C8.7 Optimal Control

Sheet 3 — HT25

Section A

1. Consider the optimization problem of finding $\tau \in [0, T]$ to minimize

$$\int_0^\tau G(s, X_s)\mathrm{d}s + \Phi(X_\tau),$$

when $\frac{\mathrm{d}}{\mathrm{d}t}X_t = F(t, X_t)$, where $X \in \mathbb{R}^m$, F is Lipschitz in x, and G is integrable.

(a) By interpreting this as an optimal control problem with controlled drift $f(t, X_t, U_t) = F(t, X_t)U_t$ and $U_t \in \{0, 1\}$, and similarly for g, write down the Hamilton–Jacobi equation describing the minimal value of this problem.

(b) Show that the Hamilton–Jacobi equation is equivalent to the linear complementarity problem

$$0 = \min\Big\{\Phi(x) - v(t, x), \partial_t v + G(t, x) + \langle \nabla v, F(t, x)\rangle\Big\}.$$

Solution:

(a) As described, we have a control problem with Hamilton–Jacobi equation $-\partial_t v = H(t, x, \nabla v)$, where the Hamiltonian is given by

$$H(t, x, q) = \min_{u\in\{0,1\}} \Big\{G(t, x)u + \langle q, F(t, x)u\rangle\} = \min\Big\{0, G(t, x) + \langle q, F(t, x)\rangle\Big\}$$

and terminal value $v(T, x) = \Phi(x)$.

(b) We know that $-\partial_t v \le 0$ and $v(T, x) = \Phi(x)$, so $v(t, x) \le \Phi(x)$. At the same time, for any (t, x), if $v(t, x) = \Phi(x)$, then this implies that x is the best attainable state (when you start at t, x), which implies $\partial_t v = 0$. Therefore, we know that

$$0 = \min\Big\{\Phi(x) - v(t, x), \partial_t v + G(t, x) + \langle q, F(t, x)\rangle\Big\}$$

which is the desired linear complementarity formulation.

Sam Cohen: cohens@maths.ox.ac.uk

2. Suppose the (un-minimized) Hamiltonian $\tilde{H}(t, x, q, u)$ is convex with respect to u, where u takes values in a finite-dimensional open set $\mathcal{U}$.

   (a) Show that Pontryagin's minimum principle can be expressed in the form:
$$\frac{\mathrm{d}}{\mathrm{d}t} X_t^* = \partial_q \tilde{H}, \qquad \frac{\mathrm{d}}{\mathrm{d}t} q_t^* = \rho q_t^* - \partial_x \tilde{H}, \qquad 0 = \partial_u \tilde{H}.$$

   (b) Show that, if $\rho = 0$ and the Hamiltonian does not depend on time, and the optimal strategy is differentiable with respect to time, then Pontryagin's minimum principle shows that, along the optimal path,
$$\tilde{H}(X_t^*, q_t^*, u_t^*) = \text{constant}$$

   (c) (*Trickier*) Suppose that:

   - the state dynamics f are Lipschitz continuous,
   - the Hamilton–Jacobi equation admits a C^1 solution v with Lipschitz continuous derivatives,
   - $\tilde{H}$ is twice differentiable and strictly convex with respect to u, in particular $\partial^2_{uu} \tilde{H}(t, u, x, q)$ has all eigenvalues above ε, for all t, u, x, q.

   Show that there exists an optimal control, and that this control is unique.

   (d) Explain why, in the setting above, Pontryagin's minimum principle is both necessary and sufficient for optimality.

Solution:

(a) This is essentially just notation – you calculate the derivatives of $\tilde{H}(t, x, q, u) = g(t, x, u) + \langle q, f(t, x, u) \rangle$ and compare with Pontryagin's minimum principle.

(b) If $\tilde{H}$ does not depend on time, assuming u^* is differentiable, then the previous results and the chain rule yield
$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} \tilde{H}(X_t^*, q_t^*, u_t^*) &= (\partial_x \tilde{H})(\frac{\mathrm{d}}{\mathrm{d}t} X_t^*) + (\partial_q \tilde{H})(\frac{\mathrm{d}}{\mathrm{d}t} q_t^*) + (\partial_u \tilde{H})(\frac{\mathrm{d}}{\mathrm{d}t} u_t^*) \\
&= (\partial_x \tilde{H})(\partial_q \tilde{H}) + (\partial_q \tilde{H})(-\partial_x \tilde{H}) + 0 \\
&= 0.
\end{aligned}$$

This connects with the interpretation, in physical problems, of the Hamiltonian as the total system energy (ie the sum of potential and kinetic energies), which is a conserved quantity.

(c) As $\tilde{H}$ is strictly convex and $\partial_u \tilde{H}$ has nonvanishing derivative, we know that there exists an interior minimizer which satisfies $\partial_u \tilde{H}(t, X_t^*, q_t^*, u) = 0$. An admissible

Sam Cohen: cohens@maths.ox.ac.uk

control is optimal if and only if this equation is satisfied, and there is at most one value of u (for each X_t^*, q_t^*) which satisfies this equation, so the optimal control (if it exists) must be unique. The challenge is to simultaneously construct the optimal control U^* and the optimal trajectory X_t^* (with which we can compute $q_t^* = \nabla v(X_t^*)$).

For t fixed, consider the map $u \mapsto \partial_u H(t,x,q,u) =: h(x,q,u)$. Suppose u satisfies $\partial_u H(t,x,q,u) = 0$ for a given x, q. Using a Taylor expansion we know that, for (x', q', u') near (x,q,u), with all derivatives being evaluated at (x,q,u),

$$\partial_u H(t,x',q',u') = (D_x h)(x'-x) + (D_q h)\cdot(q'-q) + (D_u h)\cdot(u'-u) + \text{error}$$

where the error is sublinearly small. Our assumptions guarantee that the matrix $D_u h$ has an inverse, which is uniformly bounded in (x,q,u). Therefore, if we wish u' to be optimal at x', q', we have a variation on the inverse function theorem:

$$u' = u - (D_u h)^{-1}\Big((D_x h)(x'-x) + (D_q h)\cdot(q'-q)\Big) + \text{error}.$$

(This can, of course, be made fully rigorous, by use of a contraction mapping result.) We conclude that the optimal strategy u' is a Lipschitz continuous function of x' and q'. More precisely, there exists a Lipschitz continuous function $u^*(t,\cdot,\cdot) : \mathbb{R}^m \times \mathbb{R}^{m\times m} \to \mathcal{U}$ which satisfies the optimality condition

$$\partial_u \tilde{H}(t,x,q,u^*(t,x,q)) = 0.$$

Given this, together with the smoothness of the solution v to the Hamilton–Jacobi equation, there exists a solution X^* to the state equation

$$\frac{\mathrm{d}}{\mathrm{d}t}X_t^* = f\Big(t, X_t^*, u^*(t, X_t^*, \nabla v(X^*))\Big),$$

so $U_t^* = u^*(t, X_t^*, \nabla v(X^*))$ is admissible. By construction, we know U^* satisfies the optimality condition along this trajectory, so we conclude that U^* is the unique optimal control.

(d) In the setting discussed, we know that an optimum indeed exists, and therefore it is necessary that Pontryagin's principle holds. However, as u^* is a Lipschitz function of x, q, we know that the system of equations in Pontryagin's principle admits a unique solution. It follows that there is exactly one triple (X^*, q^*, U^*) satisfying Pontryagin's principle, which implies that all solutions are optimal.

Sam Cohen: cohens@maths.ox.ac.uk

3. Suppose we have controlled state dynamics f which satisfy (for all x, u)

$$\langle x, f(x,u)\rangle \leq C(1+\|x\|^{3/2})$$

for some constant C.

(a) Using the comparison theorem for ODEs, show that for any control U, the controlled trajectory satisfies $\|X_t^U\| \leq C(1+\|x_0\|)(1+t^2)$, for some $C>0$.

The process y_t with $y_0 = 1$ and dynamics $\frac{\mathrm{d}}{\mathrm{d}t}y_t = Cy_t^{1/2}$ may provide a useful upper bound on $\|X_t^U\|^2/(1+\|x_0\|^2)$.

(b) Consider controlling this process over the horizon $[0,T]$, where we assume the cost process satisfies $|g(t,x,u)| < C(1+\|x\|)$ and we have a discount rate $\rho > 0$. Show that the total discounted cost of any control remains bounded as $T\to\infty$.

(c) Suppose the time-homogenous Hamilton–Jacobi equation

$$0 = -\rho v + H(x, \nabla v)$$

admits a bounded C^1 solution. Assuming the dynamic programming principle holds, show that this must be the value function of the infinite horizon discounted control problem. (*Hint: Compare the behaviour over finite horizons.*)

Solution:

(a) Fix a control U. We know that $\|X_t^U\|^2 = \langle X_t^U, X_t^U\rangle$ has dynamics

$$\frac{\mathrm{d}}{\mathrm{d}t}\|X_t^U\|^2 = 2\langle X_t^U, f(t, X_t, U_t)\rangle \leq C(1+\|X_t^U\|^{3/2}).$$

Dividing by $(1+\|x_0\|)^2$ simply reduces C, and gives a process which starts in the unit ball.

Consider the function y_t with $y_0 = 1$ and (Locally Lipschitz) dynamics $\frac{\mathrm{d}}{\mathrm{d}t}y_t = Cy_t^{1/2}$, which has solution $y_t = \frac{(2+Ct)^2}{4} \geq 1$. Observe that y_t^2 has dynamics $\frac{\mathrm{d}}{\mathrm{d}t}y_t^2 = 2Cy_t^{3/2}$, and that $2C\|y\|^{3/2} \geq C(1+\|y\|^{3/2})$ for all $y\geq 1$. The comparison theorem for ODEs (see Part A DE1 for a proof), then yields the desired growth bound

$$\frac{\|X_t^U\|^2}{(1+\|x_0\|)^2} \leq y_t^2 = \Big(\frac{(2+Ct)^2}{4}\Big)^2 \leq \Big(\frac{4+C^2t^2}{2}\Big)^2.$$

(b) We know from above that $\|X_t^U\| \leq C(1+\|x_0\|)(1+t^2)$. This implies that, for some constant C (which can vary from line to line)

$$\begin{aligned}\int_0^\infty |e^{-\rho t} g(t, X_t^U, U_t)|\mathrm{d}t &\leq \int_0^\infty Ce^{-\rho t}(1+\|X_t^U\|)\mathrm{d}t \\ &\leq \int_0^\infty Ce^{-\rho t}(1+\|x_0\|)(1+t^2)\mathrm{d}t \\ &\leq C(1+\|x_0\|)\frac{1}{\rho^3}\end{aligned}$$

Sam Cohen: cohens@maths.ox.ac.uk

This provides a global bound on the discounted cost of any control, as required.

(c) By dynamic programming, we know that v is the value function of the control problem on horizon $[0,T]$, if we specify that the terminal value at T is $\Phi(X_T) = v(X_T)$. (This can be seen by simply verifying that v satisfies the finite-horizon Hamilton–Jacobi equation). Furthermore, for every $T > 0$ and $\varepsilon > 0$, there exists a control U^ε such that

$$v(x) + \varepsilon \geq \int_0^T e^{-\rho t} g(s, X_s^{t,x,U^\varepsilon}, U_s^\varepsilon)\mathrm{d}s + e^{-\rho(T-t)} v(X_T^{t,x,U^\varepsilon}) \geq v(x).$$

As v is bounded, we can take $T \to \infty$ to see that there exists a control U such that

$$v(x) + \varepsilon \geq \int_0^\infty e^{-\rho t} g(s, X_s^{t,x,U^\varepsilon}, U_s^\varepsilon)\mathrm{d}s.$$

Now define the infinite-horizon discounted cost-to-go, for a general control U, which satisfies, for some $C > 0$

$$\begin{aligned}
&J(t,x,U)\\
&= \int_t^\infty e^{-\rho(s-t)} g(s, X_s^{t,x,U}, U_s)\mathrm{d}s\\
&= \int_t^T e^{-\rho(s-t)} g(s, X_s^{t,x,U}, U_s)\mathrm{d}s + e^{-\rho(T-t)} J(T, X_T^{t,x,U}, U)\\
&\geq \int_t^T e^{-\rho(s-t)} g(s, X_s^{t,x,U}, U_s)\mathrm{d}s + e^{-\rho(T-t)} v(X_T^{t,x,U}) - Ce^{-\rho(T-t)} C(1 + \|x\|)\rho^{-3}\\
&\geq v(x) - Ce^{-\rho(T-t)}\rho^{-3}(1 + \|x\|)(1 + T^2)
\end{aligned}$$

Taking $T \to \infty$ shows that $J(t,x,U) \geq v(x)$ for all t, U. A similar argument shows that

$$J(t,x,U^\varepsilon) \leq v(x) + \varepsilon + Ce^{-\rho(T-t)}\rho^{-3}(1 + \|x\|).$$

Taking $T \to \infty$, $\varepsilon \to 0$, we conclude that v is indeed the optimal value for the problem.

(*Note that we never really used the assumption that v is C^1, apart from to make sure that v is the value function over the corresponding finite horizon problem.*)

Sam Cohen: cohens@maths.ox.ac.uk

Section B

Sam Cohen: cohens@maths.ox.ac.uk

4. Consider a deterministic control problem with underlying state dynamics satisfying the scalar second order controlled ODE

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}X_t = U_t - \frac{\mathrm{d}}{\mathrm{d}t}X_t$$

with initial values $(X_0, \dot{X}_0) = (1, 0)$. Suppose we seek to minimize the value of

$$\int_0^T (X_t^2 + U_t^2)\mathrm{d}t.$$

Reexpress this problem as a vector-valued first order equation, and hence derive an equation for the value function in terms of a system of Riccati equations, and for the optimal control in feedback form.

Using a backwards Euler scheme, solve the system of Riccati equations numerically, and hence state the value, and optimal control at time $t = 0$.

Solution: We can rewrite our equation as the vector equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{X}_t = \frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} X_t \\ \dot{X}_t \end{bmatrix} = \begin{bmatrix} \dot{X}_t \\ U_t - \dot{X}_t \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix}}_{A}\mathbf{X}_t + \underbrace{\begin{bmatrix} 0 \\ 1 \end{bmatrix}}_{B}U_t$$

We observe that these are linear dynamics, and the cost is given by the Quadratic form

$$\int_0^T (X_t^2 + U_t^2)\mathrm{d}t = \int_0^T \Big(\mathbf{X}_t^\top \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}}_{Q}\mathbf{X}_t + U_t^2\Big)\mathrm{d}t$$

As this is a linear quadratic problem, we see that the value function $v(t,x) = x^\top \Sigma_t x + 2\Psi_t x + \Gamma_t$ can be obtained by solving the Riccati equations (with $R = 1$, and other parameters zero, in our calculations from lectures)

$$\begin{aligned}
K_t &= -R^{-1}B^\top \bar{\Sigma}_t, \qquad H_t = 0,\\
-\partial_t \Sigma_t &= Q + K_t^\top R K_t + 2\bar{\Sigma}_t(A + BK_t),\\
-\partial_t \Psi_t &= \Psi_t(A + BK_t),\\
-\partial_t \Gamma_t &= 0
\end{aligned}$$

with $\Sigma_T = 0$, $\Psi_T = 0$, and $\Gamma_T = 0$. The optimal control is then given by

$$U_t^* = K_t X_t.$$

Numerically, the initial value is given by $v(x_0) \approx 0.9714$ and the initial control by $U_0 \approx -0.3536$ (see `PS3_LQ.ipynb`)

Sam Cohen: cohens@maths.ox.ac.uk

5. Write down a deterministic control problem where there are exactly two (distinct) trajectories for the optimally controlled state process X.

   Solution: Example 3.2.8 in the lecture notes has this property, if we start with the initial value $X_0 = 0$.

6. Consider a deterministic optimal control problem, with optimal state trajectory X^*, and costs $\int_0^T g(t, X_t^U, U_t)\mathrm{d}t + \Phi(X_T^U)$.

   (a) Show that X^* is still optimal for the problem with costs defined by $\tilde{g}(t, x, u) = g(t, x, u) + \alpha 1_{\|x - X_t^*\| > \varepsilon}$ and $\tilde{\Phi}(x) = \Phi(x) + \alpha 1_{\|x - X_T^*\| > \varepsilon}$, for any $\alpha, \varepsilon > 0$.

   (b) Show that X^* may not be optimal for the problem with costs defined by $\tilde{g}(t, x, u) = g(t, x, u) - \alpha 1_{\|x - X_t^*\| > \varepsilon}$ and $\tilde{\Phi}(x) = \Phi(x) - \alpha 1_{\|x - X_T^*\| > \varepsilon}$, for some choices of $\alpha, \varepsilon > 0$

   (c) Suppose Pontryagin's minimum principle is satisfied for the original problem. Show that Pontryagin's minimum principle is also satisfied for both of the variations above. What conclusion do you draw?

   Solution:

   (a) We know that, in our original problem, the optimal strategy will result in state trajectory X^*. If we raise costs for other strategies, this will clearly not result in them becoming optimal. As we have not altered the costs for the optimal strategy, it therefore must remain optimal.

   (b) A simple counterexample is to suppose $g(t, x, u) = 0$, and to take ε small enough that there exists a control U with $\|X_T^U - X_T^*\| > \varepsilon$. If we then set $\alpha > \Phi(X_T^*) - \Phi(X_T^U)$, we conclude that with our pertubed costs, we know that X^U has a lower overall cost than X^*, so X^* cannot be an optimal trajectory.

   (c) In both the cases above, we do not vary our cost in a neighbourhood of X^*. As Pontryagin's principle only refers to the behaviour of costs in a neighbourhood of the optimal trajectory, it must remain valid in both cases. We conclude that (as we already knew) Pontryagin's principle is generally only a necessary condition, but not a sufficient condition, for optimality (*cf. Q2*)

Sam Cohen: cohens@maths.ox.ac.uk

7. Consider the problem of minimizing the cost $\int_0^3 (U_t^4 - X_t^2)\mathrm{d}t + X_3^2$, where X is a scalar process following the controlled dynamics $\frac{\mathrm{d}}{\mathrm{d}t}X_t = U_t - X_t + \sin(t)$ and $X_0 = 1$, where $U \in \mathcal{U} = \mathbb{R}$. Observe that this gives a convex Hamiltonian, so we believe that the optimal strategy should be unique, and should be determined by Pontryagin's minimum principle.

   Implement a forward-in-time Euler discretization of the equations in Pontryagin's minimum principle, assuming you know the value of $q_0 = \nabla v(x_0)$. By using a bisection search, or otherwise, find the value of q_0 which gives the correct value of q_T, and hence find a solution to the control problem.

   Compare the cost of your strategy with the cost of a constant control $u = 0$

   Solution: See file `PS3_Pontryagin.ipynb`.

Sam Cohen: cohens@maths.ox.ac.uk