6.231 DYNAMIC PROGRAMMING

LECTURE 8

LECTURE OUTLINE

- Deterministic continuous-time optimal control
- From the HJB equation to the Pontryagin Minimum Principle
- Examples

THE HJB EQUATION

- Continuous-time dynamic system

$$\dot{x}(t) = f\big(x(t), u(t)\big), \quad 0 \le t \le T, \quad x(0) : \text{given}$$

- Cost function

$$h\big(x(T)\big) + \int_0^T g\big(x(t), u(t)\big) dt$$

- $J^*(t, x)$: optimal cost-to-go from x at time t
- HJB equation: For all (t, x)

$$0 = \min_{u \in U} \Big[g(x, u) + \nabla_t J^*(t, x) + \nabla_x J^*(t, x)' f(x, u) \Big]$$

with the boundary condition $J^*(T, x) = h(x)$.

- Verification theorem: If we can find a solution, it must be equal to the optimal cost-to-go function. Also a (closed-loop) policy $\mu^*(t, x)$ such that

$$\mu^*(t, x) \text{ attains the min for each } (t, x)$$

is optimal.

HJB EQ. ALONG AN OPTIMAL TRAJECTORY

- Observation I: An optimal control-state trajectory pair $\{(u^*(t), x^*(t)) \mid t \in [0,T]\}$ satisfies for all $t \in [0,T]$

$$u^*(t) = \arg\min_{u \in U} \Big[g\big(x^*(t), u\big) + \nabla_x J^*\big(t, x^*(t)\big)' f\big(x^*(t), u\big) \Big]. \tag{1}$$

- Observation II: To obtain an optimal control trajectory $\{u^*(t) \mid t \in [0,T]\}$ via this equation, we don't need to know $\nabla_x J^*(t,x)$ for all (t,x) - only the time function

$$p(t) = \nabla_x J^*\big(t, x^*(t)\big), \qquad t \in [0,T].$$

- It turns out that calculating $p(t)$ is often easier than calculating $J^*(t,x)$ or $\nabla_x J^*(t,x)$ for all (t,x).
- Pontryagin's minimum principle is just Eq. (1) together with an equation for calculating $p(t)$, called the $adjoint$ equation.
- Also, Pontryagin's minimum principle is valid much more generally, even in cases where $J^*(t,x)$ is not differentiable and the HJB has no solution.

DERIVING THE ADJOINT EQUATION

- The HJB equation holds as an identity for all (t, x), so it can be differentiated [the gradient of the RHS with respect to (t, x) is identically 0].
- We need a tool for differentiation of "minimum" functions.

Lemma: Let $F(t, x, u)$ be a continuously differentiable function of $t \in \Re$, $x \in \Re^n$, and $u \in \Re^m$, and let U be a convex subset of $\Re^m$. Assume that $\mu^*(t, x)$ is a continuously differentiable function such that

$$\mu^*(t, x) = \arg\min_{u \in U} F(t, x, u), \qquad \text{for all } t, x.$$

Then

$$\nabla_t \left\{ \min_{u \in U} F(t, x, u) \right\} = \nabla_t F\big(t, x, \mu^*(t, x)\big), \text{ for all } t, x,$$

$$\nabla_x \left\{ \min_{u \in U} F(t, x, u) \right\} = \nabla_x F\big(t, x, \mu^*(t, x)\big), \text{ for all } t, x.$$

DIFFERENTIATING THE HJB EQUATION I

- We set to zero the gradient with respect to x and t of the function

$$g\big(x, \mu^*(t,x)\big) + \nabla_t J^*(t,x) + \nabla_x J^*\big(t,x\big)' f\big(x, \mu^*(t,x)\big)$$

and we rely on the Lemma to disregard the terms involving the derivatives of $\mu^*(t,x)$ with respect to t and x.

- We obtain for all (t,x),

$$\begin{aligned} 0 &= \nabla_x g\big(x, \mu^*(t,x)\big) + \nabla^2_{xt} J^*(t,x) \\ &\quad + \nabla^2_{xx} J^*(t,x) f\big(x, \mu^*(t,x)\big) + \nabla_x f\big(x, \mu^*(t,x)\big) \nabla_x J^*(t,x) \end{aligned}$$

$$0 = \nabla^2_{tt} J^*(t,x) + \nabla^2_{xt} J^*(t,x)' f\big(x, \mu^*(t,x)\big),$$

where $\nabla_x f\big(x, \mu^*(t,x)\big)$ is the matrix

$$\nabla_x f = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_1} \\ \vdots & \vdots & \vdots \\ \frac{\partial f_1}{\partial x_n} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix}$$

DIFFERENTIATING THE HJB EQUATION II

- The preceding equations hold for all (t,x). We specialize them along an optimal state and control trajectory $\big\{\big(x^*(t),u^*(t)\big)\mid t\in[0,T]\big\}$, where $u^*(t)=\mu^*\big(t,x^*(t)\big)$ for all $t\in[0,T]$.
- We have $\dot{x}^*(t)=f\big(x^*(t),u^*(t)\big)$, so the terms

$$\nabla^2_{xt}J^*\big(t,x^*(t)\big)+\nabla^2_{xx}J^*\big(t,x^*(t)\big)f\big(x^*(t),u^*(t)\big)$$

$$\nabla^2_{tt}J^*\big(t,x^*(t)\big)+\nabla^2_{xt}J^*\big(t,x^*(t)\big)'f\big(x^*(t),u^*(t)\big)$$

are equal to the total derivatives

$$\frac{d}{dt}\Big(\nabla_x J^*\big(t,x^*(t)\big)\Big),\qquad \frac{d}{dt}\Big(\nabla_t J^*\big(t,x^*(t)\big)\Big),$$

and we have

$$\begin{aligned}0=\nabla_x g\big(x,u^*(t)\big)&+\frac{d}{dt}\Big(\nabla_x J^*\big(t,x^*(t)\big)\Big)\\&+\nabla_x f\big(x,u^*(t)\big)\nabla_x J^*\big(t,x^*(t)\big)\end{aligned}$$

$$0=\frac{d}{dt}\Big(\nabla_t J^*\big(t,x^*(t)\big)\Big).$$

CONCLUSION FROM DIFFERENTIATING THE HJB

- Define
$$p(t) = \nabla_x J^*\big(t, x^*(t)\big)$$
and
$$p_0(t) = \nabla_t J^*\big(t, x^*(t)\big)$$

- We have the *adjoint equation*

$$\dot{p}(t) = -\nabla_x f\big(x^*(t), u^*(t)\big)p(t) - \nabla_x g\big(x^*(t), u^*(t)\big)$$

and
$$\dot{p}_0(t) = 0$$

or equivalently,

$$p_0(t) = \text{constant}, \qquad \text{for all } t \in [0, T].$$

- Note also that, by definition $J^*\big(T, x^*(T)\big) = h\big(x^*(T)\big)$, so we have the following boundary condition at the terminal time:

$$p(T) = \nabla h\big(x^*(T)\big)$$

NOTATIONAL SIMPLIFICATION

- Define the $Hamiltonian$ function

$$H(x,u,p) = g(x,u) + p' f(x,u)$$

- The adjoint equation becomes

$$\dot{p}(t) = -\nabla_x H\big(x^*(t), u^*(t), p(t)\big)$$

- The HJB equation becomes

$$\begin{aligned} 0 &= \min_{u\in U}\Big[H\big(x^*(t), u, p(t)\big)\Big] + p_0(t) \\ &= H\big(x^*(t), u^*(t), p(t)\big) + p_0(t) \end{aligned}$$

so since $p_0(t) =$ constant, there is a constant C such that

$$H\big(x^*(t), u^*(t), p(t)\big) = C, \qquad \text{for all } t \in [0,T].$$

PONTRYAGIN MINIMUM PRINCIPLE

- The preceding (highly informal) derivation is summarized as follows:

Minimum Principle: Let $\big\{u^*(t) \mid t \in [0,T]\big\}$ be an optimal control trajectory and let $\big\{x^*(t) \mid t \in [0,T]\big\}$ be the corresponding state trajectory. Let also $p(t)$ be the solution of the adjoint equation

$$\dot{p}(t) = -\nabla_x H\big(x^*(t), u^*(t), p(t)\big),$$

with the boundary condition

$$p(T) = \nabla h\big(x^*(T)\big).$$

Then, for all $t \in [0,T]$,

$$u^*(t) = \arg\min_{u \in U} H\big(x^*(t), u, p(t)\big).$$

Furthermore, there is a constant C such that

$$H\big(x^*(t), u^*(t), p(t)\big) = C, \qquad \text{for all } t \in [0,T].$$

2-POINT BOUNDARY PROBLEM VIEW

- The minimum principle is a necessary condition for optimality and can be used to identify candidates for optimality.

- We need to solve for $x^*(t)$ and $p(t)$ the differential equations

$$\dot{x}^*(t) = f\big(x^*(t), u^*(t)\big)$$

$$\dot{p}(t) = -\nabla_x H\big(x^*(t), u^*(t), p(t)\big),$$

with split boundary conditions:

$$x^*(0) : \text{given}, \qquad p(T) = \nabla h\big(x^*(T)\big).$$

- The control trajectory is implicitly determined from $x^*(t)$ and $p(t)$ via the equation

$$u^*(t) = \arg\min_{u \in U} H\big(x^*(t), u, p(t)\big).$$

- This 2-point boundary value problem can be addressed with a variety of numerical methods.

ANALYTICAL EXAMPLE I

$$\text{minimize} \int_0^T \sqrt{1 + \big(u(t)\big)^2}\, dt$$

subject to

$$\dot{x}(t) = u(t), \qquad x(0) = \alpha.$$

- Hamiltonian is

$$H(x, u, p) = \sqrt{1 + u^2} + pu,$$

and adjoint equation is $\dot{p}(t) = 0$ with $p(T) = 0$.

- Hence, $p(t) = 0$ for all $t \in [0, T]$, so minimization of the Hamiltonian gives

$$u^*(t) = \arg\min_{u \in \Re} \sqrt{1 + u^2} = 0, \qquad \text{for all } t \in [0, T].$$

Therefore, $\dot{x}^*(t) = 0$ for all t, implying that $x^*(t)$ is constant. Using the initial condition $x^*(0) = \alpha$, it follows that $x^*(t) = \alpha$ for all t.

ANALYTICAL EXAMPLE II

- Optimal production problem

$$\text{maximize} \int_0^T \big(1 - u(t)\big)x(t)dt$$

subject to $0 \le u(t) \le 1$ for all t, and

$$\dot{x}(t) = \gamma u(t)x(t), \qquad x(0) > 0 : \text{given.}$$

- Hamiltonian: $H(x, u, p) = (1 - u)x + p\gamma ux$.
- Adjoint equation is

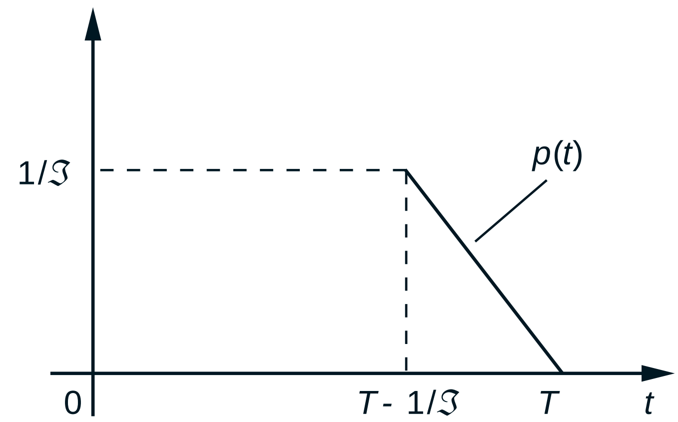
$$\dot{p}(t) = -\gamma u^*(t)p(t) - 1 + u^*(t), \quad p(T) = 0.$$

- Maximization of the Hamiltonian over $u \in [0, 1]$:

$$u^*(t) = \begin{cases} 0 & \text{if } p(t) < \frac{1}{\gamma}, \\ 1 & \text{if } p(t) \ge \frac{1}{\gamma}. \end{cases}$$

Since $p(T) = 0$, for t close to T, $p(t) < 1/\gamma$ and $u^*(t) = 0$. Therefore, for t near T the adjoint equation has the form $\dot{p}(t) = -1$.

ANALYTICAL EXAMPLE II (CONTINUED)



- For $t = T - 1/\gamma$, $p(t)$ is equal to $1/\gamma$, so $u^*(t)$ changes to $u^*(t) = 1$.
- Geometrical construction