NUMERICAL SOLUTION OF STOCHASTIC DIFFERENTIAL EQUATIONS

A thesis submitted to university of edinburgh and heriot-watt university
for the degree of Master of Science

2005

By
Seokhyun Han

Contents

Abstract

The aim of this paper is to provide a systematic treatment of time discretized numerical schemes for stochastic differential equations and to demonstrate that it is possible to improve the accuracy of the weak error of Euler's approximation by implementing Richardson's idea.

Acknowledgments

I would like to express my deep gratitude to many teaching staff who have taught and helped me in course of Graduate Certificate in Mathematical Science and MSc in Financial Mathematics. I am greatly indebted to Professor Istvan Gyongy, MSc project supervisor, who have given many interesting ideas and helpful advice. I also owe many thanks to my colleague, Tomas Lagland, for his assistance.

Chapter 1

Introduction

This paper aims to provide a systematic framework for an understanding of the basic concepts and tools needed for the development and implementation of numerical methods for stochastic differential equations, primarily time discretization methods for initial value problems of stochastic differential equations with Ito diffusions as their solutions. Recent years have witnessed that the most efficient and widely applicable approach to solving SDEs seems to be the simulation of sample paths of time discrete approximations on digital computers. This is based on a finite discretization of the time interval $[0, T]$ under consideration and generates approximate values of the sample paths step by step at the discretization times.

Ito process $X = \{X_t, t \geq 0\}$ has the form

$$X_t = X_0 + \int_0^t a(X_s)ds + \int_0^t b(X_s)dW_s$$

for $t \geq 0$ with an initial value X_0 which may be random, the drift $a(X_t)$, and the diffusion $b(X_t)$. This integral equation is often written in the differential form

$$dX_t = a(X_t)dt + b(X_t)dW_t$$

which is called an Ito stochastic differential equation. Unfortunately explicitly solvable SDEs are rare in practical applications. However, there are increasing number of numerical methods for the solution of SDEs. In connection with this issue, my intent here is to review the systematic development of efficient numerical methods for SDEs by using [2] [3]. Obviously such methods should be implementable on digital computers. We shall survey various time discrete numerical methods which are appropriate for the simulation of sample paths or functionals of Ito processes in order to estimate various statistical features of the desired solution. In dealing with this issue, we should also speak of how to judge the quality of a discrete time approximation, i.e. a criterion needs to be specified. Such a criterion should reflect the main goal of practical simulations. There are two basic types of tasks connected with the simulation of solutions of stochastic differential equations. The first occurs in situation where a good pathwise approximation is required, for instance in direct simulations, filtering or testing statistical estimators. In the second interest focuses on approximating expectations of functionals of the Ito process, such as its probability distribution and its moments. This is relevant in many practical problems because such functionals can not often be determined analytically. Furthermore, the ultimate goal of this paper will be to investigate a general idea of improving accuracy of numerical methods which is due to Richardson. He suggested to consider approximations obtained by linear combinations of approximation corresponding to different step sizes.

Before we undertake this task, it is worth taking a quick look at the main parts. Part I on Review will survey Ito formula, vector SDEs, stochastic Taylor expansions which provide a universally applicable tool for SDEs which is analogous to the deterministic Taylor formula in ordinary calculus, and Euler scheme to highlight the basic issues, types of problems, and objectives that arise when SDEs are solved numerically. In particular, we distinguish between strong and weak approximations, depending on whether good pathwise or good probability distributional approximations are sought. The object of Part II is to revisit the themes originally explored in D.Talay and L.Tubaro's paper "Expansion of the global error for numerical schemes solving stochastic differential equation" [1] in order to show how to improve the weak error of the Euler scheme.

Chapter 2

Part I - Review

In the following I will give a brief general survey of numerical solution of stochastic differential equations using P.E.Kloeden and E.Platen's two books [2] [3].

2.1 Ito formula

Let a and b be two functions from $[0,T] \times \Omega$ into $\mathbb{R}$ with measurability and integrability properties that the ordinary and stochastic integrals appearing in the following formula make sense. By a stochastic differential we mean an expression

$$dX_t(\omega) = a(t,\omega)dt + b(t,\omega)dW_t(\omega) \tag{2.1}$$

which is just a symbolical way of writing

$$X_t(\omega) - X_s(\omega) = \int_s^t a(u,\omega)du + \int_s^t b(u,\omega)dW_u(\omega)$$

w.p.1, for any $0 \leq s \leq t \leq T$. The first integral in the above is an ordinary (Riemann or Lebesgue) integral for each $\omega \in \Omega$ and the second is an Ito integral. Let $f : [0,T] \times \mathbb{R} \to \mathbb{R}$ have continuous partial derivatives $\frac{\partial f}{\partial t}$, $\frac{\partial f}{\partial x}$ and $\frac{\partial^2 f}{\partial x^2}$. A scalar transformation by f of the stochastic differential (2.1) results after some non trivial analysis in the Ito formula,

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

and in other way,

$$f(t, X_t) - f(s, X_s) = \int_s^t \left\{ \frac{\partial f}{\partial t} + a\frac{\partial f}{\partial x} + \frac{1}{2}b^2\frac{\partial^2 f}{\partial x^2} \right\} du + \int_s^t b\frac{\partial f}{\partial x} dW_u$$

w.p.1, for any $0 \le s \le t \le T$, where the integrands are all evaluated at (u, X_u).

2.2 Uniqueness and Existence of SDEs

When we have no explicit solution of an Ito stochastic equation

$$X_t = X_{t_0} + \int_{t_0}^t a(X_s)ds + \int_{t_0}^t b(X_s)dW_s \tag{2.2}$$

we need somehow to ensure the existence and uniqueness of a process $X = \{X_t, t \in [t_0, T]\}$ which satisfies it. We call $a(X_s)$ the drift coefficient and $b(X_s)$ the diffusion coefficient.We shall say that solutions of (2.2) are pathwise unique if any two such solutions $X = \{X_t, t \in [t_0, T]\}$ and $\tilde{X} = \{\tilde{X}_t, t \in [t_0, T]\}$ have, almost surely, the same sample paths on $[t_0, T]$, that is if

$$P\left(\sup_{t_0 \le t \le T} \mid X_t - \tilde{X}_t \mid > 0\right) = 0$$

In this case we call X a unique strong solution of (2.2). From a basic existence and uniqueness theorem, it follows that equation (2.2) has a unique strong solution $X = \{X_t, t \in [t_0, T]\}$ on $[t_0, T]$ with

$$\sup_{t_0 \le t \le T} E(X_t^2) < \infty$$

provided X_{t_0} is independent of $W = \{W_t, t \in [t_0, T]\}$ with $E(X_{t_0})^2 < \infty$ and the coefficients a, b satisfy Lipschitz conditions. We often call X an Ito diffusion process. Sometimes equation (2.2) may have solutions which are unique in the weaker sense that only their probability laws coincide, but not necessarily their sample paths. We shall say then that we have a unique weak solution. Existence and uniqueness of weak solutions follow if the coefficients a and b are bounded and continuous and b is nondegenerate $|b| \ge \varepsilon > 0$. Of course, any unique strong solution is always a unique weak solution.

2.3 Vector SDEs

We shall interpret a vector as a column vector and its transpose as a row vector and consider an m-dimensional Wiener process $W = \{W_t, t \geq 0\}$ with components $W_t^1, W_t^2, \ldots, W_t^m$, which are independent independent scalar Wiener processes. Then, we take a d-dimensional vector valued function $a : [t_0, T] \times \mathbb{R}^d \to \mathbb{R}^d$, the drift coefficient, and a $d \times m$-matrix valued function $b : [t_0, T] \times \mathbb{R}^d \to \mathbb{R}^{d \times m}$, the diffusion coefficient, $t_0 \in [0, T]$, to form a d-dimensional vector stochastic differential equation

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t \tag{2.3}$$

We interpret this as a stochastic integral equation

$$X_t = X_{t_0} + \int_{t_0}^{t} a(s, X_s)ds + \int_{t_0}^{t} b(s, X_s)dW_s \tag{2.4}$$

with initial value $X_{t_0} \in \mathbb{R}^d$, where the Lebesgue and Ito integrals determined component by component, with the ith component of (2.4) being

$$X_t^i = X_{t_0}^i + \int_{t_0}^{t} a^i(s, X_s)ds + \sum_{j=1}^{m} \int_{t_0}^{t} b^{i,j}(s, X_s)dW_s^j$$

If the drift and diffusion coefficients do not depend on the time variable, that is if $a(t, x) \equiv a(x)$ and $b(t, x) \equiv b(x)$, then we say that stochastic equation is autonomous. We can always write a nonautonomous equation as a vector autonomous equation of one dimension more by setting in the first component the drift coefficient equal to 1 and the diffusion coefficient as 0 to obtain as the first component of X_t the time variable $X_t^1 = t$.

There is a vector version of the Ito formula. For a sufficiently smooth transformation $f = [t_0, T] \times \mathbb{R}^d \to \mathbb{R}^k$ of the solution $X = \{X_t, t_0 \leq t \leq T\}$ of (2.3) we obtain a k-dimensional process $Y = \{Y_t = f(t, X_t), t_0 \leq t \leq T\}$ with the vector stochastic differential in component form

$$dY_t^p = \left(\frac{\partial f^p}{\partial t} + \sum_{i=1}^{d} a^i \frac{\partial f^p}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{d} \sum_{l=1}^{m} b^{i,l} b^{j,l} \frac{\partial^2 f^p}{\partial x_i \partial x_j} \right) dt + \sum_{l=1}^{m} \sum_{i=1}^{d} b^{i,l} \frac{\partial f^p}{\partial x_i} dW_t^l$$

for $p = 1, 2, \ldots, k$, where the terms are all evaluated at (t, X_t). We can sometimes use this formula to determine the solutions of certain vector stochastic differential equations in terms of known solutions of the other equations, for example linear equations.

2.4 Ito-Taylor expansion

To begin we consider the equation $X = \{X_t, t \in [t_0, T]\}$ of 1-dimensional ordinary differential equation

$$\frac{d}{dt} X_t = a(X_t)$$

with initial value X_{t_0}, for $t \in [t_0, T]$ where $0 \leq t_0 < T$, which we can write in the equivalent integral equation form as

$$X_t = X_{t_0} + \int_{t_0}^{t} a(X_s) ds \tag{2.5}$$

To justify the following constructions we require that the function a satisfies appropriate properties, for instance to be sufficiently smooth with a linear growth bound. let $f : \mathbb{R} \to \mathbb{R}$ be a continuously differentiable function. Then by the chain rule, we have

$$\frac{d}{dt} f(X_t) = a(X_t) f'(X_t) \tag{2.6}$$

which using the operator

$$Lf = af'$$

where $'$ denotes differentiation with respect to x, we can express (2.6) as the integral relation

$$f(X_t) = f(X_{t_0}) + \int_{t_0}^{t} Lf(X_s) ds \tag{2.7}$$

for all $t \in [t_0, T]$. When $f(x) \equiv x$ we have $Lf = a$, $L^2 f = La$,…and (2.7) reduces to

$$X_t = X_{t_0} + \int_{t_0}^{t} a(X_s) ds \tag{2.8}$$

that is, to equation (2.5). If we apply the relation (2.7) to the function $f = a$ in the integral in (2.8), we obtain

$$X_t = X_{t_0} + \int_{t_0}^{t} (a(X_{t_0}) + \int_{t_0}^{s} La(X_z)dz)ds \tag{2.9}$$

$$= X_{t_0} + a(X_{t_0}) + a(X_{t_0}) \int_{t_0}^{t} ds + \int_{t_0}^{t} \int_{t_0}^{s} La(X_z)dzds$$

which is the simplest nontrivial Taylor expansion for X_t. We can apply (2.7) again to the function $f = La$ in the double integral of (2.9) to derive

$$X_t = X_{t_0} + a(X_{t_0}) \int_{t_0}^{t} ds + La(X_{t_0}) \int_{t_0}^{t} \int_{t_0}^{s} dzds + R_3$$

with remainder

$$R_3 = \int_{t_0}^{t} \int_{t_0}^{s} \int_{t_0}^{z} L^2 a(X_u)dudzds$$

for $t \in [t_0, T]$. For a general $r + 1$ times continuously differentiable function $f : \mathbb{R} \to \mathbb{R}$, this method gives the classical Taylor formula in integral form

$$f(X_t) = f(X_0) + \sum_{i=1}^{r} \frac{(t - t_0)^l}{l!} L^l f(X_{t_0}) + \int_{t_0}^{t} \cdots \int_{t_0}^{s_2} L^{r+1} f(X_{s_1})ds_1 \tag{2.10}$$

for $t \in [t_0, T]$ and $r = 1, 2, 3, \dots$ since

$$\int_{t_0}^{t} \int_{t_0}^{s_1} \cdots \int_{t_0}^{s_{l-1}} ds_1 \cdots ds_l = \frac{1}{l!}(t - t_0)^l$$

for $l = 1, 2, \dots$. The Taylor formula (2.10) has proven to be a very useful tool in both theoretical and practical investigations, particularly in numerical analysis. It allows the approximation of a sufficiently smooth function in a neighborhood of a given point to any desired order of accuracy. The expansion depends on the values of the function and some of its higher derivatives at the expansion point, weighted by corresponding multiple time integrals. In addition, there is a remainder term which contains the next following multiple time integral, but with a time dependent integrand.

A stochastic counterpart of the deterministic Taylor formula for the expansion of smooth functions of an Ito process about a given value has many potential applications in stochastic analysis, for instance in the derivation of numerical methods for stochastic differential equations. There are several possibilities for such a stochastic Taylor formula. One is based on the iterated application of the Ito formula, which we shall call the Ito-Taylor expansion. We shall indicate it here for the solution X_t of the 1-dimensional Ito stochastic differential equation in integral form

$$X_t = X_{t_0} + \int_{t_0}^{t} a(X_s)ds + \int_{t_0}^{t} b(X_s)dW_s$$

for $t \in [t_0, T]$, where the second integral is an Ito stochastic integral and the coefficients a and b are sufficiently smooth real valued functions satisfying a linear growth bound. For any twice continuously differentiable function $f : \mathbb{R} \to \mathbb{R}$, the Ito formula then gives

$$\begin{aligned} f(X_t) &= f(X_0) + \int_{t_0}^{t} \left(a(X_s)f'(X_s) + \frac{1}{2}b^2(X_s)f''(X_s) \right) ds \\ &\quad + \int_{t_0}^{t} b(X_s)f'(X_s)dW_s \\ &= f(X_{t_0}) + \int_{t_0}^{t} L^0 f(X_s)ds + \int_{t_0}^{t} L^1 f(X_s)dW_s \end{aligned} \tag{2.11}$$

for $t \in [t_0, T]$. Here we have introduced the operators

$$L^0 f = af' + \frac{1}{2}b^2 f''$$

and

$$L^1 f = bf'$$

Obviously for $f(x) \equiv x$ we have $L^0 f = a$ and $L^1 f = b$, in which case (2.11) reduces to the original Ito equation for X_t, that is to

$$X_t = X_{t_0} + \int_{t_0}^{t} a(X_s)ds + \int_{t_0}^{t} b(X_s)dW_s \tag{2.12}$$

In analogy with the deterministic expansion above, if we apply the Ito formula (2.11) to the functions $f = a$ and $f = b$ in (2.12) we obtain

$$X_t = X_{t_0} + \int_{t_0}^{t} \left(a(X_{t_0}) + \int_{t_0}^{s} L^0 a(X_z)dz + \int_{t_0}^{s} L^1 a(X_z)dW_z \right) ds \tag{2.13}$$

$$
\begin{aligned}
& + \int_{t_0}^{t} \left(b(X_{t_0}) + \int_{t_0}^{s} L^0 b(X_z) dz + \int_{t_0}^{s} L^1 b(X_z) dW_z \right) dW_s \\
= \; & X_{t_0} + a(X_{t_0}) \int_{t_0}^{t} ds + b(X_{t_0}) \int_{t_0}^{t} dW_s + R
\end{aligned}
$$

with remainder

$$
\begin{aligned}
R \;=\; & \int_{t_0}^{t} \int_{t_0}^{s} L^0 a(X_z) dz ds + \int_{t_0}^{t} \int_{t_0}^{s} L^1 a(X_z) dW_z ds \\
& + \int_{t_0}^{t} \int_{t_0}^{s} L^0 b(X_z) dz dW_s + \int_{t_0}^{t} \int_{t_0}^{t} L^1 b(X_z) dW_z dW_s
\end{aligned}
$$

This is the simplest nontrivial Ito-Taylor expansion of X_t. It involves integrals with respect to both the time variable and the Wiener process, with multiple integrals with respect to both in the remainder. We can apply repeat the above procedure, for instance by applying the Ito formula (2.11) to $f = L^1 b$ in (2.13), in which case we get

$$
X_t = X_{t_0} + a(X_{t_0}) \int_{t_0}^{t} ds + b(X_{t_0}) \int_{t_0}^{t} dW_s + L^1 b(X_{t_0}) \int_{t_0}^{t} \int_{t_0}^{s} dW_z dW_s + \bar{R}
$$

with remainder

$$
\begin{aligned}
\bar{R} \;=\; & \int_{t_0}^{t} \int_{t_0}^{s} L^0 a(X_Z) dz ds + \int_{t_0}^{t} \int_{t_0}^{s} L^1 a(X_z) dW_z ds \\
& + \int_{t_0}^{t} \int_{t_0}^{s} L^0 b(X_Z) dz dW_s + \int_{t_0}^{t} \int_{t_0}^{s} \int_{t_0}^{z} L^0 L^1 b(X_u) du dW_z dW_s \\
& + \int_{t_0}^{t} \int_{t_0}^{s} \int_{t_0}^{z} L^1 L^1 b(X_u) dW_u dW_z dW_s
\end{aligned}
$$

It is possible to express the Ito-Taylor expansion for a general function f and arbitrarily many expansion terms in a succinct way. Nevertheless, its main properties are already apparent in the preceding example, with the multiple Ito integrals

$$
\int_{t_0}^{t} ds, \qquad \int_{t_0}^{t} dW_s, \qquad \int_{t_0}^{t} \int_{t_0}^{s} dW_z dW_s
$$

multiplied by certain constants and a remainder term involving the next following multiple Ito integrals, but with nonconstant integrands. The Ito-Taylor expansion can thus be considered as a generalization of both the Ito formula and the deterministic Taylor formula.

2.5 Euler Approximation

One of the simplest time discrete approximations of an Ito process is the Euler approximation. We shall consider an Ito process $X = \{X_t, t_0 \leq t \leq T\}$ satisfying the scalar stochastic differential equation

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t$$

on $t_0 \leq t \leq T$ with the initial value $X_{t_0} = X_0$. For a given discretization $t_0 = \tau_0 < \tau_1 < \cdots < \tau_n = T$ of the time interval $[t_0, T]$, an Euler approximation is a continuous time stochastic process $Y = \{Y(t), t_0 \leq t \leq T\}$ satisfying the iterative scheme

$$Y_{n+1} = Y_n + a(\tau_n, Y_n)(\tau_{n+1} - \tau_n) + b(\tau_n, Y_n)(W_{\tau_{n+1}} - W_{\tau_n})$$

for $n = 0, 1, 2, \ldots, N-1$ with initial value $Y_0 = X_0$, where we have written $Y_n = Y(\tau_n)$ for the value of the approximation at the discretization time τ_n. We shall consider equidistant discretization times

$$\tau_n = t_0 + n\Delta$$

with stepsize $\Delta = (T - t_0)/N$ for some integer N. Now, we need to generate the random increments

$$\Delta W_n = W_{\tau_{n+1}} - W_{\tau_n}$$

for $n = 0, 1, 2, \ldots, N-1$, of the Wiener process $W = \{W_t, t \geq 0\}$. We notice that these increments are independent Gaussian random variables with mean $E(\Delta W_n) = 0$ and variance $E((\Delta W_n)^2) = \Delta$. We shall apply the Euler scheme

$$Y_{n+1} = Y_n + a\Delta + b\Delta W_n$$

for $n = 0, 1, \ldots N-1$ to approximate specific stochastic differential equation.

2.6 Pathwise Approximation and Strong Convergence

We shall introduce the absolute error criterion which is appropriate for a situation where a pathwise approximation is required. The absolute error criterion is just the expectation of the absolute value of the difference between the approximation and the Ito process at the time T, that is

$$\epsilon = (E|X_T - Y(T)|^q)^{\frac{1}{q}}$$

for some $q \geq 1$, and gives a measure of the pathwise closeness at the end of the time interval $[0, T]$. We shall say that a discrete time approximation Y with maximum time step size δ converges strongly to X at time T if

$$\lim_{\delta \downarrow 0} E(|X_T - Y(T)|) = 0$$

In order to assess and compare different discrete time approximations, we need to know their rates of strong convergence. We shall say that a discrete time approximation Y converges strongly with order $\gamma > 0$ at time T if there exists a positive constant C, which does not depend on δ and a $\delta_0 > 0$ such that

$$\epsilon(\delta) = E(|X_T - Y(T)|) \leq C\delta^\gamma$$

for each $\delta \in (0, \delta_0)$. We shall investigate the strong convergence of a number of different discrete time approximations experimentally. We shall see in particular that the Euler approximation has strong order of convergence $\gamma = 0.5$.

2.7 Approximation of Moments and Weak Convergence

To introduce this weak type of convergence, we shall carry out some computer experiments to investigate the mean error

$$\mu = EY(T) - EX_T$$

for the linear stochastic differential equation

$$dX_t = aX_t dt + bX_t dW_t$$

and its Euler approximation

$$Y_{n+1} = Y_n + aY_n\Delta_n + bY_n\Delta W_n$$

for $n = 0, 1, 2, \ldots, N-1$, where $\Delta W_n = \tau_{n+1} - \tau_n$ denotes the step size and $\Delta W_n = W_{\tau_{n+1}} - W_{\tau_n}$ the increment of the Wiener process. We see that this criterion differs in its properties from the strong convergence criterion. To some extent, the above mean error is special and not appropriate for applications where the approximation of some higher moment

$$E(|X_T|^q)$$

with $q = 2, 3, \ldots$ or of some functional

$$E(g(X_T))$$

is of interest. These do not require a good pathwise approximation of the Ito process, but only an approximation of the probability distribution of X_T. We shall say that a general discrete time approximation Y with maximum time step size δ converges weakly to X at time T as $\delta \downarrow 0$ with respect to a class $\mathcal{C}$ of test functions $g : \mathbb{R}^d \to \mathbb{R}$ if we have

$$\lim_{\delta \downarrow 0} |E(g(X_T)) - E(g(Y(T)))| = 0$$

for all $g \in \mathcal{C}$. If $\mathcal{C}$ contains all polynomials this definition implies the convergence of all moments, so investigations involving it will require the existence of all moments. We shall say that a time discrete approximation Y converges weakly with order $\beta > 0$ to X at time T as $\delta \downarrow 0$ if for each polynomial g there exists a positive constant C, which does not depend on δ and a finite $\delta_0 > 0$ such that

$$|E(g(X_T)) - E(g(Y(T))))| \leq C\delta^\beta$$

for each $\delta \in (0, \delta_0)$. we shall see that the Euler approximation usually converges with weak order $\beta = 1$ in contrast with the strong order $\gamma = 0.5$.

2.8 Strong Approximations

We shall consider discrete time approximations of various strong orders that have been derived from stochastic Taylor expansions by including appropriately many terms. To describe these schemes succinctly for a general d-dimensional Ito process satisfying the stochastic differential equation

$$X_t = X_{t_0} + \int_{t_0}^{t} a(s, X_s)ds + \sum_{j=1}^{m} \int_{t_0}^{t} b^j(s, X_s)dW_s^j$$

in Ito form, for $t \in [t_0, T]$, we shall use the following generalizations of the operators:

$$L^0 = \frac{\partial}{\partial t} + \sum_{k=1}^{d} a^k \frac{\partial}{\partial x^k} + \frac{1}{2} \sum_{k,l=1}^{d} \sum_{j=1}^{m} b^{k,j} b^{l,j} \frac{\partial^2}{\partial x^k \partial x^l}$$

$$\underline{L}^0 = \frac{\partial}{\partial t} + \sum_{k=1}^{d} \underline{a}^k \frac{\partial}{\partial x^k}$$

$$L^j = \underline{L}^j = \sum_{k=1}^{d} b^{k,j} \frac{\partial}{\partial x^k}$$

for $j = 1, 2, \ldots, m$, with the corrected drift

$$\underline{a}^k = a^k - \frac{1}{2} \sum_{j=1}^{m} \underline{L}^j b^{k,j}$$

for $k = 1, 2, \ldots, d$. In addition, we shall abbreviate multiple Ito integrals by

$$I(j_1, \ldots, j_l) = \int_{\tau_n}^{\tau_{n+1}} \cdots \int_{\tau_n}^{s_2} dW_{s_1}^{j_1} \cdots dW_{s_l}^{j_l}$$

for $j_1, \ldots, j_l \in \{0, 1, \ldots, m\}$, $l = 1, 2, \ldots$, and $n = 0, 1, 2, \ldots$ with the convention that

$$W_t^0 = t$$

for all $t \in \mathbb{R}^+$. We shall also use the abbreviation

$$f = f(\tau_n, Y_n)$$

for $n = 0, 1, 2, \ldots$ in the schemes for any given function f defined on $\mathbb{R}^+ \times \mathbb{R}^d$ and usually not explicitly mention the initial value Y_0 or the step indices $n = 0, 1, \ldots$.

2.8.1 Euler scheme

The Euler scheme is simplest strong Taylor approximation, containing only the time and Wiener integrals of multiplicity one from the Ito-Taylor expansion, and usually attains the order of strong convergence $\gamma = 0.5$. In the 1-dimensional case $d = m = 1$, the Euler scheme has the form

$$Y_{n+1} = Y_n + a\Delta_n + b\Delta W_n$$

where

$$\Delta_n = \int_{\tau_n}^{\tau_{n+1}} dt = \tau_{n+1} - \tau_n$$

is the length of the time discretization subinterval $[\tau_n, \tau_{n+1}]$ and

$$\Delta W_n = \int_{\tau_n}^{\tau_{n+1}} dW_t = W_{\tau_{n+1}} - W_{\tau_n}$$

is the $N(0, \Delta_n)$ distributed increment of the Wiener process W on $[\tau_n, \tau_{n+1}]$. For the general multiple dimensional case with $d, m = 1, 2, \ldots$ the kth componenet of the Euler scheme has the form

$$Y_{n+1}^k = Y_n^k + a^k\Delta_n + \sum_{j=1}^{m} b^{k,j}\Delta W_n^j$$

where

$$\Delta W_n^j = \int_{\tau_n}^{\tau_{n+1}} dW_t^j = W_{\tau_{n+1}}^j - W_{\tau_n}^j$$

is the $N(0, \Delta_n)$ distributed increment of the jth componenet of the m-dimensional standard Wiener process W on $[\tau_n, \tau_{n+1}]$; thus $\Delta W_n^{j_1}$ and $\Delta W_n^{j_2}$ are independent for $j_1 \neq j_2$.

Theorem 2.1 *Suppose that*

(i) $E(|X_0|^2) < \infty$,

(ii) $E(|X_0 - Y_0^\delta|^2)^{1/2} \leq K_1 \delta^{1/2}$,

(iii) $|a(t,x) - a(t,y)| + |b(t,x) - b(t,y)| \leq K_2|x-y|$,

(iv) $|a(t,x)| + |b(t,x)| \leq K_3(1+|x|)$,

(v) $|a(s,x) - a(t,x)| + |b(s,x) - b(t,x)| \leq K_4(1+|x|)|s-t|^{1/2}$

for all $s,t \in [0,T]$ *and* $x,y \in \mathbb{R}^d$, *where the constants* K_1, K_2, K_3, K_4 *do not depend on* δ. *Then, for the Euler approximation* Y^δ, *the estimate*

$$E(|X_T - Y^\delta(T)|) \leq K_5 \delta^{1/2}$$

holds, where the constant K_5 *does not depend on* δ.
Proof. *refer to [2].*

In special cases, the Euler scheme may actually achieve a higher order of strong convergence. For example, when the noise is additive, that is when the diffusion coefficient has the form

$$b(t,x) \equiv b(t)$$

for all $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^d$, it turns out that the Euler scheme has order $\gamma = 1.0$ of strong convergence under appropriate smoothness assumptions on a and b. We remark that additive noise is sometimes understood to have $b(t)$ as constant. Usually the Euler scheme gives good numerical results when the drift and diffusion coefficients are nearly constant. In general, however, it is not particularly satisfactory and the use of higher order schemes is recommended.

2.8.2 Milstein scheme

If, in the 1-dimensional case with $d = m = 1$, we add to the Euler scheme the term

$$bb'I(1,1) = \frac{1}{2}bb'\{(\Delta W_n)^2 - \Delta_n\}$$

from the Ito-Taylor expansion, then we obtain the Milstein scheme

$$Y_{n+1} = Y_n + a\Delta_n + b\Delta W_n + \frac{1}{2}bb'\{(\Delta W_n)^2 - \Delta_n\}$$

We can rewrite this as

$$Y_{n+1} = Y_n + \underline{a}\Delta_n + b\Delta W_n + \frac{1}{2}bb'(\Delta W_n)^2$$

since

$$\underline{a} = a - \frac{1}{2}bb'$$

Theorem 2.2 *Suppose that*

(i) $E(|X_0|^2) < \infty$,

(ii) $E(|X_0 - Y_0^\delta|^2)^{1/2} \leq K_1\delta^{1/2}$,

(iii)

$$\begin{aligned} |\underline{a}(t,x) - \underline{a}(t,y)| &\leq K_2|x-y| \\ |b^{j_1}(t,x) - b^{j_1}(t,y)| &\leq K_2|x-y| \\ |\underline{L}^{j_1}b^{j_2}(t,x) - \underline{L}^{j_1}b^{j_2}(t,y)| &\leq K_2|x-y| \end{aligned}$$

(iv)

$$\begin{aligned} |\underline{a}(t,x)| + |\underline{L}^j\underline{a}(t,x))| &\leq K_3(1+|x|) \\ |b^{j_1}(t,x)| + |\underline{L}^j b^{j_2}(t,x)| &\leq K_3(1+|x|) \\ |\underline{L}^j\underline{L}^{j_1}b^{j_2}(t,x))| &\leq K_3(1+|x|) \end{aligned}$$

(v)

$$\begin{aligned} |\underline{a}(s,x) - \underline{a}(t,x)| &\leq K_4(1+|x|)|s-t|^{1/2} \\ |b^{j_1}(s-x) - b^{j_1}(t-x)| &\leq K_4(1+|x|)|s-t|^{1/2} \\ |\underline{L}^{j_1}b^{j_2}(s-x) - \underline{L}^{j_1}b^{j_2}(t-x)| &\leq K_4(1+|x|)|s-t|^{1/2} \end{aligned}$$

for all $s,t \in [0,T]$, $x,y \in \mathbb{R}^d$, $j = 0,\ldots,m$ and $j_1, j_2 = 1,\ldots,m$, where the constants K_1, K_2, K_3, K_4 do not depend on δ. Then for the Milstein approximation Y^δ, the estimate

$$E(|X_T - Y^\delta(T)|) \leq K_5\delta$$

holds, where the constant K_5 does not depend on δ.

Proof. *refer to [2].*

2.9 Weak Approximation

2.9.1 Weak Euler Scheme

We recall that for the general multi-dimensional case $d, m = 1, 2, \ldots$, the kth component of the Euler scheme has the form

$$Y_{n+1}^k = Y_n^k + a^k \Delta_n + \sum_{j=1}^{m} b^{k,j} \Delta W_n^j$$

with initial value $Y_0 = X_0$, where $\Delta_n = \tau_{n+1} - \tau_n$ and $\Delta W_n^j = W_{\tau_{n+1}}^j - W_{\tau_n}^j$. If, amongst other assumptions, a and b are sufficiently smooth, then the Euler approximation has order of weak convergence $\beta = 1.0$. With the weak convergence criterion

$$|E(g(X_T)) - E(g(Y(T))))| \leq C\delta^\beta$$

we have much freedom to choose other simpler random variables instead of using the Gaussian increments ΔW_n^j. Such random variables have only to coincide in their lower order moments with those of ΔW_n^1 and ΔW_n^2 to provide a sufficient accurate approximation of the probability law of the Ito diffusion. For instance, we could use two-point distributed random variables $\Delta \hat{W}_n^j$ with

$$P\left(\Delta \hat{W}_n^j = \pm\sqrt{\Delta_n}\right) = \frac{1}{2}$$

in which case becomes the simplified Euler scheme

$$Y_{n+1}^k = Y_n^k + a^k \Delta_n + \sum_{j=1}^{m} b^{k,j} \Delta \hat{W}_n^j \tag{2.14}$$

2.9.2 Order 2.0 Weak Taylor scheme

More accurate weak Taylor schemes can be derived by including further multiple stochastic integrals from the stochastic Taylor expansion. Since the objective is to obtain more information about the probability measure of the underlying Ito process rather than about its sample space, we also have the freedom to replace the multiple stochastic integrals by much simpler random variables as in (2.14). We shall consider the weak Taylor scheme obtained by adding all of the double stochastic integrals from the Ito-Taylor expansion to the Euler scheme. In the autonomous 1-dimensional case $d = m = 1$ we obtain the order 2.0 weak Taylor scheme

$$\begin{aligned} Y_{n+1} &= Y_n + a\Delta_n + b\Delta W_n + \frac{1}{2}bb'\{(\Delta W_n)^2 - \Delta_n\} \\ &+a'b\Delta Z_n + \frac{1}{2}\left(aa' + \frac{1}{2}a''b^2\right)\Delta_n^2 \\ &\left(ab' + \frac{1}{2}b''b^2\right)\{\Delta W_n \Delta_n - \Delta Z_n\} \end{aligned}$$

where ΔZ_n represents the double Ito integral

$$\Delta Z_n = \int_{\tau_n}^{\tau_{n+1}} \int_{\tau_n}^{s_2} dW_{s_1} dW_{s_2}$$

The pair of correlated Gaussian random variables $(\Delta W_n, \Delta Z_n)$ can be generated from a pair of independent standard Gaussian random variables.

Chapter 3

Part II - Improving Accuracy

In the discussion of themes below, works will be mainly taken from D.Talay and L.Tubaro [1].

3.1 Description

Let us consider the following Ito stochastic differential equation

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t \tag{3.1}$$

where X_t is a stochastic process in $\mathbb{R}^d$, W_t is a Wiener process in $\mathbb{R}^l$, $b(t, x)$ is a d-vector and $\sigma(t, x)$ is a $d \times l$-matrix. We denote the solution of (3.1) with the deterministic initial condition x at time s by $X_t^{s,x}$ or $X_s = x$. We suppose that the coefficients b and σ are smooth so that the existence and uniqueness of this solution are ensured. Let Y be a random variable independent of any increment $W_t - W_s$ of the Wiener process, and having moments of any order. The solution of (3.1) has the initial condition $X_t^{0,Y}$ or $X_0 = Y$ at time 0. Let us consider the numerical evaluation of the quantity

$$Ef(X_T)$$

where T is some fixed time, f a given smooth function from $\mathbb{R}^d \to \mathbb{R}$ by a Monte-Carlo method based on the simulation of a piecewise constant approximating process $\bar{X}_p^h, p \in \mathbb{N}$. We divide the interval $[0, T]$ in n subintervals with the same length $h = T/n$. According to our notation, $\bar{X}_n^h = \bar{X}_T^h$. We want to study the global error

$$Err(T,h) = Ef(X_T) - Ef(\bar{X}_T^h)$$

We recall that the Euler and Milshtein discretization schemes are of first order: there exists a positive constant $C(T)$, independent of h, such that

$$|Err(T,h)| \leq C(T)h$$

Let us define the following two numerical schemes:

Euler scheme

$$\begin{aligned}
\bar{X}_0^h &= X_0 \\
\bar{X}_{p+1}^h &= \bar{X}_p^h + b(ph, \bar{X}_p^h)h + \sigma(ph, \bar{X}_p^h)\Delta_{p+1}^h W
\end{aligned}$$

where $\Delta_{p+1}^h W = W_{(p+1)h} - W_{ph}$.

Milshtein scheme

$$\begin{aligned}
\bar{X}_{p+1}^h &= \bar{X}_p^h + \sum_{j=1}^{r} \sigma_j(ph, \bar{X}_p^h)\Delta_{p+1}^h W^j + b(ph, \bar{X}_p^h)h \\
&\quad + \sum_{j,k=1}^{r} \partial\sigma_j(ph, \bar{X}_p^h)\sigma_k(ph, \bar{X}_p^h) Z_{p+1}^{kj}
\end{aligned}$$

where

- σ_j denotes the jth column of σ,
- $\partial\sigma_j$ denotes the matrix whose element of the ith row and kth column is $\partial_k \sigma_j^i$,
- $\tilde{U}_{p+1}^{kj}$ is defined as the family of iid random variables satisfying common law
$$P\left(\tilde{U}_p^{kj} = \frac{1}{2}\right) = P\left(\tilde{U}_p^{kj} = -\frac{1}{2}\right) = \frac{1}{2}$$
- $(\tilde{U}_{q+1}^{kj}, \Delta_{p+1}^h W^l)_{(p,q,k,j,l)}$'s are mutually independent,

-

$$\begin{array}{rcll} Z_{p+1}^{kj} & = & \frac{1}{2}\Delta_{p+1}^h W^k \Delta_{p+1}^h W^j + \tilde{U}_{p+1}^{kj} h & , \quad k < j \\ Z_{p+1}^{kj} & = & \frac{1}{2}\Delta_{p+1}^h W^k \Delta_{p+1}^h W^j - \tilde{U}_{p+1}^{kj} h & , \quad k > j \\ Z_{p+1}^{jj} & = & \frac{1}{2}\left((\Delta_{p+1}^h W^j)^2 - h\right) & \end{array}$$

Lemma 3.1 *Suppose that the b_i's and the σ_{ij}'s are Lipschitz continuous functions. Then, for any integer k, there exists a strictly positive constant C_k such that*

$$E|\bar{X}_p^h|^k \leq e^{C_k T}$$

for any p between 0 and n.

Let us now consider the class $\mathcal{F}_T$ of functions $\phi : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ with the following properties: ϕ is of class $\mathcal{C}^\infty$, and for some positive integer s and positive $C(T)$

$$\forall \theta \in [0,T] \quad , \quad \forall x \in \mathbb{R}^d \quad : \quad |\phi(\theta, x)| \leq C(T)(1 + |x|^s)$$

A function ϕ of $\mathcal{F}_T$ will be called homogeneous if it does not depend on the time variable: $\phi(\theta, x) = \phi(x)$. We will denote by $\mathcal{L}$ the differential operator associated to (3.1):

$$\mathcal{L} = \frac{1}{2}\sum_{i,j=1}^d a_j^i(t,x)\partial_{ij} + \sum_{i=1}^d b^i(t,x)\partial_i \tag{3.2}$$

where $a(t,x) = \sigma(t,x)\sigma^*(t,x)$. If $\phi \in \mathcal{F}_T$, then the function $u(\theta; t, x)$ defined by

$$u(\theta; t, x) = E\phi(\theta, X_T^{t,x}) = E_{t,x}\phi(\theta, X_T)$$

verifies the following equation

$$\left\{ \begin{array}{rcl} \frac{\partial u}{\partial t} + \mathcal{L}u & = & 0 \\ u(\theta; T, x) & = & \phi(\theta, x) \end{array} \right.$$

Lemma 3.2 *Let us suppose that the function b and σ are $\mathcal{C}^\infty$ functions, whose derivatives of any order are bounded. For any multi-index α, there exist strictly* *positive constants $K_\alpha(T), C_\alpha(T)$ such that*

$$\forall\theta \in [0,T] \quad : \quad |\partial_\alpha u(\theta;t,x)| \leq C_\alpha(T)(1+|x|^{k_\alpha(T)})$$

Here by ∂_α we mean the mixed partial derivative of order α

$$\frac{\partial^{|\alpha|}}{\partial_1^{\alpha_1}\partial_2^{\alpha_2}\cdots\partial_r^{\alpha_r}}$$

where $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_r)$, and $|\alpha| = \alpha_1 + \cdots + \alpha_r$.

In all the sequel, we suppose that the functions b and σ are $\mathcal{C}^\infty$ functions, whose derivatives of any order are bounded. Consider a homogeneous function f of $\mathcal{F}_T$ and $u(t,x) = Ef(X_T^{t,x})$ which solves

$$\frac{\partial u}{\partial t} + \mathcal{L}u = 0 \tag{3.3}$$

$$u(T,x) = f(x) \tag{3.4}$$

then we have

$$Err(T,h) = Eu(T, \bar{X}_n^h) - Eu(0,Y) \tag{3.5}$$

Let us first consider the Euler scheme. We compute $Eu(T,\bar{X}_n^h) - Eu((n-1)h, \bar{X}_{n-1}^h)$ by performing a Taylor expansion at the point $((n-1)h, \bar{X}_{n-1}^h)$ of the form

$$\begin{aligned} u(t+\Delta t, x+\Delta x) &= u(t,x) + \Delta t\frac{\partial}{\partial t}u(t,x) + \frac{1}{2}(\Delta t)^2\frac{\partial^2}{\partial t^2}u(t,x) \\ &\quad +\Delta t \sum_{|\alpha|=1} \Delta x^\alpha \frac{\partial}{\partial t}u(t,x)\partial_\alpha u(t,x) \\ &\quad +\frac{1}{2}\Delta t \sum_{|\alpha|=2} \Delta x^\alpha \frac{\partial}{\partial t}u(t,x)\partial_\alpha u(t,x) \\ &\quad +\frac{1}{6}\Delta t \sum_{|\alpha|=3} \Delta x^\alpha \frac{\partial}{\partial t}u(t,x)\partial_\alpha u(t,x) \\ &\quad +\frac{1}{|\alpha|!}\sum_{|\alpha|=1}^{5} \Delta x^\alpha \partial_\alpha u(t,x) + \cdots \end{aligned}$$

where $\alpha = (\alpha_1, \ldots, \alpha_r)$ and Δx^α means $\Delta x^\alpha = (\Delta x_1)^{\alpha_1} \ldots (\Delta x_r)^{\alpha_r}$

We need to do the following easy and tedious computation using (3.2), (3.3), (3.4) with $\Delta t = h$ and $\Delta x = \Delta \bar{X}_n^h = \bar{X}_n^h - \bar{X}_{n-1}^h$

$$\begin{aligned}
u(T, \bar{X}_n^h) &= u((n-1)h, \bar{X}_{n-1}^h) + h\frac{\partial}{\partial t}u((n-1)h, \bar{X}_{n-1}^h) \\
&+ \frac{1}{2}h^2\frac{\partial^2}{\partial t}u((n-1)h, \bar{X}_{n-1}^h) \\
&+ h\frac{\partial}{\partial t}\left(\sum_{i=1}^{d} hb^i((n-1)h, \bar{X}_{n-1}^h) + \sum_{j=1}^{l}\sigma_j^i((n-1)h, \bar{X}_{n-1}^h)\Delta_n^h W^j\right) \\
&+ \quad \cdots
\end{aligned}$$

We get

$$Eu(T, \bar{X}_n^h) = Eu((n-1)h, \bar{X}_{n-1}^h) + h^2 E\psi_e((n-1)h, \bar{X}_{n-1}^h) + h^3 R_n^h \tag{3.6}$$

where

$$\begin{aligned}
\psi_e(t,x) &= \frac{1}{2}\sum_{i,j=1}^{d} b^i(t,x)b^j(t,x)\partial_{ij}u(t,x) \\
&+ \frac{1}{2}\sum_{i,j,k=1}^{d} b^i(t,x)a_k^j(t,x)\partial_{ijk}u(t,x) \\
&+ \frac{1}{8}\sum_{i,j,k,l=1}^{d} a_j^i(t,x)a_l^k(t,x)\partial_{ijkl}u(t,x) + \frac{1}{2}\frac{\partial^2}{\partial t^2}u(t,x) \\
&+ \sum_{i=1}^{d} b^i(t,x)\frac{\partial}{\partial t}\partial_i u(t,x) + \frac{1}{2}\sum_{i,j=1}^{d} a_j^i(t,x)\frac{\partial}{\partial t}\partial_{ij}u(t,x)
\end{aligned}$$

and there exists a constant $C(T)$ independent of h such that

$$|R_n^h| \leq C(T)$$

We can use the same expansion for $u((n-1)h, \bar{X}_{n-1}^h)$, continuing in this way n times, and we arrive to

$$Eu(T, \bar{X}_n^h) = Eu(0, Y) + h^2\sum_{j=0}^{n-1} E\psi_e(jh, \bar{X}_j^h) + h^2\mathcal{R}_n^h \tag{3.7}$$

$$\Leftrightarrow Eu(T, \bar{X}_n^h) - Eu(0, Y) = h^2 \sum_{j=0}^{n-1} E\psi_e(jh, \bar{X}_j^h) + h^2 \mathcal{R}_n^h \tag{3.8}$$

$$\Leftrightarrow Err(T, h) = h^2 \sum_{j=0}^{n-1} E\psi_e(jh, \bar{X}_j^h) + h^2 \mathcal{R}_n^h \qquad \text{by (3.5)} \tag{3.9}$$

with

$$|\mathcal{R}_n^h| \leq C(T)$$

where $\mathcal{R}_n^h = \sum_{j=1}^n R_j^h$.

Proposition 3.3 *There exists a real number $C(T)$, independent on h, such that*

$$h \sum_{j=0}^{n-1} E|\psi_e(jh, \bar{X}_j^h)| \leq C(T) \tag{3.10}$$

Proof. *By Lemma 3.2,*

$$E|\psi_e(jh, \bar{X}_j^h)| \leq C(1 + E|\bar{X}_j^h|^K)$$

and by Lemma 3.1,

$$E|\psi_e(jh, \bar{X}_j^h)| \leq K$$

where C and K are constants independent of h and j. Hence,

$$h \sum_{j=0}^{n-1} E|\psi_e(jh, \bar{X}_j^h)| \leq hnK = TK \qquad \textit{since} \quad h = T/n \qquad \blacksquare$$

Proposition 3.4 *For any function ϕ of $\mathcal{F}_T$, there exists a real number $C(T)$, independent on h, such that*

$$E\phi(\theta, \bar{X}_T^h) = E\phi(\theta, X_T) R_T(h)$$

with $|R_T(h)| \leq C(T)h$.

Proposition 3.5

$$\left| h \sum_{j=1}^{n-1} E\psi_e(jh, \bar{X}_j^h) - \int_0^T E\psi_e(s, X_s) ds \right| = \mathcal{O}(h) \tag{3.11}$$

Remark 3.6 *For the Milshtein scheme, ψ_m can be derived by performing a Taylor expansion:*

$$\begin{aligned}\psi_m(t,x) &= \psi_e(t,x) \\ &+\frac{1}{4}\sum_{i_1,i_2,j,k,l} a_k^l(t,x)\partial_l\sigma_j^{i_1}(t,x)\partial_k\sigma_j^{i_2}(t,x)\partial_{i_1,i_2}u(t,x) \\ &+\frac{1}{2}\sum_{i_1,i_2,i_3,j_1,j_2,k} \sigma_{j_1}^{i_1}(t,x)\sigma_{j_2}^{i_2}(t,x)\sigma_{j_1}^{k}(t,x)\partial_k\sigma_{j_2}^{i_3}(t,x)\partial_{i_1,i_2,i_3}u(t,x)\end{aligned}$$

and ψ_m can be substituted into (3.6),(3.7),(3.10), and (3.11) instead of ψ_e.

3.2 Main results

Theorem 3.7 *For the Euler scheme, the error is given by*

$$Err_e(T,h) = -h\int_0^T E\psi_e(s,X_s)ds + \mathcal{O}(h^2) \tag{3.12}$$

Proof.

$$Err_e(T,h) = h\left(h\sum_{j=0}^{n-1} E\psi_e(jh,\bar{X}_j^h) + \mathcal{O}(h)\right) \qquad \text{by (3.7) (3.8) and (3.9)}$$

$$\Leftrightarrow Err_e(T,h) = h\left(-\int_0^T E\psi_e(s,X_s)ds + \mathcal{O}(h)\right) \qquad \text{by Proposition 3.5}$$

$$\Leftrightarrow Err_e(T,h) = -h\int_0^T E\psi_e(s,X_s)ds + \mathcal{O}(h^2) \qquad \text{as required} \quad \blacksquare$$

Theorem 3.8 *For the Milshtein scheme, the error is given by*

$$Err_m(T,h) = -h\int_0^T E\psi_m(s,X_s)ds + \mathcal{O}(h^2)$$

Proof. *goes in the same way as that of ψ_e in Theorem 3.7 with only replacing ψ_e by ψ_m.* $\blacksquare$

Remark 3.9 *We have proved that for the Euler approximation $\bar{X}_T^h$ and smooth* *functions* f

$$Ef(\bar{X}_T^h) - Ef(X_T) = h\int_0^T E\psi_e(s, X_s)ds + g_h(T)$$

holds for all $h, h/2, h/3, h/4, \dots$ *with some functions* ψ_e *and* g_h*, where* ψ_e *is independent of the step size* $h = T/n$ *and* g_h *is a function depending on* h *such that*

$$|g_h| \leq Ch^2$$

with some constant C *independent of* h*.*

Hence, the following important consequence holds.

Theorem 3.10 *(**Richardson extrapolation**) Consider that the following new approximation*

$$Z_T^h = 2Ef(\bar{X}_T^{h/2}) - Ef(\bar{X}_T^h) \tag{3.13}$$

where $\bar{X}_T^{h/2}$ *means that we perform an approximation with the step size* $h/2$*. Then*

$$Err(T, h) = Ef(X_T) - Z_T^h \leq Ch^2$$

where C *is a constant indepedent of* h*.*

Proof. *Using* (3.12) *with* h *and* $h/2$

$$\begin{aligned} Ef(\bar{X}_T^h) &= Ef(X_T) + e_1(T)h + g_h(T)) \\ Ef(\bar{X}_T^{h/2}) &= Ef(X_T) + e_1(T)\frac{h}{2} + g_{h/2}(T) \end{aligned}$$

where

$$e_1(T) = \int_0^T E\psi_e(s, X_s)ds$$

Hence, by difference

$$Z_T^h = 2Ef(\bar{X}_T^{h/2}) - Ef(\bar{X}_T^h) = Ef(X_T) + 2g_{h/2}(T) - g_h(T)$$

$$\Leftrightarrow Z_T^h - Ef(X_T) = 2g_{h/2}(T) - g_h(T))$$

Notice that,

$$\begin{aligned}
|2g_{h/2}(T) - g_h(T))| &\leq |2g_{h/2}(T)| + |g_h(T))| \\
&\leq 2\frac{Ch^2}{4} + Ch^2 \\
&= \frac{3}{2}Ch^2
\end{aligned}$$

Therefore, $Err(T,h) \leq \frac{3}{2}Ch^2$ *as required.* ■

We have concluded from the above Theorem 3.10 that it is possible to get a result of precision of second order from results given by a first order scheme. We realize that extrapolation is an elegant and simple way to obtain weak higher order methods. The construction of such methods is based on the existence of an asymptotic expansion of the error with respect to powers of the time step size.

Theorem 3.11 *For the second order scheme, the global expansion error can be written*

$$Err(T,h) = h^2 \int_0^T E\gamma(s, X_s)ds + \mathcal{O}(h^3)$$

for some smooth function γ*.*

Theorem 3.12 *For all these schemes, it is possible to obtain an expression of the form*

$$Err(T,h) = e_r(T)h^r + e_{r+1}(T)h^{r+1} + \cdots + e_m(T)h^m + \mathcal{O}(h^{m+1})$$

where $e_1, e_2, \ldots, e_m$ *are independent of* h*, and* $\mathcal{O}(h^{m+1})$ *means a function* g *of* h *such that*

$$|g_h| \leq Ch^{m+1}$$

for all h *with a constant* C*.*

Definition 3.13 *For step sizes* $h, h/2, h/4, h/8, \ldots$*, Richardson extrapolation method is formulated recursively as follows, by using*

$$Z^h_{T,\alpha,\varrho} = \frac{1}{2^{\varrho-1} - 1}(2^{\varrho-1}Z^{h/2}_{T,\alpha-1,\varrho-1} - Z^h_{T,\alpha-1,\varrho-1})$$

$$Z^h_{T,1,2} = 2Ef(\bar{X}^{h/2}_T) - Ef(\bar{X}^h_T) \tag{3.14}$$

$$Z^h_{T,2,3} = \frac{1}{3}(4Z^{h/2}_{T,1,2} - Z^h_{T,1,2}) \tag{3.15}$$

$$Z^h_{T,3,4} = \frac{1}{7}(8Z^{h/2}_{T,2,3} - Z^h_{T,2,3}) \tag{3.16}$$

$$Z^h_{T,4,5} = \frac{1}{15}(16Z^{h/2}_{T,3,4} - Z^h_{T,3,4}) \quad \cdots\cdots\cdots \tag{3.17}$$

Remark 3.14 *Theorem 3.12 implies that there exist universal constants $\lambda_0, \lambda_1, \ldots, \lambda_m$ such that for*

$$Z^h_T = \lambda_0 Ef(\bar{X}^h_T) + \lambda_1 Ef(\bar{X}^{h/2}_T) + \lambda_2 Ef(\bar{X}^{h/4}_T) + \cdots + \lambda_m Ef(\bar{X}^{h/2^m}_T) \tag{3.18}$$

where $\sum_{i=1}^m \lambda_i = 1$, we have

$$|Z^h_T - Ef(X_T)| \leq Ch^{m+1} \tag{3.19}$$

Continuing with Theorem 3.10, we could apply the second order extraploation (3.14) in (3.15) using the Euler scheme for step sizes $h/4, h/2, h$ to obtain the third order method

$$Z^h_{T,1,3} = \frac{1}{3}(8\bar{X}^{h/4}_T - 6\bar{X}^{h/2}_T + \bar{X}^h_T) \tag{3.20}$$

where $Z^h_{T,1,3}$ implies that we could get a result of precision of third order from results given by a first order scheme. Actually, it is possible to extend the argument if the following is verified by Theorem 3.12.

$$Err(T,h) = e_1(T)h^1 + e_2(T)h^2 + \mathcal{O}(h^3) \tag{3.21}$$

where

$$e_1(T) = \int_0^T E\psi_e(s, X_s)ds$$

$$\text{and} \quad e_2(T) = \int_0^T E\gamma(s, X_s)ds \qquad \text{from Theorem 3.11}$$

Using (3.21) and (3.20), we get

$$\begin{aligned}
8Ef(\bar{X}_T^{h/4}) &= 8Ef(X_T) + 8e_1(T)\frac{h}{4} + 8e_2(T)\frac{h^2}{4^2} + 8g_{h/4}(T) \\
-6Ef(\bar{X}_T^{h/2}) &= -6Ef(X_T) - 6e_1(T)\frac{h}{2} - 6e_2(T)\frac{h^2}{2^2} - 6g_{h/2}(T) \\
Ef(\bar{X}_T^{h}) &= Ef(X_T) + e_1(T)h + e_2(T)h^2 + g_h(T)
\end{aligned}$$

Hence,

$$Z_{T,1,3}^h = Ef(X_T) + 8g_{h/4}(T) - 6g_{h/2}(T) + g_h(T)$$

$$\Leftrightarrow Err(T,h) = 8g_{h/4}(T) - 6g_{h/2}(T) + g_h(T)$$

Notice that

$$|8g_{h/4}(T) - 6g_{h/2}(T) + g_h(T)| \leq 8\frac{h^3}{4^3} + 6\frac{h^3}{2^3} + h^3 = Ch^3$$

Consequently, we have got a result of precision of order h^3 from the result given by a second order scheme. Therefore, if we use recursively Richardson extrapolation methods defined in Definition 3.13 with Theorem 3.12 in sequence, we can eventually arrive to (3.19). The choice of weights $(\lambda_0, \lambda_1, \ldots, \lambda_m)$ in the linear combination (3.18) of the outcomes from the Euler approximations with step size $h, h/2, \ldots, h/2^m$ causes the error term of order $h, h^2, \ldots, h^m$ in the leading error expansion to cancel out asymptotically. The remaining error term involve the power h^{m+1} indicating we have a $(m+1)$ order weak scheme.

3.3 Demonstration

Consider the Ito process X satisfying the linear stochastic differential equation

$$dX_t = bX_t dt + \sigma X_t dW_t$$

$$\Leftrightarrow X_t = X_0 \exp\left(\left(b - \frac{1}{2}\sigma^2\right)t + \sigma W_t\right)$$

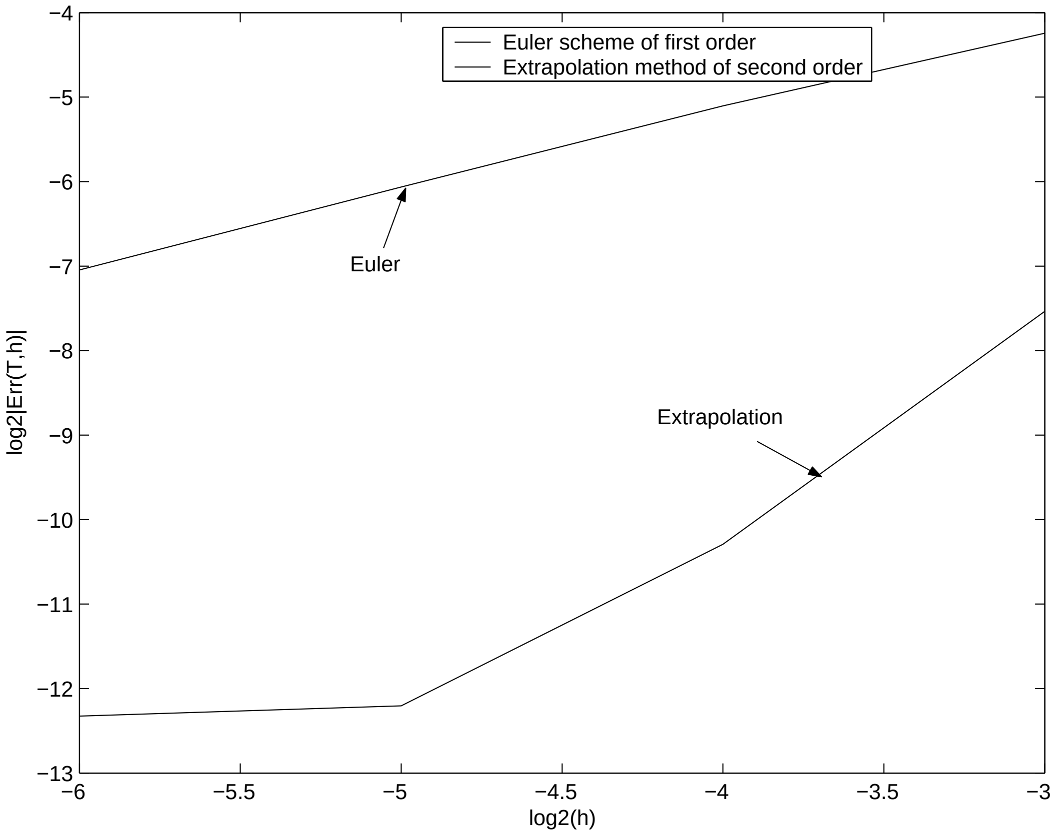

Figure 3.1: $\log_2$ of the mean error versus $\log_2 h$

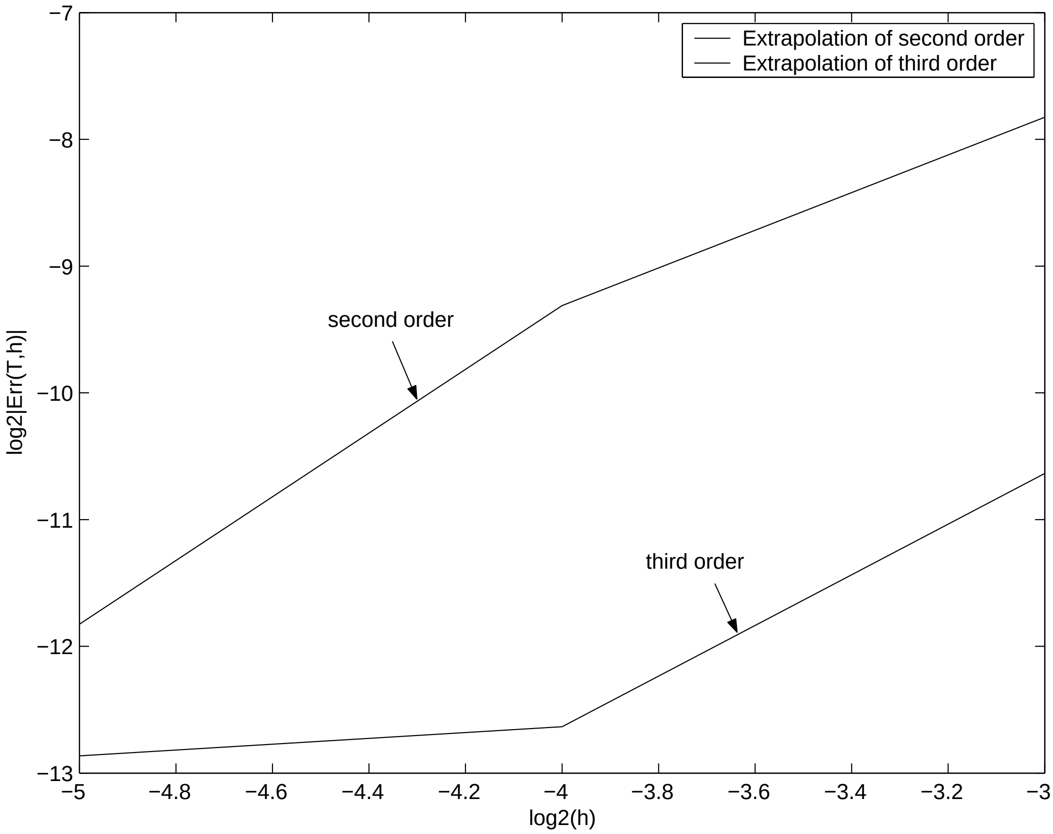

Figure 3.2: $\log_2$ of the mean error versus $\log_2 h$

with $X_0 = 0.1, b = 1.5$ and $\sigma = 0.01$ on the time interval $[0, T]$ where $T = 1$. Use the following Euler scheme

$$\begin{aligned} \bar{X}_0^h &= X_0 \\ \bar{X}_{p+1}^h &= \bar{X}_p^h + b(ph, \bar{X}_p^h)h + \sigma(ph, \bar{X}_p^h)\Delta_{p+1}^h W \end{aligned}$$

where $\Delta_{p+1}^h W = W_{(p+1)h} - W_{ph}$, in order to simulate the order 2.0 weak extrapolation

$$Z_{T,1,2}^h = 2Ef(\bar{X}_T^{h/3}) - Ef(\bar{X}_T^h)$$

and the order 3.0 weak extrapolation

$$Z_{T,1,3}^h = \frac{1}{3}(8\bar{X}_T^{h/4} - 6\bar{X}_T^{h/2} + \bar{X}_T^h)$$

for $f(x) = x$ and $h = 2^{-3}$. We generate 1000 trajectories for

$$Err(T, h) = Z_T^h - Ef(X_T)$$

We repeat the calculation for step size $h = 2^{-4}, 2^{-5}, 2^{-6}$ and 2^{-7}, and plot the results on $\log_2 |Err(T, h)|$ versus $\log_2 h$. From Figure 3.1 we see that the exponential mean error for the extrapolation method follows a steeper line than that of the Euler scheme and that the accuracy of the weak error of Euler's approximation by the extrapolation method of second order is improved. Furthermore, from Figure 3.2, we find the accuracy of the weak error become smaller by implementing the extrapolation method of third order.

Chapter 4

Part III - Implementation for Strong Approximation

Consider the Ito process X satisfying the linear stochastic differential equation

$$dX_t = bX_t dt + \sigma X_t dW_t$$

$$\Leftrightarrow X_t = X_0 \exp\left(\left(b - \frac{1}{2}\sigma^2\right) t + \sigma W_t\right)$$

with $X_0 = 0.1, b = 1.5$ and $\sigma = 0.2$ on the time interval $[0, T]$ where $T = 1$. Use the following Euler scheme

$$\begin{aligned} \bar{X}_0^h &= X_0 \\ \bar{X}_{p+1}^h &= \bar{X}_p^h + b(ph, \bar{X}_p^h)h + \sigma(ph, \bar{X}_p^h)\Delta_{p+1}^h W \end{aligned}$$

where $\Delta_{p+1}^h W = W_{(p+1)h} - W_{ph}$, in order to simulate the order 1.0 Milshtein scheme

$$\begin{aligned} \bar{Z}_0^h &= X_0 \\ \bar{Z}_{p+1}^h &= \bar{Z}_p^h + b(ph, \bar{Z}_p^h)h + \sigma(ph, \bar{Z}_p^h)\Delta_{p+1}^h W \\ &\quad + \frac{1}{2}\sigma(ph, \bar{Z}_p^h)\sigma(ph, \bar{Z}_p^h)'\{(\Delta_{p+1}^h W)^2 - h\} \end{aligned}$$

We generate 100 trajectories for

$$Err(T, h) = E|\bar{Z}_T^h - f(X_T)|$$

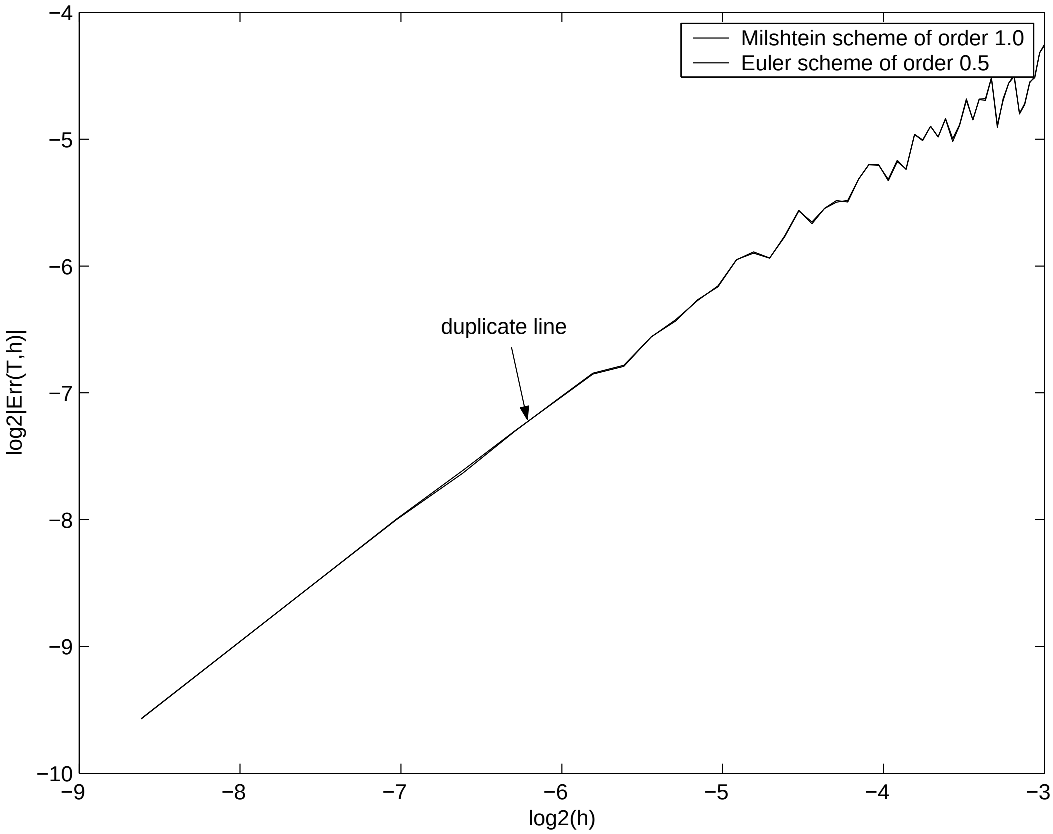

Figure 4.1: $\log_2$ of the pathwise error versus $\log_2 h$

We repeat the calculation for step size $h = 2^{-4}, 2^{-5}, 2^{-6}, 2^{-7}, 2^{-8}$ and 2^{-9}, and plot the results on $\log_2 |Err(T, h)|$ versus $\log_2 h$.

From Figure 4.1, regrettably, we see that the pathwise error of Milstein scheme is almost the same as that of Euler scheme despite of the two theoretical results(Theorem 2.1 and Theorem 2.2). Unexpectedly, Milstein scheme does not make a significant improvement on the pathwise error in case of dimension 1.

Chapter 5

Conclusion

The conclusion which can be drawn from this study of numerical solution of stochastic differential equations are these:

I. It is important that the trajectories, that is the sample paths, of the approximation should be close to those of the Ito process in some problems. We need to consider the absolute error at the final time instant T, that is

$$\epsilon(\delta) = (E|X_T - Y_N|^q)^{1/q}$$

for $k \geq 1$. The absolute error is certainly a criterion for the closeness of the sample paths of the Ito process X and the approximation Y at time T. We shall say that an approximation process Y converges in the strong sense with order $\gamma \in (0, \infty]$ if there exists a finite constant K and a positive constant δ_0 such that

$$E|X_T - Y_N| \leq K\delta^r$$

for any time discretization with maximum step size $\delta \in (0, \delta_0)$. The Euler aprroximation for SDEs

$$Y_{n+1} = Y_n + a(Y_n)\Delta_n + b(Y_n)\Delta W_n$$

has strong order $\gamma = 0.5$.

II. We may be interested only in some function of the value of the Ito process at a given final time T such as one of the first two moments EX_T and $E(X_T)^2$ or, more generally, the expectation $E(g(X_T))$ for some function g. In simulating such a functional it suffices to have a good approximation of the probability distribution of the random variable X_T rather than a close approximation of sample paths. we shall say that a time discrete approximation Y converges in the weak sense with order $\beta \in (0, \infty]$ if for any polynomial g there exists a finite constant K and a positive constant δ_0 such that

$$E(g(X_T)) - E(g(Y_N)) \leq K\delta^\beta$$

for any time discretization with maximum step size $\delta \in (0, \delta_0)$. We see that an Euler approximation of an Ito process converges with weak order $\beta = 1.0$, which is greater than its strong order of convergence $\gamma = 0.5$.

III. we obtain the Milstein scheme by adding the additional term

$$Y_{n+1} = Y_n + a\Delta_n + b\Delta W_n + \frac{1}{2}bb'\{(\Delta W_n)^2 - \Delta_n\}$$

we shall see that the Milstein scheme converges with strong order $\gamma = 1.0$ under the assumption that $E(X_0)^2 < \infty$, that a and b are twice continuously differentiable, and that a, a', b, b' and b'' satisfy a uniform Lipschitz condition.

IV. It is the purpose of this paper to attempt to represent $Err(T, h)$ as a function of h

$$Err(T, h) = e_1(T)h + e_2(T)h^2 + \cdots + e_m(T)h^m + \mathcal{O}(h^{m+1}) \tag{5.1}$$

If the above result is verified, higher order approximations of functionals can be obtained with lower order weak schemes by extrapolation methods. D.Talay and L.Tubaro [1] proposed an order 2.0 weak extrapolation method

$$Z^h_{T,1,2} = 2Ef(\bar{X}^{h/2}_T) - Ef(\bar{X}^h_T)$$

where $\bar{X}^{h/2}_T$ and $\bar{X}^h_T$ are the Euler approximation at time T for the step size $h/2$ and h respectively. Generally, if we consider

$$Z^h_T = \lambda_0 Ef(\bar{X}^h_T) + \lambda_1 Ef(\bar{X}^{h/2}_T) + \lambda_2 Ef(\bar{X}^{h/4}_T) + \cdots + \lambda_m Ef(\bar{X}^{h/2^m}_T)$$

for a sequence

$$h > h/2 > h/4 > \ldots > h/2^m$$

we can get eventually from (5.1)

$$|Z_T^h - Ef(X_T)| \leq \mathcal{O}(h^{m+1})$$

Bibliography

[1] D.Talay and L.Tubaro, “Expansion of the global error for numerical schemes solving stochastic differential equation”, *Stochastic Analysis and Applications 8(4)*, 483-509, 1990.

[2] Peter E Kloeden and Eckhard Platen, *Numerical Solution of Stochastic Differential Equations*, Springer third edition, 1999.

[3] Peter E Kloeden, Eckhard Platen and Henri Schurz, *Numerical Solution of SDE Through Computer Experiments*, Springer second edition, 1997.