A Brief Introduction to Stochastic Differential Equations

March 1, 2005

Contents

1 Introduction

Stochastic differential equations (sdes) are used with increasing frequency in a diverse range of fields. Financial engineers use sdes as the basis of stochastic volatility models. They are used for modeling neurons in computational neuroscience and for simulating protein dynamics in computational cell biology. They have a long history in physics where they are used to model Brownian motion. In chemistry sdes arise in the study of single molecule fluorescence and they also have potential applications as computational tools for quantum chemistry. They are used in the study of seismology and hydrology and in fatigue testing in engineering. This great diversity of applications is responsible for the increasing number of available texts. Amazon currently lists 143 titles dealing with the subject of stochastic differential equations. Unfortunately, most texts develop the theory in a very formal way which obscures the essential simplicity of the ideas. This document is an attempt to present these ideas in a more direct way.

2 Wiener Processes

Stochastic differential equations differ from ordinary differential equations because they are parametrized by Wiener processes in addition to time. A Wiener process W_t is a non-differentiable random function of time t obtained by sampling the normal probability density

$$\frac{1}{\sqrt{2\pi t}} e^{-W_t^2/2t}$$

at each time $t > 0$. Numerically, W_t is usually generated by sampling on some finite equidistant grid of points $t_j = j\Delta t$, for $j = 1, \ldots, K$, such that

$$W_{t_j} \quad = \quad W_0 + \sum_{l=1}^{j} \Delta W_t^l$$

where ΔW_t^l is sampled from

$$\frac{1}{\sqrt{2\pi\Delta t}} e^{-\Delta W_t^{l2}/2\Delta t}.$$

Here Δt is the spacing between times in the grid. The increments ΔW_t^l are random and therefore not equidistant, and have zero mean and variance Δt

(i.e. $\overline{\Delta W_t^{l2}} = \Delta t$). In practice it is simpler to sample a number u from the density

$$\frac{1}{\sqrt{2\pi}} e^{-u^2/2}$$

and construct ΔW_t^l through $\Delta W_t^l = u\sqrt{\Delta t}$.

3 Box-Muller Method for Normal (Gaussian) Deviates

Normally distributed random numbers u are sampled from the density

$$\frac{1}{\sqrt{2\pi}} e^{-u^2/2}.$$

Such numbers can be numerically sampled from two uniform ($[0,1]$) random numbers x_1 and x_2 through the formula

$$y = \sqrt{-2\ln x_1}\cos(2\pi x_2).$$

Random number generators for uniform random numbers are widely available.

4 Itô Differentials

Infinitesimal increments dW_t in a Wiener process W_t are sampled from

$$\frac{1}{\sqrt{2\pi dt}} e^{-dW_t^2/2dt}$$

and obey $\overline{dW_t^2} = dt$. The variance of dW_t^2 is proportional to dt^2 and hence vanishes for infinitesimal dt. This means that $dW_t^2 = dt$ in Itô calculus. Another peculiarity is that if dW_t^a and dW_t^b are statistically independent infinitesimal increments for two different Wiener processes W_t^a and W_t^b then $dW_t^a dW_t^b = 0$. These properties force us to alter the normal rules of calculus.

For example, consider the Itô differential:

$$df(t, W_t) \quad = \quad f(t + dt, W_t + dW_t) - f(t, W_t).$$

Taylor expanding $f(t+dt, W_t + dW_t)$ about t, W_t gives

$$\begin{aligned} f(t+dt, W_t + dW_t) &= f(t, W_t) + \frac{\partial f(t, W_t)}{\partial t} dt + \frac{\partial f(t, W_t)}{\partial W_t} dW_t \\ &+ \frac{1}{2} \frac{\partial^2 f(t, W_t)}{\partial W_t^2} dW_t^2 + \ldots \end{aligned}$$

and substituting into the formula for the differential and using $dW_t^2 = dt$ gives

$$df(t, W_t) = \left(\frac{\partial f(t, W_t)}{\partial t} + \frac{1}{2} \frac{\partial^2 f(t, W_t)}{\partial W_t^2}\right) dt + \frac{\partial f(t, W_t)}{\partial W_t} dW_t$$

which differs from the rule in normal calculus.

5 Itô Chain Rule

In ordinary calculus the differential of a function $f(t)$ is defined via

$$df(t) = f(t+dt) - f(t).$$

The differential of a product of two functions $f(t)$ and $g(t)$ can therefore be calculated via

$$\begin{aligned} d(f(t)g(t)) &= f(t+dt)g(t+dt) - f(t)g(t) \\ &= (f(t+dt) - f(t))g(t+dt) + f(t)(g(t+dt) - g(t)) \\ &= (f(t+dt) - f(t))(g(t+dt) - g(t)) + (f(t+dt) - f(t))g(t) \\ &+ f(t)(g(t+dt) - g(t)) \\ &= df(t)dg(t) + df(t)g(t) + f(t)dg(t). \end{aligned}$$

Since $df(t) = \frac{df(t)}{dt} dt$ and $dg(t) = \frac{dg(t)}{dt} dt$ it follows that

$$d(f(t)g(t)) = \frac{df(t)}{dt} \frac{dg(t)}{dt} dt^2 + \frac{df(t)}{dt} g(t) dt + f(t) \frac{dg(t)}{dt} dt.$$

Clearly the first term is order dt^2 while the second and third are order dt. Hence, the first term is zero when dt is infinitesimal and hence we get the usual rule

$$d(f(t)g(t)) = \left(\frac{df(t)}{dt} g(t) + f(t) \frac{dg(t)}{dt}\right) dt.$$

Now, consider functions of time and a Wiener process. Similar reasoning leads to the equality

$$\begin{aligned} d(f(t,W_t)g(t,W_t)) &= df(t,W_t)dg(t,W_t) + df(t,W_t)g(t,W_t) \\ &+ f(t,W_t)dg(t,W_t) \end{aligned}$$

but now

$$\begin{aligned} df(t,W_t) &= (\frac{\partial f(t,W_t)}{\partial t} + \frac{1}{2}\frac{\partial^2 f(t,W_t)}{\partial W_t^2})dt + \frac{\partial f(t,W_t)}{\partial W_t}dW_t \\ dg(t,W_t) &= (\frac{\partial g(t,W_t)}{\partial t} + \frac{1}{2}\frac{\partial^2 g(t,W_t)}{\partial W_t^2})dt + \frac{\partial g(t,W_t)}{\partial W_t}dW_t \end{aligned}$$

and so we obtain

$$\begin{aligned} d(f(t,W_t)g(t,W_t)) &= [(\frac{\partial f(t,W_t)}{\partial t} + \frac{1}{2}\frac{\partial^2 f(t,W_t)}{\partial W_t^2})g(t,W_t) \\ &+ f(t,W_t)(\frac{\partial g(t,W_t)}{\partial t} + \frac{1}{2}\frac{\partial^2 g(t,W_t)}{\partial W_t^2})]dt \\ &+ (\frac{\partial f(t,W_t)}{\partial W_t}g(t,W_t) + f(t,W_t)\frac{\partial g(t,W_t)}{\partial W_t})dW_t \\ &+ \frac{\partial f(t,W_t)}{\partial W_t}\frac{\partial g(t,W_t)}{\partial W_t}dW_t^2 + \text{higher order terms} \end{aligned}$$

Higher order terms proportional to $dW_t dt$ and dt^2 can be neglected but $dW_t^2 = dt$ and so

$$\begin{aligned} d(f(t,W_t)g(t,W_t)) &= [(\frac{\partial f(t,W_t)}{\partial t} + \frac{1}{2}\frac{\partial^2 f(t,W_t)}{\partial W_t^2})g(t,W_t) \\ &+ f(t,W_t)(\frac{\partial g(t,W_t)}{\partial t} + \frac{1}{2}\frac{\partial^2 g(t,W_t)}{\partial W_t^2}) \\ &+ \frac{\partial f(t,W_t)}{\partial W_t}\frac{\partial g(t,W_t)}{\partial W_t}]dt \\ &+ (\frac{\partial f(t,W_t)}{\partial W_t}g(t,W_t) + f(t,W_t)\frac{\partial g(t,W_t)}{\partial W_t})dW_t \end{aligned}$$

is the rule for the stochastic differential.

6 Itô Stochastic Differential Equations with Strong Solutions

Itô stochastic differential equations with m Wiener processes take the form

$$dX_t^j = a^j(\mathbf{X}_t, t)\ dt + \sum_{k=1}^{m} b_k^j(\mathbf{X}_t, t)\ dW_t^k,$$

where $j = 1, \ldots, n$. We assume that the coefficients $a^j(\mathbf{X}_t, t)$ and $b_k^j(\mathbf{X}_t, t)$ have regularity properties which guarantee that X_t^j are some fixed functions of the Wiener processes i.e. $X_t^j = X_j(t, W_t^1, \ldots, W_t^m)$, and that they are differentiable to high order. Such equations are said to have strong solutions. Other types of sde exist and we briefly comment on these equations below.

7 Itô Stochastic Taylor Expansion

Consider a set of sdes with strong solutions. The solutions can therefore be expanded in Taylor series. Keeping terms of order dt or less then gives

$$\begin{aligned} X_{t+dt}^j &= X_t^j + \frac{\partial X_t^j}{\partial t}\ dt + \sum_{k=1}^{m} \frac{\partial X_t^j}{\partial W_t^k}\ dW_t^k \\ &+ \frac{1}{2}\sum_{k,l=1}^{m} \frac{\partial^2 X_t^j}{\partial W_t^k \partial W_t^l}\ dW_t^k dW_t^l. \end{aligned}$$

The product of *differentials* $dW_t^k dW_t^l$ is equivalent to $\delta_{k,l} dt$ in the Itô formulation of stochastic calculus, so that

$$\begin{aligned} dX_{t+dt}^j = X_{t+dt}^j - X_t^j &= [\frac{\partial X_t^j}{\partial t} + \frac{1}{2}\sum_{k=1}^{m} \frac{\partial^2 X_t^j}{\partial W_t^{k2}}]\ dt \\ &+ \sum_{k=1}^{m} \frac{\partial X_t^j}{\partial W_t^k}\ dW_t^k. \end{aligned}$$

Comparison to the original sdes allows us to identify the first derivatives

$$\frac{\partial X_t^j}{\partial W_t^k} = b_k^j(\mathbf{X}_t, t)$$

$$
\begin{aligned}
\frac{\partial X_t^j}{\partial t} &= a^j(\mathbf{X}_t, t) - \frac{1}{2}\sum_{k=1}^{m} \frac{\partial^2 X_t^j}{\partial W_t^{k2}} \\
&= a^j(\mathbf{X}_t, t) - \frac{1}{2}\sum_{k=1}^{m}\sum_{i=1}^{n} b_k^i(\mathbf{X}_t, t)\frac{\partial b_k^j(\mathbf{X}_t, t)}{\partial X_t^i}.
\end{aligned}
$$

From these first order derivatives, expressed in terms of a^j and b_k^j, higher order derivatives can be computed. Thus a Taylor expansion of the solutions

$$
\begin{aligned}
X_{t+\Delta t}^j &= X_t^j + \frac{\partial X_t^j}{\partial t}\Delta t + \sum_{k=1}^{m} \frac{\partial X_t^j}{\partial W_t^k}\ \Delta W_t^k \\
&+ \frac{1}{2}\sum_{k,l=1}^{m} \frac{\partial^2 X_t^j}{\partial W_t^k \partial W_t^l}\Delta W_t^k \Delta W_t^l + \ldots
\end{aligned}
$$

can be obtained for finite displacements Δt and ΔW_t^k.

8 Runge-Kutta methods for sdes

This Taylor expansion of strong solutions of sdes can be employed to develop Runge-Kutta algorithms and other integration schemes.

For simplicity consider the case of autonomous equations. Define $f_j(\mathbf{X}_t)$ for $j = 1, \ldots, n$ via

$$
\begin{aligned}
f_j(\mathbf{X}_t) &= \frac{\partial X_t^j}{\partial t}\Delta t + \sum_{k=1}^{m} \frac{\partial X_t^j}{\partial W_t^k}\Delta W_t^k \\
&= [a^j(\mathbf{X}_t) - \frac{1}{2}\sum_{k=1}^{m}\sum_{i=1}^{n} b_k^i(\mathbf{X}_t)\frac{\partial b_k^j(\mathbf{X}_t)}{\partial X_t^i}]\Delta t \\
&+ \sum_{k=1}^{m} b_k^j(\mathbf{X}_t)\Delta W_t^k.
\end{aligned}
$$

The displaced solution $\mathbf{X}_{t_{i+1}} = \mathbf{X}(t_i{+}\Delta t, W_{t_i}^1{+}\Delta W_{t_i}^1, \ldots, W_{t_i}^m{+}\Delta W_t^m)$ can be expanded in a Taylor series about the initial solution $\mathbf{X}_{t_i} = \mathbf{X}(t_i, W_{t_i}^1, \ldots, W_{t_i}^m)$ using the formula

$$
\mathbf{X}_{t_{i+1}} = \exp\{\Delta t\frac{\partial}{\partial t} + \sum_{k=1}^{m} \Delta W_t^k \frac{\partial}{\partial W_t^k}\}\mathbf{X}_{t_i}.
$$

It is convenient to rewrite the Taylor expansion in terms of $f_j(\mathbf{X}_t)$ using

$$\begin{aligned}\frac{\partial}{\partial t} &= \sum_{l=1}^{n} \frac{\partial X^l_{t_i}}{\partial t}\frac{\partial}{\partial X^l_{t_i}} = \sum_{l=1}^{n}[a^l(\mathbf{X}_t) - \frac{1}{2}\sum_{k=1}^{m}\sum_{i=1}^{n} b^i_k(\mathbf{X}_t)\frac{\partial b^l_k(\mathbf{X}_t)}{\partial X^i_t}]\frac{\partial}{\partial X^l_{t_i}} \\ \frac{\partial}{\partial W^k_t} &= \sum_{l=1}^{n} \frac{\partial X^l_{t_i}}{\partial W^k_t}\frac{\partial}{\partial X^l_{t_i}} = \sum_{l=1}^{n} b^l_k(\mathbf{X}_t)\frac{\partial}{\partial X^l_{t_i}}.\end{aligned}$$

It then follows that

$$\exp\{\Delta t\frac{\partial}{\partial t} + \sum_{k=1}^{m}\Delta W^k_t\frac{\partial}{\partial W^k_t}\} = \exp\{\sum_{l=1}^{n} f_l\frac{\partial}{\partial X^l_{t_i}}\}$$

and so the Taylor expansion can be equivalently expressed as

$$\mathbf{X}_{t_{i+1}} = \sum_{k=0}^{\infty}\frac{1}{k!}\left[\sum_{l=1}^{n} f_l\frac{\partial}{\partial X^l_{t_i}}\right]^k \mathbf{X}_{t_i}.$$

By expressing the Taylor expansion in this form we can treat ordinary and stochastic differential equations in a unified way. All classical Runge-Kutta methods reproduce this Taylor expansion up to a certain number of terms. The terms following these serve as an estimate of the error. If the order of the error for some Runge-Kutta method is δt to the exponent l for odes then the same method will be order $l/2$ for sdes.

Now Runge-Kutta methods can be developed in a straightforward way. Consider the following four stage method

$$\begin{aligned}K^1_j &= f_j(\mathbf{X}_{t_i}) \\ K^2_j &= f_j(\mathbf{X}_{t_i} + \alpha\mathbf{K}^1) \\ K^3_j &= f_j(\mathbf{X}_{t_i} + \beta\mathbf{K}^2) \\ K^4_j &= f_j(\mathbf{X}_{t_i} + \gamma\mathbf{K}^3) \\ \mathbf{X}_{t_{i+1}} &= \mathbf{X}_{t_i} + c_1\mathbf{K}^1 + c_2\mathbf{K}^2 + c_3\mathbf{K}^3 + c_4\mathbf{K}^4\end{aligned}$$

where α, β, γ and c_i are to be determined. Taylor expansion of $\mathbf{X}_{t_{i+1}}$ about $\mathbf{X}_{t_i}$ to order Δt^2 then gives the result

$$\begin{aligned}\mathbf{X}^j_{t_{i+1}} &= \mathbf{X}^j_{t_i} + (c_1+c_2+c_3+c_4)f_j + (\alpha c_2 + \beta c_3 + \gamma c_4)\sum_{k=1}^{n}\frac{\partial f_j}{\partial X^k_{t_i}} f_k \\ &+ \frac{1}{2}(\alpha^2 c_2 + \beta^2 c_3 + \gamma^2 c_4)\sum_{k,l=1}^{n}\frac{\partial^2 f_j}{\partial X^k_{t_i}\partial X^l_{t_i}} f_k f_l\end{aligned}$$

$$
\begin{aligned}
&+ \quad \frac{1}{6}(\alpha^3 c_2 + \beta^3 c_3 + \gamma^3 c_4) \sum_{k,l,m=1}^{n} \frac{\partial^3 f_j}{\partial X_{t_i}^k \partial X_{t_i}^l \partial X_{t_i}^m} f_k f_l f_m \\
&+ \quad (\alpha\beta c_3 + \beta\gamma c_4) \sum_{k,l=1}^{n} \frac{\partial f_j}{\partial X_{t_i}^k} \frac{\partial f_k}{\partial X_{t_i}^l} f_l \\
&+ \quad \frac{1}{2}(\alpha^2 \beta c_3 + \beta^2 \gamma c_4) \sum_{k,l,m=1}^{n} \frac{\partial f_j}{\partial X_{t_i}^k} \frac{\partial^2 f_k}{\partial X_{t_i}^l \partial X_{t_i}^m} f_l f_m \\
&+ \quad (\beta^2 \alpha c_3 + \gamma^2 \beta c_4) \sum_{k,l,m=1}^{n} \frac{\partial^2 f_j}{\partial X_{t_i}^k \partial X_{t_i}^l} \frac{\partial f_k}{\partial X_{t_i}^m} f_l f_m \\
&+ \quad \alpha\beta\gamma c_4 \sum_{k,l,m=1}^{n} \frac{\partial f_j}{\partial X_{t_i}^k} \frac{\partial f_k}{\partial X_{t_i}^l} \frac{\partial f_l}{\partial X_{t_i}^m} f_m .
\end{aligned}
$$

Here it is understood that f_j and their derivatives are evaluated at $\mathbf{X}_{t_i}$. The requirement that this expansion reproduce the exact Taylor expansion then yields the conditions

$$
\begin{aligned}
c_1 + c_2 + c_3 + c_4 &= 1 \\
\alpha c_2 + \beta c_3 + \gamma c_4 &= 1/2 \\
\alpha^2 c_2 + \beta^2 c_3 + \gamma^2 c_4 &= 1/3 \\
\alpha\beta c_3 + \beta\gamma c_4 &= 1/6 \\
\alpha^3 c_2 + \beta^3 c_3 + \gamma^3 c_4 &= 1/4 \\
\alpha^2 \beta c_3 + \beta^2 \gamma c_4 &= 1/12 \\
\alpha\beta^2 c_3 + \beta\gamma^2 c_4 &= 1/8 \\
\alpha\beta\gamma c_4 &= 1/24.
\end{aligned}
$$

It can be readily verified that $c_1 = 1/6 = c_4$, $c_2 = 1/3 = c_3$, $\alpha = 1/2 = \beta$, $\gamma = 1$ is a solution.

This procedure generalized to the non-autonomous case gives the classic fourth order Runge-Kutta scheme with four stages

$$
\begin{aligned}
K_j^1 &= f_j(\mathbf{X}_{t_i}, t_i) \\
K_j^2 &= f_j(\mathbf{X}_{t_i} + \frac{1}{2}\mathbf{K}^1, t_i + \frac{1}{2}\Delta t) \\
K_j^3 &= f_j(\mathbf{X}_{t_i} + \frac{1}{2}\mathbf{K}^2, t_i + \frac{1}{2}\Delta t) \\
K_j^4 &= f_j(\mathbf{X}_{t_i} + \mathbf{K}^3, t_{i+1}) \\
\mathbf{X}_{t_{i+1}} &= \mathbf{X}_{t_i} + \frac{1}{6}(\mathbf{K}^1 + 2\mathbf{K}^2 + 2\mathbf{K}^3 + \mathbf{K}^4)
\end{aligned}
$$

where

$$
\begin{aligned}
f_j(\mathbf{X}_t, t) &= \frac{\partial X_t^j}{\partial t}\Delta t + \sum_{k=1}^{m} \frac{\partial X_t^j}{\partial W_t^k}\Delta W_t^k \\
&= [a^j(\mathbf{X}_t, t) - \frac{1}{2}\sum_{k=1}^{m}\sum_{i=1}^{n} b_k^i(\mathbf{X}_t, t)\frac{\partial b_k^j(\mathbf{X}_t, t)}{\partial X_t^i}]\Delta t \\
&+ \sum_{k=1}^{m} b_k^j(\mathbf{X}_t, t)\Delta W_t^k.
\end{aligned}
$$

Here t_i is the initial time and $t_{i+1} = t_i + \Delta t$.

Thus, it is possible to develop Runge-Kutta type schemes which reproduce the Taylor expansion of the solution to some given order. Fixed step-size Runge-Kutta methods are neither accurate nor efficient for general systems of equations. We need some means of controlling the local error. Hence ANISE, which is based on this sort of Taylor expansion, is implemented with variable step sizes such that local error can be internally minimized.

9 Weak Equations

Unfortunately, there are sdes that do not have strong solutions. Consider for example the case of a scalar equation with one Wiener process. If the derivatives $\frac{\partial X_t}{\partial W_t^k}$ and $\frac{\partial X_t}{\partial t}$ are constructed by the recipe given above and

$$
\frac{\partial^2 X_t}{\partial t \partial W_t} \neq \frac{\partial^2 X_t}{\partial W_t \partial t}
$$

or if these derivatives do not exist then the sde is not really a true differential but rather a Pfaffian differential and its solutions will depend on the path of integration. This situation should be familiar from thermodynamics where both heat and work have this property. ANISE may not work for these equations. The free program weak4.f available from our website may work in this case.