Derivative Securities -- Spring 2007 — Section 6

Notes by Robert V. Kohn, modified by Steve Allen
Courant Institute of Mathematical Sciences

Stochastic differential equations and the Black-ScholesPDE. We derived the Black-Scholes formula by using no-arbitrage-based (risk-neutral) valuation in a discrete-time, binomial tree setting, then passing to a continuum limit. We started that way because binomial trees are very explicit and transparent. However the power of the discrete framework as a conceptual tool is rather limited. Therefore we now begin developing the more powerful continuous-time framework, via the Ito calculus and the Black-Scholes differential equation. This material is discussed in many places. Baxter & Rennie emphasize risk-neutral expectation, avoiding almost completely the discussion of PDE's. The "student guide" by Wilmott, Howison, & Dewynne takes almost the opposite approach: it emphasizes PDE's, avoiding almost completely the discussion of risk-neutral expectation. Neftci's book provides a good introduction to Brownian motion, the Ito calculus, stochastic differential equations, and their relation to option pricing, at a level that should be accessible to students in this class. (Students taking Stochastic Calculus will learn this material and much more over the course of the semester.) A brief survey of stochastic calculus (similar in spirit to what's here, but with more detail and more examples) can be found at the top of Bob Kohn's Spring 2003 PDE for Finance course notes (on his web page).

****

W**hy work in continuous time?** Our discrete-time approach has the advantage of being very clear and explicit. However there is a different approach, based on Taylor expansion, the Ito calculus, and the "Black-Scholes differential equation." It has its own advantages:

- Passing to the continuous time limit is clearly legitimate for describing the forward price process. But is it legitimate for describing the value of the option, as determined by arbitrage? This is less clear, since a continuous-time hedging strategy is unattainable in practice. A continuous time process can be the limit of a trinomial branching process, for which the arbitrage hedging argument does not apply, just as easily as it can be the limit of a binomial hedging process. In what sense can we "approximately replicate" the option by trading at discrete times? The Black-Scholes differential equation will help us answer these questions.
- The differential equation approach gives fresh insight and computational flexibility. Imagine trying to understand the implications of compound interest without using the differential equation $df/dt = rf$! (Especially: imagine how stuck you'd be if r depended on f.)
- Differential-equation-based methods lead to efficient computational schemes (and even explicit solution formulas in some cases) not only for European options, but also for more complicated instruments such as barrier options.

Brownian motion. Recall our discussion of the lognormal hypothesis for forward price dynamics. It says that log $[s(t_2)/s(t_1)]$ is a Gaussian random variable with mean $\mu(t_2 - t_1)$ and variance $\sigma^2 (t_2 - t_1)$, and disjoint intervals give rise to independent random variables.

A time-dependent random variable is called a *stochastic process*. The lognormal hypothesis is related to Brownian motion $z(t)$, also known as the Wiener process, which satisfies:

(a) $z(t_2)$ - $z(t_1)$ is a Gaussian random variable with mean value 0 and variance t_2 - t_1;

(b) distinct intervals give rise to independent random variables;

(c) $z(0) = 0$.

It can be proved that these properties determine a unique stochastic process, i.e. they uniquely determine the probability distribution of any expression involving $z(t_1), z(t_2),\cdots,$ $z(t_N)$. Also: for almost any realization, the function $t \rightarrow z(t)$ is continuous but not differentiable. The process $z(t)$ can be viewed as a limit of suitably scaled random walks (we showed this in Section 4). Another important fact: writing $z(t_2)\;\; z(t_1) = \Delta z$ and t_2 - t_1= Δt,

$$E [\mid \Delta z|^j] = C_j \mid \Delta t|^{j/2}, j = 1,2,3),\cdots.$$

Our lognormal hypothesis can be reformulated as the statement that

$$F_{\mathrm{t}} = F_0 \mathrm{e}^{[\mu t + \sigma z(t)]}.$$

Stochastic differential equations and Ito's lemma. Let's first review ordinary differential equations. Consider the ODE $dy / dt = f(y,t)$ with initial condition $y(0) = y_0$. It is a convenient mnemonic to write the equation in the form

$$dy = f(y,t)\, dt.$$

This reminds us that the solution is well approximated by its (explicit) finite difference approximation $y([j+1]\, \delta t) - y(j\, \delta t) = f(y(j\delta t), j\delta t)\delta t$, which we sometimes write more schematically as

$$\Delta y = f(y,t)\, \Delta t.$$

An extremely useful aspect of ODE's is the ability to use chain rule. From the ODE for $y(t)$ we can easily deduce a new ODE satisfied by any function of $y(t)$. For example, $w(t)=e^{y(t)}$ satisfies $dz/dt = e^y\, dy/dt = w\, f(\log w,t)$. In general, $w = A(y(t))$ satisfies $dw/dt = A'(y)\, dy/dt$. The mnemonic for this is

$$dA(y) = \frac{dA}{dy} dy = \frac{dA}{dy} f(y, t)\, dt.$$

It reminds us of the proof, which boils down to the fact that (by Taylor expansion)

$$\Delta A = A'(y)\Delta y + \text{error of order } |\Delta y|^2.$$

In the limit as the timestep tends to 0 we can ignore the error term, because $|\Delta y|^2 \le C\,|\Delta t|^2$ and the sum of $1/\Delta t$ such terms is small, of order $|\Delta t|$.

OK, now stochastic differential equations. We consider only the simplest class of stochastic differential equations, namely

$$dy = f(y,t)dt + g(y,t)dz, \quad y(0) = y_0,$$

where $z(t)$ is Brownian motion. The solution is a stochastic process, the limit of the processes obtained by the (explicit) finite difference scheme

$$y([j+1]\delta t) - y(j\delta t) = f(y(j\delta t), j\delta t)\delta t + g(y(j\delta t), t)\,(z([j+1]\delta t) - z(j\delta t)),$$

which we usually write more schematically as

$$\Delta y = g(y, t)\Delta z + f(y, t)\Delta t.$$

Put differently (this is how the rigorous theory begins): we can understand the stochastic differential equation by rewriting it in integral form:

$$y(t') = y(t) + \int_t^{t'} f(y(\tau),\tau)d\tau + \int_t^{t'} g(y(\tau),\tau)dz(\tau)$$

where the first integral is a standard Riemann integral, and the second one is a *stochastic integral*:

$$\int_t^{t'} g(y(\tau),\tau)dz(\tau) = \lim_{\Delta\tau\to 0} \sum g(y(\tau_i),\tau_i)\left[z(\tau_{i+1}) - z(\tau_i)\right]$$

where $t = \tau_0 < \cdot\cdot\cdot < \tau_N = t'$.

It's easy to see that when μ and σ are constant, $y(t) = \mu t + \sigma z(t)$ solves

$$dy = \mu dt + \sigma dz$$

The analogue of the chain rule calculation done above for ODE's is known as Ito's lemma. It says that if $dy = \mu dt + \sigma dz$ then $w = A(y)$ satisfies the stochastic differential equation

$$dw = A'(y)dy + \tfrac{1}{2}A''(y)\sigma^2 dt = A'(y)\sigma dz + [A'(y)\,\mu + \tfrac{1}{2}A''(y)\,\sigma^2]dt.$$

Here is a heuristic justification: carrying the Taylor expansion of $A(y)$ to second order gives

$$\Delta A = A'(y)\Delta y + \tfrac{1}{2} A''(y)(\Delta y)^2 + \text{error of order } |\Delta y|^3$$
$$= A'(y)\,(\mu\Delta t + \sigma\Delta z) + \tfrac{1}{2} A''(y)\sigma^2(\Delta z)^2 + \text{errors of order } (|\Delta y|^3 + |\Delta z||\Delta t| + |\Delta t|^2).$$

One can show that the error terms are negligible in the limit $\Delta t \to 0$. For example, the sum of $1/\Delta t$ terms of order $|\Delta z|\,|\Delta t|$ has expected value of order $\sqrt{\Delta t}$. Thus

$$\Delta A \approx A'(y)\,(\mu\Delta t + \sigma\Delta z) + \tfrac{1}{2} A''(y)\sigma^2(\Delta z)^2$$

Now comes the subtle part of Ito's Lemma: the assertion that we can replace $(\Delta z)^2$ in the preceding expression by Δt . This is sometimes mistakenly justified by saying "$(\Delta z)^2$ behaves deterministically as $\Delta t \to 0$" - which is certainly not true; in fact $(\Delta z)^2 = u^2\Delta t$ where u is a Gaussian random variable with mean value 0 and variance 1 .

So why can we substitute Δt for $(\Delta x)^2$? This can be thought of as an extension of the law of large numbers. When we solve a difference equation (to approximate a differential equation) we must add the terms corresponding to different time intervals. So we're really interested in *sums* of the form

$$\sum_{j=1}^{N} A''(y(t_j))\sigma^2(t_j)(\Delta z)_j^2$$

with $(\Delta z)_j = z(t_{j+1}) - z(t_j)$ and $N = T/\Delta t$. If A'' and σ^2 were constant then, since the $(\Delta z)_j$ are independent, $\sum_{j=1}^{N}(\Delta z)_j^2 = \sum_{j=1}^{N} u_j^2\Delta t$ would have mean value $N\Delta t = T$ and variance of order $N(\Delta t)^2 = T\,\Delta t$. Thus the sum would have standard deviation $\sqrt{T\Delta t}$, i.e. it is asymptotically deterministic. The rigorous argument is different, of course, since in truth $A''(y)\sigma^2$ is not constant; but the essential idea is similar.

The version of Ito's Lemma stated and justified above is not the most general one - though it has all the main ideas. Similar logic applies, for example, if A is a function of both y and t. Then $z = A(y,t)$ satisfies

$$dw = \frac{\partial A}{\partial y}dy + \frac{\partial A}{\partial t}dt + \frac{1}{2}\frac{\partial A^2}{\partial y^2}\sigma^2 dt = \frac{\partial A}{\partial y}\sigma dz + \left[\frac{\partial A}{\partial y}\mu + \frac{\partial A}{\partial t} + \frac{1}{2}\frac{\partial^2 A}{\partial y^2}\sigma^2\right]dt$$

Let's apply Ito's lemma to find the stochastic differential equation for the forward price process $F(t)$. The lognormal hypothesis says $F = e^y$ where $dy = \mu dt + \sigma dz$. Therefore $dF = e^y(\mu dt + \sigma dz) + \tfrac{1}{2} e^y\sigma^2 dt$, i.e.

$$\tfrac{1}{F}dF = (\mu + \tfrac{1}{2}\sigma^2)dt + \sigma dz\,.$$

The Black-Scholes partial differential equation. Consider a European option with payoff $V(F_T)$ at maturity T. We have a formula for its value at time t from Section 4:

$$\text{value at time } t = e^{-r(T-t)} \frac{1}{\sigma\sqrt{2\Pi(T-t)}} \int_{-\infty}^{\infty} f(F_t e^x) \exp\left[\frac{-(x-\left[r-\frac{1}{2}\sigma^2\right](T-t))^2}{2\sigma^2(T-t)}\right] dx$$

Notice that the value is a function of the present time t and the present forward price F_t, i.e. it can be expressed in the form:

$$\text{value at time } t = V(\mathrm{F}_t, t).$$

for a suitable function $V(F, t)$ defined for $F > 0$ and $t < T$. It's obvious from the interpretation of V that

$$V(F, T) = V_T F_T.$$

The Black-Scholes differential equation says that

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 F^2 \frac{\sigma^2 V}{\partial F^2} - rV = 0$$

It offers an alternative procedure for evaluating the value of the option, by solving the PDE "backwards in time" numerically, using $t = T$ as the initial time.

Recall that in the setting of binomial trees we had two ways of evaluating the value of an option: one by expressing it as a weighted sum over all paths; the other by working backward through the tree. Evaluating the integral formula is the continuous-time analogue of summing over all paths. Solving the Black-Scholes PDE is the continuous-time analogue of working backward through the tree. Recall also that working backward through the tree was a little more flexible - for example it didn't require that the interest rate be constant. Similarly, the Black-Scholes equation can easily be solved numerically even when the interest rate and volatility are (deterministic) functions of time.

Where does the equation come from? We'll give a justification, based on Ito's formula. Examining this derivation you'll be able to see how the Black-Scholes PDE generalizes to more complicated market models (for example when the volatility and drift depend on forward price). However for simplicity we'll present the arguments in the usual constant-volatility, constant-drift setting

$$dF = \sigma F\, dw + (\mu + \tfrac{1}{2}\sigma^2)F dt$$

and we'll continue to assume that the interest rate is constant.

Derivation by considering a hedging strategy. Remember that when hedging in the discrete-time setting, we rebalance the portfolio so that it contains ϕ units of futures, then we let the futures price jump to the new value. (I write ϕ not Δ to avoid confusion, because we have been using Δ for increments.) The analogous procedure in the continuous-time setting is to rebalance at successive at time intervals of length δt, then pass to the limit $\delta t \to 0$. Suppose that after rebalancing at time $j\delta t$ the portfolio contains $\phi = \phi(s(j\delta t), j\delta t)$ units of futures. Consider the value of the option less the value of the future position during the next time interval:

$$\Pi = V - \phi F.$$

Its increment $d\Pi = \Pi([j+1]\delta t) - \Pi(j\delta t)$ is approximately

$$dV - \phi dF = \frac{\partial V}{\partial F} dF + \frac{\partial V}{\partial t} dt + \tfrac{1}{2}\frac{\partial^2 V}{\partial F^2}\sigma^2 F^2 dt - \phi dF$$

$$= \left(\frac{\partial V}{\partial F} - \phi\right)\sigma F dz + \left(\frac{\partial V}{\partial F} - \phi\right)\left(\mu + \tfrac{1}{2}\sigma^2\right)F dt + \left(\frac{\partial V}{\partial t} + \tfrac{1}{2}\frac{\partial^2 V}{\partial F^2}\sigma^2 F^2\right)dt$$

Note that we do not differentiate ϕ because it is being held fixed during this time interval. We know enough to expect that the right choice of ϕ is $\phi(F,t) = \partial V / \partial F$. But if we didn't already know, we'd discover it now: this is the choice that eliminates the dz term on the right hand side of the last equation. Fixing ϕ this way, we see that

$$dV - \phi dF = \left(\frac{\partial V}{\partial t} + \tfrac{1}{2}\frac{\partial^2 V}{\partial F^2}\sigma^2 F^2\right)dt \text{ is deterministic.}$$

Now, the principle of no arbitrage says that a portfolio whose return is deterministic must grow at the risk-free rate. In other words, for this choice of ϕ we must have

$$dV - \phi\, dF = rV\, dt.$$

Combining these equations gives

$$\left(\frac{\partial V}{\partial t} + \tfrac{1}{2}\frac{\partial^2 V}{\partial F^2}\sigma^2 F^2\right) = rV$$

with $\phi = \partial V / \partial F$. This is precisely the Black-Scholes PDE.

Reduction of the Black-Scholes PDE to the linear heat equation.

Black and Scholes originally proved their now-famous formula by reducing the Black-Scholes PDE to the heat equation, a PDE whose structure had been studied for 200 years and for which several well-known solution techniques had been developed. Shortly after the original Black-Scholes paper, a more economically intuitive method was devised to derive the Black-Scholes formula from the Black-Scholes PDE: first, it was shown that the Black-Scholes PDE led to a derivative value equal to the expectation under the risk-neutral distribution discounted at the risk-free rate, and second, the Black-Scholes formula was derived from this result. We have already studied this second derivation in section 5. Next week, when we study martingales, we will derive the first result. For right now, we will show how to reduce the Black-Scholes PDE to the heat equation, but we will not look at the solution techniques for the heat equation (if you are interested in how this is done, you can take a look at Chapter 5 of Wilmott-Dewynne-Howison or Chapter 11 of Steele.)

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; this discussion follows the book by Dewynne, Howison, and Wilmott.

Recall that the Black-Scholes PDE is

$$V_t + \tfrac{1}{2}\sigma^2 F^2 V_{FF} - rV = 0$$

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

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

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

$$v_\tau - v_{xx} + v_x + kv = 0$$

with $k = r / \tfrac{1}{2}\sigma^2$. 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

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

To get an equation without u or u_x we should set

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

These equations are solved by

$$\alpha = \frac{1}{2}; \qquad \beta = k - \frac{1}{4}$$

Thus,

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

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

In addition to giving a proof of the integral formula for the value of an option (using the fundamental solution of the linear heat equation), this 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.