Derivative Securities

Topics in this section: (a) further discussion of SDE's, including some examples and applications; (b) reduction of Black-Scholes PDE to the linear heat equation; and (c) discussion of what happens when you hedge discretely rather than continuously in time.

When I taught this class in Fall 2000 I discussed barrier options at this point. This time around I prefer to postpone that discussion. But you now have enough background to read about barrier options if you like; see Section 7 of my Fall 2000 notes, or the discussion in the "student guide" by Dewynne, Howison, and Wilmott.

****

Further discussion of stochastic differential equations. Several students requested more information on examples of SDEs' and how they can be used. Therefore the discussion that follows goes somewhat beyond the bare minimum we'll be using in this class. (Everything here is, however, relevant to financial applications.) For simplicity, we restricted the discussion to problems with a "single source of randomness," i.e. scalar SDE's of the form

$$dy = f(y(t),t)\,dt + g(y(t),t)\,dw \tag{1}$$

where w is a scalar-valued Brownian motion. The main things we will use about stochastic integrals and SDE's are the following:

(1) *Ito's lemma.* We discussed in the Section 6 notes the fact that if A is a smooth function of two variables and y solves (1) then $z = A(t, y(t))$ solves the SDE

$$dz = A_t\,dt + A_y\,dy + \tfrac{1}{2}A_{yy}\,dy\,dy = (A_t + A_y f + \tfrac{1}{2}A_{yy}g^2)\,dt + A_y g\,dw.$$

We'll also sometimes use this generalization: if y_1 and y_2 solve SDE's using the *same* Brownian motion w, say

$$dy_1 = f_1\,dt + g_1\,dw \quad \text{and} \quad dy_2 = f_2\,dt + g_2\,dw,$$

and $A(t, y_1, y_2)$ is a smooth function of three variables, then $z = A(t, y_1(t), y_2(t))$ solves the SDE

$$dz = A_t\,dt + A_1\,dy_1 + A_2\,dy_2 + \tfrac{1}{2}A_{11}\,dy_1\,dy_1 + A_{12}\,dy_1\,dy_2 + \tfrac{1}{2}A_{22}\,dy_2\,dy_2$$

with the understanding that

$$A_{ij} = \partial^2 A/\partial y_i \partial y_j \quad \text{and} \quad dy_i\,dy_j = g_i g_j\,dt.$$

The heuristic Taylor-expansion-based explanation is exactly parallel to the one sketched in Section 6.

(2) *A stochastic integral* $\int_a^b F\,dw$ *has mean value zero.* We used (and explained) this assertion at the end of Section 6, but perhaps we didn't emphasize it enough. The explanation is easy. The integrand $F = F(t, y(t))$ can be any function of t and $y(t)$. (The key point: its value at time t should depend only on information available at time t.) The stochastic integral is the limit of the Riemann sums

$$\sum F(t_j, y(t_j))[w(t_{j+1}) - w(t_j)]$$

and each term of this sum has mean value zero, since the increment $w(t_{j+1}) - w(t_j)$ has mean value 0 and is independent of $F(t_j, y(t_j))$.

(3) *Calculating the variance of a* dw *integral.* We just showed that $\int_a^b F\,dw$ has mean value 0. What about its variance? The answer is simple:

$$E\left[\left(\int_a^b F(s, y(s))\,dw\right)^2\right] = \int_a^b E[F^2(s, y(s))]\,ds. \tag{2}$$

The justification is easy. Just approximate the stochastic integral as a sum. The square of the stochastic integral is approximately

$$\begin{aligned}&\left(\sum_{i,j} F(s_i, y(s_i))[w(s_{i+1}) - w(s_i)]\right)\left(\sum_{i,j} F(s_j, y(s_j))[w(s_{j+1}) - w(s_j)]\right)\\ &\quad = \sum_{i,j} F(s_i, y(s_i))F(s_j, y(s_j))[w(s_{i+1}) - w(s_i)][w(s_{j+1}) - w(s_j)] \quad .\end{aligned}$$

For $i \neq j$ the expected value of the i, jth term is 0 (for example, if $i < j$ then $[w(s_{j+1}) - w(s_j)]$ has mean value 0 and is independent of $F(s_i, y(s_i))$, $F(s_j, y(s_j))$, and $[w(s_{i+1}) - w(s_i)]$). For $i = j$ the expected value of the i, jth term is $E[F^2(s_i, y(s_i))][s_{i+1} - s_i]$. So the expected value of the squared stochastic integral is approximately

$$\sum_i E[F^2(y(s_i), s_i)][s_{i+1} - s_i],$$

and passing to the limit $\Delta s \to 0$ gives the formula (2).

The following examples have been extracted from the "Stochastic Calculus Primer" posted at the top of my Spring 2003 PDE for Finance notes.

Log-normal dynamics with time-dependent drift and volatility. Suppose

$$dy = \mu(t)ydt + \sigma(t)ydw \tag{3}$$

where $\mu(t)$ and $\sigma(t)$ are (deterministic) functions of time. What stochastic differential equation describes $\log y$? Ito's lemma gives

$$\begin{aligned} d(\log y) &= y^{-1}dy - \tfrac{1}{2}y^{-2}dydy \\ &= \mu(t)dt + \sigma(t)dw - \tfrac{1}{2}\sigma^2(t)dt. \end{aligned}$$

Remembering that $y(t) = e^{\log y(t)}$, we see that

$$y(t_1) = y(t_0)e^{\int_{t_0}^{t_1}(\mu-\sigma^2/2)ds+\int_{t_0}^{t_1}\sigma dw}.$$

When μ and σ are constant in time we recover the formula (which we already knew):

$$y(t_1) = y(t_0)e^{(\mu-\sigma^2/2)(t_1-t_0)+\sigma(w(t_1)-w(t_0))}.$$

Stochastic stability. Consider once more the solution of (3). It's natural to expect that if μ is negative and σ is not too large then y should tend (in some average sense) to 0. This can be seen directly from the solution formula just derived. But an alternative, instructive approach is to consider the second moment $\rho(t) = E[y^2(t)]$. From Ito's formula,

$$d(y^2) = 2ydy + dydy = 2y(\mu ydt + \sigma ydw) + \sigma^2 y^2 dt.$$

Taking the expectation, we find that

$$E[y^2(t_1)] - E[y^2(t_0)] = \int_{t_0}^{t_1}(2\mu+\sigma)E[y^2]ds$$

or in other words

$$d\rho/dt = (2\mu+\sigma)\rho.$$

Thus $\rho = E[y^2]$ can be calculated by solving this deterministic ODE. If the solution tends to 0 as $t \to \infty$ then we conclude that y tends to zero in the mean-square sense. When μ and σ are constant this happens exactly when $2\mu+\sigma < 0$. When they are functions of time, the condition $2\mu(t) + \sigma(t) \leq -c$ is sufficient (with $c > 0$) since it gives $d\rho/dt \leq -c\rho$.

An example related to Girsanov's theorem. Suppose $\gamma(t)$ depends only on information up to time t. (For example, it could have the form $\gamma(t) = F(t, y(t))$ where y solves an SDE of the form (1).) Then

$$E\left[e^{\int_a^b \gamma(s)dw - \frac{1}{2}\int_a^b \gamma^2(s)ds}\right] = 1.$$

In fact, this is the expected value of $e^{z(b)}$, where

$$dz = -\tfrac{1}{2}\gamma^2(t)dt + \gamma(t)dw, \quad z(a) = 0.$$

Ito's lemma gives

$$d(e^z) = e^z dz + \tfrac{1}{2}e^z dzdz = e^z\gamma dw.$$

So

$$e^{z(b)} - e^{z(a)} = \int_a^b e^z \gamma dw.$$

The right hand side has expected value zero, so

$$E[e^{z(b)}] = E[e^{z(a)}] = 1.$$

Notice the close relation with the previous example "lognormal dynamics": all we've really done is identify the conditions under which $\mu = 0$ in (3).

[Comment for those taking Stochastic Calculus: this example is related to Girsanov's theorem, which gives the relation between the measure on path space associated with drift γ and the measure on path space associated with no drift. The expression

$$e^{\int_a^b \gamma(s)dw - \frac{1}{2}\int_a^b \gamma^2(s)ds}$$

is the Radon-Nikodym derivative relating these measures. The fact that it has expected value 1 reflects the fact that both measures are probability measures.]

The Ornstein-Uhlenbeck process. You should have learned in calculus that the deterministic differential equation $dy/dt + Ay = f$ can be solved explicitly when A is constant. Just multiply by e^{At} to see that $d(e^{At}y)/dt = e^{At}f$ then integrate both sides in time. So it's natural to expect that linear stochastic differential equations can also be solved explicitly. We focus on one important example: the "Ornstein-Uhlenbeck process," which solves

$$dy = -cydt + \sigma dw, \quad y(0) = x$$

with c and σ constant. (This is *not* a special case of (3), because the dw term is not proportional to y.) Ito's lemma gives

$$d(e^{ct}y) = ce^{ct}ydt + e^{ct}dy = e^{ct}\sigma dw$$

so

$$e^{ct}y(t) - x = \sigma\int_0^t e^{cs}dw,$$

or in other words

$$y(t) = e^{-ct}x + \sigma\int_0^t e^{c(s-t)}dw(s).$$

Now observe that $y(t)$ is a Gaussian random variable – because when we approximate the stochastic integral as a sum, the sum is a linear combination of Gaussian random variables. (We use here that a sum of Gaussian random variables is Gaussian; also that a limit of Gaussian random variables is Gaussian.) So $y(t)$ is entirely described by its mean and variance. They are easy to calculate: the mean is

$$E[y(t)] = e^{-ct}x$$

since the "dw" integral has expected value 0. To calculate the variance we use the formula (2). It gives

$$\begin{aligned} E\left[(y(t) - E[y(t)])^2\right] &= \sigma^2 E\left[\left(\int_0^t e^{c(s-t)}dw(s)\right)^2\right] \\ &= \sigma^2\int_0^t e^{2c(s-t)}ds \\ &= \sigma^2\frac{1-e^{-2ct}}{2c}. \end{aligned}$$

We close this example with a brief discussion of the relevance of the Ornstein-Uhlenbeck process. One of the simplest interest-rate models in common use is that of Vasicek, which supposes that the (short-term) interest rate $r(t)$ satisfies

$$dr = a(b - r)dt + \sigma dw$$

with a, b, and σ constant. Interpretation: r tends to revert to some long-term average value b, but noise keeps perturbing it away from this value. Clearly $y = r - b$ is an Ornstein-Uhlenbeck process, since $dy = -aydt + \sigma dw$. Notice that $r(t)$ has a positive probability of being negative (since it is a Gaussian random variable); this is a reminder that the Vasicek model is not very realistic. Even so, its exact solution formulas provide helpful intuition.

****

Reduction of the Black-Scholes PDE to the linear heat equation. The linear heat equation $u_t = u_{xx}$ is the most basic example of a parabolic PDE; its properties and solutions are discussed in every textbook on PDE's. The Black-Scholes equation is really just this standard equation written in special variables. This fact is very well-known; my discussion follows the book by Dewynne, Howison, and Wilmott.

Recall that the Black-Scholes PDE is

$$V_t + \tfrac{1}{2}\sigma^2 s^2 V_{ss} + rsV_s - rV = 0;$$

we assume in the following that r and σ are constant. Consider the preliminary change of variables from (s, t) to (x, τ) defined by

$$s = e^x, \quad \tau = \tfrac{1}{2}\sigma^2(T - t),$$

and let $v(x, \tau) = V(s, t)$. An elementary calculation shows that the Black-Scholes equation becomes

$$v_\tau - v_{xx} + (1 - k)v_x + kv = 0$$

with $k = r/(\frac{1}{2}\sigma^2)$. We've done the main part of the job: reduction to a constant-coefficient equation. For the rest, consider $u(x, t)$ defined by

$$v = e^{\alpha x + \beta\tau} u(x, \tau)$$

where α and β are constants. The equation for v becomes an equation for u, namely

$$(\beta u + u_\tau) - (\alpha^2 u + 2\alpha u_x + u_{xx}) + (1 - k)(\alpha u + u_x) + ku = 0.$$

To get an equation without u or u_x we should set

$$\beta - \alpha^2 + (1 - k)\alpha + k = 0, \quad -2\alpha + (1 - k) = 0.$$

These equations are solved by

$$\alpha = \frac{1-k}{2}, \quad \beta = -\frac{(k+1)^2}{4}.$$

Thus,

$$u = e^{\frac{1}{2}(k-1)x + \frac{1}{4}(k+1)^2\tau} v(x,\tau)$$

solves the linear heat equation $u_\tau = u_{xx}$.

What good is this? Well, it can be used to give another proof of the integral formula for the value of an option (using the fundamental solution of the linear heat equation). It can also be used to understand the sense in which the value of an option at time $t < T$ is obtained by "smoothing" the payoff. Indeed, the solution of the linear heat equation at time t is obtained by "Gaussian smoothing" of the initial data.

****

Discrete-time hedging. My discussion of this topic follows the beginning of a paper by H. E. Leland, *Option pricing and replication with transaction costs*, J. Finance 40 (1985) 1283-1301 (available online through JSTOR). A thoughtful, quite readable discussion of this topic is the paper by E. Omberg, *On the theory of perfect hedging*, Advances in Futures and Options Research 5 (1991) 1-29 (not available online; I'll put a copy on reserve in the CIMS library in the green box with my name).

Suppose an investment bank sells an option and tries to replicate it dynamically, but the bank trades only at evenly spaced time intervals $j\delta t$. (Now δt is positive, not infinitesimal). The bank follows the standard trading strategy of rebalancing to hold $\phi = \partial V/\partial s$ units of stock each time it trades, where V is the value assigned by the Black-Scholes theory. As we shall see in a moment, this strategy is no longer self-financing – but it is *nearly so*, in a suitable stochastic sense, in the limit $\delta t \to 0$.

People often ask, when examining the derivation of the Black-Scholes PDE by examination of the hedging strategy, "Why do we apply Ito's lemma to $V(s(t),t)$ but not to Δ, even though the choice of Δ also depends on $s(t)$?" The answer, of course, is that the hedge portfolio is held fixed from t to $t+\delta t$. The following discussion – in which δt is small but not infinitesimal – should help clarify this point.

OK, let's return to that investment bank. The question is: how much additional money will the bank have to spend over the life of the option as a result of its discrete-time (rather than continuous-time) hedging? We shall answer this by considering each discrete time interval, then adding up the results.

The bank holds a short position on the option and a long position in the replicating portfolio. The value of its position just after rebalancing at any time $t = j\delta t$ is (by hypothesis)

$$0 = -V(s(t),t) + \phi s(t) + [V(s(t),t) - \phi s(t)] = \text{short option} + \text{stock position} + \text{bond position}$$

with $\phi = \frac{\partial V}{\partial s}(s(t), t)$. The value of its position just before the next rebalancing is

$$-V(s(t+\delta t), t+\delta t) + \phi s(t+\delta t) + [V(s(t), t) - \phi s(t)]e^{r\delta t}.$$

The cost (or benefit) of rebalancing at time $t + \delta t$ is minus the value of the preceding expression. Put differently: it is the difference between the two preceding expressions. So it equals

$$\delta V - \phi \delta s - [V - \phi s](e^{r\delta t} - 1).$$

If we estimate δV by Taylor expansion keeping just the terms one normally keeps in Ito's lemma, we get (remembering that $\phi = \partial V/\partial s$)

$$\frac{\partial V}{\partial s}\delta s + \tfrac{1}{2}\frac{\partial^2 V}{\partial s^2}(\delta s)^2 + \frac{\partial V}{\partial t}\delta t - \frac{\partial V}{\partial s}\delta s - rV\delta t + rs\frac{\partial V}{\partial s}\delta t.$$

Notice that the first and fourth terms cancel. Also notice that the substitution $(\delta s)^2 = \sigma^2 s^2 \delta t$ leads to an expression that vanishes, according to the Black-Scholes equation. Thus, the failure to be self-financing is attributable to two sources: (a) errors in the approximation $(\delta s)^2 \approx \sigma^2 s^2 \delta t$, and (b) higher order terms in the Taylor expansion. Our task is to estimate the associated costs.

Collecting the information obtained so far: if the investment bank re-establishes the "replicating portfolio" demanded by the Black-Scholes analysis at each multiple of δt then it incurs cost

$$\tfrac{1}{2}\frac{\partial^2 V}{\partial s^2}(\delta s)^2 + \frac{\partial V}{\partial t}\delta t - rV\delta t + rs\frac{\partial V}{\partial s}\delta t$$

at each time step, plus an error of magnitude $|\delta t|^{3/2}$ due to higher order terms in the Taylor expansion. Using the Black-Scholes PDE, this cost has the alternative expression

$$\tfrac{1}{2}\frac{\partial^2 V}{\partial s^2}[(\delta s)^2 - \sigma^2 s^2 \delta t] \quad \text{plus an error of order } |\delta t|^{3/2}.$$

It can be shown that when $ds = (\mu + \frac{1}{2}\sigma^2)s\,dt + \sigma s\,dw$,

$$\delta s = \sigma s u \sqrt{\delta t} + (\mu + \tfrac{1}{2}\sigma^2)s\delta t \quad \text{plus an error of order } |\delta t|^{3/2}$$

where u is Gaussian with mean 0 and variance 1 (this is closely related to our our discussion of Ito's lemma). Therefore

$$(\delta s)^2 = \sigma^2 s^2 u^2 \delta t \quad \text{plus an error of order } |\delta t|^{3/2}.$$

Thus neglecting the error terms, the cost of refinancing at any given timestep is

$$\tfrac{1}{2}\frac{\partial^2 V}{\partial s^2}\sigma^2 s^2 (u^2 - 1)\delta t$$

where u is Gaussian with mean value 0 and variance 1. This expression is obviously random; its expected value is 0 and its standard deviation is of order δt. Moreover the contributions associated with different time intervals are independent. Notice that the distribution of refinancing costs is *not* Gaussian, since it is proportional to $u^2 - 1$ not u.

Pulling this together: since the expected value of $u^2 - 1$ is zero, the *expected cost* of refinancing at any given timestep is at most of order $|\delta t|^{3/2}$, due entirely to the "error terms." However the *actual cost* (or benefit) of refinancing is larger, a random variable of order δt. But the picture changes when we consider many time intervals. Over $n = T/\delta t$ intervals, the terms $\frac{1}{2}\frac{\partial^2 V}{\partial s^2}\sigma^2 s^2(u^2 - 1)\delta t$ accumulate to a sum

$$\sum_{j=1}^{n} \tfrac{1}{2}\sigma^2 s^2(t_j)\frac{\partial^2 V}{\partial s^2}(s(t_j), t_j)(u_j^2 - 1)\delta t$$

with mean 0 and standard deviation of order $\sqrt{n\delta t^2} = \sqrt{T\delta t}$; the sum is still random, but it's small, statistically speaking, if δt is close to zero, by a sort of law-of-large-numbers. (Notice the resemblance of this argument to our explanation of Ito's lemma. That's no accident: we are in essence deriving Ito's formula all over again.) We've been ignoring the error terms – but they cause no trouble, because they too accumulate to a term of order $\sqrt{\delta t}$, because $n(\delta t)^{3/2} = T\sqrt{\delta t}$.

Final conclusion: the errors of refinancing tend to self-cancel, by a sort of law-of-large-numbers, since their mean value is 0. The net effect, when δt is small, is random but small — in the sense that its mean and standard deviation are of order $\sqrt{\delta t}$.

We have argued that the cost of refinancing tends to zero as $\delta t \to 0$. A recent article by A. Lo, D. Bertsimas, and L. Kogan goes further, examining the statistical distribution of refinancing costs when δt is small. (The relation between their work and the preceding discussion is like the relation between the central limit theorem and the law of large numbers.) The reference is: J. Financial Economics 55 (2000) 173-204 (available online through sciencedirect.com).