Chapter 10

Diffusions and Stochastic Differential Equations

A diffusion process is a stochastic process $X = (X(t) : t \geq 0)$ satisfying a stochastic differential equations (SDE). Such diffusions are Markov processes that evolve in continuous time and take values in a continuous state space.

10.1 Stochastic Differential Equations

A common approach to modeling deterministic dynamical systems is to postulate that the state $x = (x(t) : t \geq 0)$ satisfies a deterministic ordinary differential equation (ODE).

$$\begin{cases} \frac{d}{dt}x(t) = \mu(x(t)) \\ x(0) = x_0 \end{cases} \tag{10.1}$$

A natural stochastic analog to (10.1) is

$$\begin{cases} \frac{d}{dt}X(t) = \mu(X(t)) + \sigma(X(t))\xi(t) \\ X(0) = x_0 \end{cases} \tag{10.2}$$

where $(\xi(t) : t \geq 0)$ is a unit variance "white noise" process for which $E\left[\xi(t)\right] = 0$ and

$$\operatorname{cov}\left(\xi(s), \xi(t), =\right)\delta_{st}$$

where δ_{st} is one if $s = t$ and is zero otherwise. Note that for $t_1 < t_2 < t_3$,

$$\operatorname{cov}\left(\int_{t_1}^{t_2} \xi(s)ds, \int_{t_2}^{t_3} \xi(u)du, =\right) 0, \tag{10.3}$$

where as

$$\operatorname{var}\left(\int_0^t \xi(s)ds\right) = t. \tag{10.4}$$

Such a process $(\xi(t) : t \geq 0)$ has highly irregular sample paths and is very difficult to work with directly. (Imagine trying to simulate ξ!) As a consequence, it is mathematically easier to work with its (smoother) integral. This suggests writing (10.2) in the form

$$X(t) - X(0) = \int_0^t \mu(X(s))ds + \int_0^t \sigma(X(s))\xi(s)ds \tag{10.5}$$

In this integrated version, we must make mathematical sense of the stochastic integral involving the "integrator" $\xi(s)ds$. From a notational standpoint, it is standard to write

$$dX(t) = \mu(X(t))dt + \sigma(X(t))\xi(t)dt \tag{10.6}$$

in place of (10.5). The equation (10.6) is what is known as a stochastic differential equation (SDE). The rigorous meaning of (10.6) is the integral equation (10.5).

10.2 Brownian Motion

Brownian motion plays a key role in the theory of stochastic integration. A standard Brownian motion is a Gaussian process $B = (B(t) : t \geq 0)$ satisfying:

- $E[B(t)] = 0 \text{ for } t \geq 0$
- $\text{cov}\left((,B)(s), B(t)\right) = \min(s,t) \text{ for } s,t \geq 0$
- B has continuous sample paths

A Brownian motion with drift μ and variance σ^2 is a process $Z = (Z(t) : t \geq 0)$ taking the form

$$Z(t) = \mu t + \sigma B(t)$$

for $t \geq 0$. Note that

$$Z(t) \stackrel{D}{=} N(\mu t, \sigma^2 t)$$

A Brownian motion has stationary independent increments:

- For $t_1 < t_2 < \ldots < t_n$, $Z(t_1) - Z(0)$, $Z(t_2) - Z(t_1), \ldots Z(t_n) - Z(t_{n-1})$ are independent random variables (i.e. independent increments)
- $Z(t+s) - Z(t) \stackrel{D}{=} Z(s) - Z(0)$ (i.e. stationary increments)

As a consequence, it is easily verified that

$$\text{cov}\left(B(t_2) - B(t_1), B(t_3) - B(t_2), =\right) 0$$

and

$$\text{var}\left(B(t) - B(0)\right) = t$$

Given the similarity with (10.3) and (10.4), this suggests that B can be viewed as "integrated white noise", so that we can rigorously define

$$\int_0^t \xi(s)ds$$

to be $B(T) - B(0)$ (= B(t)).

Remark 10.1: This is something of an over simplification. To write

$$B(t) = \int_0^t \xi(s)ds$$

would require that B is differentiable almost everywhere (in time). But

$$\frac{B(t+h) - B(t)}{h} \stackrel{D}{=} N\left(0, h^{-\frac{1}{2}}\right)$$

so no limit exists as $h \to 0$. Hence, B is non-differentiable at t. This over simplification comes from to the fact that white noise does not exist as a well-define stochastic process. On the other hand, Brownian motion is well-defined, so this suggests that mathematically, we should replace (10.5) with

$$X(t) - X(0) = \int_0^t \mu(X(s))ds + \int_0^t \sigma(X(s))dB(s) \tag{10.7}$$

and (10.6) by

$$dX(t) = \mu(X(t))dt + \sigma(X(t))dB(t) \tag{10.8}$$

10.3 Stochastic Integrals

The integral

$$\int_0^t \mu(X(s))ds$$

can be defined via a standard Riemann approximation. On the other hand,

$$\int_0^t \sigma(X(s))dB(s)$$

must be defined differently, since the integrator here is a non-differentiable stochastic process (namely, B). The most commonly accepted definition of the stochastic integral is to define it as a limit of approximations of the form

$$\sum_{k=0}^{n-1} \sigma\left(X\left(\frac{kt}{n}\right)\right)\left[B\left(\frac{(k+1)t}{n}\right) - B\left(\frac{kt}{n}\right)\right]$$

as $n \to \infty$. This leads to the so-called "Itô integral" definition for

$$\int_0^t \sigma(X(s))dB(s)$$

Remark 10.2: Because of the non-differentiability of B, it turns out that the approximation

$$\sum_{k=0}^{n-1} \sigma\left(X\left(\frac{(k+1)t}{n}\right)\right)\left[B\left(\frac{(k+1)t}{n}\right) - B\left(\frac{kt}{n}\right)\right]$$

converges to a different limit as $n \to \infty$. Hence, care must be taken in working with stochastic integrals.

10.4 The Infinitesimal Drift and Variance of a Diffusion

Under modest conditions on $\mu(\cdot)$ and $\sigma(\cdot)$, there exists a solution $X = (X(t) : t \geq 0)$ to the SDE

$$dX(t) = \mu(X(t))dt + \sigma(X(t))dB(t)$$

The diffusion X is a Markov process with continuous sample paths and is time-homogeneous in the sense that

$$P_x\{X(t+h) \in \cdot | X(u) : 0 \leq u \leq t\} = P(h, X(t), \cdot\}$$

where

$$P\{h, x, B\} = P_x\{X(h) \in B\}$$

Note that when $h > 0$ is small,

$$X(t) - X(0) = \int_0^h \mu(X(s))ds + \int_0^h \sigma(X(s))dB(s) \approx \mu(X(0))h + \sigma(X(0))[B(h) - B(0)] \tag{10.9}$$

So

$$E_x\left[X(h) - x\right] = \mu(x)h + o(h)$$

and

$$E_x\left[(X(h) - x)^2\right] = \sigma^2(x)h + o(h)$$

as $h \to \infty$. As a consequence, $\mu(x)$ is called the infinitesimal drift of the diffusion X at x and $\sigma^2(x)$ is the infinitesimal variance of X at x. Hence, a diffusion / SDE is formulated (from a modeling viewpoint) by specifying its infinitesimal mean and variance functions.

10.5 Computing Expectations for Diffusions

Expectations and probabilities can be computed in the diffusion setting by solving ordinary or partial differential equations. To determine the appropriate differential equation, we use an analog to "first transition analysis" in the discrete time Markov chain setting. We illustrate this idea via several examples.

Example 10.1: Computing Exit Probabilities from an Interval

For $a < x < b$, compute

$$u(x) = P_x\{X(T) = a\}$$

where $T = \inf\{t \geq 0 : X(t) \notin (a, b)\}$, is the exit time from (a, b). Note that $u(a) = 1$ and $u(b) = 0$. For $h > 0$ and small,

$$u(x) = E_x\left[u(X(h))\right] + o(h) \tag{10.10}$$

Assuming $u(\cdot)$ is twice continuously differentiable,

$$\begin{aligned} E_x\left[u(X(h))\right] &= u(x) + u'(x)E_x\left[X(h) - x\right] + \frac{u''(x)}{2}E_x\left[(X(h) - x)^2\right] + o(h) \\ &= u(x) + \mu(x)u'(x)h + \frac{\sigma^2(x)}{2}u''(x)h + o(h) \end{aligned}$$

as $h \to \infty$. Plugging this into (10.10), we get

$$0 = \mu(x)u'(x)h + \frac{\sigma^2(x)}{2}u''(x)h + o(h)$$

Dividing by h and letting $h \to 0$ we find that:

$$0 = \mu(x)u'(x) + \frac{\sigma^2(x)}{2}u''(x)$$

subject to $u(a) = 1$ and $u(b) = 0$. For example, if $\mu = 0$ and $\sigma^2 = 1$ (so X is just standard Brownian motion),

$$u(x) = \frac{b - x}{b - a}$$

Example 10.2: Computing the Mean Exit Time from an Interval

Let $u(x) = E_x[T]$ where T is as in Example 1. For $h > 0$ and small,

$$u(x) = h + E_x[u(X(h))] + o(h) \tag{10.11}$$

Assuming $u(\cdot)$ is twice continuously differentiable,

$$E_x[u(X(h))] = u(x) + \mu(x)u'(x)h + \frac{\sigma^2(x)}{2}u''(x)h + o(h)$$

as $h \to 0$. Plugging this into (10.11), subtracting $u(x)$ from each side, dividing by h and sending $h \to 0$, we get

$$-1 = \mu(x)u'(x) + \frac{\sigma^2(x)}{2}u''(x)$$

subject to the (obvious) boundary conditions that $u(a) = u(b) = 0$.

Example 10.3: Computing the Mean Reward up to the Exit From an Interval

Let

$$u(x) = E_x\left[\int_0^T r(X(s))ds\right]$$

For $h > 0$ and small,

$$u(x) = r(x)h + E_x[u(X(h))] + o(h)$$

Assuming $u(\cdot)$ is twice continuously differentiable,

$$E_x[u(X(h))] = u(x) + \mu(x)u'(x)h + \frac{\sigma^2(x)}{2}u''(x)h + o(h)$$

as $h \to 0$. This leads to the ordinary differential equation (ODE)

$$-r(x) = \mu(x)u'(x) + \frac{\sigma^2(x)}{2}u''(x)$$

subject to $u(a) = u(b) = 0$.

Example 10.4: Computing the Infinite Horizon Discounted Reward

Let

$$u(x) = E_x\left[\int_0^\infty e^{-\alpha t}r(X(t))dt\right]$$

for $\alpha > 0$. For $h > 0$ and small,

$$\begin{aligned} u(x) &= E_x\left[\int_0^h e^{-\alpha s}r(X(s))ds + e^{-\alpha h}\int_0^\infty e^{-\alpha s}r(X(s+h))dx\right] \\ &= r(x)h + e^{-\alpha h}E_x[u(X(h))] + o(h) \\ &= r(x)h + (1-\alpha h)E_x[u(X(h))] + o(h) \end{aligned}$$

Assuming $u(\cdot)$ is twice continuously differentiable,

$$E_x[u(X(h))] = u(x) + \mu(x)u'(x)h + \frac{\sigma^2(x)}{2}u''(x)h + o(h)$$

as $h \to 0$. This leads to the ODE:

$$-r(x) = \mu(x)u'(x) + \frac{\sigma^2(x)}{2}u''(x) - \alpha u(x)$$

If $r(\cdot)$ is bounded, the solution $u(\cdot)$ must be bounded.

Example 10.5: Computing a Transient Expectation

Let

$$u(x,t) = E_x\left[r(X(t))\right]$$

For $h > 0$ and small,

$$\begin{aligned} u(x,t) &= E_x\left[r(X(t))|X(u) : 0 \leq u \leq h\right] \\ &= E_x\left[r(X(t))|X(h)\right] \\ &= E_x\left[u(t-h, X(h))\right] \end{aligned}$$

Assuming that $u(\cdot)$ is smooth,

$$E_x\left[u(t-h,X(h))\right] - u(t,x) = -\frac{\partial}{\partial t}u(t,x)h + \frac{\partial}{\partial x}u(t,x)\mu(x)h + \frac{1}{2}\frac{\partial^2}{\partial x^2}u(t,x)\sigma^2(x)h + o(h)$$

Hence, we arrive at the partial differential equation (PDE)

$$u_t(t,x) = \mu(x)u_x(t,x) + \frac{\sigma^2(x)}{2}u_{xx}(t,x)$$

subject to $u(0,x) = r(x)$.

10.6 Multi-dimensional Diffusions

Suppose that X_1 and X_2 jointly satisfy a coupled system of SDEs (B_1, B_2 independent standard Brownian motion).

$$dX_1(t) = \mu_1(X_1(t),X_2(t))dt + \sigma_{11}(X_1(t),X_2(t))dB_1(t) + \sigma_{12}(X_1(t),X_2(t))dB_2(t)$$

$$dX_2(t) = \mu_2(X_1(t),X_2(t))dt + \sigma_{21}(X_1(t),X_2(t))dB_1(t) + \sigma_{22}(X_1(t),X_2(t))dB_2(t)$$

The same analysis as followed above shows that

$$\begin{aligned} E_{x,y}\left[X_1(h) - x\right] &= \mu_1(x,y)h + o(h) \\ E_{x,y}\left[(X_1(h) - x)^2\right] &= \left(\sigma_{11}^2(x,y) + \sigma_{12}^2(x,y)\right)h + o(h) \\ E_{x,y}\left[X_2(h) - y\right] &= \mu_2(x,y)h + o(h) \\ E_{x,y}\left[(X_2(h) - x)^2\right] &= \left(\sigma_{21}^2(x,y) + \sigma_{22}^2(x,y)\right)h + o(h) \\ E_{x,y}\left[(X_1(h) - x)(X_2(h) - y)\right] &= \left(\sigma_{11}(x,y)\sigma_{21}(x,y) + \sigma_{22}(x,y)\sigma_{12}(x,y)\right)h + o(h) \end{aligned}$$

Let $K \subseteq \mathbb{R}^2$ and let $(x,y) \in K$. To compute:

$$u(x,y) = E_{x,y}\left[T\right]$$

where $T = \inf\{t \geq 0 : (X_1(t), X_2(t)) \notin K\}$, we solve the PDE

$$\begin{aligned}\mu_1(x,y)u_x(x,y) + \mu_2(x,y)u_y(x,y) \\ + \frac{\sigma_{11}^2(x,y)+\sigma_{12}^2(x,y)}{2}u_{xx}(x,y) + \frac{\sigma_{22}^2(x,y)+\sigma_{21}^2(x,y)}{2}u_{yy}(x,y) \\ + \left(\sigma_{11}(x,y)\sigma_{21}(x,y) + \sigma_{22}(x,y)\sigma_{12}(x,y)\right)u_{xy}(x,y) = -1\end{aligned}$$

Subject to $u(x,y) = 0$ on the boundary of K. Examples 1 through 4 lead to "elliptic PDEs" in two variables; Example 5 leads to a "parabolic PDE" in two spacial variables.

More generally, if $X_1, \ldots, X_d$ jointly satisfy a coupled system of d SDEs, Example 1 through 4 lead to elliptic PDEs in d variables; Example 5 leads to a parabolic PDE in d spacial variables. Thus, the full force of numerical PDEs can be brought to bear on solving such problems.

Conversely, in solving elliptic and parabolic PDEs, we can represent the solutions to such PDEs as expectations of diffusion processes. Hence, one means of solving such PDEs is via the Monte Carlo method. This is an attractive solution methodology when dealing with high-dimensional elliptic and parabolic PDEs.