PDE for Finance Notes – Section 3

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.

Announcements

(1) There will be no lectures March 10 [I'm out of town], March 17 [spring break], and March 31 [Passover].

(2) A mechanism has been set up for students in this course to send each other email. To receive such messages, send email to the address majordomo@cs.nyu.edu with a blank subject line and the single text line *subscribe* g63_2706_001_sp99. To send a message to every subscriber, send it to g63_2706_001_sp99@*cs.nyu.edu*. For further information on the listserver software, send email to majordomo@cs.nyu.edu with a blank subject line and the single text line *help*. If your system automatically puts a signature at the end of your email, put "end" on a separate line after "help" to avoid the listserver software trying to process your signature as a series of commands.

Introduction to stochastic dynamic programming. Stochastic dynamic programming is like deterministic dynamic programming except the equation of state is a stochastic differential equation, and the goal is to maximize or minimize the expected utility or cost. To see what new issues this raises, we set up and briefly discuss two examples, namely: (a) perturbation of a deterministic problem by small noise, and (b) :w Merton's optimal consumption problem for one stock and one bond. Here we basically follow bits of Chapter VI of Fleming and Rishel.

Building the machinery to do these and other examples more carefully will take several lectures. Moreover we have yet to connect these ideas with martingales and risk-neutral measures. But before undertaking these tasks, we spend some time getting used to probabilistic thinking by examining some discrete-time stochastic dynamic programming problems, namely: (1) optimal control of execution costs, and (2) when to sell an asset. For topic (1) we follow a recent article by Dmitris Bertsimas and Andrew Lo, "Optimal control of execution costs," *J. Financial Markets* 1 (1998) 1-50 (copy available in the Green box reserve). For topic (2) we follow section 2.4 of Bertsekas.

****

Perturbation of a deterministic problem by small noise. We've discussed at length the deterministic dynamic programming problem with state equation

$$dy/ds = f(y(s), \alpha(s)) \text{ for } t < s < T, \quad y(t) = x,$$

controls $\alpha(s) \in A$, and objective

$$\max_{\alpha} \left\{ \int_t^T h(y(s), \alpha(s))\, ds + g(y(T)) \right\}.$$

Its value function satisfies the HJB equation

$$u_t + H(\nabla u, x) = 0 \text{ for } t < T, \quad u(x,T) = g(x),$$

with Hamiltonian

$$H(p,x) = \max_{a \in A} \{f(x,a) \cdot p + h(x,a)\}. \tag{1}$$

Let us show (heuristically) that when the state is perturbed by a little noise, the value function of resulting stochastic control problem solves the perturbed HJB equation

$$u_t + H(\nabla u, x) + \frac{1}{2}\epsilon^2 \Delta u = 0 \tag{2}$$

where H is still given by (1), and $\Delta u = \sum_i \frac{\partial^2 u}{\partial x_i^2}$. This gives another explanation why the viscosity solution of the deterministic HJB is the proper notion of weak solution.

Our phrase "perturbing the state by a little noise" means this: we replace the ODE governing the state by the stochastic differential equation (SDE)

$$dy = f(y, \alpha) ds + \epsilon dw,$$

keeping the initial condition $y(t) = x$. Here dw is a standard, vector-valued Brownian motion (each component w_i is a scalar-valued Brownian motion, and different components are independent). If you're completely new to stochastic differential equations, you can read about them e.g. in Neftci. If you know even the statement of Ito's lemma, the following should be accessible.

The evolution of the state is now stochastic, hence so is the value of the utility. Our goal in the stochastic setting is to maximize the *expected* utility. The value function is thus

$$u(x,t) = \max_{\alpha} E_{y(t)=x} \left\{ \int_t^T h(y(s), \alpha(s))\, ds + g(y(T)) \right\}.$$

There is some subtlety to the question: what is the class of admissible controls? Of course we still restrict $\alpha(s) \in A$. But since the state is random, it's natural for the control to be random as well – however its value at time s should depend only on the past and present, not on the future (which is after all unknown to the controller). Such controls are called "non-anticipating." A simpler notion, sufficient for most purposes, is to restrict attention to *feedback* controls, i.e. to assume that $\alpha(s)$ is a deterministic function of s and $y(s)$. One can show (under suitable hypotheses, when the state equation is a stochastic differential equation) that these two different notions of "admissible control" lead to the same optimal value.

Courage. Let's look for the HJB by applying the usual heuristic argument, based on the principle of dynamic programming applied to a short time interval:

$$u(x,t) \approx \max_{a \in A} \left\{ h(x,a)\Delta t + E_{y(t)=x} u(y(t+\Delta t), t+\Delta t) \right\}.$$

The term $h(x,a)\Delta t$ approximates $\int_t^{t+\Delta t} h(y(s), a)\, ds$, because $y(s) = x+$ terms tending to 0 with Δt. It is deterministic. The expression $u(y(t+\Delta t), t+\Delta t)$ is the optimal expected utility starting from time $t + \Delta t$ and spatial point $y(t + \Delta t)$. We must take its expected value, because $y(t + \Delta t)$ is random. (If you think carefully you'll see that the Markov property of the process $y(s)$ is being used here.)

We're almost in familiar territory. In the deterministic case the next step was to express $u(y(t + \Delta t), t + \Delta t)$ using the state equation and the Taylor expansion of u. Here we do something analogous: use Ito's lemma and the stochastic differential equation. Ito's lemma says the process $\phi(s) = u(y(s), s)$ satisfies

$$\begin{aligned} d\phi &= \frac{\partial u}{\partial s} ds + \sum_i \frac{\partial u}{\partial y_i} dy_i + \frac{1}{2} \sum_{i,j} \frac{\partial^2 u}{\partial y_i \partial y_j} dy_i dy_j \\ &= u_t(y(s), s) ds + \nabla u \cdot (f(y(s), \alpha(s)) ds + \epsilon dw) + \frac{1}{2} \epsilon^2 \Delta u \, ds. \end{aligned}$$

The real meaning of this statement is that

$$\begin{aligned} u(y(t'), t') - u(y(t), t) &= \int_t^{t'} [u_t(y(s), s) + \nabla u(y(s), s) \cdot (f(y(s), \alpha(s)) + \frac{1}{2} \epsilon^2 \Delta u(y(s), s)] ds \\ &\quad + \int_t^{t'} \epsilon \nabla u(y(s), s) \cdot dw. \end{aligned}$$

The expected value of the second integral is 0, so

$$E_{y(t)=x}[u(y(t + \Delta t), t + \Delta t)] - u(x, t) \approx [u_t(x, t) + \nabla u(x, t) \cdot f(x, a) + \frac{1}{2} \epsilon^2 \Delta u(x, t)] \Delta t.$$

(If you don't know enough about the Ito integral to follow this, don't be afraid – my main purpose is to motivate you to learn it.)

Assembling these ingredients, we have

$$u(x, t) \approx \max_{a \in A} \left\{ h(x, a) \Delta t + u(x, t) + [u_t(x, t) + \nabla u(x, t) \cdot f(x, a) + \frac{1}{2} \epsilon^2 \Delta u(x, t)] \Delta t \right\}.$$

This is almost identical to the relation we got in the deterministic case. The only difference is the new term $\frac{1}{2} \epsilon^2 \Delta u(x, t) \Delta t$ on the right. It doesn't depend on a, so the optimal a is unchanged – it still maximizes $h(x, a) + f(x, a) \cdot \nabla u$ – and we conclude, as asserted, that u solves (2).

Optimal portfolio selection and consumption. This is the simplest of a class of problems solved by Robert Merton in his paper "Optimal consumption and portfolio rules in a continuous-time model", *J. Economic Theory* 3, 1971, 373-413 (reprinted in his book *Continuous Time Finance*.) Consider a world with one risky asset and one risk-free asset. The risk-free asset grows at a constant risk-free rate r, i.e. its price per share satisfies $dp_1/dt = p_1 r$. The risky asset executes a geometric Brownian motion with constant drift $\mu > r$ and volatility σ, i.e. its price per share solves the stochastic differential equation $dp_2 = \mu p_2 dt + \sigma p_2 dw$.

The control problem is this: an investor starts with initial wealth x at time t. His control variables are

$$\begin{aligned}\alpha_1(s) &= \text{fraction of total wealth invested in the risky asset at time } s\\ \alpha_2(s) &= \text{rate of consumption at time } s.\end{aligned}$$

It is natural to restrict these controls by $0 \leq \alpha_1(s) \leq 1$ and $\alpha_2(s) \geq 0$. We ignore transaction costs. The state is the investor's total wealth y as a function of time; it solves

$$dy = (1-\alpha_1)yrdt + \alpha_1 y(\mu dt + \sigma dw) - \alpha_2 dt$$

so long as $y(s) > 0$. We denote by τ the first time $y(s) = 0$ if this occurs before time T, or $\tau = T$ (a fixed horizon time) otherwise. The investor seeks to maximize the discounted total utility of his consumption. We therefore consider the value function

$$u(x,t) = \max_{\alpha_1,\alpha_2} E_{y(t)=x} \int_t^\tau e^{-\rho s} h[\alpha_2(s)]\, ds$$

where $h[\cdot]$ is a specified utility function (monotone increasing and concave, with $h(0) = 0$). We shall specialize below to $h(\alpha_2) = \alpha_2^\gamma$ with $0 < \gamma < 1$.

Note: Our $u(x,t)$ is utility of consumption discounted to time 0. It might seem more natural to consider the value of utility discounted to time t, i.e. to use the discount factor $e^{-\rho(s-t)}$ instead of $e^{-\rho s}$. No need to redo the problem: this alternative value function is $e^{\rho t}u(x,t)$.)

We find the HJB by a heuristic argument very similar to the one presented above. The principle of dynamic programming applied on a short time interval gives:

$$u(x,t) \approx \max_{a_1,a_2} \left\{ e^{-\rho t} h(a_2)\Delta t + E_{y(t)=x} u(y(t+\Delta t), t+\Delta t) \right\}.$$

To evaluate the expectation term, we use Ito's lemma again. Using the state equation

$$dy = [(1-\alpha_1)yr + \alpha_1 y\mu - \alpha_2]dt + \alpha_1 y\sigma dw$$

and skipping straight to the conclusion, we have

$$u(y(t'),t') - u(y(t),t) = \int_t^{t'} [u_t + u_y[(1-\alpha_1)yr + \alpha_1 y\mu - \alpha_2] + \frac{1}{2}u_{yy}y^2\alpha_1^2\sigma^2]dt + \int_t^{t'} \alpha_1\sigma y u_y dw.$$

The expected value of the second integral is 0, so

$$E_{y(t)=x}[u(y(t+\Delta t), t+\Delta t)] - u(x,t) \approx [u_t + u_y[(1-\alpha_1)yr + \alpha_1 y\mu - \alpha_2 + \frac{1}{2}u_{yy}y^2\alpha_1^2\sigma^2]\Delta t.$$

Assembling these ingredients,

$$u(x,t) \approx \max_{a_1,a_2} \left\{ e^{-\rho t} h(a_2)\Delta t + u(x,t) + [u_t + u_x[(1-a_1)xr + a_1 x\mu - a_2] + \frac{1}{2}u_{xx}x^2 a_1^2\sigma^2]\Delta t \right\}.$$

Cleaning up, and taking the limit $\Delta t \to 0$, we get

$$u_t + \max_{a_1,a_2} \left\{ e^{-\rho t} h(a_2) + [(1-a_1)xr + a_1 x\mu - a_2]u_x + \frac{1}{2}x^2 a_1^2 \sigma^2 u_{xx} \right\} = 0.$$

This is the relevant HJB equation. It is to be solved for $t < T$, with $u(x, T) = 0$ since we have associated no utility associated to final-time wealth.

That looks pretty horrible, but it isn't really so bad. Let us assume that $u_x > 0$ (practically obvious – how would you prove it?) and $u_{xx} < 0$ (not quite so obvious – reflects the concavity of the utility function). Then optimal a_1 (ignoring the constraint $0 \leq a_1 \leq 1$) is

$$a_1^* = -\frac{(\mu - r)u_x}{\sigma^2 x u_{xx}}$$

which is positive. We proceed, postponing till later the verification that $a_1^* \leq 1$. The optimal a_2 satisfies

$$h'(a_2^*) = e^{\rho t} u_x;$$

we can sure this a_2^* is positive by assuming that $h'(0) = \infty$.

When $h(a_2) = a_2^\gamma$ with $0 < \gamma < 1$ we can get an explicit solution. Indeed, let us look for a solution of the form

$$u(x, t) = g(t) x^\gamma.$$

The associated a_1^* and a_2^* are

$$a_1^* = \frac{(\mu - r)}{\sigma^2(1 - \gamma)}, \quad a_2^* = \left[e^{\rho t} g(t)\right]^{1/(\gamma - 1)} x.$$

We assume henceforth that $\mu - r < \sigma^2(1 - \gamma)$ so that $a_1^* < 1$. Substuting these values into the HJB equation gives, after some arithmetic,

$$\frac{dg}{dt} + \nu\gamma g + (1 - \gamma) g (e^{\rho t} g)^{1/(\gamma - 1)} = 0$$

with

$$\nu = r + \frac{(\mu - r)^2}{2\sigma^2(1 - \gamma)}.$$

We must solve this with $g(T) = 0$. The substitution $f = (e^{\rho t} g)^{1/(\gamma - 1)}$ leads to a linear differential equation for $f(s)$. It is readily solved to give

$$g(t) = e^{-\rho t} \left[\frac{1 - \gamma}{\rho - \nu\gamma} \left(1 - e^{-\frac{(\rho - \nu\gamma)(T - t)}{1 - \gamma}}\right)\right]^{1 - \gamma}.$$

Of course this solution is meaningful only if $\rho > \nu\gamma$.

We have solved the HJB equation; but have we found the value function? The answer is yes, by a suitable verification argument. See the Section 3 addendum for the verification argument.

It should not be surprising that we have to place some restrictions on the parameters to get this solution. When these restrictions fail, inequalities that previously didn't bother us become important (such as a restriction on borrowing).

Discrete-time stochastic optimal control. An honest treatment of continuous-time stochastic dynamic programming requires mastery of certain tools, such as stochastic differential equations and the Ito calculus. The discrete-time setting is more accessible – an honest treatment requires nothing but basic probability. Moreover some very interesting problems can be done this way. Let's sample a couple of them.

Optimal control of execution costs. This example is developed in the recent article by Dmitris Bertsimas and Andrew Lo, "Optimal control of execution costs," *J. Financial Markets* 1 (1998) 1-50. Robert Almgren and Neil Chriss have a related working paper "Optimal liquidation." Both articles are on reserve in the green box.

The problem is this: an investor wants to buy a large amount of some specific stock. If he buys it all at once he'll drive the price up, thereby paying much more than necessary. Better to buy part of the stock today, part tomorrow, part the day after tomorrow, etc. till the full amount is in hand. But how best to break it up?

Here's a primitive model. It's easy to criticize (we'll do this below), but it's a good starting point – and a nice example of stochastic optimal control. Suppose the investor wants to buy $S_{\rm tot}$ shares of stock over a period of N days. His control variable is S_i, the number of shares bought on day i. Obviously we require $S_1 + \ldots + S_N = S_{\rm tot}$.

We need a model for the impact of the investor's purchases on the market. Here's where this model is truly primitive: we suppose that the price P_i the investor achieves on day i is related to the price P_{i-1} on day $i-1$ by

$$P_i = P_{i-1} + \theta S_i + \sigma e_i \tag{3}$$

where e_i is a Gaussian random variable with mean 0 and variance 1 (independent of S_i and P_{i-1}. Here θ and σ are fixed constants.

And we need a goal. Following Bertsimas and Lo we focus on minimizing the expected total cost:

$$\min E\left[\sum_i^N P_i S_i\right].$$

To set this up as a dynamic programming problem, we must identify the *state*. There is a bit of art here: the principle of dynamic programming requires that we be prepared to start the optimization at any day $i = N, N-1, N-2, \ldots$ and when $i = 1$ we get the problem at hand. Not so hard here: the state on day i is described by the most recent price P_{i-1} and the amount of stock yet to be purchased $W_i = S_{\rm tot} - S_1 - \ldots - S_{i-1}$. The state equation is easy: P_i evolves as specified above, and W_i evolves by

$$W_{i+1} = W_i - S_i.$$

Dynamic programming finds the optimal control by starting at day N, and working backward one day at a time. The relation that permits us to work backward is the one-time-step version of the principle of dynamic programming (known as the Bellman equation). In this case it says:

$$V_i(P_{i-1}, W_i) = \min_s E\left[P_i s + V_{i+1}(P_i, W_{i+1})\right].$$

Here $V_i(P, W)$ is the value function:

$$\begin{aligned} V_i(P, W) &= \text{optimal expected cost of purchasing } W \text{ shares} \\ &\quad\ \text{starting on day } i, \text{ if the most recent price was } P. \end{aligned}$$

(The subscript i plays the role of time.)

To find the solution, we begin by finding $V_N(P, W)$. Since $i = N$ the investor has no choice but to buy the entire lot of W shares, and his price is $P_N = P + \theta W + e_N$, so his expected cost is

$$V_N(P, W) = E\left[(P + \theta W + e_N)W\right] = PW + \theta W^2.$$

Next let's find $V_{N-1}(P, W)$. The Bellman equation gives

$$\begin{aligned} V_{N-1}(P, W) &= \min_s E\left[(P + \theta s + e_{N-1})s + V_N(P + \theta s + e_{N-1}, W - s\right] \\ &= \min_s E\left[(P + \theta s + e_{N-1})s + (P + \theta s + e_{N-1})(W - s) + \theta(W - s)^2\right] \\ &= \min_s \left[(P + \theta s)s + (P + \theta s)(W - s) + \theta(W - s)^2\right] \\ &= \min_s \left[W(P + \theta s) + \theta(W - s)^2\right]. \end{aligned}$$

The optimal s is $W/2$, giving value

$$V_{N-1}(P, W) = PW + \frac{3}{4}\theta W^2.$$

Thus: starting at day $N = 1$ (so there are only 2 trading days) the investor should split his purchase in two equal parts, buying half the first day and half the second day. His impact on the market costs him, on average, an extra $\frac{3}{4}\theta W^2$ over the no-market-impact value PW.

Proceeding similarly for day $N - 2$ etc., a pattern quickly becomes clear: starting at day $N - i$ with the goal of purchasing W shares, if the most recent price was P, the optimal trade on day i (the optimal s) is $W/(i + 1)$, and the expected cost of all W shares is

$$V_{N-i}(P, W) = WP + \frac{i + 2}{2(i + 1)}\theta W^2.$$

This can be proved by induction (the inductive step is very similar to our calculation of V_{N-1}).

Notice the net effect of this calculation is extremely simple: no matter when he starts, the investor should divide his total goal W into equal parts – as many as there are trading days – and purchase one part each day. Taking $i = N - 1$ we get the answer to our original question: if the most recent price is P and the goal is to buy S_{tot} over N days, then this optimal strategy leads to an expected total cost

$$V_1(P, S_{\text{tot}}) = PS_{\text{tot}} + \frac{\theta}{2}(1 + \frac{1}{N})S_{\text{tot}}^2.$$

There's something unusual about this conclusion. The investor's optimal strategy is not influenced by the random flucutations of the prices. It's always the same, and can be fixed in advance. That's extremely unusual in stochastic control problems: the optimal control can usually be chosen as a *feedback control*, i.e. a deterministic function of the state – but since the state depends on the fluctuations, so does the control.

I warned you it was easy to criticize this model. Some comments:

1. The variance of the noise in the price model never entered our analysis. That's because our hypothetical investor is completely insensitive to risk – he cares only about the expected result, not about its variance. No investor is like this. The paper by Almgren and Chriss considers the trade-off between return and risk (expected cost and variance of cost).

2. The price law (3) is certainly wrong: it has the ith trade S_i increasing not just the ith price P_i but also *every subsequent price*. A better law would surely make the impact of trading *temporary*. Bertismas and Lo consider one such law, for which the problem still has a closed-form solution derived by methods similar to those used above.

The take-home messages: (a) discrete-time stochastic dynamic programming is fun and easy, though of course closed-form solutions aren't always available; and (b) there's lots more to be done concerning modeling and control of execution costs. A food-for-thought question: (c) what is the continuous-time analogue of the example worked out above?

When to sell an asset. This is an *optimal stopping* problem. It's also interesting because the state is described by a discrete variable as well as a continuous one.

The problem is this: you have an asset (e.g. a house) you wish to sell. One offer arrives each week (yes, this example is oversimplified). The offers are independent draws from a single, known distribution. You must sell the house by the end of N weeks. If you sell it earlier, you'll invest the cash (risk-free), and its value will increase by factor $(1 + r)$ each week. Your goal is to maximize the expected cash on hand at the end of the Nth week.

The control, of course, is the decision (taken each week) to sell or not to sell. The state in a given week consists of (a) the current offer, and (b) whether the house is already sold. We use w to denote the current offer, and s, n to denote sold vs. not-sold. The state evolution is summarized by the following diagram:

The value function is

$$J_i(w,\cdot) \quad = \quad \text{expected week-}N\text{ cash produced by current and future sales,}$$
$$\text{if the current week is } i \text{ and the current offer is } w.$$

The $\cdot$ in $J_i(w,\cdot)$ stands for n or s. Obviously $J_i(w,s) = 0$ for all i and all w.

We start as usual with the final time, which as we've indexed things is $N-1$. If the house isn't already sold you have no choice but to sell it, realizing

$$J_{N-1}(w,n) = w.$$

The key to working backward is the Bellman equation, which in this setting says:

$$J_i(w,n) = \max\left\{(1+r)^{N-1-i}w, E[J_{i+1}(w',n)]\right\}.$$

Here w' is an independent trial from the specified distribution (the next week's offer); the first choice corresponds to the decision "sell now", the second choice ot the decision "don't sell now".

It's convenient to work not with J_i but rather

$$V_i(w) = (1+r)^{i-N+1}J_i(w),$$

which is the present value (at week i) of the income from future sales. Evidently we have $V_{N-1}(w) = w$ and

$$V_i(w,n) = \max\left\{w, (1+r)^{-1}E[V_{i+1}(w',n)]\right\}.$$

The optimal decision in week i is:

$$\begin{array}{rcl} \text{accept offer } w & \text{if} & w \geq \alpha_i \\ \text{reject offer } w & \text{if} & w \leq \alpha_i \end{array}$$

with

$$\alpha_i = (1+r)^{-1}E[V_{i+1}(w')].$$

To complete the solution to the problem we must find the sequence of real numbers $\alpha_0, \ldots \alpha_{N-2}$. Since

$$V_{i+1}(w,n) = \begin{cases} w & \text{if } w > \alpha_{i+1} \\ \alpha_{i+1} & \text{if } w \leq \alpha_{i+1} \end{cases}$$

we have

$$\begin{aligned} \alpha_i &= \frac{1}{1+r}\int_0^{\alpha_{i+1}} \alpha_{i+1}\, dP(w) + \frac{1}{1+r}\int_{\alpha_{i+1}}^{\infty} w\, dP(w) \\ &= \frac{1}{1+r}\alpha_{i+1}P(\alpha_{i+1}) + \frac{1}{1+r}\int_{\alpha_{i+1}}^{\infty} w\, dP(w) \end{aligned}$$

where $P(\lambda) = \text{prob}\{w < \lambda\}$ is the distribution function of w. This relation, with the initialization $\alpha_{N-1} = 0$, permits one to calculate the α's one by one (numerically). It can be shown that they are monotone in i: $\alpha_0 > \alpha_1 > \ldots$ (see Bertsekas). This is natural, since early in the sales process it makes no sense to take a low offer, but later on it may be a good idea to avoid being forced to take a still lower one on week N. One can also show that after many steps of the recursion relation for α_i, the value of α_i approaches the fixed point α_* which solves

$$\alpha_* = \frac{1}{1+r}\alpha_* P(\alpha_*) + \frac{1}{1+r}\int_{\alpha_*}^{\infty} w\, dP(w).$$

Thus when the horizon is very far away, the optimal policy is to reject offers below α_* and accept offers above α_*.