PDE in Finance, Spring 2008,
`http://www.math.nyu.edu/faculty/goodman/teaching/PDEfin/index.html`
Written by Jonathan Goodman. Please report errors, sources of confusion, and typos on the class Blackboard site.
Last updated March 28, 2008.

Section 6: Deterministic Optimal Control

1 Introduction to optimal control

Suppose the state of the system at time t is $X(t) \int R^n$. Also at time t we can choose a *control*, $\alpha(t) \in R^m$. We want to choose α to optimize some performance measure. These notes discuss only one PDE approach. The PDE is the *Hamilton Jacobi Bellman Pontriagin* equation. We usually leave out the twentieth century authors and call it simply Hamilton Jacobi. The unknown is the *value function*, $u(x,t)$, which is the optimal performance starting from $X(t) = x$. We derive the Hamilton Jacobi equation is a backward equation. It's derivation is similar to the derivation of the backward Kolmogorov equation.

One new feature is that the actual value function may not have all the derivatives called for by the PDE. The simplest case is an optimal hitting time problem (Section 3 below) in one dimension, for which the PDE is

$$|\partial_x u(x)| = 1 \, , \tag{1}$$

and the boundary conditions are $u(-1) = u(1) = 0$. The value function is $u(x) = 1 - |x|$. This function is not differentiable when $x = 0$, so it is unclear what the PDE (1) means at that point. This is a serious issue, because the function $v(x) = -u(x) = |x| - 1$ seems to satisfy the PDE every bit as well as the actual value function. This is resolved by the notion of *viscosity solution*. We will see that $u(x)$ satisfies (1) in the viscosity sense even at $x = 0$. The non-solution $v(x)$ satisfies (1) at every point except at $x = 0$. That single failure is enough to disqualify it as a solution.

The link between non-smoothness and seeming non-uniqueness is as follows. The viscosity solution is unique. The value function is a viscosity solution (hence *the* viscosity solution). A solution is *classical* if the derivatives called for in the PDE all exist in the usual sense at every point, and the PDE is satisfied at every point. A classical solution is a viscosity solution; the difference between classical and viscosity is only at points where the classical derivatives do not exist. All this is explained (to some extent) below.

The theory of viscosity solutions helps us find computational methods for solving a Hamilton Jacobi equation. If we simply approximate the derivatives that appear in the PDE using finite differences, we may get something that works for classical solutions but picks the wrong non-classical solution. The viscosity solution theory allows us to construct finite difference methods that are guaranteed to converge to the correct viscosity solution which is the value function we seek.

Here we discuss only deterministic optimal control problems. These are problems in which giving initial data $X(0) = x_0$ and control $\alpha(t)$ for $t \in [0, T]$ determines $X(t)$ exactly. *Stochastic control* means that there is randomness in the X dynamics, for example, as an SDE involving X and α. The Hamilton Jacobi equation for deterministic control typically involves only first derivatives of u. Stochastic control typically gives second order PDEs. In all cases, we expect very annoying nonlinearities, as in (1). Only the simplest problems have explicit solutions. There seem to be no general methods, such as Green's functions or Fourier integrals, that describe general solutions to these nonlinear PDEs.

2 Example: optimal consumption

We give a specific example before describing the general formulation, the simplest optimal consumption problem. Here, $X(t)$ is the wealth of a person we call Alice[1] at time t. Alice will die at time T and wants to consume all her wealth by then. The control is $\alpha(t)$, the consumption rate at time t. Whatever she does not consume grows with deterministic rate r, so

$$dX(t) \;=\; rX(t)dt \;-\; \alpha(t)dt \;. \tag{2}$$

The performance measure it total discounted utility,

$$J \;=\; \int_0^T e^{-\rho t}\, h(\alpha(s))\, ds \;. \tag{3}$$

2.1 Utility functions

The function $h(\alpha)$ in (3) is a *utility function*. It is supposed to measure the utility to Alice of consumption level α. In financial applications we always assume $h(\alpha)$ is an increasing function of α (avarice and/or gluttony). The *marginal utility* is $h'(\alpha)$. This measures how much extra utility Alice gets from $d\alpha$ of extra consumption rate[2]. Most often we assume that the marginal utility is a decreasing function of α, i.e. that $h''(\alpha) \leq 0$. This models the idea that additional consumption has more utility (value) to you if your consumption is small than if it is large. It is harder to notice the extra $d\alpha$ when α already is large.

In specific computations, the most common utility functions are *power laws* $h(\alpha) = C\alpha^\gamma$. We may, and do, take $C = 1$ without loss of generality. $h(\alpha)$ is increasing and concave if $\gamma > 0$ and $\gamma < 1$. The limits $\gamma \to 0$ may be modeled either as $h(\alpha) = 1 - e^{-\alpha}$ (the exponential utility), or $h(\alpha) = \log(\alpha)$ (the log utility). Exponential and log utilities lead to very strange policies. People consider them because the solution structure of the equations may be simpler.

[1] People working on cryptography always talk about Bob sending a coded message to Alice.
[2] It might be more accurate to call $h'(\alpha)d\alpha$ the marginal utility, but we don't.

If $h''(\alpha) = 0$, Alice is *risk neutral*. She is *risk averse* if the strict inequality $h''(\alpha) < 0$ holds. This terminology makes sense in the context of stochastic control. Let W be a random variable with mean $\overline{W} = E[W]$, and *expected utility* $\overline{h} = E[h(W)]$. The monetary value to Alice of W is supposed to be the w_0 so that $h(w_0) = \overline{h}$. If the utility of the known amount w_0 is equal to the expected utility of W, then Alice would be willing to exchange the unknown amount W for the known amount w_0. She is risk neutral (meaning indifferent to risk) if she values W at it's expected value. It is easy to see that if $h''(w) = 0$, then $w_0 = \overline{W}$.

Risk aversion refers to $w_0 < \overline{W}$. Many financial optimization problems do not make sense unless Alice (the agent) is risk averse. *Jensen's inequality* is a theorem in probability that says that $w_0 < \overline{W}$ if W is genuinely random and $h'' < 0$. To prove it, let $M = h'(\overline{W})$ (for *marginal*). Then $h(w) < h(\overline{W}) + M\left(w - \overline{W}\right)$, so

$$\begin{aligned} E\left[h(W)\right] &\leq E\left[h(\overline{W}) + M\left(W - \overline{W}\right)\right] \\ &= E\left[h(\overline{W})\right] + E\left[M\left(W - \overline{W}\right)\right] \\ &= E\left[h(\overline{W})\right] \ . \end{aligned}$$

The inequality on the top line will be strict unless $W = \overline{W}$ (W is not random). But, if $E\left[h(W)\right] = h(w_0) < h(\overline{W})$ and h is an increasing function of w, then $w_0 < \overline{W}$.

Utility theory, the belief (or model) that people optimize their expected utility, is deeply flawed. The half the field of *behavioral finance* is devoted to collecting data of choices of people that cannot come from any utility function. For example, utility maximization would not allow Alice to trade a deterministic dollar for a lottery ticket whose expected return is less than a dollar. Moreover, to use utility theory in a practical setting, we have to specify the utility function. The idea of being interviewed to determine your utility function is comical.

Nevertheless, there is a deep theorem, the *von Neumann Morgenstern theorem* that says that any choice function that is "rational" in some sense (the von Neumann Morgenstern axioms) is given by some utility function. If someone is trying to make rational decisions, for example by using PDE methods, it makes sense that they would use a utility function. They also should understand how the results depend on the particular utility function chosen. You might think of this as being like the efficient frontier theorems in the basic theory of investments. Different levels of risk aversion (different utility functions) give you different efficient strategies.

It sometimes is convenient to use a quadratic utility function $h(w) = w - C(w - w_1)^2$. This is concave but not monotone for large w. It has the advantage that optimizing a quadratic performance measure with a linear dynamical model gives a linear problem that may be solved explicitly.

2.2 Back to the Optimal consumption problem

The optimization problem will be to find the minimum of J in (3) subject to the dynamics (2) and the constraint $X(t) > 0$ for all $t \in [0, T)$, and starting with initial wealth $X(0) = x_0$. The parameters are the discount rate, ρ, and the rate of return, r. The *value function* is

$$u(x,t) \;=\; \min_{\alpha[t,T]} \int_t^T e^{-\rho(s-t)} h(\alpha(s))\, ds \;, \quad X(t) = x \;. \tag{4}$$

This value function is the solution to the same optimization problem, but starting at time t with wealth $X(t) = x$ given. The notation $\alpha[t, T]$ refers to a general function defined for $s \in [t, T]$. The optimal such function is the optimal control we are seeking.

The value function satisfies a PDE, which we derive using arguments similar to those that give the Kolmogorov backward equation. We calculate what will happen if we use control α, not necessarily optimal, for time interval $[t, t + \Delta t]$ and the optimal control thereafter. The integral in (4) is

$$\int_t^{t+\Delta t} e^{-\rho(s-t)} h(\alpha)\, ds \;+\; \int_{t+\Delta t}^T e^{-\rho(s-t)} h(\alpha(s))\, ds \;.$$

We calculate both terms up to terms of order Δt. The first is simply

$$h(\alpha)\Delta t \;.$$

For the second term, first observe that

$$e^{-\rho(s-t)} \;=\; e^{-\rho(s-(t+\Delta t))}\, e^{-\rho\Delta t} \;\approx\; (1 - \rho\Delta t)\, e^{-\rho(s-(t+\Delta t))} \;,$$

so

$$\begin{aligned} \int_{t+\Delta t}^T e^{-\rho(s-t)} h(\alpha(s))\, ds \;&\approx\; (1 - \rho\Delta t) \int_{t+\Delta t}^T e^{-\rho(s-(t+\Delta t))} h(\alpha(s))\, ds \\ &\approx\; \int_{t+\Delta t}^T e^{-\rho(s-(t+\Delta t))} h(\alpha(s))\, ds \;-\; \rho\Delta t u(x,t) \;. \end{aligned}$$

Comparing the first term on the right to (4), if we use the optimal $\alpha(s)$ for $s \in [t + \Delta t, T]$, it is $u(X(t + \Delta t), t + \Delta t)$. Altogether, the integral to order Δt is

$$\begin{aligned} &\int_t^T e^{-\rho(s-t)} h(\alpha(s))\, ds \\ &\quad=\; u(X(t + \Delta t), t + \Delta t) \;+\; h(\alpha)\Delta t \;-\; \rho u(x,t)\Delta t \;+\; O(\Delta t)^2) \;. \end{aligned}$$

Recall the assumptions: we apply control α in the interval $[t, t + \Delta t]$, which takes us to (using the dynamics (2))

$$X(t + \Delta t) \;=\; x \;+\; \Delta x \;=\; x \;+\; rx\Delta t \;-\; \alpha\Delta t \;.$$

Starting from time $t + \Delta t$ we apply the optimal control for whatever state we find ourselves in. We use this and the Taylor expansion (as in the derivation of backward equations). To simplify notation, we leave out the (x, t) arguments everywhere.

$$\int_t^T e^{-\rho(s-t)} h(\alpha(s))\, ds$$
$$\approx\ u\ +\ \partial_t u\, \Delta t\ +\ \partial_x u\, (rx - \alpha)\Delta t\ +\ h(\alpha)\Delta t\ -\ \rho u \Delta t\ . \tag{5}$$

The optimal value function (4) comes from optimizing the right hand side (5) with respect to α. Collecting terms that involve α this gives

$$\max_{\alpha}\ -\partial_x u(x,t)\alpha\ +\ h(\alpha)\ .$$

We denote the solution $\alpha_*(x, t)$. For the power law utility $h(\alpha) = \alpha^\gamma$, we have the explicit formula

$$\alpha_*(x,t)\ =\ \left(\frac{1}{\gamma}\partial_x u(x,t)\right)^{\frac{1}{\gamma-1}}\ . \tag{6}$$

There are two ways to express our conclusions. One is to use our expansion of (4) together with the optimal α to get (again leaving out (x, t) arguments)

$$u\ =\ u\ +\ \Delta t\, (\partial_t u\ +\ (rx - \alpha_*)\, \partial_x u\ +\ h(\alpha_*)\ -\ \rho u)\ +\ O(\Delta t^2)\ .$$

Following the derivation of other backward equations, we deduce from this one form of the Hamilton Jacobi equation for this case

$$0\ =\ \partial_t u\ +\ (rx - \alpha_*)\, \partial_x u\ +\ h(\alpha_*)\ -\ \rho u\ , \tag{7}$$

where α_* solves

$$\max_{\alpha}\ -\alpha\partial_x u\ +\ h(\alpha)\ .$$

The other way combines the choice of α with the PDE (7):

$$0\ =\ \partial_t u\ +\ \max_{\alpha}\Big\{\ (rx - \alpha)\, \partial_x u\ +\ h(\alpha)\ -\ \rho u\ \Big\}\ . \tag{8}$$

To understand this last form, note that $u(x, t)$ is the value function given the optimal control. We achieve the optimal $\partial_t u$ only by using the optimal α. Any other choice of α is (as engineers often say) suboptimal.

The formula (6) gives the optimal α as a function of x and t. To find this function, we need to compute $u(x, t)$ by solving the Hamilton Jacobi equation (7) or (8). You can't find the optimal control without knowing the value function and you can't find the value function without at least implicitly learning the optimal control. If $X(t)$ is the state at time t, the corresponding control is

$$\alpha_*(t)\ =\ \left(\frac{1}{\gamma}\partial_x u(X(t),t)\right)^{\frac{1}{\gamma-1}}\ .$$

Considering the optimal $\alpha_*(t)$ to be a function of the state at time t (and t) is *feedback* or *closed loop* control. The closed loop dynamics is the original dynamics (2) under the optimal control. In this case, it is

$$dX(t) \;=\; rX(t)dt \;-\; \alpha_*(t)dt \;=\; \left(rX(t) \;-\; \left(\frac{1}{\gamma} \partial_x u(X(t),t) \right)^{\frac{1}{\gamma-1}} \right) dt \; .$$

For deterministic optimal control problems, the *calculus of variations* gives a way to compute the optimal control and corresponding optimal trajectory from a specific starting point $x_0 = X(0)$ without computing the value function. This works even for high dimensional models where the *curse of dimensionality* prevents us from computing the value function. There seems to be no analogue of this *open loop* control for stochastic systems. Avoiding the curse for high dimensional stochastic systems is an active research area both in finance and control engineering.

For the power law utility, there is an explicit solution to the control problem, which is based on the ansatz that the value function also is a power law:

$$u(x,t) \;=\; A(t) x^\gamma \; . \tag{9}$$

A little thought should convince you that the value function should be a power law. Ask what we would do if x_0 were replaced by λx_0 for some multiple, λ. Clearly, the optimal control for the multiple problem would be a scaled version of the original. Even if we can't make a complete rigorous justification for the ansatz (9), if we find a solution of that form, a suitable uniqueness theorem will tell us that we have solved the problem.

From the ansatz (9) and (6), we find

$$\alpha_* \;=\; \left(\frac{1}{\gamma} A \gamma x^{\gamma-1} \right)^{\frac{1}{\gamma-1}} \;=\; A(t)^{\frac{1}{\gamma-1}} x \; .$$

Putting this into (7) gives (after canceling the common x^γ factors)

$$\dot{A} \;=\; (\rho - \gamma r)\, A \;-\; (1-\gamma)\, A^{\frac{\gamma}{\gamma-1}} \; . \tag{10}$$

Let's examine (10) before solving it. We expect $A > 0$ so that the value function will be positive. Further, we hope that $A \to 0$ as $t \to T$ so that $v(x,T) = 0$. The power in the second term of (10) is $\frac{\gamma}{\gamma-1} < 0$ for γ in the range $(0,1)$. Thus, the second term pushes goes to infinity and strongly pushes A toward zero once A is small. We will determine $A(0)$ so that $A = 0$ when $t = T$.

Solving (10) uses a trick some will remember from their ODE class. Mathematically, (10) takes the form

$$\frac{dx}{dt} \;=\; ax - bx^{-p} \; .$$

Simple separation of variables gives

$$\frac{dx}{ax - bx^{-p}} \;=\; dt \;\; , \quad \int \frac{dx}{ax - bx^{-p}} \;=\; t + C \; .$$

For general values of p, we cannot work this integral in closed form. On the other hand, we might be able to change the value of p by multiplying the whole equation by x^q, with q unknown for now. Note that

$$x^q \dot{x} \;=\; \frac{1}{q+1}\frac{d}{dt}x^{q+1}\;,$$

so

$$\frac{1}{q+1}\frac{d}{dt}x^{q+1} \;=\; ax^{q+1} - bx^{q-p}\;.$$

The choice $q = p$ seems promising. It simplifies the second term. Moreover, if $y = x^{q+1} = x^{p+1}$, we get

$$\frac{d}{dt}\,y \;=\; (p+1)ay \;-\; (p+1)b\;.$$

This is a linear equation in y that we can solve.

To implement this strategy for (10), note that $p = \frac{\gamma}{1-\gamma}$, so $p+1 = \frac{1}{1-\gamma}$. We multiply by (10) $A^{\frac{\gamma}{1-\gamma}}$, define $B = A^{\frac{1}{1-\gamma}}$, and get

$$\dot{B} \;=\; \frac{\rho - \gamma r}{1-\gamma}B \;-\; 1\;. \tag{11}$$

The solution of $\dot{B} = aB - 1$ is (we remember from ODE) an exponential plus a constant

$$B(t) \;=\; be^{at} \;+\; c\;.$$

Substituting gives $ac = 1$. Therefore, the solution of (11) is

$$B(t) \;=\; be^{\frac{\rho-\gamma r}{1-\gamma}t} \;+\; \frac{1-\gamma}{\rho-\gamma r}\;.$$

This gives a formula for A and then for u:

$$u(x,t) \;=\; x^\gamma\left(be^{\frac{\rho-\gamma r}{1-\gamma}t} \;+\; \frac{1-\gamma}{\rho-\gamma r}\right)^{1-\gamma}\quad. \tag{12}$$

The parameter b is determined so that $B(T) = 0$. The nature of the algebra, and of the solution, depends on the sign of $\rho - \gamma r$, which is explored in Assignment 6, question 1.

3 The general problem

If we don't look too carefully, the general theory is nearly the same as the example. The state of the system at time t is described by $X(t) \in R^n$. The dynamics are

$$dX \;=\; a(X,\alpha)dt\;. \tag{13}$$

$X(0) = x_0$ is given. We seek to maximize

$$J \;=\; \int_0^T h(X(s), \alpha(s), s)\, ds \;+\; V(X(T)) \;. \tag{14}$$

The *running payout* (the negative of the running cost for an engineer) is a function of the state at time s, the control, and s (the latter to accommodate exponential weighting as in the example). We also allow for a final time payout that was not present in the example. The value function is

$$u(x,t) \;=\; \max_{\alpha[t,T]} \int_t^T h(X(s), \alpha(s), s)\, ds \;+\; V(X(T)) \;, \tag{15}$$

with the condition that $X(t) = x$.

This value function satisfies a Hamilton Jacobi equation. If we apply control α in the time interval $[t, t+\Delta t]$ and the optimal control thereafter, the integral in (15) becomes

$$h(x, \alpha, t)\Delta t \;+\; u(x + \Delta x, t + \Delta t) \;.$$

The dynamics (13) give

$$\Delta x \;=\; a(x, \alpha)\Delta t \;.$$

The usual Taylor expansion gives

$$h(x, \alpha, t)\Delta t \;+\; u(x,t) \;+\; a(x, \alpha)\nabla u(x,t) \;+\; \partial_t u(x,t)\Delta t \;.$$

The backward equation argument then gives

$$0 \;=\; \partial_t u \;+\; \max_\alpha \Big\{ h(\alpha, x) \;+\; a(x, \alpha)\nabla u \Big\} \;. \tag{16}$$

The quantity in braces is the *Hamiltonian*. We give the formula for it using the standard notation $p = \nabla u$:

$$H(x,p) \;=\; \max_\alpha \Big\{ h(\alpha, x) \;+\; a(x, \alpha)p \Big\} \;. \tag{17}$$

Then (16) is written as a general Hamilton Jacobi equation

$$\partial_t u \;=\; -H(x, \nabla u) \;, \tag{18}$$

where H is given by (17) in the present case.

The value function is found by solving the Hamilton Jacobi equation (18) with final conditions $u(x,T) = V(x)$. The closed loop optimal control $\alpha_*(x,t)$ is the $\alpha_*(x,t)$ that solves

$$\max_\alpha \Big\{ h(\alpha, x) \;+\; a(x, \alpha)\nabla u(x,t) \Big\} \;.$$

The closed loop dynamics are

$$dX \;=\; a(X(t), \alpha_*(X(t), t))\, dt \;.$$

There are mathematical subtleties that we have not mentioned, many of which are related.

1. The maximum may not be attained, either in the large problem (15) or in the small problem (17). There are two ways a maximum can fail to be attained

   (a) The maximum is infinite. Arbitrarily large values are possible. In this case, either we made a mistake in formulating the problem or we won the lottery.

   (b) There is an optimal performance that can be approached arbitrarily closely but not actually achieved. For example, in (17), there may be a sequence α_n so that $h(\alpha_n, x) + a(x, \alpha_n)p \to H$ as $n \to \infty$ while at the same time $h(\alpha, x) + a(x, \alpha)p < H$ for any α. In this case we write $H = \sup_\alpha h(\alpha, x) + a(x, \alpha)p$, with sup for *supremum* instead of max for maximum. If the maximum in (15) is not attained, the family of approximating optimal controls probably become increasingly complex as $n \to \infty$ so that $\lim_{n\to\infty} \alpha_n(t)$ does not exist in the proper sense. A large branch of modern control theory is devoted to understanding such examples.

2. The final value problem (18) may not have a classical solution. Instead, the solution may *break down* at some time t_*. As $t \downarrow t_*$ the solution $u(x,t)$ may develop a singularity at time t_*. This means that

$$u(x, t_*) = \lim_{t \downarrow t_*} u(x, t)$$

   is not a differentiable function of t. Of course, the value function (15) is defined (possibly as infinity or as a sup rather than max), so it is the value function that may become singular at time t_*. Now there is a general theory of *viscosity solutions* to Hamilton Jacobi equations, which contains the following results:

   (a) The definition of a viscosity solution.

   (b) The proof that there is a viscosity solution.

   (c) A comparison principle that shows that he viscosity solution is unique.

   (d) A proof that any classical solution is this unique viscosity solution.

   (e) A proof that the viscosity solution is the value function (15).

   (f) Computational procedures that are guaranteed to converge to the viscosity solution as $\Delta x \to 0$ and $\Delta t \to 0$ (the mesh sizes for the computation).

3. The minimum problem (17) might not have a unique minimizer. There might be more than one value of α for which $h(\alpha, x) + a(x, \alpha)p = \max$. In the bigger picture, there might be two or more completely different controls and trajectories that achieve the optimal performance in (15).