_0 , é uma expressão da forma

$$dX_t = f(t, X_t) dt + g(t, X_t) dW_t, \quad t \in [t_0, T], \quad (2.8)$$

onde $(X_t, t \in [t_0, T])$ é um processo estocástico tal que $X_{t_0} = X_0$, q.c.. As funções f e g denominam-se os coeficientes ⁴ da equação.

Claro que a expressão (2.8) é meramente simbólica, pelo que devemos começar por explicitar as condições que permitem interpretá-la como uma equação integral, atribuindo-lhe, deste modo, um significado matemático. Vejamos, então, o que se entende por uma solução da EDE (2.8), (c.f. [11]).

Definição 2.2.1. *O processo estocástico $(X_t, t \in [t_0, T])$ diz-se uma **solução forte** da EDE (2.8) se:*

- $(X_t, t \in [t_0, T])$ tem trajetórias contínuas q.c.;
- $(X_t, t \in [t_0, T])$ é um processo adaptado a $(\mathcal{F}_t)_{t \in [t_0, T]}$;
- $\int_{t_0}^T |f(s, X_s)| ds < +\infty$ q.c.;
- $\int_{t_0}^T g^2(s, X_s) ds < +\infty$ q.c.;
- para cada $t \in [t_0, T]$, X_t verifica a equação integral

$$X_t = X_0 + \int_{t_0}^t f(s, X_s) ds + \int_{t_0}^t g(s, X_s) dW_s \quad q.c..$$

Note-se que as soluções fortes não são as únicas soluções de uma EDE. Com efeito, definem-se dois tipos de soluções: as soluções fortes e as soluções fracas. Fixados os coeficientes f e g e a variável aleatória real X_0 , a diferença entre os dois conceitos reside essencialmente no facto de, na noção de solução forte, o processo de Wiener ser especificado à partida enquanto que, na noção de solução fraca, ele ser também uma incógnita da equação.

Determinar uma solução fraca de uma EDE pressupõe, assim, a construção de dois processos estocásticos: um processo de Wiener standard, $(\widetilde{W}_t, t \in [t_0, T])$, e um

⁴De deriva e de difusão, respetivamente.

processo $(\tilde{X}_t, t \in [t_0, T])$ que verifica a equação (2.8) quando o processo de Wiener subjacente à equação é $(\tilde{W}_t, t \in [t_0, T])$. Este tipo de soluções não será aqui abordado, uma vez que nos dedicaremos apenas às soluções fortes.

Naturalmente, a primeira questão que se coloca é a da existência e da unicidade da solução forte de uma EDE. Tal como nas EDO's, a resposta a esta questão, patente na Proposição 2.2.1, passa por efetuar, por um lado, uma restrição sobre o crescimento dos coeficientes (que garante a existência de pelo menos uma solução) e, por outro lado, exigir o seu caráter lipschitziano (que assegura a unicidade da solução).

Proposição 2.2.1. *Suponhamos que a variável aleatória real X_0 é mensurável em relação a \mathcal{F}_{t_0} e admite momento simples de segunda ordem. Suponhamos ainda que os coeficientes f e g verificam as seguintes condições:*

(i)

$$\exists K > 0 : \forall t \in [t_0, T], \forall x \in \mathbb{R}, |f(t, x)|^2 + |g(t, x)|^2 \leq K(1 + |x|^2); \quad (2.9)$$

(ii) $\exists L > 0 : \forall t \in [t_0, T], \forall x \in \mathbb{R},$

$$|f(t, x) - f(t, y)| + |g(t, x) - g(t, y)| \leq L|x - y|. \quad (2.10)$$

Então, a equação (2.8) admite uma única solução forte, $(X_t, t \in [t_0, T])$. Além disso, esta solução é tal que

$$\sup_{t_0 \leq t \leq T} E(|X_t|^2) < +\infty.$$

Demonstração: Veja-se, por exemplo, [19], pág. 131. ■

Observação 2.2.1.

(i) *A unicidade da solução significa que se $(X_t, t \in [t_0, T])$ e $(X'_t, t \in [t_0, T])$ forem duas soluções fortes da equação (2.8), então estes processos são indistinguíveis, isto é, existe $\Omega' \in \mathcal{A}$, com $P(\Omega') = 1$, tal que*

$$\forall \omega \in \Omega', \forall t \in [t_0, T], X_t(\omega) = X'_t(\omega).$$

(ii) *Tal como é referido em [19], existem versões mais gerais desta Proposição, em que, por um lado, se prescinde da condição sobre a existência de $E(|X_0|^2)$ e, por outro lado, se substitui a condição de Lipschitz por hipóteses mais fracas. Para as possíveis generalizações, pode consultar-se, por exemplo, [18].*

No que se segue, admitiremos que X_0 verifica as hipóteses da proposição anterior. Iremos apresentar agora duas propriedades do processo solução que serão usadas no próximo capítulo.

Proposição 2.2.2. *Suponhamos que as funções f e g verificam as condições (2.9) e (2.10) e que*

$$E(|X_0|^{2n}) < +\infty,$$

para algum $n \in \mathbb{N}$, $n \geq 2$.

Então, existem constantes positivas C e D , dependentes apenas de n , K , L , T e t_0 , tais que

(i)

$$\sup_{t_0 \leq t \leq T} E(|X_t|^{2n}) \leq C (1 + E(|X_0|^{2n})); \quad (2.11)$$

(ii)

$$\forall s, t \in [t_0, T], \quad s \leq t, \quad E(|X_t - X_s|^{2n}) \leq D (1 + E(|X_0|^{2n})) (t - s)^n. \quad (2.12)$$

Demonstração: A propriedade (i) decorre diretamente da primeira desigualdade estabelecida em [19] no Teorema 4.5.4 (pág. 136). Decalcando o raciocínio efetuado pelos autores na segunda parte do mesmo Teorema, obtemos (ii). ■

2.3. Solução exata de uma equação diferencial estocástica

2.3.1. Equações diferenciais estocásticas lineares

À semelhança do que se verifica nas EDO's, é impossível, na grande maioria dos casos, determinar explicitamente a solução de uma EDE. Uma das exceções a esta impossibilidade diz respeito às EDE's lineares, cuja resolução é conseguida sem grandes dificuldades, tal como teremos oportunidade de constatar.

Definição 2.3.1. *A EDE (2.8) diz-se **linear** se as funções f e g são lineares relativamente ao segundo argumento, isto é, se são da forma*

$$f(t, x) = c_1(t)x + c_2(t), \quad g(t, x) = \sigma_1(t)x + \sigma_2(t), \quad (t, x) \in [t_0, T] \times \mathbb{R},$$

com c_i e σ_i , $i = 1, 2$, funções reais deterministas.

Definição 2.3.2.

- (i) Se as funções c_2 e σ_2 forem identicamente nulas, a equação diz-se uma **EDE linear homogénea**;
- (ii) Se a função σ_1 for identicamente nula, a equação diz-se uma **EDE linear com ruído aditivo**; caso contrário, diz-se uma **EDE linear com ruído multiplicativo**.

A próxima proposição mostra que a continuidade das funções c_i e σ_i , $i = 1, 2$ é suficiente para garantir a existência e unicidade da solução de uma EDE linear.

Proposição 2.3.1. *Se as funções c_i e σ_i , $i = 1, 2$, forem contínuas em $[t_0, T]$, então uma EDE linear admite uma única solução forte.*

Demonstração: De acordo com a Proposição (2.2.1), basta verificar que, sob esta hipótese, as funções $f(t, x) = c_1(t)x + c_2(t)$ e $g(t, x) = \sigma_1(t)x + \sigma_2(t)$, $(t, x) \in [t_0, T] \times \mathbb{R}$, satisfazem (2.9) e (2.10).

Ora, sendo c_i e σ_i , $i = 1, 2$ contínuas em $[t_0, T]$, então são limitadas neste intervalo, pelo que a segunda desigualdade é trivialmente verificada. Relativamente à primeira, basta invocar este facto e a desigualdade $(a + b)^2 \leq 2(a^2 + b^2)$.

■

Antes de nos debruçarmos sobre a resolução de uma EDE linear na sua forma mais geral, analisaremos dois casos particulares que ilustram algumas das técnicas habitualmente utilizadas.

O primeiro caso diz respeito a uma equação com ruído aditivo.

Consideremos, então, a equação

$$dX_t = (c_1(t)X_t + c_2(t)) dt + \sigma_2(t) dW_t, \quad t \in [t_0, T], \quad (2.13)$$

com σ_2 não identicamente nula, sob a condição inicial $X_{t_0} = X_0$ q.c..

Começemos por determinar a solução da EDO que se obtém substituindo X_t e W_t por funções reais diferenciáveis, $x(t)$ e $w(t)$, respetivamente:

$$x'(t) = c_1(t)x(t) + c_2(t) + \sigma_2(t)w'(t), \quad t \in [t_0, T], \quad x(t_0) = x_0 \in \mathbb{R}.$$

Supondo, para simplificar, que w' é uma função contínua, a solução pretendida é, como sabemos, dada por

$$x(t) = \mu(t) \left(x_0 + \int_{t_0}^t \frac{c_2(s)}{\mu(s)} ds + \int_{t_0}^t \frac{\sigma_2(s)}{\mu(s)} w'(s) ds \right),$$

com $\mu(t) = \exp \left(\int_{t_0}^t c_1(s) ds \right)$ o correspondente fator integrante.

Escrevendo o segundo integral como um integral de Riemann-Stieltjes, vem

$$x(t) = \mu(t) \left(x_0 + \int_{t_0}^t \frac{c_2(s)}{\mu(s)} ds + \int_{t_0}^t \frac{\sigma_2(s)}{\mu(s)} dw(s) \right), \quad t \in [t_0, T].$$

A expressão anterior sugere-nos a forma da solução da EDE considerada, conforme mostra o resultado seguinte.

Proposição 2.3.2. *Suponhamos que as funções c_i , $i = 1, 2$, e σ_2 são contínuas em $[t_0, T]$. A solução forte da equação (2.13) é dada por*

$$X_t = \mu(t) \left(X_0 + \int_{t_0}^t \frac{c_2(s)}{\mu(s)} ds + \int_{t_0}^t \frac{\sigma_2(s)}{\mu(s)} dW_s \right) \quad q.c., \quad t \in [t_0, T],$$

com $\mu(t) = \exp \left(\int_{t_0}^t c_1(s) ds \right)$.

Demonstração: Uma vez que, sob as condições da hipótese, estão asseguradas a existência dos integrais envolvidos na expressão de X_t , bem como a unicidade da solução forte da equação, a demonstração reduz-se apenas a constatar que o processo estocástico $(X_t, t \in [t_0, T])$, assim definido, verifica a equação.

Para tal, determinemos o seu diferencial estocástico, recorrendo ao lema de Itô.

De facto, podemos escrever $X_t = F(t, Y_t)$ com

$$F(t, x) = \mu(t)x, \quad (t, x) \in [t_0, T] \times \mathbb{R},$$

e

$$Y_t = X_0 + \int_{t_0}^t \frac{c_2(s)}{\mu(s)} ds + \int_{t_0}^t \frac{\sigma_2(s)}{\mu(s)} dW_s \quad q.c., \quad t \in [t_0, T],$$

um processo de Itô.

Notando que, para todo o $(t, x) \in [t_0, T] \times \mathbb{R}$, se tem

$$F_t(t, x) = \mu'(t)x = c_1(t)\mu(t)x, \quad F_x(t, x) = \mu(t) \quad e \quad F_{xx}(t, x) = 0,$$

obtemos, por aplicação de (2.3),

$$\begin{aligned} F(t, Y_t) - F(t_0, Y_{t_0}) &= \int_{t_0}^t \left(c_1(s)\mu(s)Y_s + \mu(s)\frac{c_2(s)}{\mu(s)} \right) ds \\ &+ \int_{t_0}^t \mu(s)\frac{\sigma_2(s)}{\mu(s)} dW_s \quad q.c., \quad t \in [t_0, T], \end{aligned}$$

isto é,

$$X_t - X_0 = \int_{t_0}^t (c_1(s)X_s + c_2(s)) ds + \int_{t_0}^t \sigma_2(s) dW_s \quad \text{q.c.}, \quad t \in [t_0, T],$$

o que mostra que o processo estocástico $(X_t, t \in [t_0, T])$ satisfaz a equação (2.13). ■

Ilustramos a aplicação desta proposição com os seguintes exemplos.

Exemplo 2.3.1. *A solução forte da equação de Langevin*

$$dX_t = -cX_t dt + \sigma dW_t, \quad t \in [0, T], \quad (2.14)$$

com c e σ constantes reais, $\sigma \neq 0$, é o processo estocástico definido por

$$X_t = e^{-ct} \left(X_0 + \sigma \int_0^t e^{cs} dW_s \right) \quad \text{q.c.}, \quad t \in [0, T],$$

usualmente designado, na literatura, por processo de Ornstein-Uhlenbeck. △

Exemplo 2.3.2. *Consideremos a equação*

$$dX_t = \left(\frac{2}{1+t} X_t + b(1+t)^2 \right) dt + c(1+t)^2 dW_t, \quad t \in [0, T], \quad X_0 = 1 \quad \text{q.c.},$$

onde b e c são constantes reais, com $c \neq 0$.

Temos, evidentemente, $\mu(t) = (1+t)^2$, pelo que

$$X_t = (1+t)^2 \left(1 + \int_0^t b ds + \int_0^t c dW_s \right) = (1+t)^2 (1 + bt + cW_t) \quad \text{q.c.}, \quad t \in [0, T].$$
△

Apresentamos, agora, um método que permite resolver a EDE linear homogénea

$$dX_t = c(t)X_t dt + \sigma(t)X_t dW_t, \quad t \in [t_0, T], \quad (2.15)$$

com σ não identicamente nula e $X_{t_0} = X_0$ q.c..

Analisemos, em primeiro lugar, o caso em que $X_0 = 1$, q.c..

Adotando um procedimento análogo ao anterior, começamos por determinar a solução da equação determinista

$$x'(t) = a(t)x(t) + b(t)x(t)w'(t), \quad t \in [t_0, T], \quad x(0) = 1,$$

onde a , b e w são funções reais, com a e b contínuas e w continuamente diferenciável.

Uma vez que a solução procurada é dada por

$$x(t) = \exp \left(\int_{t_0}^t (a(s) + b(s)w'(s)) ds \right) = \exp \left(\int_{t_0}^t a(s) ds + \int_{t_0}^t b(s) dw(s) \right), \quad t \in [t_0, T],$$

consideramos então, para (2.15), uma solução da forma

$$X_t = \exp \left(\int_{t_0}^t a(s) ds + \int_{t_0}^t b(s) dW_s \right) \quad q.c., \quad t \in [t_0, T],$$

com a e b funções reais adequadas, de modo a que $(X_t, t \in [t_0, T])$ satisfaça a referida equação.

Para tal, escrevemos $X_t = F(Y_t)$, $t \in [t_0, T]$, com $F(x) = e^x$, $x \in \mathbb{R}$ e

$$Y_t = \int_{t_0}^t a(s) ds + \int_{t_0}^t b(s) dW_s \quad q.c., \quad t \in [t_0, T],$$

vindo, por (2.3),

$$dX_t = \left(a(t) + \frac{1}{2}b^2(t) \right) X_t dt + b(t)X_t dW_t, \quad t \in [t_0, T]. \quad (2.16)$$

Nestas condições, as funções em causa podem ser encontradas identificando os coeficientes de (2.15) e (2.16):

$$a(t) + \frac{1}{2} b^2(t) = c(t) \quad e \quad b(t) = \sigma(t),$$

o que conduz a

$$X_t = \exp \left(\int_{t_0}^t (c(s) - \frac{1}{2} \sigma^2(s)) ds + \int_{t_0}^t \sigma(s) dW_s \right) \quad q.c., \quad t \in [t_0, T].$$

Desta expressão, decorre o resultado que se segue:

Proposição 2.3.3. *Suponhamos que as funções c e σ são contínuas em $[t_0, T]$. A solução da EDE (2.15), com a condição inicial $X_{t_0} = X_0$ q.c., é dada por*

$$X_t = X_0 \exp \left(\int_{t_0}^t \left(c(s) - \frac{1}{2} \sigma^2(s) \right) ds + \int_{t_0}^t \sigma(s) dW_s \right) \quad q.c., \quad t \in [t_0, T].$$

Demonstração: Claro que no caso de $X_0 = 0$ q.c., a solução da equação é, trivialmente, $X_t = 0$ q.c., $t \in [t_0, T]$.

Admitamos, então, que $X_0 \neq 0$ q.c.. Invocando o raciocínio efetuado anteriormente, podemos afirmar que o processo estocástico $(\tilde{X}_t, t \in [t_0, T])$, com $\tilde{X}_t = \frac{X_t}{X_0}$, está bem definido e constitui a única solução forte da equação

$$d\tilde{X}_t = c(t)\tilde{X}_t dt + \sigma(t)\tilde{X}_t dW_t, \quad t \in [t_0, T], \quad \tilde{X}_{t_0} = 1 \quad q.c.,$$

verificando, portanto, a igualdade

$$\tilde{X}_t - 1 = \int_{t_0}^t c(s)\tilde{X}_s ds + \int_{t_0}^t \sigma(s)\tilde{X}_s dW_s \quad q.c., \quad t \in [t_0, T]. \quad (2.17)$$

Ora, sendo X_0 mensurável relativamente a \mathcal{F}_{t_0} , tem-se

$$X_0 \int_{t_0}^t \sigma(s)\tilde{X}_s dW_s = \int_{t_0}^t \sigma(s)X_s dW_s.$$

Consequentemente, multiplicando ambos os membros de (2.17) por X_0 , concluímos que

$$X_t - X_0 = \int_{t_0}^t c(s)X_s ds + \int_{t_0}^t \sigma(s)X_s dW_s \quad q.c., \quad t \in [t_0, T],$$

o que conduz ao resultado anunciado, uma vez que as condições da hipótese asseguram a unicidade da solução forte da equação dada.

■

Vamos exemplificar a aplicação desta proposição com a equação de Black-Scholes.

Exemplo 2.3.3. Consideremos a equação

$$dX_t = \mu X_t dt + \sigma X_t dW_t, \quad t \in [0, T], \quad (2.18)$$

onde μ e σ são números reais, com $\sigma > 0$.

Fixada a variável aleatória real X_0 , q.c. limitada, a solução forte desta equação é dada por

$$X_t = X_0 \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W_t\right) \quad q.c., \quad t \in [t_0, T],$$

processo que é conhecido por movimento browniano geométrico.

△

Dedicamo-nos, por fim, à resolução de uma EDE linear geral

$$dX_t = (c_1(t)X_t + c_2(t)) dt + (\sigma_1(t)X_t + \sigma_2(t)) dW_t, \quad t \in [t_0, T], \quad (2.19)$$

com a condição inicial $X_{t_0} = X_0$ q.c..

Procuremos uma solução da forma

$$X_t = U_t V_t,$$

com $(U_t, t \in [t_0, T])$ e $(V_t, t \in [t_0, T])$ processos de Itô verificando

$$dU_t = c_1(t)U_t dt + \sigma_1(t)U_t dW_t, \quad t \in [t_0, T], \quad U_{t_0} = 1 \quad q.c. \quad (2.20)$$

e

$$dV_t = A_t dt + B_t dW_t, \quad t \in [t_0, T], \quad V_{t_0} = X_0 \text{ q.c.}, \quad (2.21)$$

onde $(A_t, t \in [t_0, T])$ e $(B_t, t \in [t_0, T])$ são processos estocásticos convenientemente escolhidos.

Tal como vimos na Proposição 2.3.3, a solução da equação (2.20) é dada por

$$U_t = \exp \left(\int_{t_0}^t \left(c_1(s) - \frac{1}{2} \sigma_1^2(s) \right) ds + \int_{t_0}^t \sigma_1(s) dW_s \right) \text{ q.c.}, \quad t \in [t_0, T]. \quad (2.22)$$

Por outro lado, determinando o diferencial do produto $U_t V_t$, vem, de acordo com (2.4),

$$\begin{aligned} d(U_t V_t) &= U_t dV_t + V_t dU_t + \sigma_1(t) B_t U_t dt \\ &= U_t (A_t dt + B_t dW_t) + V_t (c_1(t) U_t dt + \sigma_1(t) U_t dW_t) + \sigma_1(t) B_t U_t dt \\ &= (U_t A_t + c_1(t) X_t + \sigma_1(t) B_t U_t) dt + (U_t B_t + \sigma_1(t) X_t) dW_t. \end{aligned}$$

Se compararmos esta expressão com a do diferencial estocástico de X_t , dada por (2.19), e igualarmos os coeficientes, obtemos

$$A_t = \frac{c_2(t) - \sigma_1(t) \sigma_2(t)}{U_t} \quad \text{e} \quad B_t = \frac{\sigma_2(t)}{U_t}.$$

Consequentemente, o processo estocástico $(V_t, t \in [t_0, T])$ é definido por

$$V_t = X_0 + \int_{t_0}^t \frac{c_2(s) - \sigma_1(s) \sigma_2(s)}{U_s} ds + \int_{t_0}^t \frac{\sigma_2(s)}{U_s} dW_s \text{ q.c.}, \quad t \in [t_0, T], \quad (2.23)$$

pelo que

$$X_t = U_t V_t = U_t \left(X_0 + \int_{t_0}^t \frac{c_2(s) - \sigma_1(s) \sigma_2(s)}{U_s} ds + \int_{t_0}^t \frac{\sigma_2(s)}{U_s} dW_s \right) \text{ q.c.}, \quad t \in [t_0, T].$$

Proposição 2.3.4. *Suponhamos que as funções c_i e $\sigma_i, i = 1, 2$ são contínuas em $[t_0, T]$. A solução da equação (2.19) é dada por*

$$X_t = U_t \left(X_0 + \int_{t_0}^t \frac{c_2(s) - \sigma_1(s) \sigma_2(s)}{U_s} ds + \int_{t_0}^t \frac{\sigma_2(s)}{U_s} dW_s \right) \text{ q.c.}, \quad t \in [t_0, T],$$

onde $(U_t, t \in [t_0, T])$ é a solução da equação homogénea que lhe está associada:

$$dU_t = c_1(t) U_t dt + \sigma_1(t) U_t dW_t, \quad t \in [t_0, T],$$

com a condição inicial $U_{t_0} = 1$ q.c..

Demonstração: Começamos por notar que, sendo $X_t = U_t V_t, t \in [t_0, T]$, com $(U_t, t \in [t_0, T])$ e $(V_t, t \in [t_0, T])$ definidos por (2.22) e (2.23), respetivamente, o

processo estocástico $(X_t, t \in [t_0, T])$ está bem definido, pois o mesmo acontece com $(U_t, t \in [t_0, T])$ (seguindo o que vimos na Proposição 2.3.3) e com $(V_t, t \in [t_0, T])$. Com efeito, neste último caso, a continuidade das funções c_2 e σ_i , $i = 1, 2$, bem como o facto das trajetórias de $(U_t, t \in [t_0, T])$ serem q.c. contínuas, garantem que

$$\int_{t_0}^t \left(\frac{\sigma_2(s)}{U_s} \right)^2 ds < +\infty \quad \text{q.c.} \quad \text{e} \quad \int_{t_0}^t \left| \frac{c_2(s) - \sigma_1(s)\sigma_2(s)}{U_s} \right| ds < +\infty \quad \text{q.c.}$$

Além disso, ambos os processos integrandos são adaptados a $(\mathcal{F}_t)_{t \in [t_0, T]}$.

Por outro lado, a demonstração reduz-se, uma vez mais, a mostrar que $(X_t, t \in [t_0, T])$ é solução da equação. Para tal, basta decalcar o raciocínio efetuado na página anterior, calculando o seu diferencial estocástico através de (2.4) e substituir, na expressão obtida, os diferenciais de $(U_t, t \in [t_0, T])$ e $(V_t, t \in [t_0, T])$.

■

Observação 2.3.1. Analogamente ao que se passa nas EDO's lineares de primeira ordem, constatamos que o processo solução de (2.19) admite a representação

$$X_t = U_t X_0 + Z_t \quad \text{q.c.}, \quad t \in [t_0, T],$$

onde $(U_t X_0, t \in [t_0, T])$ é a solução geral da correspondente equação homogénea, sob a mesma condição inicial, e $(Z_t, t \in [t_0, T])$ é a solução da equação não homogénea

$$dZ_t = (c_2(t) - \sigma_1(t)\sigma_2(t)) dt + \sigma_2(t) dW_t, \quad t \in [t_0, T],$$

com a condição inicial $Z_{t_0} = 0$ q.c..

2.3.2. Equações diferenciais estocásticas redutíveis

No que diz respeito às EDE's não lineares a situação é substancialmente mais difícil, pelo que nos limitaremos, à semelhança do que encontramos na literatura (c.f. [8] e [19]), a apresentar uma técnica que permite reduzir certo tipo de equações à forma linear, usando uma mudança de variável apropriada.

Mais concretamente, dada a equação

$$dX_t = f(t, X_t)dt + g(t, X_t)dW_t, \quad t \in [t_0, T], \quad X_{t_0} = X_0 \quad \text{q.c.}, \quad (2.24)$$

a ideia de base consiste em efetuar uma transformação da forma $Y_t = h(t, X_t)$, $t \in [t_0, T]$, com h uma função suficientemente regular, construída de tal modo que $(Y_t, t \in [t_0, T])$ satisfaça a equação

$$dY_t = (c_1(t)Y_t + c_2(t)) dt + (\sigma_1(t)Y_t + \sigma_2(t)) dW_t, \quad t \in [t_0, T], \quad Y_{t_0} = Y_0 \quad \text{q.c.}, \quad (2.25)$$

para determinadas funções deterministas c_i e σ_i , $i = 1, 2$. Se esta última equação admitir uma solução, podemos construir, por inversão, uma solução da equação inicial.

Note-se que este procedimento pressupõe que h seja uma função invertível, no sentido de que exista uma função real, h^* , definida sobre $[t_0, T] \times \mathbb{R}$, tal que, para todos $x, y \in \mathbb{R}$ e $t \in [t_0, T]$, se tenha $y = h(t, h^*(t, y))$ e $x = h^*(t, h(t, x))$.

Definição 2.3.3. A equação (2.24) diz-se **reduzível** se existir uma função função real, h , definida sobre $[t_0, T] \times \mathbb{R}$, invertível, contínua e com derivadas parciais

$$h_t = \frac{\partial h}{\partial t}, \quad h_x = \frac{\partial h}{\partial x} \quad \text{e} \quad h_{xx} = \frac{\partial^2 h}{\partial x^2}$$

contínuas, tal que o processo estocástico $(Y_t, t \in [t_0, T])$ verifica uma equação da forma (2.25) (c.f. [8]).

Naturalmente, se a equação (2.24) for reduzível, há que determinar as funções h , c_i e σ_i , $i = 1, 2$, pretendidas, o que é conseguido através do lema de Itô. De facto, escrevendo

$$\begin{aligned} dY_t &= \left(h_t(t, X_t) + f(t, X_t)h_x(t, X_t) + \frac{1}{2}g^2(t, X_t)h_{xx}(t, X_t) \right) dt + \\ &+ g(t, X_t)h_x(t, X_t)dW_t \quad t \in [t_0, T], \end{aligned}$$

constatamos que $(Y_t, t \in [t_0, T])$ satisfaz (2.25) se, para todo o $t \in [t_0, T]$,

$$\begin{cases} h_t(t, h^*(t, Y_t)) + f(t, h^*(t, Y_t))h_x(t, h^*(t, Y_t)) + \frac{1}{2}g^2(t, h^*(t, Y_t))h_{xx}(t, h^*(t, Y_t)) = \\ \quad = c_1(t)Y_t + c_2(t) \\ g(t, h^*(t, Y_t))h_x(t, h^*(t, Y_t)) = \sigma_1(t)Y_t + \sigma_2(t), \end{cases}$$

uma vez que $X_t = h^*(t, Y_t)$.

Concluimos, assim, que as incógnitas h , c_i e σ_i , $i = 1, 2$ serão obtidas resolvendo o sistema

$$\begin{cases} h_t(t, x) + f(t, x)h_x(t, x) + \frac{1}{2}g^2(t, x)h_{xx}(t, x) = c_1(t)h(t, x) + c_2(t) \\ g(t, x)h_x(t, x) = \sigma_1(t)h(t, x) + \sigma_2(t), \end{cases}$$

o que se reveste, em geral, de alguma complexidade.

Por esta razão, iremos ilustrar a técnica no caso em que as funções envolvidas (f , g , h , c_i e σ_i , $i = 1, 2$) não dependem de t , remetendo para [8] o caso geral e para [19] o caso em que c_1 e σ_1 são identicamente nulas.

Admitiremos, desde já, que $g(x) \neq 0$, para todo o $x \in \mathbb{R}$.

Nestas condições, o sistema anterior reduz-se a

$$\begin{cases} f(x)h'(x) + \frac{1}{2} g^2(x)h''(x) = c_1 h(x) + c_2 \\ g(x)h'(x) = \sigma_1 h(x) + \sigma_2. \end{cases} \quad (2.26)$$

Ora, uma solução da segunda equação é dada por

$$h(x) = K e^{\sigma_1 r(x)} - \frac{\sigma_2}{\sigma_1}, \text{ se } \sigma_1 \neq 0 \quad (2.27)$$

e

$$h(x) = \sigma_2 r(x), \text{ se } \sigma_1 = 0, \quad (2.28)$$

onde K é uma constante real não nula e $r(x) = \int \frac{1}{g(x)} dx$.

Analisemos detalhadamente o caso em que $\sigma_1 \neq 0$, pois o caso correspondente a $\sigma_1 = 0$ é, em tudo, semelhante.

Substituindo as expressões de h , h' e h'' na primeira equação de (2.26), obtemos a equação

$$\left(\sigma_1 s(x) + \frac{\sigma_1^2}{2} - c_1 \right) e^{\sigma_1 r(x)} = \frac{c_2 \sigma_1 - c_1 \sigma_2}{\sigma_1 K}, \quad (2.29)$$

onde $s(x) = \frac{f(x)}{g(x)} - \frac{g'(x)}{2}$, $x \in \mathbb{R}$.

Derivando (2.29), vem ainda

$$\left(s'(x) + \frac{1}{g(x)} \left(\sigma_1 s(x) + \frac{\sigma_1^2}{2} - c_1 \right) \right) \sigma_1 e^{\sigma_1 r(x)} = 0, \quad x \in \mathbb{R}.$$

Por fim, simplificando esta igualdade e derivando novamente, somos conduzidos a

$$\sigma_1 s'(x) + (gs')'(x) = 0, \quad x \in \mathbb{R}.$$

Portanto, o sistema (2.26) implica que

$$\forall x \in \mathbb{R}, \quad s'(x) = 0 \quad \vee \quad \left(\frac{(gs')'}{s'} \right)'(x) = 0. \quad (2.30)$$

De igual modo se verifica que, se $\sigma_1 = 0$, a escolha $h(x) = \sigma_2 r(x)$, $x \in \mathbb{R}$, conduz à condição $(gs')'(x) = 0$, $x \in \mathbb{R}$, que também implica (2.30).

Estamos agora em condições de enunciar a seguinte proposição, que esclarece o problema da redutibilidade de uma EDE não linear da forma considerada.

Proposição 2.3.5. *Suponhamos que as funções f e g são deriváveis até às ordens 3 e 4, respetivamente, e que $g(x) \neq 0$, para todo o $x \in \mathbb{R}$.*

A condição (2.30) é necessária e suficiente para que a EDE

$$dX_t = f(X_t) dt + g(X_t) dW_t, \quad t \in [t_0, T], \quad X_t = X_0, \quad q.c.$$

seja redutível.

Demonstração: *Atendendo ao raciocínio efetuado nas páginas anteriores, basta mostrar que a condição é suficiente.*

Para tal, vejamos que:

- *se $s'(x) = 0$, para todo o $x \in \mathbb{R}$, então podemos tomar h da forma (2.27), com $\sigma_1 \in \mathbb{R} \setminus \{0\}$ arbitrária e $\sigma_2 = 0$ garantindo, deste modo, que (2.26) é verificado;*
- *se $s'(x) \neq 0$ e $(gs')'(x) = 0$, para todo o $x \in \mathbb{R}$, basta tomar h da forma (2.28), para $\sigma_2 \in \mathbb{R} \setminus \{0\}$ arbitrária;*
- *se $s'(x) \neq 0$ e $(gs')'(x) \neq 0$ mas $\left(\frac{(gs')'}{s'}\right)'(x) = 0$, para todo o $x \in \mathbb{R}$, então podemos tomar h da forma (2.27), com $\sigma_2 = 0$ e $\sigma_1 = -\frac{(gs')'(x)}{s'(x)}$.*

Mas, em todos os casos, está também garantida a existência de constantes c_i , $i = 1, 2$ tais que a primeira equação de (2.26) é verificada, o que termina a demonstração.

■

Ilustramos a aplicação deste resultado com o seguinte exemplo.

Exemplo 2.3.4. *Consideremos a equação*

$$dX_t = \frac{1}{2}X_t dt + \sqrt{1 + X_t^2} dW_t, \quad t \in [0, T]$$

com X_0 nas condições habituais.

Sendo $f(x) = \frac{1}{2}x$ e $g(x) = \sqrt{1 + x^2}$, trata-se de uma equação redutível, uma vez que $s(x) = 0$, $x \in \mathbb{R}$.

Tomemos, então, de acordo com o que foi dito,

$$h(x) = K e^{\sigma_1 r(x)} = K e^{\sigma_1 \arg \sinh(x)},$$

com $K, \sigma_1 \in \mathbb{R} \setminus \{0\}$ arbitrárias.

Para determinar c_1 e c_2 , recorremos à primeira equação de (2.26), substituindo h'

e h'' no primeiro membro.

Ora,

$$h' = \frac{\sigma_1}{g}h \quad e \quad h'' = \frac{\sigma_1}{g^2}h(\sigma_1 - g'),$$

pelo que

$$fh' + \frac{1}{2}g^2h'' = \sigma_1h \left(\frac{f}{g} + \frac{1}{2}(\sigma_1 - g') \right).$$

Por outro lado,

$$\frac{f}{g} = \frac{g'}{2},$$

donde

$$fh' + \frac{1}{2}g^2h'' = \sigma_1h \left(\frac{g'}{2} + \frac{\sigma_1}{2} - \frac{g'}{2} \right) = \frac{\sigma_1^2}{2}h,$$

e, portanto $c_1 = \frac{\sigma_1^2}{2}$ e $c_2 = 0$.

Assim, o processo estocástico $Y_t = h(X_t)$, $t \in [0, T]$ verifica a EDE

$$dY_t = \frac{\sigma_1^2}{2}Y_t dt + \sigma_1Y_t dW_t,$$

com $Y_0 = h(X_0) = K \exp(\sigma_1 \arg \sinh(X_0))$.

Então, temos

$$Y_t = Y_0 e^{\sigma_1 W_t}, \quad t \in [0, T],$$

e, conseqüentemente,

$$X_t = h^{-1}(Y_t) = \sinh \left(\frac{1}{\sigma_1} \left(\ln \left(\frac{Y_0}{K} \right) + \sigma_1 W_t \right) \right) = \sinh(\arg \sinh(X_0) + W_t).$$

△

Capítulo 3

Métodos Numéricos para resolver Equações Diferenciais Estocásticas

3.1. Introdução

Com já foi referido, o estudo dos métodos numéricos conheceu um desenvolvimento considerável nas últimas décadas.

Desde o trabalho pioneiro de Milstein [26] até à atualidade, surgiram diversas abordagens para a construção de tais métodos, como, por exemplo, as que foram propostas por Boyce [3], que passa por conceber a equação como uma equação diferencial aleatória, por Kushner et al. [23], em que a solução da equação é aproximada por uma cadeia de Markov, ou a que é descrita por Karatzas e Shreve [17], obtida a partir da fórmula de Feynman-Kac.

Contudo, tanto quanto pudemos perceber, destacam-se, sobretudo, as que utilizam técnicas análogas às usadas no caso das EDO's, baseadas na truncatura da expansão de Itô-Taylor, assim como as versões estocásticas dos métodos de Runge-Kutta tradicionais.

Em conformidade com esta analogia, surgiram naturalmente as classes dos métodos de passo único e de passo múltiplo, assim como as dos métodos explícitos e implícitos. Para uma revisão bibliográfica veja-se [19], [27] e [31].

Neste trabalho, limitar-nos-emos a apresentar os métodos de Euler-Maruyama e de Milstein explícitos e de passo único. Apesar da sua simplicidade, são bastante populares e permitem ilustrar algumas das dificuldades subjacentes à generalização da abordagem clássica ao caso estocástico, quer ao nível teórico (na análise da qualidade das aproximações construídas), quer a nível prático (na implementação propriamente dita).

A fim de simplificar a exposição, tomaremos, no que se segue, $t_0 = 0$ e, de acordo com as convenções adotadas, consideraremos uma EDE da forma (2.8), com a con-

dição inicial $X_0 = x_0$ q.c., onde x_0 é um número real.

Admitiremos ainda que os coeficientes f e g verificam as condições (2.9) e (2.10), que asseguram a existência e a unicidade da solução (forte) da equação, denotada por $(X_t, t \in [0, T])$, como anteriormente.

No contexto que aqui nos interessa, a construção de uma aproximação numérica de $(X_t, t \in [0, T])$ passa genericamente por efetuar a discretização do intervalo $[0, T]$ segundo uma partição

$$\pi : 0 = \tau_0 < \tau_1 < \dots < \tau_N = T,$$

de diâmetro $\delta = \max_{n=0,1,\dots,N} (\tau_{n+1} - \tau_n)$ e definir $N + 1$ variáveis aleatórias reais $\bar{X}_{\tau_n}^\pi$, $n = 0, 1, \dots, N$, segundo um algoritmo recursivo da forma

$$\begin{cases} \bar{X}_{\tau_0}^\pi = F(0, x_0) \\ \bar{X}_{\tau_{n+1}}^\pi = F_{n+1}(\bar{X}_{\tau_0}^\pi, \dots, \bar{X}_{\tau_n}^\pi) + G_n(Z_{n+1}), \quad n = 0, 1, \dots, N - 1, \end{cases} \quad (3.1)$$

com F_j e G_j funções reais mensuráveis e Z_j uma variável aleatória real mensurável em relação a \mathcal{F}_{τ_j} , $j = 1, \dots, N$, de tal modo que $\bar{X}_{\tau_n}^\pi$ é uma variável aleatória real mensurável em relação a \mathcal{F}_{τ_n} , $n = 1, \dots, N$ (c.f. [20]).

A sequência $\bar{X}_{\tau_n}^\pi$, $n = 0, 1, \dots, N$ constitui, assim, a aproximação numérica procurada que, à partida, deverá ser tanto “melhor” quanto menor for o diâmetro da partição.

Ora, como seria de esperar, para avaliar a qualidade de aproximações numéricas deste tipo, levantam-se questões de índole diversa, uma vez que, para além dos problemas numéricos envolvidos (à semelhança do que se passa nos métodos numéricos para as EDO's), surgem agora os problemas relacionados com os critérios de “proximidade” entre variáveis aleatórias. Não cabendo, no presente trabalho, uma descrição de todas as perspetivas possíveis (nomeadamente, a consistência e a estabilidade numérica), abordaremos apenas a questão da convergência de tais aproximações.

No que diz respeito a esta questão, distinguem-se, fundamentalmente dois aspetos: o primeiro prende-se com a ideia de “proximidade” entre as trajetórias do processo solução e da correspondente aproximação numérica, enquanto que o segundo está relacionado com a “proximidade” das respetivas leis de probabilidade e, em particular, dos seus momentos.

Quanto ao primeiro tipo de convergência, usualmente designado por **convergência forte**, a medida de qualidade da aproximação proposta por [19] é baseada em $\epsilon(\delta) = E(|X_T - \bar{X}_T^\pi|)$ e define-se do seguinte modo:

Definição 3.1.1. Diz-se que uma aproximação numérica da forma (3.1)

(i) **converge fortemente** para $(X_t, t \in [0, T],$ em $T,$ se

$$\lim_{\delta \rightarrow 0} E(|X_T - \bar{X}_T^\pi|) = 0;$$

(ii) **converge fortemente** para $(X_t, t \in [0, T],$ **com ordem** $\gamma > 0,$ em $T,$ se existem uma constante positiva $C,$ que não depende de $\delta,$ e $\delta_0 \in \mathbb{R}^+$ tais que

$$E(|X_T - \bar{X}_T^\pi|) \leq C\delta^\gamma,$$

para todo o $\delta \in (0, \delta_0).$

Observação 3.1.1. Note-se que uma majoração suficientemente precisa desta quantidade em função de δ é, efetivamente, de grande utilidade na aproximação das trajetórias do processo solução no ponto $T,$ pois permite também, com base na desigualdade de Markov, controlar o erro $|X_T(\omega) - \bar{X}_T^\pi(\omega)|,$ para cada trajetória.

Mas, quer em termos teóricos, quer em termos práticos, há toda a conveniência em trabalhar com medidas de qualidade mais exigentes, majorando convenientemente o erro global dado por

$$E\left(\sup_{n=0,1,\dots,N} |X_{\tau_n} - \bar{X}_{\tau_n}^\pi|\right).$$

Para atingir este objetivo, a estratégia usual consiste em construir um processo estocástico a tempo contínuo $(\bar{X}_t^\pi, t \in [0, T]),$ cujas variáveis, nos instantes $\tau_n,$ $n = 0, 1, \dots, N,$ coincidam com as da sequência associada à partição e procurar um majorante de

$$E\left(\sup_{0 \leq t \leq T} |X_t - \bar{X}_t^\pi|\right).$$

Em [19], são sugeridos dois procedimentos para a construção de $(\bar{X}_t^\pi, t \in [0, T]):$ uma via consiste em definir $\bar{X}_t^\pi, t \in [0, T],$ considerando $\bar{X}_t^\pi = \bar{X}_{\tau_n}^\pi,$ para $t \in [\tau_n, \tau_{n+1}[,$ $n = 0, 1, \dots, N - 1;$ em alternativa, \bar{X}_t^π poderá ser definida por interpolação linear

$$\bar{X}_t^\pi = \bar{X}_{\tau_n}^\pi + \frac{t - \tau_n}{\tau_{n+1} - \tau_n} (\bar{X}_{\tau_{n+1}}^\pi - \bar{X}_{\tau_n}^\pi), \quad t \in [\tau_n, \tau_{n+1}[,$$
¹

o que possui a vantagem de gerar trajetórias contínuas e simples de implementar computacionalmente.

Contudo, tal como é referido em [19], em ambos os casos, as trajetórias geradas não

¹Claro que $\bar{X}_T^\pi = \bar{X}_{\tau_N}^\pi,$ uma vez que $\tau_N = T.$

refletem o caráter irregular das trajetórias do processo $(X_t, t \in [0, T])$ no que se refere à (não) diferenciabilidade, que decorre da irregularidade das trajetórias do processo de Wiener subjacente à EDE.

Sob o ponto de vista teórico, é habitual recorrer a interpolações que são igualmente baseadas em integrais estocásticos, como veremos nas próximas secções.

Relativamente à aproximação dos momentos, tem lugar a noção de **convergência fraca**, definida como se segue (c.f. [19]):

Definição 3.1.2. *Seja \mathcal{C} uma determinada classe de funções reais de variável real.*

Diz-se que uma aproximação numérica da forma (3.1)

(i) converge fracamente para $(X_t, t \in [0, T])$, em T , em relação à classe \mathcal{C} , se

$$\lim_{\delta \rightarrow 0} |E(\varphi(X_T)) - E(\varphi(\bar{X}_T^\pi))| = 0$$

para qualquer $\varphi \in \mathcal{C}$; ²

(ii) converge fracamente para $(X_t, t \in [0, T])$ com ordem $\beta > 0$, em T , em relação à classe \mathcal{C}^3 , se, existem uma constante positiva C , que não depende de δ , e $\delta_0 \in \mathbb{R}^+$ tais que

$$|E(\varphi(X_T)) - E(\varphi(\bar{X}_T^\pi))| \leq C\delta^\beta,$$

para todo o $\delta \in (0, \delta_0)$.

Note-se que este tipo de convergência não exige que o processo $(X_t, t \in [0, T])$ e a sequência $\bar{X}_{T_n}^\pi$, $n = 0, 1, \dots, N$, sejam gerados à custa do mesmo processo de Wiener, uma vez que apenas pressupõe a “proximidade” das leis de probabilidade de X_T e \bar{X}_T^π .

Tal como foi referido na introdução, iremos estudar apenas as propriedades de convergência forte dos métodos de Euler-Maruyama e de Milstein.

Observação 3.1.2. *Repare-se que, se a função φ for, por exemplo, lipschitziana, então toda a aproximação numérica que convirja fortemente para $(X_t, t \in [0, T])$, em T , com ordem de convergência γ , convergirá também em sentido fraco, com a mesma ordem de convergência. No entanto, de um modo geral, é possível obter ordens de convergência fraca mais elevadas (c.f. [19]).*

²Desde que exista a esperança das variáveis aleatórias intervenientes.

³Em [19], os autores restringem esta classe ao espaço $\mathcal{C}_P^{2(\beta+1)}$, constituído pelas funções reais de variável real de classe $\mathcal{C}^{2(\beta+1)}$, cujas derivadas até à ordem $2(\beta + 1)$ (incluindo a própria função) têm crescimento polinomial.

3.2. Expansão de Itô-Taylor

À semelhança do que se passa para as EDO's, os métodos de Euler-Maruyama e de Milstein baseiam-se na denominada expansão de Itô-Taylor, que não é mais do que a generalização da “fórmula” de Taylor no âmbito do cálculo estocástico. Apresentaremos, agora, uma versão simplificada desta expansão.

A expansão de Itô-Taylor obtém-se por iteração da “fórmula” de Itô e envolve conceitos de integração múltipla (c.f. [19]).

Aplicando a “fórmula” de Itô (2.3) a uma função real de variável real, ψ , nas condições do lema de Itô, vem, para $t \in [0, T]$ arbitrariamente fixo,

$$\begin{aligned} \psi(X_{t+h}) &= \psi(X_t) + \int_t^{t+h} \left(f(s, X_s)\psi'(X_s) + \frac{1}{2}g(s, X_s)^2\psi''(X_s) \right) ds \\ &\quad + \int_t^{t+h} (g(s, X_s)\psi'(X_s)) dW_s \\ &= \psi(X_t) + \int_t^{t+h} \mathcal{L}^1\psi(X_s)ds + \int_t^{t+h} \mathcal{L}^2\psi(X_s)dW_s, \quad h > 0, t+h \in]0, T], \end{aligned} \tag{3.2}$$

onde \mathcal{L}^1 e \mathcal{L}^2 denotam os seguintes operadores

$$\mathcal{L}^1 = \frac{\partial}{\partial t} + f \frac{\partial}{\partial x} + \frac{1}{2}g^2 \frac{\partial^2}{\partial x^2} \quad e \quad \mathcal{L}^2 = g \frac{\partial}{\partial x}$$

Assim, após a aplicação da “fórmula” de Itô a $\mathcal{L}^1\psi$ e $\mathcal{L}^2\psi$ obtemos

$$\begin{aligned} \psi(X_{t+h}) &= \psi(X_t) + \int_t^{t+h} \left(\mathcal{L}^1\psi(X_t) + \int_t^s \mathcal{L}^1\mathcal{L}^1\psi(X_r)dr + \int_t^s \mathcal{L}^2\mathcal{L}^1\psi(X_r)dW_r \right) ds \\ &\quad + \int_t^{t+h} \left(\mathcal{L}^2\psi(X_t) + \int_t^s \mathcal{L}^1\mathcal{L}^2\psi(X_r)dr + \int_t^s \mathcal{L}^2\mathcal{L}^2\psi(X_r)dW_r \right) dW_s \\ &= \psi(X_t) + \mathcal{L}^1\psi(X_t) \int_t^{t+h} ds + \mathcal{L}^2\psi(X_t) \int_t^{t+h} dW_s + R_1, \end{aligned}$$

ou, mais abreviadamente,

$$\psi(X_{t+h}) = \psi(X_t) + \mathcal{L}^1\psi(X_t)I_{(0),h} + \mathcal{L}^2\psi(X_t)I_{(1),h} + R_1.$$

Voltando a iterar sobre os termos de R_1 , somos ainda conduzidos a

$$\begin{aligned} R_1 &= \mathcal{L}^1\mathcal{L}^1\psi(X_t) \int_t^{t+h} ds \int_t^s dr + \mathcal{L}^2\mathcal{L}^1\psi(X_t) \int_t^{t+h} ds \int_t^s dW_r \\ &\quad + \mathcal{L}^1\mathcal{L}^2\psi(X_t) \int_t^{t+h} dW_s \int_t^s dr + \mathcal{L}^2\mathcal{L}^2\psi(X_t) \int_t^{t+h} dW_s \int_t^s dW_r + R_2 \\ &= \mathcal{L}^1\mathcal{L}^1\psi(X_t)I_{(0,0),h} + \mathcal{L}^2\mathcal{L}^1\psi(X_t)I_{(0,1),h} + \mathcal{L}^1\mathcal{L}^2\psi(X_t)I_{(1,0),h} \\ &\quad + \mathcal{L}^2\mathcal{L}^2\psi(X_t)I_{(1,1),h} + R_2, \end{aligned}$$

onde agora $I_{(i,j),h}$, $i, j = 0, 1$, denota um integral duplo, com i e j tomando os valores 0 ou 1 consoante o integral correspondente seja de Lebesgue ou de Itô, respetivamente.

Deste modo, a expansão de Itô-Taylor de segunda ordem de $\psi(X_{t+h})$ é dada por ⁴

$$\begin{aligned} \psi(X_{t+h}) &= \psi(X_t) \\ &+ \mathcal{L}^1\psi(X_t)I_{(0),h} + \mathcal{L}^2\psi(X_t)I_{(1),h} \\ &+ \mathcal{L}^1\mathcal{L}^1\psi(X_t)I_{(0,0),h} + \mathcal{L}^2\mathcal{L}^1\psi(X_t)I_{(0,1),h} + \mathcal{L}^1\mathcal{L}^2\psi(X_t)I_{(1,0),h} \\ &+ \mathcal{L}^2\mathcal{L}^2\psi(X_t)I_{(1,1),h} + R_2, \quad h > 0, t+h \in]0, T]. \end{aligned} \tag{3.3}$$

Nas secções seguintes, consideraremos ψ a função identidade sobre \mathbb{R} e usaremos a expansão de Itô-Taylor de X_t para motivar os métodos numéricos apresentados.

3.3. Método de Euler-Maruyama

Um dos métodos mais simples para construir uma aproximação numérica de $(X_t, t \in [0, T])$ é o **método de Euler** estocástico ou **método de Euler-Maruyama**, que se baseia em (3.2) e resulta das aproximações

$$\int_t^{t+h} f(s, X_s)ds \simeq f(t, X_t)h$$

e

$$\int_t^{t+h} g(s, X_s)dW_s \simeq g(t, X_t)(W_{t+h} - W_t),$$

analogamente ao caso determinista.

A sequência $\bar{X}_{\tau_n}^\pi$, $n = 0, 1, \dots, N$, é pois, definida iterativamente por:

$$\begin{aligned} \bar{X}_0^\pi &= x_0 \\ \bar{X}_{\tau_{n+1}}^\pi &= \bar{X}_{\tau_n}^\pi + f(\tau_n, \bar{X}_{\tau_n}^\pi)(\tau_{n+1} - \tau_n) + g(\tau_n, \bar{X}_{\tau_n}^\pi)(W_{\tau_{n+1}} - W_{\tau_n}). \end{aligned} \tag{3.4}$$

Na sequência do que foi afirmado na secção 3.1, é conveniente, em termos teóricos, considerar interpolações que se exprimem à custa de integrais estocásticos.

Neste sentido, uma forma natural de proceder consiste em definir, para $t \in [\tau_n, \tau_{n+1}]$,

$$\begin{aligned} \bar{X}_t^\pi &= \bar{X}_{\tau_n}^\pi + f(\tau_n, \bar{X}_{\tau_n}^\pi)(t - \tau_n) + g(\tau_n, \bar{X}_{\tau_n}^\pi)(W_t - W_{\tau_n}) \\ &= \bar{X}_{\tau_n}^\pi + \int_{\tau_n}^t f(\tau_n, \bar{X}_{\tau_n}^\pi)ds + \int_{\tau_n}^t g(\tau_n, \bar{X}_{\tau_n}^\pi)dW_s. \end{aligned}$$

⁴Sob reserva de existência dos integrais envolvidos.

Portanto, substituindo sucessivamente as expressões de X_{τ_j} , $j = 1, \dots, n$, podemos escrever

$$\begin{aligned}\bar{X}_t^\pi &= \bar{X}_{\tau_n}^\pi + \int_{\tau_n}^t f(\tau_n, \bar{X}_{\tau_n}^\pi) ds + \int_{\tau_n}^t g(\tau_n, \bar{X}_{\tau_n}^\pi) dW_s \\ &= x_0 + \sum_{i=0}^{n-1} (f(\tau_i, \bar{X}_{\tau_i}^\pi)(\tau_{i+1} - \tau_i) + g(\tau_i, \bar{X}_{\tau_i}^\pi)(W_{\tau_{i+1}} - W_{\tau_i})) \\ &\quad + \int_{\tau_n}^t f(\tau_n, \bar{X}_{\tau_n}^\pi) ds + \int_{\tau_n}^t g(\tau_n, \bar{X}_{\tau_n}^\pi) dW_s \\ &= x_0 + \sum_{i=0}^{n-1} \left(\int_{\tau_i}^{\tau_{i+1}} f(\tau_i, \bar{X}_{\tau_i}^\pi) ds + \int_{\tau_i}^{\tau_{i+1}} g(\tau_i, \bar{X}_{\tau_i}^\pi) dW_s \right) \\ &\quad + \int_{\tau_n}^t f(\tau_n, \bar{X}_{\tau_n}^\pi) ds + \int_{\tau_n}^t g(\tau_n, \bar{X}_{\tau_n}^\pi) dW_s.\end{aligned}$$

Seja agora $\pi_t = \sup \{ \tau_n \in \pi : \tau_n \leq t \}$, $t \in [0, T]$. Como, para $s \in [\tau_i, \tau_{i+1}]$, se tem $\pi_s = \tau_i$, \bar{X}_t^π admite ainda a expressão

$$\begin{aligned}\bar{X}_t^\pi &= x_0 + \sum_{i=0}^{n-1} \left(\int_{\tau_i}^{\tau_{i+1}} f(\pi_s, \bar{X}_{\pi_s}^\pi) ds + \int_{\tau_i}^{\tau_{i+1}} g(\pi_s, \bar{X}_{\pi_s}^\pi) dW_s \right) \\ &\quad + \int_{\tau_n}^t f(\pi_s, \bar{X}_{\pi_s}^\pi) ds + \int_{\tau_n}^t g(\pi_s, \bar{X}_{\pi_s}^\pi) dW_s,\end{aligned}$$

ou seja,

$$\bar{X}_t^\pi = x_0 + \int_0^t f(\pi_s, \bar{X}_{\pi_s}^\pi) ds + \int_0^t g(\pi_s, \bar{X}_{\pi_s}^\pi) dW_s.$$

Esta expressão de \bar{X}_t^π será particularmente útil na próxima proposição, que estabelece a ordem de convergência do método de Euler-Maruyama. Antes, porém, apresentamos um resultado auxiliar que será necessário na demonstração da referida proposição.

Lema de Burkholder-Davis-Gundy: *Se $(H_t, t \in [0, T])$ é um processo estocástico que pertence à classe $L^2(\Omega, [0, T])$, então o processo estocástico definido por*

$$Z_t = \int_0^t |H_s| dW_s, \quad t \in [0, T],$$

é tal que

$$\forall p \geq 1, \exists B > 0 : E \left(\left(\sup_{0 \leq t \leq T} |Z_t| \right)^p \right) \leq BE \left(\left(\int_0^T |H_s|^2 ds \right)^{\frac{p}{2}} \right).$$

Demonstração: Veja-se, por exemplo, [18].

■

Proposição 3.3.1. *Suponhamos que os coeficientes f e g verificam as condições da Proposição 2.2.1, bem como a seguinte condição:*

$$\exists A \in \mathbb{R}^+ : \forall x \in \mathbb{R}, \forall s, t \in [0, T],$$

$$|f(t, x) - f(s, x)| + |g(t, x) - g(s, x)| \leq A(1 + |x|)|t - s|^{\frac{1}{2}}. \quad (3.5)$$

Então, existe uma constante $C^* > 0$, dependente apenas de T e de x_0 , tal que, para toda a partição π de $[0, T]$, se tem

$$E \left(\sup_{0 \leq t \leq T} |X_t - \bar{X}_t^\pi| \right) \leq C^* \delta^{\frac{1}{2}}.$$

Demonstração: No que se segue e para simplificar a escrita, denotaremos \bar{X}_t^π apenas por \bar{X}_t , sempre que não for necessário explicitar a dependência da partição.

Seja $Z_t = \sup_{0 \leq u \leq t} |X_u - \bar{X}_u|$, para $t \in [0, T]$ arbitrariamente fixo.

Para todo o $u \in [0, T]$, podemos escrever

$$X_u - \bar{X}_u = \int_0^u (f(s, X_s) - f(\pi_s, \bar{X}_{\pi_s})) ds + \int_0^u (g(s, X_s) - g(\pi_s, \bar{X}_{\pi_s})) dW_s, \quad q.c.,$$

pelo que

$$\begin{aligned} Z_t &\leq \sup_{0 \leq u \leq t} \left| \int_0^u (f(s, X_s) - f(\pi_s, \bar{X}_{\pi_s})) ds + \sup_{0 \leq u \leq t} \left| \int_0^u (g(s, X_s) - g(\pi_s, \bar{X}_{\pi_s})) dW_s \right| \right|, \quad q.c. \\ &\leq \int_0^t |f(s, X_s) - f(\pi_s, \bar{X}_{\pi_s})| ds + \sup_{0 \leq u \leq t} \left| \int_0^u (g(s, X_s) - g(\pi_s, \bar{X}_{\pi_s})) dW_s \right|, \quad q.c.. \end{aligned}$$

e, portanto

$$E(Z_t^2) \leq 2E(U_1^2) + 2E(U_2^2) \quad q.c., \quad (3.6)$$

com

$$U_1 = \int_0^t |f(s, X_s) - f(\pi_s, \bar{X}_{\pi_s})| ds$$

e

$$U_2 = \sup_{0 \leq u \leq t} \left| \int_0^u (g(s, X_s) - g(\pi_s, \bar{X}_{\pi_s})) dW_s \right|.$$

A estratégia da demonstração consiste, assim, em majorar $E(U_1^2)$ e $E(U_2^2)$ convenientemente.

No que diz respeito à esperança de U_1^2 , temos, tendo em conta a desigualdade de Cauchy-Schwarz,

$$E(U_1^2) \leq tE \left(\int_0^t |f(s, X_s) - f(\pi_s, \bar{X}_{\pi_s})|^2 ds \right),$$

ou ainda,

$$E(U_1^2) \leq t \int_0^t E \left(|f(s, X_s) - f(\pi_s, \bar{X}_{\pi_s})|^2 \right) ds, \quad (3.7)$$

pelo teorema de Fubini-Tonelli.

Relativamente a $E(U_2^2)$, a desigualdade de Burkholder-Davis-Gundy e novamente o teorema de Fubini-Tonelli conduzem a

$$E(U_2^2) \leq B \int_0^t E \left(|g(s, X_s) - g(\pi_s, \bar{X}_{\pi_s})|^2 \right) ds. \quad (3.8)$$

Por forma a majorar as funções integrandas de ambos os integrais, notemos que, para todo o $s \in [0, t]$,

$$\begin{aligned} |f(s, X_s) - f(\pi_s, \bar{X}_{\pi_s})|^2 &\leq 3|f(s, X_s) - f(s, X_{\pi_s})|^2 + 3|f(s, X_{\pi_s}) - f(\pi_s, X_{\pi_s})|^2 + \\ &+ 3|f(\pi_s, X_{\pi_s}) - f(\pi_s, \bar{X}_{\pi_s})|^2. \end{aligned}$$

Com base nas hipóteses formuladas sobre os coeficientes f e g , (3.5) e (2.10), os termos anteriores são majorados q.c., por

$$3L^2|X_s - X_{\pi_s}|^2, \quad 3A^2(1 + |X_{\pi_s}|)^2|s - \pi_s| \quad e \quad 3L^2|X_{\pi_s} - \bar{X}_{\pi_s}|,$$

respetivamente.

Como

$$|X_{\pi_s} - \bar{X}_{\pi_s}|^2 \leq \left(\sup_{0 \leq u \leq s} |X_u - \bar{X}_u|^2 \right) \leq Z_s^2$$

vem, tomando esperanças em ambos os membros da desigualdade obtida,

$$\begin{aligned} E \left(|f(s, X_s) - f(\pi_s, \bar{X}_{\pi_s})|^2 \right) &\leq 3L^2 E(|X_s - X_{\pi_s}|^2) + 3A^2 E((1 + |X_{\pi_s}|)^2) |s - \pi_s| + \\ &+ 3L^2 E(Z_s^2). \end{aligned}$$

Ora,

$$E(|X_s - X_{\pi_s}|^2) \leq D(s - \pi_s)(1 + |x_0|^2),$$

por (2.12) da Proposição 2.2.2, (ii), uma vez que $X_0 = x_0$, q.c., com $x_0 \in \mathbb{R}$.

Por outro lado,

$$E((1 + |X_{\pi_s}|)^2) \leq 2 + 2E(|X_{\pi_s}|^2),$$

sendo esta última expressão majorada ainda por $2 + 2C(1 + |x_0|^2)$, tendo em conta (2.11) da Proposição 2.2.2, (i).

Constatamos, então, que a função integranda do integral que figura no segundo membro de (3.7) é majorada por

$$3L^2D(1 + |x_0|^2)\delta + 6A^2(1 + C + C|x_0|^2)\delta + 3L^2E(Z_s^2),$$

pois $0 \leq s - \pi_s \leq \delta$.

Concluimos, assim, que

$$E(U_1) \leq 3(LT)^2(1 + |x_0|^2)\delta + 6(AT)^2(1 + C + C|x_0|^2)\delta + 3L^2T \int_0^t E(Z_s^2)ds,$$

isto é,

$$E(U_1) \leq C_1(T, x_0)\delta + C_2(T) \int_0^t E(Z_s^2)ds.$$

Analogamente, a partir de (3.8), podemos afirmar que

$$E(U_2) \leq C_3(T, x_0)\delta + C_4 \int_0^t E(Z_s^2)ds.$$

Portanto, de (3.6), obtemos

$$E(Z_t^2) \leq C_5(T, x_0)\delta + C_6 \int_0^t E(Z_s^2)ds,$$

com $C_5(T, x_0)$ e C_6 constantes reais positivas dependentes apenas de T e x_0 .

Para terminar a demonstração, resta aplicar a desigualdade de Gronwall à função $\phi(t) = E(Z_t^2)$, $t \in [0, T]$, vindo

$$\forall t \in [0, T], \quad E(Z_t^2) \leq C_5(T, x_0)\delta e^{C_6 T},$$

o que conduz à desigualdade enunciada, com $C^*(T, x_0) = \sqrt{C_5(T, x_0)e^{C_6 T}}$, pois $E(Z_t) \leq E(Z_t^2)$.

■

Nas condições da proposição anterior, tem-se, então, para toda a partição, π , de $[0, T]$, com diâmetro δ ,

$$E \left(\sup_{n=0,1,\dots,N} |X_{\tau_n} - \bar{X}_{\tau_n}^\pi| \right) \leq C^* \delta^{\frac{1}{2}},$$

o que mostra, em particular, que este método conduz a aproximações numéricas fortemente convergentes, com ordem $\frac{1}{2}$.

Observação 3.3.1.

- (i) Se os coeficientes f e g não dependerem de t , a condição (3.5) é trivialmente verificada e a demonstração anterior é consideravelmente mais simples, no entanto não se consegue melhorar a ordem de convergência.

(ii) Encontramos em [29], por exemplo, uma generalização da Proposição anterior. De facto, substituindo a hipótese (3.5) pela condição mais geral $\forall \alpha \in]0, 1[, \exists A > 0 : \forall x \in \mathbb{R}, \forall s, t \in [0, T],$

$$|f(t, x) - f(s, x)| + |g(t, x) - g(s, x)| \leq A(1 + |x|)|t - s|^\alpha,$$

a autora conclui que

$$\forall p \in [1, +\infty[, E \left(\sup_{0 \leq t \leq T} |X_t - \bar{X}_t^\pi|^p \right) \leq C^*(T, x_0, p) \delta^{\beta p},$$

com $\beta = \min(\frac{1}{2}, \alpha)$.

Mais ainda, em [14] e [33], são estabelecidos resultados mais gerais, no sentido de serem exigidas hipóteses menos restritivas sobre os coeficientes da equação.

3.4. Método de Milstein

Para construir a aproximação numérica proposta por este método, adiciona-se aos três primeiros termos de (3.3), a parcela

$$\mathcal{L}^2 \mathcal{L}^2 \psi(X_t) I_{(1,1),h} = gg_x \int_t^{t+h} (W_s - W_t) dW_s,$$

recordando que consideramos ψ a função identidade.

Mas,

$$\begin{aligned} \int_t^{t+h} (W_s - W_t) dW_s &= \int_t^{t+h} W_s dW_s - \int_t^{t+h} W_t dW_s = \\ &= Y_{t+h} - Y_t - W_t(W_{t+h} - W_t), \end{aligned} \quad (3.9)$$

onde $Y_t = \int_0^t W_s dW_s$, $t \in [0, T]$.

Então, $(Y_t, t \in [0, T])$ é a (única) solução forte da equação

$$dY_t = W_t dW_t \text{ q.c.}, \quad t \in [0, T],$$

com a condição inicial $Y_0 = 0$ q.c. e, portanto,

$$Y_t = \frac{1}{2} W_t^2 - \frac{1}{2} t \text{ q.c.}, \quad t \in [0, T],$$

como sabemos.

Consequentemente, substituindo esta expressão em (3.9) e simplificando, vem

$$\int_t^{t+h} (W_s - W_t) dW_s = \frac{1}{2} (W_{t+h} - W_t)^2 - \frac{1}{2} h.$$

A aproximação numérica $\bar{X}_{\tau_n}^\pi$, $n = 0, 1, \dots, N$, construída pelo método de Milstein é, assim, definida iterativamente por:

$$\begin{aligned}\bar{X}_0^\pi &= x_0 \\ \bar{X}_{\tau_{n+1}}^\pi &= \bar{X}_{\tau_n}^\pi + f(\tau_n, \bar{X}_{\tau_n}^\pi)(\tau_{n+1} - \tau_n) + g(\tau_n, \bar{X}_{\tau_n}^\pi)(W_{\tau_{n+1}} - W_{\tau_n}) \\ &\quad + \frac{1}{2}g(\tau_n, \bar{X}_{\tau_n}^\pi)g_x(\tau_n, \bar{X}_{\tau_n}^\pi) \left((W_{\tau_{n+1}} - W_{\tau_n})^2 - (\tau_{n+1} - \tau_n) \right).\end{aligned}\quad (3.10)$$

O desempenho deste método é superior ao de Euler-Maruyama,⁵ uma vez que a sua ordem de convergência forte é um, tal como afirma a proposição seguinte (c.f. [19]).

Proposição 3.4.1. *Suponhamos que os coeficientes f e g verificam as condições do lema de Itô, bem como as condições da Proposição 3.3.1.*

Suponhamos ainda, que existem constantes positivas L' , K' e A' tais que

$$\forall x \in \mathbb{R}, \forall t \in [0, T], |gg_x(t, x) - gg_x(t, y)| \leq L'|x - y|, \quad (3.11)$$

$$\forall x \in \mathbb{R}, \forall t \in [0, T], |gg_x(t, x)|^2 \leq K'|1 + |x|^2|, \quad (3.12)$$

$$\forall x \in \mathbb{R}, \forall t \in [0, T], |f(t, x)| + |L^j f(t, x)| + |g(t, x)| + |L^j g(t, x)| \leq A'(1 + |x|)|t - s|^{\frac{1}{2}}. \quad (3.13)$$

Então, a aproximação numérica dada por (3.10) converge fortemente para a solução $(X_t, t \in [0, T])$ da EDE (2.8), com ordem de convergência $\gamma = 1$.

Demonstração: Veja-se [19], pág.350.

■

3.5. Aplicações

Apresentam-se, por fim, alguns exemplos que ilustram o desempenho dos métodos referidos, comparando as aproximações numéricas construídas com a trajetória da solução exata da equação.

Antes de mais, começámos por simular as trajetórias do processo de Wiener standard.

Para tal, considerámos uma partição do intervalo $[0, T]$, de diâmetro $\Delta = \frac{T}{M}$, baseada em instantes equidistantes, $t_n = n\Delta$, $n = 0, 1, \dots, M$, $M \in \mathbb{N}$.

⁵O método de Milstein coincide com o método de Euler-Maruyama quando $g_x = 0$.

Dado que $W_0 = 0$ q.c., podemos escrever

$$W_{t_{n+1}} = W_{t_n} + (W_{t_{n+1}} - W_{t_n}), \quad n = 0, 1, \dots, M - 1,$$

e, deste modo, gerar as trajetórias de $(W_t, t \in [0, T])$, com base na independência dos seus acréscimos e no facto de

$$W_{t_{n+1}} - W_{t_n} \sim N(0, \Delta), \quad n = 1, \dots, M.$$

Por outro lado, para construir as aproximações numéricas pretendidas, considerámos uma outra partição de $[0, T]$, tomando como diâmetro $\delta = \frac{T}{N}$, definida à custa dos instantes $\tau_n = n\delta$, $n = 0, 1, \dots, N$, $N \in \mathbb{N}$.

De modo a garantir que $\{\tau_0, \tau_1, \dots, \tau_N\} \subseteq \{t_0, t_1, \dots, t_M\}$, tomámos $\delta = R\Delta$, com $R \in \mathbb{N}$ (c.f. [13]). Os acréscimos subjacentes à construção das aproximações são, assim, dados por

$$W_{\tau_{n+1}} - W_{\tau_n} = W_{t_{(n+1)R\Delta}} - W_{t_{nR\Delta}} = \sum_{k=nR-R+1}^{nR} (W_k - W_{k-1}).$$

Em qualquer um dos casos, as trajetórias (a tempo contínuo), no intervalo $[0, T]$, são obtidas por interpolação linear, fazendo, sem perda de generalidade, $T = 1$ nos dois primeiros exemplos e $T = 5$ no terceiro exemplo.

Nas trajetórias da solução exata e das aproximações numéricas, os segmentos de reta referentes à interpolação linear encontram-se a cheio e a tracejado, respetivamente. As trajetórias da solução exata e das aproximações de Euler-Maruyama e de Milstein estão representadas a verde, azul e magenta, respetivamente.

Exemplo 3.5.1. *Seja $X = (X_t, t \in [0, 1])$ o processo de Itô que satisfaz a equação de Black-Scholes*

$$dX_t = 1.5 X_t dt + X_t dW_t, \quad t \in [0, 1], \quad (3.14)$$

com condição inicial $X_0 = 1$ q.c..

De acordo com o exemplo (2.3.3), a solução exata da equação é dada por

$$X_t = \exp(t + W_t) \quad \text{q.c.}, \quad t \in [0, 1]$$

pelo que podemos comparar uma aproximação obtida a partir de (3.4) com a solução exata.

Obtemos as soluções exatas, para a mesma trajetória do processo de Wiener gerada, a partir de

$$X_{t_n} = \exp\left(t_n + \sum_{i=1}^n (W_i - W_{i-1})\right).$$

Capítulo 3 Métodos Numéricos para resolver Equações Diferenciais Estocásticas

Para obter as aproximações de Euler-Maruyama e de Milstein consideramos $\delta = 2^{-2}$ e $\delta = 2^{-4}$, $\bar{X}_0 = X_0 = 1$ e procedemos recursivamente de acordo com (3.4) e (3.10), respectivamente.

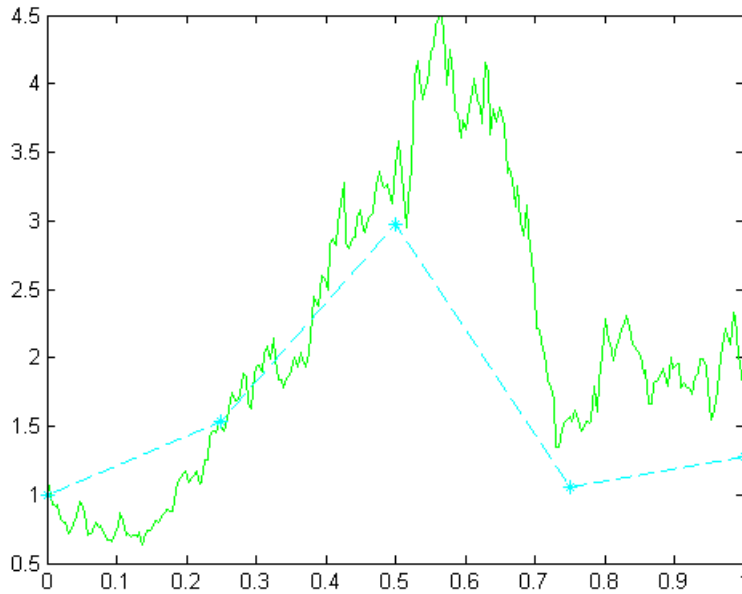


Figura 3.1: Aproximação de Euler e solução exata com $\delta = 2^{-2}$

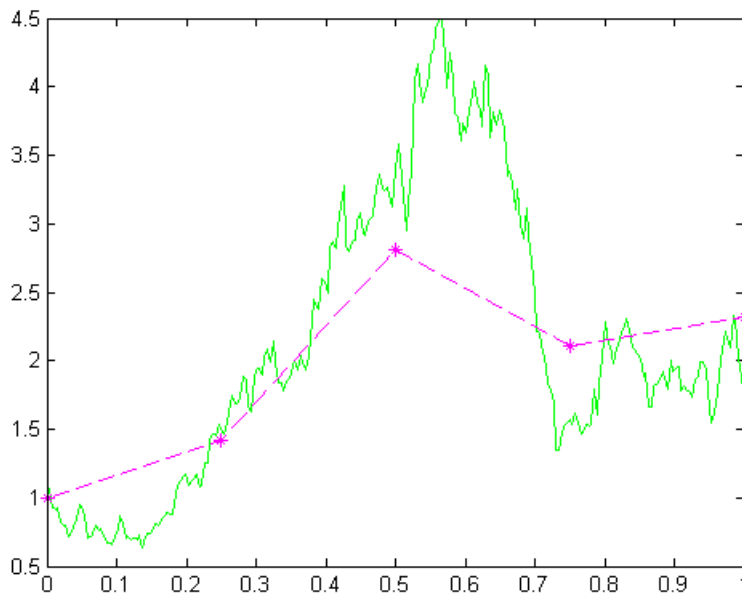


Figura 3.2: Aproximação de Milstein e solução exata com $\delta = 2^{-2}$

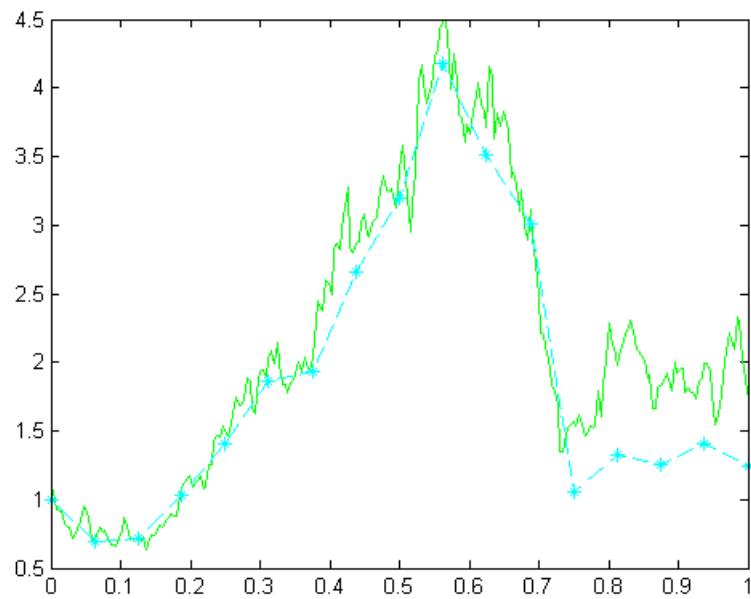


Figura 3.3: Aproximação de Euler e solução exata com $\delta = 2^{-4}$

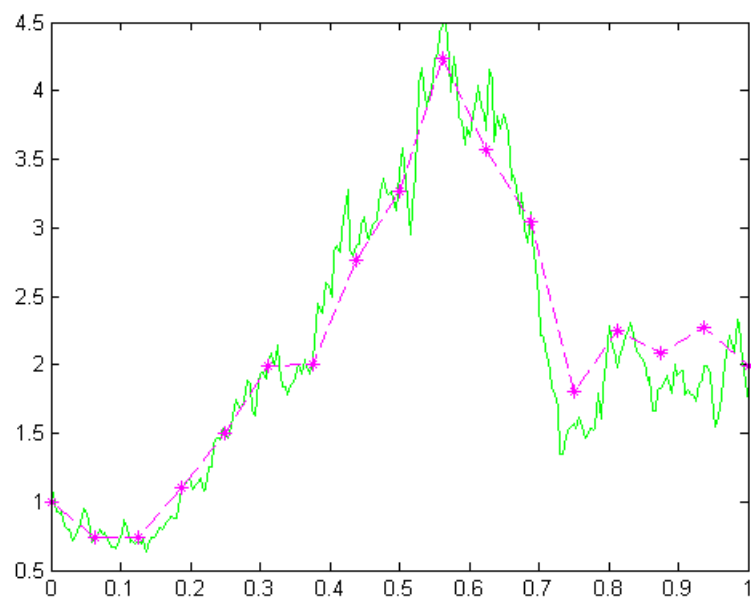


Figura 3.4: Aproximação de Milstein e solução exata com $\delta = 2^{-4}$

A partir da visualização das figuras, é notório que a qualidade das aproximações aumenta consideravelmente ao aumentar o diâmetro de $\delta = 2^{-2}$ para $\delta = 2^{-4}$, bem como, que as aproximações obtidas pelo método de Milstein apresentam uma

qualidade superior relativamente às obtidas pelo método de Euler-Maruyama.

Simulando agora, $L = 25$ trajetórias diferentes do processo $(X_t, t \in [0, 1])$ e as correspondentes aproximações de Euler e de Milstein para a mesma trajetória do processo de Wiener gerada, podemos estimar o erro absoluto médio através do erro empírico

$$\hat{\epsilon} = \frac{1}{L} \sum_{k=1}^L |X_{T,k} - \bar{X}_{T,k}|,$$

onde $X_{T,k}$ e $\bar{X}_{T,k}$ denotam os valores exatos e aproximados, em T , para a k -ésima trajetória, respetivamente.

Assim, pelo método de Euler, vem

δ	2^{-4}	2^{-5}	2^{-6}
$\hat{\epsilon}$	0.3212	0.2981	0.2463

Tabela 3.1: Erros no método de Euler-Maruyama

Por observação da tabela, constatamos que o erro absoluto diminui à medida que diminui o diâmetro δ , confirmando o que observámos nas figuras.

No caso do método de Milstein, temos

δ	2^{-4}	2^{-5}	2^{-6}
$\hat{\epsilon}$	0.1535	0.0770	0.0429

Tabela 3.2: Erros no método de Milstein

Confirmamos, agora, que o erro absoluto é menor em relação ao do método de Euler-Maruyama para qualquer δ considerado e, analogamente ao método de Euler-Maruyama, que este diminui ao diminuir o diâmetro da partição.

△

Exemplo 3.5.2. Seja $X = (X_t, t \in [0, 1])$ o processo de Itô que satisfaz a equação

$$dX_t = \frac{1}{2} X_t dt + \sqrt{1 + X_t^2} dW_t, t \in [0, 1] \quad (3.15)$$

com condição inicial $X_0 = 1$ q.c..

Como vimos no Exemplo 2.3.4 esta equação é redutível e tem solução explícita

$$X_t = \sinh(\arg \sinh(1) + W_t) \text{ q.c., } t \in [0, 1].$$

Comparemos a solução exata com as aproximações obtidas a partir de (3.4) e (3.10), tomando $\delta = 2^{-4}$.

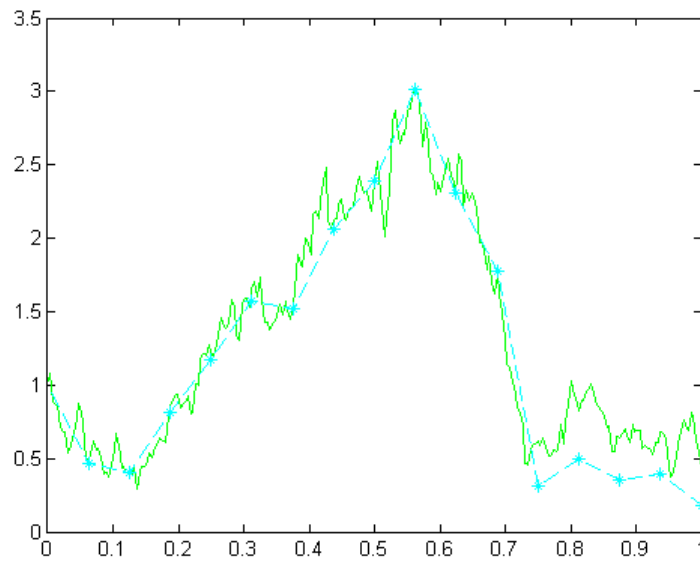


Figura 3.5: Aproximação de Euler e solução exata da equação redutível

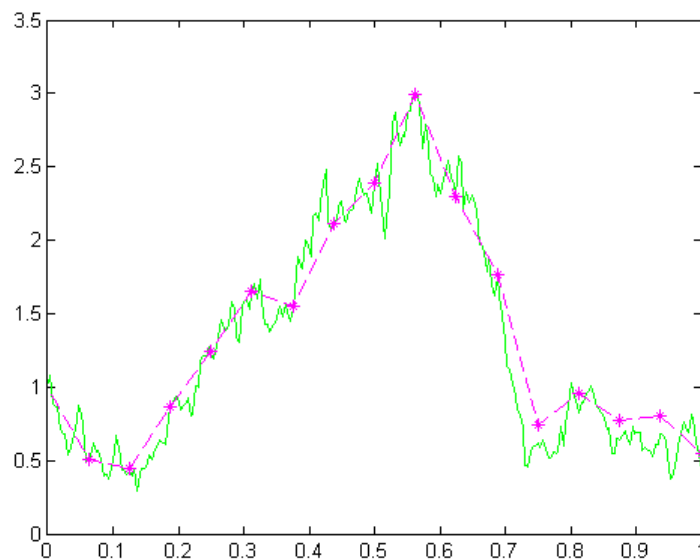


Figura 3.6: Aproximação de Milstein e solução exata da equação redutível

Analogamente ao exemplo anterior, a observação das figuras permite constatar que, as aproximações de Milstein apresentam qualidade superior às de Euler-Maruyama.

△

3.5.1. Aplicação de métodos numéricos na modelação de taxas de juro

A taxa de juro variável pode ser calibrada através do modelo de Vasicek (Vasicek, 1977) que propõe a dinâmica

$$dr_t = k(\theta - r_t)dt + \sigma dW_t, \quad t \in [0, T], \quad (3.16)$$

com condição inicial r_0 , onde k , θ e σ são constantes positivas que representam, respetivamente, a velocidade de reversão à média, a média e o desvio-padrão.

Esta equação é uma variante da equação de Langevin (2.14) que tem como solução o denominado processo de Ornstein-Uhlenbeck.

De acordo com o Exemplo 2.3.1, a solução desta EDE é dada por

$$r_t = \theta + e^{-kt} \left((r_0 - \theta) + \sigma \int_0^t e^{ks} dW_s \right).$$

Por outro lado, podemos calcular o integral $\int_0^t e^{ks} dW_s$, recorrendo à integração por partes

$$\begin{aligned} \int_0^t e^{ks} dW_s(\omega) &= \left[e^{ks} W_s(\omega) \right]_0^t - \int_0^t W_s(\omega) d(e^{ks}) = \\ &= e^{kt} W_t(\omega) - W_0(\omega) - \int_0^t k e^{ks} W_s(\omega) ds \end{aligned} \quad (3.17)$$

embora não consigamos escrever o integral de Riemann envolvido de uma forma mais explícita.

Note-se que, como a função integranda é determinista, $\int_0^t e^{ks} dW_s$ tem distribuição

$$N \left(0, \int_0^t (e^{ks})^2 ds \right) \equiv N \left(0, \frac{1}{2k} (e^{2kt} - 1) \right),$$

pelo que

$$r_t \sim N \left(\theta + e^{-kt}(r_0 - \theta), \frac{\sigma^2}{2k} (1 - e^{-2kt}) \right).$$

Na ausência de flutuações do mercado ($\sigma = 0$), há uma tendência para que a taxa reverta para o valor de referência θ , isto é, converge para θ quando $t \rightarrow +\infty$. Quando há flutuações aleatórias do mercado, esta reversão é perturbada, mantendo-se, no entanto, esta tendência. Tal facto estilizado denomina-se reversão para a média, isto é, o valor esperado da taxa de juro tende, quando $t \rightarrow +\infty$, para o valor θ , sem que a sua variância cresça demasiadamente.

A maior desvantagem apontada ao modelo de Vasicek é a possibilidade de valores negativos para a taxa de juro, r_t . Contudo, a facilidade de manuseamento analítico conseguida com a densidade Gaussiana é dificilmente atingida para outros modelos

estocásticos para a taxa de juro.

Salienta-se que, neste caso, os métodos numéricos são fundamentais, dado que, para obter a trajetória da solução exata necessitaríamos de aproximar o integral de Riemann que figura no segundo membro de (3.17), cuja função integranda, é por sua vez, uma trajetória do processo de Wiener.

Este modelo pode ser calibrado através de dados históricos, no entanto, tal tarefa reveste-se de alguma complexidade e foge ao âmbito deste trabalho, pelo que iremos fixar os valores k , θ e σ .

Consideremos então a equação, onde t está expresso em anos

$$dr_t = k(\theta - r_t)dt + \sigma dW_t, \quad t \in [0, 5],$$

com condição inicial $r_0 = 5.5\%$, assumindo $k = 15\%$, $\theta = 4\%$ e $\sigma = 1\%$.

Para obtermos uma aproximação da trajetória de r_t , recorreremos ao método de Euler-Maruyama, com passo $\delta = 2^{-4}$, obtendo-se

$$r_{\tau_{n+1}} = r_{\tau_n} + k(\theta - r_{\tau_n})(\tau_{n+1} - \tau_n) + \sigma(W_{\tau_{n+1}} - W_{\tau_n}).$$

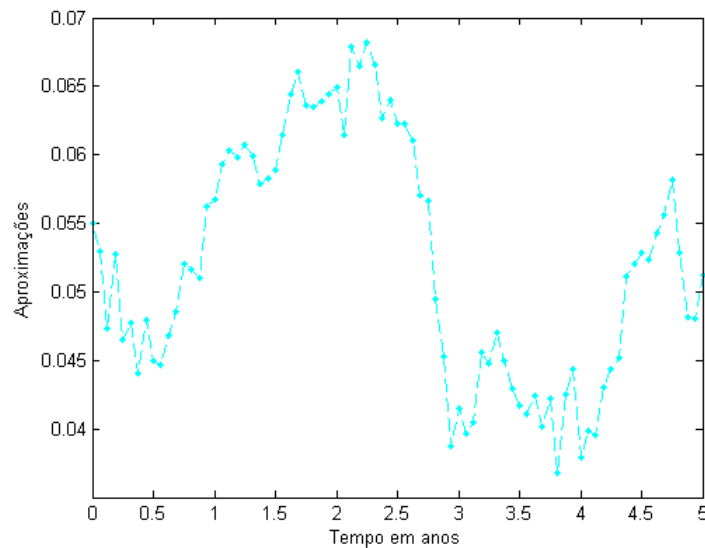


Figura 3.7: Aproximação de Euler-Maruyama do Modelo de Vasicek

As trajetórias de 100 aproximações do processo em causa estão representadas na figura que se segue:

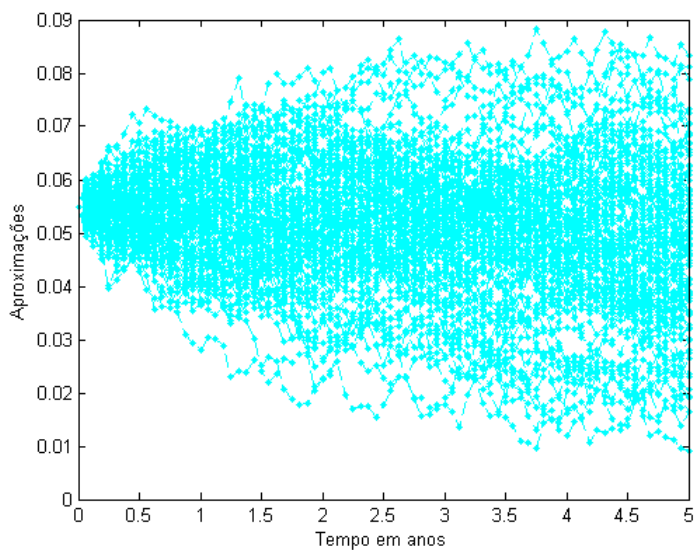


Figura 3.8: *Simulação de 100 aproximações de Euler-Maruyama do Modelo de Vasicek*

A figura reflete a característica estilizada do modelo que tem a ver com o facto deste reverter para a taxa média (4%).

Neste caso particular, os métodos de Euler-Maruyama e de Milstein coincidem dado que a equação do modelo de Vasicek é uma EDE com ruído aditivo.

Capítulo 4

Conclusão

Podemos dizer que existem duas classes gerais de equações diferenciais estocásticas cuja solução explícita é determinada sem grandes dificuldades: as equações lineares e as equações redutíveis. Contudo, no caso da última classe, as condições exigidas aos coeficientes da equação são bastante restritivas, mesmo quando eles não dependem do tempo. A nossa constatação é comprovada pela lista de equações resolúveis apresentadas por [19], onde não encontramos quaisquer exemplos de equações redutíveis com coeficientes dependentes do tempo. Este facto reforça a utilidade da aplicação de métodos numéricos para construir aproximações da solução de uma equação deste tipo.

Do que pudemos observar na literatura, os métodos numéricos mais populares para aproximar numericamente a solução de uma equação diferencial estocástica são baseados na expansão de Itô-Taylor, segundo um procedimento de truncatura semelhante ao que se efetua nas equações diferenciais ordinárias. Dois exemplos simples mas típicos de tal procedimento são os métodos de Euler-Maruyama e de Milstein, abordados nesta dissertação. Com efeito, enquanto que no primeiro se retêm apenas os três primeiros termos da expansão, inclui-se, no segundo, um termo adicional que envolve um integral duplo.

Situando-nos no contexto estocástico, diversificam-se, naturalmente, os critérios de qualidade exigidos às aproximações construídas. Constatámos que os mais usuais na literatura correspondem às noções de convergência forte e fraca, de índole completamente diferente, já que a convergência forte assegura a proximidade entre as trajetórias da solução exata e das aproximações numéricas, e a convergência fraca garante, sobretudo, a proximidade dos seus momentos.

Apesar de nos termos limitado a analisar a qualidade das aproximações obtidas segundo o critério da convergência forte, verificámos, quer a nível teórico, quer a nível prático, que o método de Milstein, com ordem de convergência 1, conduz (exceptuando, evidentemente, o caso das equações com ruído aditivo) a melhores

aproximações do que o método de Euler-Maruyama, cuja ordem de convergência é igual a 0.5. Esta conclusão vai de encontro ao esperado, no sentido de que a inclusão de mais termos na expansão de Itô-Taylor conduziria a melhores aproximações. Existem, de facto, vários métodos de Taylor estocásticos com ordem de convergência superior a 1, que são estudados detalhadamente em [19]. Contudo, também aqui se verifica que, ao aumentarmos o número de termos considerados, somos obrigados a trabalhar com derivadas de ordens superiores dos coeficientes da equação, o que dificulta a implementação dos métodos. Para ultrapassar estas dificuldades, foram propostos (c.f. [19]) métodos do tipo Runge-Kutta, que não recorrem ao cálculo de tais derivadas.

Por último, gostaríamos de salientar que, face à diversidade (acrescida) dos problemas que se colocam no contexto estocástico, apenas foi “levantada a ponta do icebergue”. Para além de questões relativas à consistência e à estabilidade dos métodos, o problema da convergência fraca, embora menos exigente do que o analisado, reveste-se de grande importância, uma vez que, em muitas situações práticas, nomeadamente em Finanças, se pretende apenas aproximar os momentos das variáveis do processo solução.

Bibliografia

- [1] E. Allen. **Modeling with Itô Stochastic Differential Equations**. Springer, 2007.
- [2] L. Arnold. **Stochastic differential equations: Theory and Applications**. John Wiley and Sons, New York, 1974.
- [3] N. E. Boyce. **Approximate solution of random ordinary differential equations**, *Advances in Applied Probability*, 10, pp 172-184, 1978.
- [4] K. Burrage, P. M. Burrage, and T. Tian. **Numerical methods for strong solutions of stochastics differential equations: an overview**. The Royal Society, 2003.
- [5] V. Capasso, and D. Bakstein. **An Introduction to Continuous-Time Stochastic Processes**. Theory, Models, and Applications to Finance, Biology, and Medicine. Birkhäuser, Boston, 2005.
- [6] L. Cobb. **Stochastic Differential Equations for the Social Sciences**, 1998. Revised and extended from Chapter 2 of Cobb and Thrall (eds), **Mathematical Frontiers of the Social and Policy Sciences**. Westview press, 1981.
- [7] R. Cont, and P. Tankov. **Financial Modeling with Jump Processes**. Chapman and Hall/CRC, 2004.
- [8] T. C. Gard. **Introduction to Stochastic Differential Equations**. Marcel Dekker, 1988.
- [9] P. Glasserman. **Monte Carlo Methods in Financial Engineering**. Springer-Verlag, New York, 2004.
- [10] M. Grigoriu. **Stochastic Calculus**. Applications in Science and Engineering. Birkhäuser, Boston, 2002.

- [11] D. Gusak, A. Kukush, A. Kulik, Y. Mishura, and A. Pilipenko. **Theory of Stochastic Processes**. Springer-Verlag, New York, 2010.
- [12] D. Henderson, and P. Plaschko. **Stochastic Differential Equations in Science and Engineering**. World Scientific Publishing, 2006.
- [13] D. J. Higham. **An Algorithmic Introduction to Numerical Simulation of Stochastic Differential Equations**. SIAM Rev., 43: 525-546, 2001.
- [14] M. Higham, A. M. Stuart, and X. Mao. **Strong convergence of Euler-type methods in nonlinear SDE's**. SIAM J. Num. Anal., 40, 1041-1063, 2002.
- [15] S.M. Iacus. **Simulation and Inference for Stochastic Differential Equations with R Examples**. Springer Science+Business Media, LLC 2008.
- [16] R. Jarrow, and P. Protter. **A short history of stochastic integration and mathematical finance: the early years, 1880-1970**. IMS Lecture Notes Monograph Vol.45 1-17, 2004.
- [17] I. Karatzas, and S. E. Shreve. **Brownian Motion and Stochastic Calculus** 2nd Edition. Springer-Verlag, 1991.
- [18] F. Klebaner. **Introduction to Stochastic Calculus and Applications** 2nd Edition. Imperial College Press, London, 2005.
- [19] P. E. Kloeden, and E. Platen. **Numerical Solution of Stochastic Differential Equations**. Springer-Verlag, Berlin, 1992.
- [20] R. Korn, E. Korn, and G. Kroisandt. **Monte Carlo Methods and Models in Finance**. Chapman and Hall, 2010.
- [21] R. Kress. **Numerical Analysis**. Springer-Verlag, New York, 1998.
- [22] H-H. Kuo. **Introduction to Stochastic Integration**. Springer-Verlag, New York, 2006.
- [23] H. J. Kushner, and P. Dupuis. **Numerical methods for stochastic control problems in continuous time**. Springer-Verlag, Berlin, 1072.
- [24] The *MathWorks*TM. *MATLAB*[®].
- [25] T. Mikosch. **Elementary Stochastic Calculus with Finance in View**. World Scientific Publishing, 1998.

- [26] G.N. Milstein. **Approximate integration of stochastic differential equations**, Theory Probab. Appl. 19: 557-562, 1974.
- [27] G. N. Milstein. **Numerical Integration of Stochastic Differential Equations**. Kluwer, Dordrecht, 1995.
- [28] E. Platen, and D. Heath. **A Benchmark Approach to Quantitative Finance**. Springer-Verlag, Berlin, 2006.
- [29] M. Sanz-Solé. **An Introduction to Stochastic Calculus**. 2010. Acedido em Janeiro de 2012 em <http://www.mat.ub.edu/~sanz/cursos/lecturenotes-sc2008.pdf>.
- [30] T. Sauer. **Numerical Solution of Stochastic Differential Equations in Finance**. Handbook of Computational Finance, pp. 529-550. EDS. J.-C. Duan, W. Härdle, J. Gentle. Springer, Berlin-Heidelberg, 2012.
- [31] H. Schurz. **A Brief Introduction to Numerical Analysis of Stochastic Differential Equations without Tears**. Acedido em Janeiro de 2012 em www.math.siu.edu/preprints/schurz/siuppr1.pdf.
- [32] Z. Schuss. **Theory and Applications of Stochastic Differential Equations**. John Wiley and Sons, 1980.
- [33] X. Zhang. **Euler-Maruyama approximations for SDE's with nonlinear-lipschitz coefficients and applications**. Jour. Math. Anal., 316, 447-458, 2006.

Apêndice A

Algoritmos utilizados na implementação dos métodos numéricos

A.1. Algoritmos dos exemplos (3.5.1) e (3.5.2)

Método de Euler-Maruyama

EDE $dX = \mu * X dt + \sigma * X dW$

```
randn('state',100)
delta=2^(-8);
N=1/delta;
X_0=1.0;
mu=1.5;
sigma=1.0;
t_0=0;
T=1;
dW=sqrt(delta)*randn(1,N);
W=cumsum(dW);
X=X_0*exp((mu-(1/2)*sigma^2)*([delta:delta:T])+sigma*W);
plot([0:delta:T],[X_0,X],'g-');
hold;
R=2^4;
Delta=R*delta;
L=N/R;
Y=zeros(1,L);
Ytemp=X_0;
for j=1:L
    Winc=sum(dW(R*(j-1)+1:R*j));
```

Apêndice A Algoritmos utilizados na implementação dos métodos numéricos

```
Ytemp=Ytemp+mu*Ytemp*Delta+sigma*Ytemp*Winc;
Y(j)=Ytemp;
end
plot([0:Delta:T],[X_0,Y],'c--*')

Método de Milstein
EDE dX=mu*Xdt+sigma*XdW

randn('state',100)
delta=2^(-8);
N=1/delta;
X_0=1.0;
mu=1.5;
sigma=1.0;
t_0=0;
T=1;
dW=sqrt(delta)*randn(1,N);
W=cumsum(dW);
X=X_0*exp((mu-(1/2)*sigma^2)*([delta:delta:T])+sigma*W);
plot([0:delta:T],[X_0,X],'g-');
hold;
R=2^4;
Delta=R*delta;
L=N/R;
Y=zeros(1,L);
Ytemp=X_0;
for j=1:L
    Winc=sum(dW(R*(j-1)+1:R*j));
    Ytemp=Ytemp+mu*Ytemp*Delta+sigma*Ytemp*Winc+(1/2)*sigma^2*Ytemp*(Winc^2-Delta);
    Y(j)=Ytemp;
end
plot([0:Delta:T],[X_0,Y],'m--*')
```

A.1 Algoritmos dos exemplos (3.5.1) e (3.5.2)

Simulação de 25 trajetórias do processo de Itô, X , para determinar o erro absoluto do método de Euler-Maruyama com $\delta=2^{-4}$, 2^{-5} e 2^{-6} .

EDE $dX=\mu*Xdt+\sigma*XdW$

```
randn('state',100)
delta=2^(-9);
N=1/delta;
X_0=1.0;
mu=1.5;
sigma=1.0;
t_0=0;
T=1;
M=25;
Xerr=zeros(M,3);
for m=1:M
    dW=sqrt(delta)*randn(1,N);
    W=cumsum(dW);
    X=X_0*exp((mu-(1/2)*sigma^2)+sigma*W(end));
    for p=1:3
        R=2^(6-p);
        Delta=R*delta;
        L=N/R;
        Y=X_0;
        for j=1:L
            Winc=sum(dW(R*(j-1)+1:R*j));
            Y=Y+mu*Y*Delta+sigma*Y*Winc;
        end
        Xerr(m,p)=abs(Y-X);
    end
end
Erro=mean(Xerr);
```

Apêndice A Algoritmos utilizados na implementação dos métodos numéricos

Simulação de 25 trajetórias do processo de Itô, X , para determinar o erro absoluto do método de Milstein com $\delta=2^{-4}$, 2^{-5} e 2^{-6} .

EDE $dX=\mu*Xdt+\sigma*XdW$

```
randn('state',100)
delta=2^(-9);
N=1/delta;
X_0=1.0;
mu=1.5;
sigma=1.0;
t_0=0;
T=1;
M=25;
Xerr=zeros(M,3);
for m=1:M
    dW=sqrt(delta)*randn(1,N);
    W=cumsum(dW);
    X=X_0*exp((mu-(1/2)*sigma^2)+sigma*W(end));
    for p=1:3
        R=2^(6-p);
        Delta=R*delta;
        L=N/R;
        Y=X_0;
        for j=1:L
            Winc=sum(dW(R*(j-1)+1:R*j));
            Y=Y+mu*Y*Delta+sigma*Y*Winc+(1/2)*sigma^2*Y*(Winc^2-Delta);
        end
        Xerr(m,p)=abs(Y-X);
    end
end
Erro=mean(Xerr);
```

Método de Euler-Maruyama

EDE $dX=1/2Xdt+\text{sqrt}(1+X^2)dW$

```

randn('state',100)
delta=2^(-8);
N=1/delta;
X_0=1.0;
t_0=0;
T=1;
dW=sqrt(delta)*randn(1,N);
W=cumsum(dW);
X=sinh(log(1+sqrt(2))+W);
plot([0:delta:T],[X_0,X],'g-');
hold;
R=2^4;
Delta=R*delta;
L=N/R;
Y=zeros(1,L);
Ytemp=X_0;
for j=1:L
    Winc=sum(dW(R*(j-1)+1:R*j));
    Ytemp=Ytemp+(1/2)*Ytemp*Delta+sqrt(1+Ytemp^2)*Winc;
    Y(j)=Ytemp;
end
plot([0:Delta:T],[X_0,Y],'c--*')

```

Método de Milstein

EDE $dX=1/2Xdt+\text{sqrt}(1+X^2)dW$

```

randn('state',100)
delta=2^(-8);
N=1/delta;
X_0=1.0;
t_0=0;

```

```
T=1;
dW=sqrt(delta)*randn(1,N);
W=cumsum(dW);
X=sinh(log(1+sqrt(2))+W);
plot([0:delta:T],[X_0,X],'g-');
hold;
R=2^4;
Delta=R*delta;
L=N/R;
Y=zeros(1,L);
Ytemp=X_0;
for j=1:L
    Winc=sum(dW(R*(j-1)+1:R*j));
    Ytemp=Ytemp+(1/2)*Ytemp*Delta+sqrt(1+Ytemp^2)*Winc+
    +(1/2)*sqrt(1+Ytemp^2)*Ytemp*(1+Ytemp^2)^(-0.5)*(Winc^2-Delta);
    Y(j)=Ytemp;
end
plot([0:Delta:T],[X_0,Y],'m--*')
```

A.2. Algoritmos para as aproximações do Modelo de Vasicek

Modelo de Vasicek

EDE $dX=k*(teta-X) dt + sigma dW$

```
randn('state',100)
delta=2^(-6);
N=5/delta;
X_0=0.055;
k=0.15;
teta=0.04;
sigma=0.01;
t_0=0;
T=5;
dW=sqrt(delta)*randn(1,N);
R=2^2;
```

A.2 Algoritmos para as aproximações do Modelo de Vasicek

```
Delta=R*delta;
L=N/R;
Y=zeros(1,L);
Ytemp=X_0;
for j=1:L
    Winc=sum(dW(R*(j-1)+1:R*j));
    Ytemp=Ytemp+k*(teta-Ytemp)*Delta+sigma*Winc;
    Y(j)=Ytemp;
end
plot([0:Delta:T],[X_0,Y],'c--.');
xlabel('Tempo em anos')
ylabel('Aproximações')
```

100 simulações do Modelo de Vasicek

EDE $dX=k*(teta-X) dt + sigma dW$

```
for i=1:100
    delta=2^(-6);
    N=5/delta;
    X_0=0.055;
    k=0.15;
    teta=0.04;
    sigma=0.01;
    t_0=0;
    T=5;
    dW=sqrt(delta)*randn(1,N);
    R=2^2;
    Delta=R*delta;
    L=N/R;
    Y=zeros(1,L);
    Ytemp=X_0;
    for j=1:L
        Winc=sum(dW(R*(j-1)+1:R*j));
        Ytemp=Ytemp+k*(teta-Ytemp)*Delta+sigma*Winc;
```

Apêndice A Algoritmos utilizados na implementação dos métodos numéricos

```
    Y(j)=Ytemp;
end
plot([0:Delta:T],[X_0,Y],'c--.'); hold on
xlabel('Tempo em anos')
ylabel('Aproximações')
end
```