5 An Introduction to Itô Calculus

In this final Section, we introduce a further type of stochastic process known as an Itô process. We have already met two such processes – Brownian motion and geometric Brownian motion. Here, we will formulate such models as stochastic differential equations (S.D.E.s) and develop a framework for working with these processes.

5.1 The Itô Integral

Assume that stock price follows a G.B.M. with parameters μ and σ^2. Suppose that we are interested in the behaviour of stock price in the interval $[t, t + \Delta t]$. Using the relationship between G.B.M. and B.M., we can express the change in the logarithm of an assets price as

$$\log(S_{t+\Delta t}) - \log(S_t) = (\mu - \sigma^2/2)\Delta t + \sigma(W_{t+\Delta t} - W_t)$$

where W_t is standard Brownian motion. Assuming zero volatility (i.e. $\sigma^2 = 0$) we have

$$\frac{\log(S_{t+\Delta t}) - \log(S_t)}{\Delta t} = \mu$$
$$\Rightarrow \frac{d\log(S_t)}{dt} = \mu \qquad \text{upon taking } \Delta t \to 0.$$

In other words, the change in the logarithm of an asset's price is constant. Note that solving the above ordinary differential equation (ODE) system gives

$$S_t = S_0 e^{\mu t}$$

which we may have anticipated, since this is also the expectation of S_t when σ is non zero. What if we don't assume zero volatility? Clearly,

$$\frac{\log(S_{t+\Delta t}) - \log(S_t)}{\Delta t} = \mu - \sigma^2/2 + \sigma\frac{W_{t+\Delta t} - W_t}{\Delta t}.$$

We would like to take the limit (as $\Delta t \to 0$) of both sides and write dW_t/dt on the RHS. However, although sample paths of W_t are continuous functions of t, they are almost everywhere non-differentiable. To see this, note the difference quotient

$$\frac{W_{t+\Delta t} - W_t}{\Delta t}$$

is not bounded in the neighbourhood of t. Intuitively, the variance of the difference quotient explodes as $\Delta t \to 0$. Consequently, dW_t/dt has no meaning in the usual sense. In particular, the Riemann integral

$$\int_0^t g(\tau)\frac{dW_\tau}{d\tau}d\tau$$

has no meaning since $dW_\tau/d\tau$ is not defined. Therefore, a new definition of a stochastic integral is needed.

Let $g(t)$ be a function of time, where $t \in [a, b]$. We define the Itô stochastic integral for functions g satisfying

$$\int_a^b g^2(t)\,dt < \infty.$$

Definition 5.1 (Itô Stochastic Integral)

Partition $[a, b]$ as $a = t_0 < t_1 < \ldots < t_{n-1} < t_n = b$ and set $\Delta t = t_{i+1} - t_i = (b-a)/n$, $\Delta W_{t_i} = W_{t_{i+1}} - W_{t_i}$. The Itô stochastic integral of g is defined as

$$\int_a^b g(t)\,dW_t = l.i.m._{n \to \infty} \sum_{i=0}^{n-1} g(t_i)\Delta W_{t_i}\,,$$

where $l.i.m.$ denotes mean square convergence. In other words, if $G_n = \sum_{i=0}^{n-1} g(t_i)\Delta W_{t_i}$ and $\mathcal{I} = \int_a^b g(t)\,dW_t$ then $l.i.m._{n \to \infty} G_n = \mathcal{I}$ means that

$$\lim_{n \to \infty} \mathrm{E}[(G_n - \mathcal{I})^2] = 0\,.$$

Simple Itô integrals

Simple Itô stochastic integrals can be verified directly from the Definition 5.1. For example taking $g(t) = 1$ gives

$$\begin{aligned}
\int_a^b dW_t &= l.i.m._{n \to \infty} \sum_{i=0}^{n-1} \Delta W_{t_i} \\
&= l.i.m._{n \to \infty}(W_{t_1} - W_{t_0} + W_{t_2} - W_{t_1} + \ldots + W_{t_n} - W_{t_{n-1}}) \\
&= l.i.m._{n \to \infty}(W_{t_n} - W_{t_0}) \\
&= W_b - W_a\,.
\end{aligned}$$

where the last line follows since $t_0 = a$ and $t_n = b$ irrespective of the partition. Similarly, taking $a = 0$ and $b = t$ gives

$$\int_0^t dW_\tau = W_t - W_0 = W_t$$

since $W_0 = 0$ and we used τ as a dummy variable for time since t is the upper limit of the integral. Now take $g(t) = k$ for some constant k to see that

$$\int_0^t k\,dW_\tau = k(W_t - W_0) = kW_t.$$

Comments

- Solutions of Itô integrals are stochastic processes.

- Note that a general Itô stochastic integral of the form

$$X_t = X_0 + \int_0^t g(\tau)dW_\tau$$

  can be written in differential form as

$$dX_t = g(t)dW_t\,.$$

  This latter form is called the stochastic differential representation of X_t. In the following section we will generalise this construct.

5.2 Stochastic Differential Equations (S.D.E.s)

Definition 5.2

A stochastic process X_t is said to satisfy an Itô stochastic differential equation (S.D.E.), if we may write

$$dX_t = \alpha(t, X_t)\, dt + \beta(t, X_t)\, dW_t \,.$$

Comments

- The corresponding integral representation is

$$X_t = X_0 + \int_0^t \alpha(\tau, X_\tau)\, d\tau + \int_0^t \beta(\tau, X_\tau)\, dW_\tau \,.$$

- Note that the first integral above is the usual Riemann integral and the second is an Itô stochastic integral.

- The α term is known as the *drift* coefficient and β^2 as the *diffusion* coefficient.

- If $\alpha(t, X_t)$ and $\beta(t, X_t)$ do not depend explicitly on t then the S.D.E. is said to be time-homogeneous.

Simple S.D.E.s

- Suppose that $X_0 = 0$, $\alpha(t, X_t) = 0$ and $\beta(t, X_t) = 1$. Then

$$dX_t = dW_t.$$

Integrating both sides of the above between 0 and t gives

$$\int_0^t dX_\tau \;=\; \int_0^t dW_\tau$$
$$\Rightarrow X_t - X_0 \;=\; W_t - W_0$$

for which we see that using $X_0 = 0$ and $W_0 = 0$ gives $X_t = W_t$, a standard Brownian motion process.

- Take $X_0 = x_0$, $\alpha(t, X_t) = a$ and $\beta(t, X_t) = b$ where x_0, a and b are real numbers. The resulting S.D.E. is

$$dX_t = a\, dt + b\, dW_t \,.$$

Integrating both sides of the above between 0 and t gives

$$\int_0^t dX_\tau \;=\; \int_0^t a\, d\tau + b \int_0^t dW_\tau$$
$$\Rightarrow X_t - X_0 \;=\; at + b(W_t - W_0)$$

for which we obtain

$$X_t = x_0 + a\, t + b\, W_t$$

and we recognise X_t as a generalised Brownian motion process.

- We would like to be able to write down a S.D.E. for the stock price process S_t, assuming it follows a geometric B.M. We know that $S_t = \exp(X_t)$ where X_t is a generalised Brownian motion with initial condition $\log S_0$, drift $\mu - \sigma^2/2$ and diffusion coefficient σ^2 and we know dX_t. Hence, we just need a way of differentiating $\exp(X_t)$ in a sensible way. Essentially, we need a *chain rule* for S.D.E.s. This chain rule is the *Itô Formula*.

5.2.1   Itô Formula

Consider an S.D.E. of the form

$$dX_t = \alpha(t, X_t)\, dt + \beta(t, X_t)\, dW_t \,.$$

Let $G(t, x)$ be a real valued function with continuous partial derivatives

$$G_t = G_t(t, x) = \frac{\partial G}{\partial t}\,, \qquad G_x = G_x(t, x) = \frac{\partial G}{\partial x}\,, \qquad G_{xx} = G_{xx}(t, x) = \frac{\partial^2 G}{\partial x^2}\,,$$

then

$$
\begin{aligned}
dG(t, X_t) \quad = \quad & \left(G_t(t, X_t) + \alpha(t, X_t) G_x(t, X_t) + \frac{1}{2} \beta^2(t, X_t) G_{xx}(t, X_t) \right) dt + \\
& + \beta(t, X_t) G_x(t, X_t)\, dW_t \,.
\end{aligned}
$$

Sketch proof (not examinable)

To give an indication of why this is the case, write

$$dG(t, x) \approx G(t + \Delta t, x + \Delta x) - G(t, x)\,.$$

Now take a Taylor series expansion of $G(t + \Delta t, x + \Delta x)$ about (t, x) to give

$$dG(t, x) \quad \approx \quad \Delta t\, G_t + \Delta x\, G_x + \frac{1}{2}(\Delta t)^2\, G_{tt} + \frac{1}{2}\Delta t \Delta x\, G_{tx} + \frac{1}{2}(\Delta x)^2\, G_{xx} + \ldots$$

where $G_{tt} = \partial^2 G/\partial t^2$ and $G_{tx} = \partial^2 G/\partial t \partial x$. Now replace Δx with $\alpha \Delta t + \beta \Delta W$ and $(\Delta x)^2$ by $\alpha^2(\Delta t)^2 + 2\alpha\beta\Delta t \Delta W + \beta^2(\Delta W)^2$ to obtain

$$
\begin{aligned}
dG(t, x) \quad \approx \quad & \Delta t\, G_t + (\alpha \Delta t + \beta \Delta W)\, G_x + \frac{1}{2}(\Delta t)^2\, G_{tt} + \frac{1}{2}(\alpha(\Delta t)^2 + \beta \Delta t \Delta W)\, G_{tx} + \\
& + \frac{1}{2}(\alpha^2(\Delta t)^2 + 2\alpha\beta\Delta t \Delta W + \beta^2(\Delta W)^2)\, G_{xx} + \ldots
\end{aligned}
$$

Now approximate $(\Delta W)^2$ by Δt (and note in fact that $\mathrm{E}[(\Delta W^2)] = \Delta t$) to write the preceding expression as

$$dG(t, x) \quad \approx \quad \left(G_t + \alpha G_x + \frac{\beta^2}{2} G_{xx} \right) \Delta t + \beta G_x \Delta W + o(\Delta t)\,.$$

Itô's formula then follows by letting $\Delta t \to 0$. The above sketch proof is not examinable.

Comments

- We can write down Itô's formula for a function of Brownian motion W_t by noting that W_t satisfies an S.D.E. with $X_t = W_t$, $\alpha(t, X_t) = 0$ and $\beta(t, X_t) = 1$. Hence for a function $G(t, W_t)$ we obtain

$$dG(t, W_t) = \left(G_t(t, W_t) + \frac{1}{2} G_{xx}(t, W_t) \right) dt + G_x(t, W_t) \, dW_t \, . \quad (1)$$

This is known as the *special case of Itô formula* and is appropriate for deriving the S.D.E. satisfied by a function of t and W_t.

Example 5.1 (Two applications of Itô formula)

Deriving the S.D.E. satisfied by G.B.M.

Consider again the task of formulating geometric Brownian motion (to model stock price S_t) as an S.D.E. We can write $S_t = \exp(X_t)$ where

$$X_t = X_0 + \left(\mu - \frac{\sigma^2}{2} \right) t + \sigma \, W_t$$

and $X_0 = \log(S_0)$. The S.D.E. satisfied by the X_t process is then given by

$$dX_t = \left(\mu - \frac{\sigma^2}{2} \right) dt + \sigma \, dW_t \, .$$

Hence, we derive an S.D.E. for S_t by applying Itô's formula with $G(t, x) = e^x$. Identify

$$G_t = 0 \, , \qquad G_x = e^x \, , \qquad G_{xx} = e^x \, .$$

Hence

$$G_t(t, X_t) = 0 \, , \qquad G_x(t, X_t) = e^{X_t} \, , \qquad G_{xx}(t, X_t) = e^{X_t}$$

and applying Itô's formula gives

$$
\begin{aligned}
d\left(e^{X_t} \right) &= \left(\left[\mu - \frac{\sigma^2}{2} \right] e^{X_t} + \frac{1}{2} \sigma^2 e^{X_t} \right) dt + \sigma e^{X_t} \, dW_t \\
\Rightarrow d\left(e^{X_t} \right) &= \mu e^{X_t} \, dt + \sigma e^{X_t} \, dW_t \, .
\end{aligned}
$$

Now, writing $S_t = e^{X_t}$, we obtain the usual S.D.E. for geometric Brownian motion as

$$dS_t \;\; = \;\; \mu S_t \, dt + \sigma S_t \, dW_t \, .$$

Solving the S.D.E. satisfied by G.B.M.

We can solve the above S.D.E. explicitly (in the sense that we can obtain a closed form expression for S_t). Take the S.D.E. satisfied by G.B.M. and write

$$
\begin{aligned}
\frac{1}{S_t} dS_t &\;\; = \;\; \mu \, dt + \sigma \, dW_t \, , \\
\Rightarrow \int_0^t \frac{1}{S_\tau} dS_\tau &\;\; = \;\; \int_0^t \mu \, d\tau + \sigma \int_0^t dW_\tau \, ,
\end{aligned}
$$

which suggests that the solution involves $\log(S_t)$. We can check this hunch via Itô formula. Apply the Itô formula to $\log(S_t)$ by taking $G(t,x) = \log(x)$ and identify

$$G_t(t, S_t) = 0\,, \qquad G_x(t, S_t) = \frac{1}{S_t}\,, \qquad G_{xx}(t, S_t) = -\frac{1}{S_t^2}\,.$$

Hence,

$$
\begin{aligned}
d\left(\log(S_t)\right) &= \left(\mu S_t \frac{1}{S_t} - \frac{1}{2} S_t^2 \frac{\sigma^2}{S_t^2}\right) dt + \sigma S_t \frac{1}{S_t}\, dW_t \\
&= \left(\mu - \frac{1}{2}\sigma^2\right) dt + \sigma\, dW_t\,.
\end{aligned}
$$

Our hunch is just about correct, except for the extra term involving σ^2 appearing on the RHS. Nevertheless, we may still proceed by integrating both sides of the preceding equation to give

$$
\begin{aligned}
\int_0^t d\left(\log(S_\tau)\right) &= \int_0^t \left(\mu - \frac{1}{2}\sigma^2\right) d\tau + \int_0^t \sigma\, dW_\tau \\
\Rightarrow \log\left(\frac{S_t}{S_0}\right) &= \left(\mu - \frac{1}{2}\sigma^2\right) t + \sigma\, W_t \\
\Rightarrow S_t &= S_0 \exp\left\{\left(\mu - \frac{1}{2}\sigma^2\right) t + \sigma\, W_t\right\}
\end{aligned}
$$

which is exactly what we would expect.

Example 5.2

Using Itô's formula, solve

$$\int_0^t W_\tau\, dW_\tau\,.$$

Solution

We would expect the solution to the above integral to contain a W_t^2 term. Therefore we find $d(W_t^2)$ via Itô's formula for functions of Brownian motion, given by equation (1).

Set $G(t,x) = x^2$ and calculate

$$G_t(t, W_t) = 0\,, \qquad G_x(t, W_t) = 2W_t\,, \qquad G_{xx}(t, W_t) = 2\,.$$

Hence, we obtain

$$
\begin{aligned}
d(W_t^2) &= dt + 2W_t\, dW_t \\
\Rightarrow W_t\, dW_t &= \frac{1}{2} d(W_t^2) - \frac{1}{2} dt \\
\Rightarrow \int_0^t W_\tau\, dW_\tau &= \frac{1}{2} \int_0^t d(W_\tau^2) - \frac{1}{2} \int_0^t d\tau \\
&= \frac{1}{2} W_t^2 - \frac{1}{2} t\,.
\end{aligned}
$$

Example 5.3

Show that $Y_t = t^3 + t^2 W_t^4$ is an Itô process by writing it in the form

$$dY_t = \alpha_t\, dt + \beta_t\, dW_t$$

for suitable choices of the processes α_t and β_t.

Solution

We use the special case of Itô's formula (1) with $G(t, x) = t^3 + t^2 x^4$. Hence

$$G_t(t, W_t) = 3t^2 + 2tW_t^4\,,\ \ G_x(t, W_t) = 4t^2 W_t^3\,,\ \ G_{xx}(t, W_t) = 12t^2 W_t^2$$

and we get

$$d(t^3 + t^2 W_t^4) = dY_t = \left(3t^2 + 2tW_t^4 + 6t^2 W_t^2\right) dt + 4t^2 W_t^3\, dW_t$$

which is in the form required.

5.3 Solving S.D.E.s numerically

Consider the task of generating a skeleton path on $[0, T]$ of a process satisfying an S.D.E. of the form given by

$$dX_t = \alpha(t, X_t)\, dt + \beta(t, X_t)\, dW_t\,.$$

This would typically involve solving the S.D.E. analytically to obtain the transition density, and then simulating using this density. For example, a Gaussian transition density is obtained for generalised B.M. and a log-normal transition density is obtained for geometric B.M. However, for all but the most trivial of cases, solving the S.D.E. is impossible. We therefore seek to approximate a skeleton sample path of X_t by a single realisation of a numerical solution; although many types of numerical solution exist, only the simplest scheme, the Euler-Maruyama method, is considered here. For a small time increment Δt, the first order Euler-Maruyama approximation of the above S.D.E. is

$$\Delta X_t = \alpha(t, X_t)\, \Delta t\ +\ \beta(t, X_t)\, \Delta W_t\,,$$

where $\Delta W_t = (W_{t+\Delta t} - W_t) \sim \mathrm{N}(0, \Delta t)$. A sample path is then constructed by dividing the time interval $[0, T]$ into n equidistant points, $0 = t_0 < t_1 < \ldots < t_n = T$ (so that $\Delta t = t_{i+1} - t_i$ and $\Delta W_{t_i} = W_{t_i + \Delta t} - W_{t_i}$). Denoting the numerical solution at times $t_0, \ldots, t_n$ by $X_0, \ldots, X_n$ and using the equation directly above yields the recursion,

$$X_i = X_{i-1} + \alpha(t_{i-1}, X_{i-1})\Delta t + \beta(t_{i-1}, X_{i-1})\, \Delta W_{t_{i-1}}\,,\ \ X_{t_0} = X_0\,,\ i = 1, \ldots, n\,.$$

Algorithmically:-

1. Initialise with X_0. Put $i := 1$

2. Simulate $X_i \sim N(X_{i-1} + \alpha(t_{i-1}, X_{i-1})\Delta t, \beta^2(t_{i-1}, X_{i-1})\Delta t)$

3. If $t_i = T$, stop otherwise put $i := i + 1$ and go to step 2.

As an example, consider the S.D.E. formulation of a geometric Brownian motion process (denoted by X_t here),

$$dX_t = \mu X_t \, dt + \sigma X_t \, dW_t.$$

To make the distinction between the exact solution and an approximate solution (obtained via Euler-Maruyama), consider a time interval $[t, t + \Delta t]$ and suppose that we have $X_t = x_t$ at time t. The analytic solution over this time interval is

$$X_{t+\Delta t} = x_t \exp \left\{ (\mu - \sigma^2/2)\Delta t + \sigma(W_{t+\Delta t} - W_t) \right\}$$

for which we see

$$X_{t+\Delta t} | X_t = x_t \sim LN \left(\log x_t + (\mu - \sigma^2/2)\Delta t, \sigma^2 \Delta t \right).$$

The Euler-Maruyama approximation gives

$$X_{t+\Delta t} = x_t + \mu x_t \Delta t + \sigma x_t (W_{t+\Delta t} - W_t)$$

for which we see

$$X_{t+\Delta t} | X_t = x_t \sim N \left(x_t + \mu x_t \Delta t, \sigma^2 x_t^2 \Delta t \right).$$

Although beyond the scope of the course, taking the power series characterisation of the exponential function and truncating terms can be used to obtain the Euler-Maruyama approximation from the analytic solution.

The following suite of R functions generate a numerical solution to the S.D.E. formulation of geometric Brownian motion:

```
itosim=function(T=20,dt=0.01,x0=40,afun=alpha,bfun=beta)
{
 n=T/dt
 xvec=vector("numeric",n+1)
 xvec[1]=x0
 for(i in 2:(n+1))
 {
  t=i*dt
  xvec[i]=xvec[i-1]+afun(xvec[i-1],t)*dt+bfun(xvec[i-1],t)*rnorm(1,0,sqrt(dt))
 }
 xvec
}

alpha=function(x,t)
{
 mu=0.1
 mu*x
```

```
}

beta=function(x,t)
{
 sigma=0.2
 sigma*x
}
```

Note that this flexible setup allows us to simulate from any (univariate) S.D.E. simply by re-writing the `alpha` and `beta` functions. The main `itosim` function remains unchanged. We plot the numerical solution with
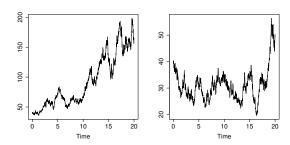
```
plot(ts(itosim(),start=0,deltat=0.01))
```


Figure 1: 2 simulated realisations of a geometric B.M. with $x_0 = 40$, $\mu = 0.1$ and $\sigma = 0.2$ found by numerically solving the S.D.E.

Comment

- Taking Δt smaller and smaller gives increased accuracy at greater computational expense.

5.3.1 Assessing the accuracy of the Euler-Maruyama scheme (not examinable)

Suppose that we're interested in the accuracy of the numerical scheme at some (maturity) time T. Let $X_T^{\Delta t}$ denote the Euler-Maruyama approximation of X_T at time T, based on a discretisation (step-size) of Δt.

The numerical scheme is *strongly convergent* if

$$\lim_{\Delta t \to 0} \mathrm{E}\left(|X_T - X_T^{\Delta t}|\right) = 0.$$

The numerical scheme is *weakly convergent* if

$$\lim_{\Delta t \to 0} \left|\mathrm{E}\{g(X_T)\} - \mathrm{E}\{g(X_T^{\Delta t})\}\right| = 0$$

for polynomials $g(\cdot)$.

The Euler-Maruyama scheme is strongly and weakly convergent subject to suitable conditions on the drift and diffusion coefficient. We may also wonder

how good the approximation is for a particular Δt. The numerical scheme is strongly convergent with order δ if

$$\mathrm{E}\left(|X_T - X_T^{\Delta t}|\right) \leq C_T(\Delta t)^\delta$$

where the constant C_T depends on T and the considered SDE. The numerical scheme is weakly convergent with order δ if

$$\left|\mathrm{E}\{g(X_T)\} - \mathrm{E}\{g(X_T^{\Delta t})\}\right| \leq C_T^g(\Delta t)^\delta$$

where C_T^g depends on T, g and the considered SDE. If a numerical scheme is convergent with order δ and we make the step size k times smaller, the approximation error will decrease by a factor k^δ.

The Euler-Maruyama scheme is weakly convergent with order 1 and strongly convergent with order 0.5. Therefore, the order equal 1 means that if we want to decrease the error 100 times, we have to make the step 100 times smaller. The order equal to 0.5 means that if we want to decrease the error 100 times, we have to make the step $100^2 = 10000$ times smaller. Consequently, the Euler-Maruyama scheme is useful for pricing options whose payoff is not path-dependent. Where the payoff is path-dependent, a higher order scheme (e.g. the Milstein scheme) should be used.

5.4 Models of Interest Rate

Interest rate derivatives are instruments whose payoffs are dependent on the level of interest rates. Such derivatives can be more difficult to value than equity or exchange rate derivatives because

- the behaviour of an individual interest rate is more complicated than that of a stock price,

- interest rates are used for discounting as well as for defining the payoff from the derivative.

In this Section, we will consider several models which provide a description of how short-term interest rate changes through time. The short rate r_t is the rate that applies to an infinitesimally short period of time at time t.

We will describe the short rate as an Itô process of the form

$$dr_t = \alpha(r_t)\, dt + \beta(r_t)\, dW_t$$

where the drift and diffusion coefficients are independent of time. We will consider three models

1. $\alpha(r_t) = \mu r_t$, $\beta(r_t) = \sigma r_t$ (Rendleman and Bartter model),

2. $\alpha(r_t) = a(b - r_t)$, $\beta(r_t) = \sigma$ (Vasicek model),

3. $\alpha(r_t) = a(b - r_t)$, $\beta(r_t) = \sigma \sqrt{r_t}$ (Cox, Ingersoll and Ross model).

5.4.1 Rendleman and Bartter Model

In the Rendleman and Bartter model, the short rate is governed by the S.D.E.

$$dr_t = \mu r_t\, dt + \sigma r_t\, dW_t\,.$$

This means that r_t follows a geometric Brownian motion and is the same type of process that we assumed for stock price in Section 1. This assumption is a good starting point; for example, the model ensures that interest rates are positive. However, some key properties are also lacking, such as mean reversion.

Definition 5.3 (mean reversion)

When interest rates are high, the economy tends to slow down and there is a low demand for funds from borrowers. As a result, rates go down. When rates are low, there is a high demand from borrowers and rates tend to rise. This is known as *mean reversion*. The effect is negative drift when r_t is high and positive drift when r_t is low. Overall, interest rates appear to be pulled back to some long-run average level over time.

Comment

The Rendleman and Bartter model does not incorporate mean reversion since plainly,

$$\mathrm{E}\,(r_t) = r_0 \exp\,(\mu t) \to \infty \qquad \text{as } t \to \infty$$

for μ positive. If μ is negative,

$$\mathrm{E}\,(r_t) = r_0 \exp\,(\mu t) \to 0 \qquad \text{as } t \to \infty.$$

Neither scenario is compatible with the definition of mean reversion.

5.4.2 Vasicek Model

Here, the short rate is modelled by

$$dr_t = a(b - r_t)\, dt + \sigma\, dW_t\,,$$

where a, b and σ are positive real numbers. Note that the model incorporates mean reversion and r_t is pulled to a level b at rate a. The process is also known as a Gaussian Ornstein-Uhlenbeck process.
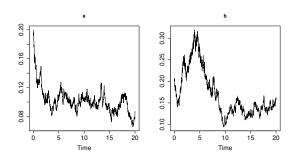

Figure 2: 2 simulated skeleton paths from the Vasicek model with $b = 0.1$, $\sigma = 0.02$ and (a) $a = 1$, and (b) $a = 0.1$

To see that the Vasicek model is mean reverting, we apply Itô's formula with $G(t, x) = xe^{at}$. We obtain

$$G_t = xae^{at}\,, \qquad G_x = e^{at}\,, \qquad G_{xx} = 0\,.$$

Hence,

$$G_t(t, r_t) = ar_t e^{at}\,, \qquad G_x(t, r_t) = e^{at}\,, \qquad G_{xx}(t, r_t) = 0$$

Thus,

$$
\begin{aligned}
d\left(r_t e^{at}\right) &= \left(ar_t e^{at} + a(b - r_t)e^{at}\right) dt + \sigma e^{at} dW_t \\
&= abe^{at} dt + \sigma e^{at} dW_t \\
\Rightarrow r_t e^{at} - r_0 &= \int_0^t abe^{a\tau} d\tau + \sigma \int_0^t e^{a\tau} dW_\tau \\
\Rightarrow r_t e^{at} &= r_0 + [be^{a\tau}]_0^t + \sigma \int_0^t e^{a\tau} dW_\tau \\
\Rightarrow r_t &= r_0 e^{-at} + b(1 - e^{-at}) + \sigma e^{-at} \int_0^t e^{a\tau} dW_\tau
\end{aligned}
$$

To proceed, we need (the first part) of the following *result* for square-integrable functions $g(t)$.

Result 5.1

(i) $\mathrm{E}\left(\int_0^t g(\tau) dW_\tau\right) = 0$

12

(ii) $\mathrm{E}\left(\left[\int_0^t g(\tau)dW_\tau \right]^2 \right) = \int_0^t g^2(\tau)d\tau.$

Note that part (ii) is known as the *Itô isometry*. Justification of this result is left as an exercise. Combining the two results and noting that Itô integrals involve combinations of normal random variables leads to part (iii) that

(iii)
$$\int_0^t g(\tau)dW_\tau \sim N\left(0, \int_0^t g^2(\tau)d\tau \right).$$

Now, returning to the Vasicek model and using the fact that $\mathrm{E}\left(\int_0^t g(\tau)dW_\tau \right) = 0$, we have
$$\mathrm{E}(r_t) = r_0 e^{-at} + b(1 - e^{-at})$$

which is just b in the limit as $t \to \infty$. The Vasicek model is mean reverting.

As an interesting exercise, we may also find the variance of r_t. We have using (ii) of Result 5.1 that

$$\begin{aligned} Var(r_t) &= \sigma^2 e^{-2at} \int_0^t e^{2a\tau} d\tau \\ &= \frac{\sigma^2}{2a} \left(1 - e^{-2at} \right). \end{aligned}$$

Part (iii) of Result 5.2 then gives

$$r_t | r_0 \sim N\left(r_0 e^{-at} + b(1 - e^{-at}), \frac{\sigma^2}{2a} \left(1 - e^{-2at} \right) \right).$$

This distribution can be sampled recursively to give a skeleton path of the process (that is, a realisation of the process at discrete times). Two such realisations can be seen in Figure 2. The effect of mean reversion is apparent, as is the effect of a, since the larger the value of a results in stronger reverting behaviour. It should also be clear that the Vasicek model allows for negative interest rates, which is an undesirable property.

5.4.3   Cox, Ingersoll and Ross Model

In the model proposed by Vasicek, the short term rate can become negative. Whilst this is not considered impossible by some (in practical terms), others consider it to be unsatisfactory. Cox, Ingersoll and Ross propose an alternative model where rates are always nonnegative,

$$dr_t = a(b - r_t)\, dt + \sigma \sqrt{r_t}\, dW_t\,,$$

where a, b and σ are positive real numbers. The process is mean reverting and the standard deviation of the change in interest rate in a short time period is proportional to $\sqrt{r_t}$. Consequently, as short rate increases, so does the standard deviation. Although the S.D.E. can be solved analytically, this task is beyond the scope of the course. Therefore, skeleton paths were generated using the Euler-Maruyama scheme and plotted in Figure 3. Finally, note that the process is non-negative. To see this intuitively (rather than formally), note that the
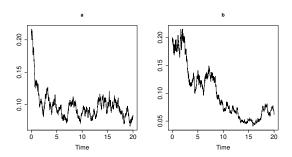
Figure 3: 2 simulated skeleton paths from the CIR model with $b = 0.1$, $\sigma = 0.02$ and (a) $a = 1$, and (b) $a = 0.1$

infinitesimal variance $\sigma^2 r_t$ tends to zero as the process approaches zero (from above). We expect this to be the case for non-negative processes, otherwise fluctuations around zero due to the non-negligible stochastic noise term near zero will send the process negative.

5.5 Stochastic Volatility Models (not examinable)

We have seen in Section 3 that it may be unreasonable to assume that the volatility of a stock is constant over time. Here, we use the S.D.E. formalism to construct a class of models whose volatility is also a stochastic process. Such models are called *stochastic volatility (S.V.) models*. We will consider models of the form

$$
\begin{aligned}
dS_t &= (\mu - q)S_t\, dt + S_t\sqrt{V_t}\, dW_{1,t} \\
dV_t &= a(b - V_t)\, dt + \sigma\sqrt{V_t}\, dW_{2,t}
\end{aligned}
$$

where S_t is an asset's price, V_t is a volatility process, μ, q, a, b, σ are constants and $W_{1,t}$, $W_{2,t}$ are uncorrelated Brownian motions. Clearly, V_t is of C.I.R. type and S_t (conditional on V_t) is of G.B.M. type. Note that we can write this model in the form

$$
dX_t = \alpha(X_t)\, dt + \beta(X_t)\, dW_t
$$

by setting

$$
X_t = \begin{pmatrix} S_t \\ V_t \end{pmatrix}, \ \alpha(X_t) = \begin{pmatrix} (\mu - q)S_t \\ a(b - V_t) \end{pmatrix}, \ \beta(X_t) = \begin{pmatrix} S_t\sqrt{V_t} & 0 \\ 0 & \sigma\sqrt{V_t} \end{pmatrix}, \ dW_t = \begin{pmatrix} dW_{1,t} \\ dW_{2,t} \end{pmatrix}.
$$

By adopting the notation of Section 5.3 for a partition of $[0, T]$, the Euler approximation for this model is given by

$$
\begin{aligned}
S_i &= S_{i-1} + (\mu - q)S_{i-1}\,\Delta t + S_{i-1}\sqrt{V_{i-1}}\,\Delta W_{1,t} \\
V_i &= V_{i-1} + a(b - V_{i-1})\,\Delta t + \sigma\sqrt{V_{i-1}}\,\Delta W_{2,t}
\end{aligned}
$$

for $i = 1, \ldots, n$ with $\Delta W_{1,t} \sim \mathrm{N}(0, \Delta t)$ and $\Delta W_{2,t} \sim \mathrm{N}(0, \Delta t)$ uncorrelated. Since the volatility process is independent of the price process, we simulate a skeleton path from the S.V. model with the following algorithm:

1. Initialise with S_0, V_0.

2. For $i = 1, \ldots, n$, draw $V_i \sim \mathrm{N}(V_{i-1} + a(b - V_{i-1})\,\Delta t\,,\ \sigma^2 V_{i-1}\Delta t)$.

3. For $i = 1, \ldots, n$, draw $S_i \sim \mathrm{N}(S_{i-1} + (\mu - q)S_{i-1}\,\Delta t\,,\ S_{i-1}^2 V_{i-1}\Delta t)$.

Figure 4 shows a simulated skeleton path from the S.V. model. It is easy to see the effect of the volatility realisation on the stock price realisation.

Naturally, the S.V. model can be used to price options. Since typically no closed form pricing formulae exist when it is assumed that stock price follows the S.V. model, the Monte Carlo methods of Section 3 must be used.

Example 5.4

Suppose that stock price follows a S.V. model with parameters $q = 0.0$, $a = 0.2$, $b = 0.1$ and $\sigma = 0.02$. Suppose further that the initial price of stock is £40, the risk free interest rate is $r = 0.05$, and we wish to find the fair price of the Asian call option with strike price $K = 50$ and maturity $T = 1$ year.

We estimate the fair price via Monte Carlo by performing the following steps:

1. Simulate daily prices $S_0^d, S_1^d, \ldots, S_{252}^d$ in the risk-neutral world. For the S.V. model, this means setting $\mu = r$.
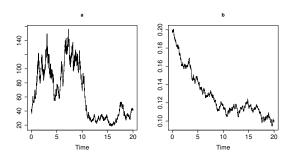
Figure 4: 1 simulated skeleton path from the S.V. model with $\mu = 0.1$, $q = 0.0$, $a = 0.2$, $b = 0.1$ and $\sigma = 0.02$. Plot (a) shows the S_t process and plot (b) shows the V_t process.

2. Calculate the payoff at time $T = 1$:

$$\left(\sum_{i=1}^{252} \frac{S_i^d}{252} - K \right)^+ .$$

3. Repeat steps 1 and 2 to get N sample values of the payoff.

4. Calculate the mean of these sample payoffs to get an estimate of the expected payoff.

5. Discount the expected payoff at the risk-free interest rate ($r = 0.05$ here) to get an estimate of the value of the option.

The following R function performs the algorithm:

```
sv=function(T=1,dt=1/252,s0=c(50,0.1),r=0.05,a=0.2,b=0.1,sigma=0.02,k=50,N=1000)
{
 n=T/dt
 sdt=sqrt(dt)
 s=matrix(0,ncol=2,nrow=n+1)
 payoff=vector("numeric",len=N)
 for(j in 1:N)
 {
  s[1,]=s0
  for(i in 2:(n+1))
  {
   s[i,2]=s[i-1,2]+a*(b-s[i-1,2])*dt+sigma*sqrt(s[i-1,2])*rnorm(1,0,sdt)
  }
  for(i in 2:(n+1))
  {
   s[i,1]=s[i-1,1]+r*s[i-1,1]*dt+sqrt(s[i-1,2])*s[i-1,1]*rnorm(1,0,sdt)
  }
  payoff[j]=max(mean(s[,1])-k,0)
 }
 exp(-r*T)*mean(payoff)
}
```

A single execution of this function gave an estimate of £3.80 for the fair price.

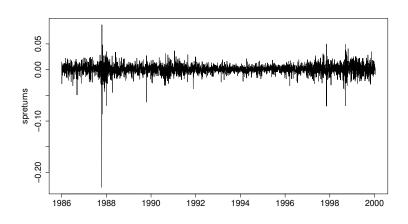## 5.6 Stochastic Volatility Models with Jumps (not examinable)

Whilst the stochastic volatility remedies the constant volatility assumption of G.B.M., it does not allow for any discontinuities in the price process. Such jumps can correspond to a crash in the market (e.g. Black Monday October 1987, 9/11 etc). We therefore extend the S.V. model of the previous Section to incorporate jumps in both prices and volatilities. The discretised model is

$$
\begin{aligned}
S_i &= S_{i-1} + (\mu - q)S_{i-1}\,\Delta t + S_{i-1}\sqrt{V_{i-1}}\,\Delta W_{1,t} + Z_i^s J_i \\
V_i &= V_{i-1} + a(b - V_{i-1})\,\Delta t + \sigma\sqrt{V_{i-1}}\,\Delta W_{2,t} + Z_i^v J_i
\end{aligned}
$$

defined for $i = 1, \ldots, n$. Now $J_i \sim Bernouilli(\Delta t)$ represent the jump times and Z_i^s and Z_i^v are the jump sizes. Typically Z_i^v follows an *Exponential* distribution and Z_i^s follows a *Normal* distribution.

Consider the S&P500 index data of Section 3. The returns are plotted overleaf as well as the estimated volatility process (using a Bayesian estimation method beyond the scope of this course). Clearly, fitting the above model to this dataset gives a jump in volatility on Black Monday.
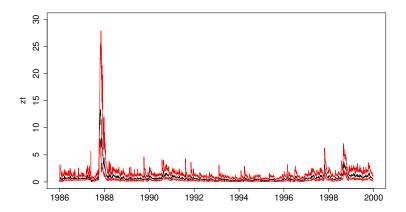
Figure 5: S&P500 index data (left) and filtered volatilities (right) — 2.5 and
97.5 percentiles (red lines) and 50 percentiles (black line).