PDE for Finance Notes – Section 5

Notes by Robert V. Kohn, Courant Institute of Mathematical Sciences. For use only in connection with the NYU course PDE for Finance, G63.2706, Spring 1999.

Reminders: No lecture March 31 [Passover]. Last lecture April 28. There will be an in-class final exam on May 5.

More on stochastic differential equations; then the link to partial differential equations. We discuss some some specific stochastic differential equations that are relevant to finance. This provides, among other things, practice using Ito's lemma. Then we develop the link with PDE, specifically the backward Kolmogorov equation and the Feynman-Kac formula. These explain, for example, the equivalence of the two basic formulas for the value of a European option – (a) as the expected discounted payoff (relative to the the risk-neutral probability), and (b) as the solution of the Black-Scholes partial differential equation. Mainly I'll be drawing from Chapter 5 of Fleming and Rishel, with supplementary material from Lamberton and Lapeyre and Oksendal.

Two excellent books have come to my attention. *Basic stochastic processes* by Zdzislaw Brzezniak and Tomasz Zastawniak (Springer-Verlag, new) discusses conditional expectations, martingales, and the Ito calculus starting at a very basic level with lots of exercises; it should be accessible to all students in this class. *Stochastic differential equations* by Bernt Oksendal (Springer-Verlag, 5th edition, 1998, $34.95 paperback) is more difficult and more comprehensive but also excellent. It's about at the level of Arnold's book, but with substantial emphasis on applications to finance. Brzezniak and Zastawniak is in the Courant library (not on reserve); Oksendal is in the Courant library (on reserve). Oksendal recently went out of stock at Springer, but will be back in stock within a couple of weeks I'm told.

****

Stochastic differential equations. We've been discussing stochastic differential equations of the form

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

where w is brownian motion. In the vector-valued case this means

$$dy_i = f_i(y,t)dt + \sum_j g_{ij}(y,t)dw_j$$

where each component of w is an independent Brownian motion. When considering optimal control, it is natural to assume the coefficients f and g depend not only on y and t but also on some (nonanticipating) control $\alpha(t)$. Here we'll concentrate on the case when there is no such dependence, i.e. f and g are functions of y and t alone. (However some of what we do extends to the case when f and g are random but nonanticipating.)

We interpret (1) as meaning

$$y(t_1) - y(t_0) = \int_{t_0}^{t_1} f(y(s),s)\,ds + \int_{t_0}^{t_1} g(y(s),s)\,dw.$$

Of course some conditions are needed on f and g for this to make sense. In particular: we need $E[g^2(y(s),s)]$ to be bounded (or at least square-integrable) as a function of s. This condition is natural, since the theory of stochastic integrals is an L^2-based theory (the stochastic integral is the limit, in the mean-square sense, of certain sums). Following our usual custom of glossing over technical details, we shall assume such conditions without explicitly mentioning them in what follows. Careful statements of most results can be found in Fleming and Rishel.

We already discussed the fact that integrals "dw" have mean value 0:

$$E\left[\int_{t_0}^{t_1} g(y(s),s)\,dw\right] = 0.$$

We'll also need this formula for the variance:

$$E\left[\left(\int_{t_0}^{t_1} g(y(s),s)\,dw\right)^2\right] = \int_{t_0}^{t_1} E[g^2(y(s),s)]\,ds.$$

This is easy to see from the approximation of the stochastic integral as a sum: the square of the stochastic integral is approximately

$$\left(\sum_{i=1}^{N-1} g(y(s_i),s_i)[w(s_{i+1})-w(s_i)]\right)\left(\sum_{j=1}^{N-1} g(y(s_j),s_j)[w(s_{j+1})-w(s_j)]\right)$$
$$= \sum_{i,j=1}^{N-1} g(y(s_i),s_i)g(y(s_j),s_j)[w(s_{i+1})-w(s_i)][w(s_{j+1})-w(s_j)] \qquad .$$

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 $g(y(s_i),s_i)$, $g(y(s_j),s_j)$, and $[w(s_{i+1})-w(s_i)]$). For $i=j$ the expected value of the i,jth term is $E[g^2(y(s_i),s_i)][s_{i+1}-s_i]$. So the expected value of the squared stochastic integral is approximately

$$\sum_{i=1}^{N-1} E[g^2(y(s_i),s_i)][s_{i+1}-s_i],$$

and passing to the limit $\Delta s \to 0$ gives the desired assertion.

Let's practice using Ito's lemma by doing a few interesting things with it.

Redoing an example from Section 4. We showed that

$$\int_{t_0}^{t_1} w\,dw = \tfrac{1}{2}(w^2(t_1)-w^2(t_0)) - \tfrac{1}{2}(t_1-t_0)$$

by directly calculating the stochastic integral as a limit of sums. Ito's lemma gives a much easier proof of the same result: applying it to $\phi(w)=w^2$ gives

$$d(w^2) = 2wdw + dwdw = 2wdw + dt$$

which means $w^2(t_1) - w^2(t_0) = 2\int_{t_0}^{t_1} w\,dw + (t_1 - t_0)$.

Log-normal dynamics. Suppose

$$dy = \mu(t)ydt + \sigma(t)ydw \tag{2}$$

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

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

In particular, if μ and σ are constant in time we get

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

Stochastic stability. Consider once more the solution of (2). 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,

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

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. In financial terms, Girsanov's theorem gives the relation between the "subjective" and "risk-neutral" price processes. We'll discuss it later, and when we do so the following fact will have a natural interpretation:

$$E\left[e^{\int_{t_0}^{t_1}\gamma(s)dw-\frac{1}{2}\gamma^2(s)ds}\right] = 1.$$

In fact, this is the value of e^z, where

$$dz = -\tfrac{1}{2}\gamma^2(t)dt + \gamma(t)dw, \quad z(t_0) = 0.$$

Ito's lemma gives

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

So

$$e^{z(t_1)} - e^{z(t_0)} = \int_{t_0}^{t_1} e^z \gamma dw.$$

The right hand side has mean value zero, so

$$E[e^{z(t_1)}] = E[e^{z(t_0)}] = 1.$$

Notice the close relation with the previous examples: all we've really done is identify the conditions under which $\mu = 0$ in (2).

The Ornstein-Uhlenbeck process. You should have learned in Calculus that the deterministic differential equation $dy/dt + Ay = f$ can be solved explicitly. 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 (2), 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. The variance is

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

We digress to discuss 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.

Historically, the Ornstein-Uhlenbeck process was introduced by physicists Ornstein and Uhlenbeck, who believed that a diffusing particle had brownian *acceleration* not brownian velocity. Their idea was that the position $x(t)$ of the particle at time t should satisfy

$$\begin{aligned} dx &= vdt \\ \epsilon dv &= -vdt + dw \end{aligned}$$

with $\epsilon > 0$ small. As $\epsilon \to 0$, the resulting $x_\epsilon(t)$ converges to a brownian motion process. Formally: when $\epsilon = 0$ we recover $0 = -vdt + dw$ so that $dx = (dw/dt)dt = dw$. Honestly: we claim that $|x_\epsilon(t) - w(t)|$ converges to 0 (uniformly in t) as $\epsilon \to 0$. In fact, writing the equations for the Ornstein-Uhlenbeck process as

$$\begin{aligned} dx_\epsilon &= v_\epsilon dt \\ dw &= v_\epsilon dt + \epsilon dv_\epsilon \end{aligned}$$

then subtracting, we see that

$$d(x_\epsilon - w) = \epsilon dv_\epsilon.$$

Now use our explicit solution formula for the Ornstein Uhlenbeck process to represent v_ϵ in terms of stochastic integrals, ultimately concluding that $\epsilon v_\epsilon(t) \to 0$ as $\epsilon \to 0$. (Details left to the reader.)

****

The link to PDE. We've already seen a link between stochastic differential equations and PDE, in our discussion of optimal control. Our present goal is a little different, though the methods we'll use to attain it are closely related. Technical terms: we'll be discussing the *backward Kolmogorov equation* associated with a diffusion process, and the *Feynman-Kac formula.*

If you know a little finance, you know that the value of a European option can be determined in two different ways: (a) as the expected discounted value of the payoff (with respect to the risk-neutral probability), and (b) as the solution of the Black-Scholes partial differential equation. The backward Kolmogorov equation and the Feynman-Kac formula explain why these two representations are equivalent. (In other words, they give a means of passing from one to the other.) However we will not assume any specific familiarity with finance in what follows.

The backward Kolmogorov equation. Here's the most basic version. Suppose $y(t)$ solves the scalar stochastic differential equation

$$dy = f(y,t)dt + g(y,t)dw,$$

and let

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

be the expected value of some payoff Φ at maturity time $T > t$, given that $y(t) = x$. Then u solves

$$u_t + f(x,t)u_x + \tfrac{1}{2}g^2(x,t)u_{xx} = 0 \text{ for } t < T, \text{ with } u(x,T) = \Phi(x). \tag{3}$$

Sounds familiar, right? It's just like our discussion of stochastic control – except that there is no control, hence no need to maximize over anything.

This is a special case of arguments we've done before. Let's review the explanation anyway. For any function $\phi(y,t)$, Ito's lemma gives

$$\begin{aligned} d(\phi(y(t),t)) &= \phi_y dy + \tfrac{1}{2}\phi_{yy}dydy + \phi_t dt \\ &= (\phi_t + f\phi_y + \tfrac{1}{2}g^2\phi_{yy})dt + g\phi_y dw. \end{aligned}$$

Choosing $\phi = u$, the solution of (3), we get

$$u(y(T),T) - u(y(t),t) = \int_t^T (\phi_t + f\phi_y + \tfrac{1}{2}g^2\phi_{yy})dt + \int_t^T g\phi_y dw.$$

Taking the expected value and using the PDE gives

$$E_{y(t)=x}\left[\Phi(y(T))\right] - u(x,t) = 0$$

which is precisely our assertion.

That was the simplest case. It can be jazzed up in many ways. We discuss some of them:

Vector-valued diffusion. Suppose y solves a vector-valued stochastic differential equation

$$dy_i = f_i(y,t)dt + \sum_j g_{ij}(y,t)dw_j,$$

where each component of w is an independent Brownian motion. Then

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

solves

$$u_t + \mathcal{L}u = 0 \text{ for } t < T, \text{ with } u(x,T) = \Phi(x),$$

where $\mathcal{L}$ is the differential operator

$$\mathcal{L}u(x,t) = \sum_i f_i \frac{\partial u}{\partial x_i} + \tfrac{1}{2} \sum_{i,j,k} g_{ik} g_{jk} \frac{\partial^2 u}{\partial x_i \partial x_j}.$$

The justification is just as in the scalar case, using the multidimensional version of Ito's lemma. The operator $\mathcal{L}$ is called the "infinitesimal generator" of the diffusion process $y(t)$.

The Feynman-Kac formula. We discuss the scalar case first, for clarity. Consider as above the solution of

$$dy = f(y,t)dt + g(y,t)dw$$

but suppose we are interested in a suitably "discounted" final-time payoff of the form:

$$u(x,t) = E_{y(t)=x}\left[e^{-\int_t^T b(y(s))ds} \Phi(y(T)) \right] \tag{4}$$

for some specified function $b(y)$. Then u solves

$$u_t + f(x,t)u_x + \tfrac{1}{2} g^2(x,t) u_{xx} - b(x)u = 0 \tag{5}$$

instead of (3). (Its final-time condition is unchanged: $u(x,T) = \Phi(x)$.) If you know some finance you'll recognize that when y is log-normal and b is the interest rate, (5) is precisely the Black-Scholes partial differential equation.

To explain (5), we must calculate the stochastic differential $d[z_1(s)\phi(y(s),s)]$ where $z_1(s) = e^{-\int_t^s b(y(r))dr}$. The multidimensional version of Ito's lemma gives

$$d[z_1(s)z_2(s)] = z_1 dz_2 + z_2 dz_1 + dz_1 dz_2.$$

We apply this with z_1 as defined above and $z_2(s) = \phi(y(s),s)$. Ito's lemma (or ordinary differentiation) gives

$$dz_1(s) = -z_1 b(y(s))ds$$

and we're already familiar with the fact that

$$\begin{aligned} dz_2(s) &= (\phi_s + f\phi_y + \tfrac{1}{2} g^2 \phi_{yy})ds + g\phi_y dw \\ &= (\phi_s + \mathcal{L}\phi)ds + g\phi_y dw. \end{aligned}$$

Notice that $dz_1 dz_2 = 0$. Applying the above with $\phi = u$, the solution of the PDE (5), gives

$$\begin{aligned} d\left(e^{-\int_t^s b(y(r))dr} u(y(s),s) \right) &= z_1 dz_2 + z_2 dz_1 \\ &= z_1 \left[(u_s + \mathcal{L}u)ds + g u_y dw \right] - z_1 u b ds \\ &= z_1 g u_y dw. \end{aligned}$$

The right hand side has expected value 0, so

$$E_{y(t)=x}[z_1(T)z_2(T)] = z_1(t)z_2(t) = u(x,t)$$

as asserted.

A moment's thought reveals that vector-valued case is no different. The discounted expected payoff (4) solves the PDE

$$u_t + \mathcal{L}u - bu = 0$$

where $\mathcal{L}$ is the infinitesimal generator of the diffusion y.

Running payoff. Suppose we are interested in

$$u(x,t) = E_{y(t)=x}\left[\int_t^T \Psi(y(s),s)ds\right]$$

for some specified function Ψ. Then u solves

$$u_t + \mathcal{L}u + \Psi(x,t) = 0.$$

The final-time condition is $u(x,T) = 0$, since we have included no final-time term in the "payoff." The proof is hardly different from before: by Ito's lemma,

$$\begin{aligned} d[u(y(t),t)] &= (u_t + \mathcal{L}u)dt + \nabla u \cdot g \cdot dw \\ &= -\Psi(y(t),t)dt + \nabla u \cdot g \cdot dw. \end{aligned}$$

Integrating and taking the expectation gives

$$E_{y(t)=x}\left[u(y(T),T)\right] - u(x,t) = E_{y(t)=x}\left[-\int_t^T \Psi(y(s),s)ds\right].$$

This gives the desired assertion, since $u(y(T),T) = 0$.

Boundary value problems and exit times. The preceding examples use stochastic integration from time t to a fixed time T, and they give PDE's that must be solved for all $x \in R^n$. It's also interesting to consider integration from time t to the first time y exits from some specified region. The resulting PDE must be solved on this region, with suitable boundary data.

Let D be a region in R^n. Suppose y is an R^n-valued diffusion solving

$$dy = f(y,s)ds + g(y,s)dw \text{ for } s > t, \text{ with } y(t) = x$$

with $x \in D$. Let

$$\begin{aligned} \tau(x) &= \text{the first time y(s) exits from } D, \text{ if} \\ &\ \text{prior to } T; \text{ otherwise } \tau(x) = T. \end{aligned}$$

This is an example of a "stopping time" (key feature: the statement "$\tau(x) < t$" is $\mathcal{F}_t$-measurable; in other words, knowledge of events up to time t determines whether or not the process has exited from D before time t). Suppose we are interested in

$$u(x,t) = E_{y(t)=x}\left[\int_t^{\tau(x)} \Psi(y(s),s)ds + \Phi(y(\tau(x)),\tau(x))\right].$$

Then u solves

$$u_t + \mathcal{L}u + \Psi = 0 \text{ for } x \in D$$

with boundary condition

$$u(x,t) = \Phi(x,t) \text{ for } x \in \partial D$$

and final-time condition

$$u(x,T) = \Phi(x,T) \text{ for all } x \in D.$$

The justification is entirely parallel to our earlier examples. The only change is that we integrate, in the final step, to the stopping time τ rather than the final time T. (This is permissible for any stopping time satisfying $E[\tau] < \infty$.)

Elliptic boundary-value problems. Now suppose f and g in the stochastic differential equation don't depend on t, and for $x \in D$ let

$$\tau(x) = \text{the first time } y(s) \text{ exits from } D.$$

(Unlike the previous example, we do not impose a final time T). Suppose furthermore the process does eventually exit from D, (more precisely: assume $E[\tau(x)] < \infty$, for all $x \in D$). Then

$$u(x) = E_{y(0)=x}\left[\int_0^{\tau(x)} \Psi(y(s))ds + \Phi(y(\tau(x)))\right]$$

solves

$$\mathcal{L}u + \Psi = 0 \text{ for } x \in D,$$

with boundary condition

$$u = \Phi \text{ for } x \in \partial D.$$

The justification is again entirely parallel to our earlier examples. Notice the analogy with the "least arrival time" problems of deterministic optimal control.

Examples. These results are already interesting for the simplest diffusion process: Brownian motion itself. For example: consider n-dimensional Brownian motion starting at x. What is the mean time it takes to exit from a ball of radius R, for $R > |x|$? Answer: apply the last example with $f = 0$, $g =$ identity matrix, $\Psi = 1$, $\Phi = 0$. It tells us the mean exit time is the solution $u(x)$ of

$$\tfrac{1}{2}\Delta u + 1 = 0$$

in the ball $|x| < R$, with $u = 0$ at $|x| = R$. The (unique) solution is

$$u(x) = \frac{1}{n}(R^2 - |x|^2).$$

Second example, this time scalar. Consider the scalar lognormal process

$$dy = \mu y dt + \sigma y dw$$

with μ and σ constant. Starting from $y(0) = x$, what is the mean exit time from a specified interval (a,b) with $a < x < b$? Answer: the mean exit time $u(x)$ solves

$$\mu x u_x + \tfrac{1}{2}\sigma^2 x^2 u_{xx} + 1 = 0 \text{ for } a < x < b$$

with boundary conditions $u(a) = u(b) = 0$. The solution is

$$u(x) = \frac{1}{\frac{1}{2}\sigma^2 - \mu} \left(\log(x/a) - \frac{1 - (x/a)^{1-2\mu/\sigma^2}}{1 - (b/a)^{1-2\mu/\sigma^2}} \log(b/a) \right)$$

(readily verified by checking the equation and boundary conditions).

Oksdendal has a nice discussion of basic properties of Brownian motion (transience and recurrence) using methods similar to the above (Oksendal example 7.4.2, see also exercises 7.4 and 7.9).