Lecture 6: Optimal Control Theory and Hamilton-Jacobi Equations

Goals:

1. A scheme for "guessing" solutions to optimal control problems.
2. Make the connection between optimal control and Hamilton-Jacobi equations.

1. Examples

A typical problem is of the form:

\[
\min_{y} \int_{0}^{T} h(y(s), x(s)) \, ds + g(x(T))
\]

- \[y(s) = x(y(s), x(s)) \, x(0) = y_0\]
- \[x(0) \in \mathbb{R}^n\]

1. \[x(0, \infty) \to A \subset \mathbb{R}^n\] is the control.

2. \[P \times T = \int_0^T h(y(s), x(s)) \, ds + g(x(T))\] is the payoff functional.

3. \[h = \text{running payoff} \to \text{fuel, consumption, time.}\]

4. \[g = \text{terminal payoff} \to \text{parking penalty.}\]

5. \[y = \text{state of the system.}\]

A. Bolza-Young

\[
\min_{v(0)} \int_{0}^{T} (v_x^2 - 1) + v_x^2 \, dx
\]
Let $\frac{dy}{dt} = \alpha$, $y(0) = 0$ be the state equation. The running cost is $h(y, \alpha) = (\alpha^2 + 1)^2 + y^2$.

B.) **Minimal Arrival Time**

- **Objective:** Minimize the time at which $y(s)$ reaches target $\gamma$

\[
\begin{align*}
\dot{y}(s) &= f(y(s), \alpha(s)), \quad y(0) = y_0 = 0 \\
h &= \begin{cases} 
1 & \text{if } y(t) \neq 0, \text{ for all } t \leq s \\
0 & \text{if } y(t) = 0 \text{ for some } t > s 
\end{cases}
\end{align*}
\]

\[T = \infty \iff y(\infty) = 0 \iff \alpha \in \mathbb{R} \quad \text{or} \quad (0, 1) \]

\[\mathcal{R} = \mathbb{R}^2 \times S^1\]

**Parallel Parking:**
\[
\begin{align*}
\dot{x} &= \alpha_1 \cos(\phi) + \alpha_2 \cos(\phi) \\
\dot{y} &= \alpha_1 \sin(\phi) + \alpha_2 \sin(\phi) \\
\phi &= \alpha_1 \quad \text{or} \quad \alpha_2 \\
\end{align*}
\]

**Note:** We want to go sideways!!
C. Example Rocket Railroad

\[ x(t) = \text{position}, \quad \dot{x}(t) = \text{velocity} \]
\[ v(t) = \dot{x}(t) \]

Goal: Park at the origin in minimal time
\[ \dot{x}(t) = v(t) \]
\[ v(t) = \alpha(t) \]

\[ -1 \leq \alpha \leq 1 \]

\[ P[\alpha] = -T \quad \text{with} \quad T = \text{first time when } x(t) = 0, \quad v(t) = 0. \]
Suppose we only need \( \alpha = \pm 1 \).

Case 1:

If \( \alpha = 1 \) then we also know:
\[ v(0) = v \quad \Rightarrow \quad \frac{dx}{dv} = \frac{v^2 - v_0^2}{2} = x - x_0 \]

Case 2:

If \( \alpha = 2 \) then:
\[ \frac{dx}{dv} = -v \quad \Rightarrow \quad v^2 = -2x + c \]

Optimal trajectory is a "bang-bang" control. Fully accelerate and then jam on brakes. \( x = (t)v \).
Dynamic Programming

A typical trick in applied math is to embed the problem in a larger class of problems and solve all of them at once.

Let's define the value function:
\[
u(x, t) = \min_{x(s)} \int_{t}^{T} h(y(s), x(s)) \, ds + g(y(T)).
\]

\[
\dot{y}(s) = s(y(s), x(s)) \quad \text{for} \quad t < s < T,
\]
\[y(t) = x.
\]

The dynamic programming principle is that:
\[
u(x, t) = \min_{x(s) \in X} \left\{ \int_{t}^{s} h(y(s), x(s)) \, ds + \psi(x, x(\alpha(\tau)), t') \right\}
\]

**Interpretation:**

It is better to be smart at the beginning than wait until the end. The optimal strategy should do something between \( t \) and \( t' \). Starting from \( t' \), it should solve the same problem with new starting time and position.

We can derive a P.D.E. for \( u \). We will suppose that \( x' = x + \Delta x \) and take the limit as \( \Delta x \to 0 \). Let's do the formal argument:

1. Let's guess that over the interval \([t, t+\Delta t]\), that \( x(t) = a \) a constant.

2. \( u(x, t) \approx \min_{a \in X} \left\{ h(x, a) \Delta t + u(x + \frac{f(x, a)}{\Delta t} \Delta t, t + \Delta t) \right\} \)

where we have applied
\[
\dot{x} = f(y(s), a) = x + s(y(a), x) s.
\]

\( y(t) = x \)

and dropped terms of order \( \Delta t \).
3. Now, if we assume that $u$ is differentiable then

$$u(x + s, y(x), a) \Delta t, y(x) + u(x, y(x)) + \nabla u \cdot s = a \Delta t + u(x, y(x))$$

4. $u(x, t) = \min_{a \in A} \{ h(x, a) \Delta t + u(x, y(x)) + \nabla u \cdot s \Delta t + u(x, y(x)) \}$

$$\Rightarrow u(x, t) = \min_{a \in A} \{ h(x, a) + \nabla u \cdot s(t, a) \} = 0.$$

This is the Hamilton–Jacobi–Bellman equation.

We obtain initial conditions by noticing that:

$$u(x, 0) = g(x).$$

Notice, if we define

$$H(x, \nabla u) = \min_{a \in A} \{ h(x, a) + \nabla u \cdot s(t, a) \}$$

we get the (P.D.E)

$$u_t + H(x, \nabla u) = 0.$$

3. **Hopf–Lax Formula**

Recall we can obtain "calculus of variations" problems if we take:

$$\begin{align*}
\dot{x} &= f(x, y(x), \dot{y}(x)) \\
y(x) &= x
\end{align*}$$

Assume that $h(x, a) = h(a)$

$$u_t + \max_{a \in A} \{ -h(x, a) + \nabla u \cdot a \} = 0$$

$$u_t + H(x, \nabla u) = 0$$

$$H = \max_{a \in A} \{ -h(x, a) + \nabla u \cdot a \} = 0$$

$-h = H^* \Rightarrow$ the Legendre transform.

* Hamilton–Jacobi functions generate equations generate calculus of variations problems
* Calculus of variations generate H–J equations for the value function
* The Hamiltonian is the Legendre transform of the Lagrangian
Now, let's assume that $h$ is independent of $y$. Then
\[ u(x, t) = \min_{y \in \mathbb{R}} \left( \sum_{i=1}^{n} x_i (s_i) + y (T-t) \right) \]
\[ y = x \iff \frac{y}{T-t} = \frac{x}{T-x} \]
and we assume that $h$ is convex. Then,
\[ c = \left( \frac{\sum_{i=1}^{n} x_i (s_i) ds}{T-t} \right) \]
\[ y(x) = \frac{1}{T-t} \left( \sum_{i=1}^{n} y_i (s_i) ds \right) \]
Jensen’s Inequality.

With equality if and only if $y(x) = \text{constant}$.

Therefore, the optimal trajectory must have $y(x) = \text{constant}$.

This gives us the Hopf–Lax formula:
\[ u(x, t) = \max_{z \in \mathbb{R}} \left( (T-t) h \left( \frac{z-x}{T-t} \right) \right) \]

Summary

1. If we want to solve
\[ u_t + H(Du) = 0 \]
\[ u(x, 0) = \phi(x) \]
2. Let $h = -h^*$
3. We just need to solve the 1-D optimization problem:
\[ u(x, t) = \max_{z \in \mathbb{R}} \left( h \left( \frac{z-x}{t} \right) \right) \]
\[ u(x, t) = \min_{z \in \mathbb{R}} \left( h \left( \frac{x-z}{t} \right) \right) \]
4. **Example - Dog in the sand**

\[ I[f] = \int_0^l \sqrt{1 + \left( \frac{dy}{dx} \right)^2} \, dx \]

\[ A = \exists y \in W^{1,1} : y(0) = 0, y(1) = 1 \]

- C: top speed of dog
- \( n(x) \): the effect of the sand on the dog's speed.

Define a value function:

\[ U(x, y) = \inf_{\alpha \in A} \int_x^y \frac{1 + \alpha^2}{c_n(s)} \, ds \]

\[ \frac{dy}{ds} = \alpha(s) \quad y(0) = 0, \quad A = \exists \alpha : y(1) = 1 \]

\[
\Rightarrow a = \frac{t y_n c_n(x)}{\sqrt{1 - B^2 c_n^2(x)}}
\]

\( \Rightarrow \) **Getting the sign correct** is not so trivial. 

We clearly want \( a > 0 \) and we expect \( y_n < 0 \) so we must take \( - \) root!
\[ U_x + \frac{u_x^2 c^2 n^2}{1 - u_x^2 c^2 n^2} = 0 \]

\[ U_x (\sqrt{1 - u_x^2 c^2 n^2}) c \cdot n(x) + 1 - u_x^2 c^2 n^2 = 0 \]

\[ U_x \cdot c \cdot n(x) = -\frac{(1 - u_x^2 c^2 n^2)}{\sqrt{1 - u_x^2 c^2 n^2}} \]

\[ \nabla U \cdot c \cdot n(x)^2 = (1 - u_x^2 c^2 n^2) \]

This is an eikonal equation for the value function.

Interpretation:

The contours of the value function are wavefronts, and the outward normal is the tangent to the optimal curve. We see this by noticing that

\[ a = \frac{\pm \nabla U}{U_x} \]

The reduced speed of light in silica bends the light. This is why Snell's law actually works, i.e., light takes the shortest path of time.
5. **Issues**

1. $u(x, t)$ may not be $C^1$, bringing into question our derivation.

2. The P.D.E. may have many solutions, for example, the *Beltrami-Young* condition.

3. How can we use the P.D.E. to actually find the optimal control?

*Answering 1. and 2. requires the viscosity method to select distinguished solutions to the P.D.E. The viscosity solution is the value function.*

\[
\begin{align*}
\mathcal{A} &= \{(x, x) \mid x \in \mathbb{R}\} \\
\gamma &= \{(x_t, x) \mid (x_t, x) \in \mathbb{R}^2, x = x_t \}
\end{align*}
\]

Theorem:

\[
u(t, x) = \inf_{\gamma \in \mathcal{A}} \left\{ \int_{t}^{T} h(y(s), x(s)) \, ds + \mathcal{L}(x_t, x, t) \right\}
\]

\[
\gamma(t, x) = \frac{\partial}{\partial x} \mathcal{L}(x_t, x, t)
\]

*Proof:*

Let $\nu(t, x)$ be the R.H.S. of the above equation. From definition of $u(t, x)$, we have \(\forall \varepsilon > 0\), \(\exists \varepsilon > 0\) such that

\[
u(t, x) + \varepsilon \geq \mathcal{L}(x_t, x, t)
\]

with \(\mathcal{L}(x_t, x, t)\) denoting the payoff functional for $x_t$ with initial value problem.

\[
u(t, x) = \text{inf}_{\gamma \in \mathcal{A}} \left\{ \int_{t}^{T} h(y(s), x(s)) \, ds + \mathcal{L}(x_t, x, t) \right\}
\]

\[
u(t, x) + \varepsilon > \mathcal{L}(x_t, x, t)
\]

The other inequality is easy:

\[
u(t, x) = \text{inf}_{\gamma \in \mathcal{A}} \left\{ \int_{t}^{T} h(y(s), x(s)) \, ds + \mathcal{L}(x_t, x, t) \right\} = u(x, t)
\]

\[
u(t, x) + \varepsilon \geq \mathcal{L}(x_t, x, t)
\]
From this rigorous proof of the dynamic programming principle we see that C1 continuity/regularity is enough to derive HJB equations.

Now, to design an optimal control, we select \( \alpha(s) \) to be the value where the minimum of the HJB equation is obtained.

**Steps:**

1. Solve the HJB equations and compute \( v(x, t) \) for each \( x \in \mathbb{R}^n \) and \( t \in [0, T] \).

2. Use \( v(x, t) \) to define an optimal feedback control:

   For each \( x \in \mathbb{R}^n \) and \( t \in [0, T] \) set
   \[
   \alpha(x, t) = a \in A
   \]

   to be where the minimum is obtained
   \[
   u_x + f(x, x(t), \alpha(x(t), s)) \cdot \nabla_x v(x(t), t) + h(x, x(t), s) = 0.
   \]

3. Next solve the O.D.E.

   \[
   x^*(s) = f(x^*(s), \alpha(x^*(s), s)),
   \]

   \[
   x(t) = x
   \]

4. Define the feedback control

   \[
   \alpha^*(s) = \alpha(x^*(s), s).
   \]

**Theorem:** The control \( \alpha^*(s) \) defined by this construction is optimal.

\[
\mathbb{E}[\alpha^*]_{x,t} = \sum_{x} h(x^*(s), \alpha^*(s)) \cdot ds + g(x^*(T)).
\]

\[
\mathbb{E}[\alpha^*]_{x,t} = \int_{s}^{T} (-u(x^*(s), s) + x^*(s)) \cdot \nabla_x u(x^*(s), s) ds + g(x^*(T)).
\]

\[
\mathbb{E}[\alpha^*]_{x,t} = \mathbb{E}[\alpha^*]_{x,t} = u(x^*(T), s).
\]
6. Example:

\[ u(x, t) = \max_{x(t)} \int_{0}^{t} e^{-g(s-t)} x(s) \, ds \]

\[ \frac{dx}{dt} = ry - \alpha(s), \quad y(t) = x \]

\[ x(t) = \alpha(s) \]

\[ \alpha(s) \geq 0, \quad \gamma(s) \geq 0 \]

- \( r > 0 \) is an interest rate
- \( \alpha(s) \) is consumption
- \( u(x, t) \) is utility
- \( -y(s) \) is money

**Step #1**

Let's find the HJB equation

\[ v(x, t) = \max \left\{ \left. a \frac{d}{dx} \Delta t + e^{-g(s-t)} \Delta t \right| u(x, s', x'), \right. \right. \]

\[ v(x, t) = \max_{a \geq 0} \left\{ a \frac{d}{dx} \Delta t + e^{-g(s-t)} \Delta t \right. \right. \]

\[ = \max_{a \geq 0} \left\{ a \Delta t + (1-x) \Delta t \right. \right. \]

\[ = \max_{a \geq 0} \left\{ a \Delta t + (1-x) \Delta t \right. \right. \]

\[ = \max_{a \geq 0} \left\{ a \Delta t + (1-x) \Delta t \right. \right. \]

\[ \Rightarrow v_t + \max_a \left( a^2 - \gamma(s) u + (x-a) u_x \right) \]

**Step #2**

Let's try to find the optimal policy. Let's show that

\[ u(x, t) = g(t) x \]

\[ \Rightarrow g(t) = u(1, t) \]
Let \( \lambda x(s) \) be a control for problem starting at \( \lambda x \), where \( x(s) \) is the optimal choice starting from \( x \).

\[
\frac{dx}{ds} = ry_x - \lambda x(s)
\]

This is clear then that \( y_x(x) = \lambda y(x) \). Therefore,

\[
u_1(\lambda x, x) = \lambda^q v(x, t)
\]

Replacing \( \lambda \) with \( \lambda^q \) gives

\[
u(\lambda^{-1} x, x) = \lambda^-q v(x, t)
\]

\[
u(x, t) = \lambda^-q v(x, t)
\]

This gives us the result,

\[
u(\lambda x, t) = \lambda^q v(x, t)
\]

**Step 3:**

Now let's solve the HJB equation. From step #2 we have that \( \nabla x \geq 0 \), i.e., it is increasing in \( x \).

Maximizing \( \nabla x \), we get that

\[
a(t) = \frac{1}{g(x)}
\]

\[
\Rightarrow a(t) = g(t)^{-1} x, \text{ since } v(x, t) = x^q g(t).
\]

Substituting back into the P.D.E., we get that

\[
x + x^q - g(x) x^q + (q+1)(1-q) + q x^q = 0
\]

Letting \( H(t) = g(x)^{1-q} \), we have the linear O.D.E.

\[
H(t) = -N H(t), \text{ with } N = \frac{q}{1-q}
\]

The solution satisfying \( u(x, T) = 0 \) gives

\[
H(t) = N^{-1} \left( 1 - e^{-N(T-t)} \right)
\]

The optimal control is then

\[
u(t) = \frac{1}{x} \left( 1 - e^{-u(T-t)} \right)
\]