MS-E2148 Dynamic optimization

Lecture 8

Contents

- Deterministic, continuous time problems
- Hamilton-Jacobi-Bellman equation
- Material Bertsekas 3.1, 3.2 and Kirk 3.11

Recap

- DP-algorithm can be applied in many discrete time problems

$$\dot{x}(t) = f(x(t), u(t), t), \qquad 0 \leq t \leq T, \tag{1}$$

where initial state $x(0)$ and final time T are known

- $x(t) \in \mathbb{R}^n$, $\dot{x}(t) \in \mathbb{R}^n$, $u(t) \in U \subset \mathbb{R}^m$, $f : \mathbb{R}^{n+m+1} \to \mathbb{R}^n$
- The state variable x, its time derivative $\dot{x}$, and control u are vectors; the **system equations** (1) are first-order differential equations (n equations)
- Assume that f_i are continuously differentiable in x and continuous in u

- The objective is to find the admissible control that minimizes the cost

$$h(x(T), T) + \int_0^T g(x(t), u(t), t)dt \qquad (2)$$

- g is the (instant) cost function, h is the final/terminal cost
- Functions h and g are continuously differentiable in x and g is continuous in u

Hamilton-Jacobi-Bellman equation

- Let us derive the corresponding equations for DP in continuous time by discretizing the problem
- The result is a partial differential equation which gives the solution for the optimal cost-to go

- Time interval $[0, T]$ is split into N parts: $\delta = \frac{T}{N}$
- Discrete time state and controls are

$$\begin{aligned} x_k &= x(k\delta), \\ u_k &= u(k\delta), \end{aligned} \qquad k = 0, 1, ..., N$$

Hamilton-Jacobi-Bellman equation

- The continuous-time system (1) is approximated with Euler discretization:

$$x_{k+1} = x_k + f(x_k, u_k, k\delta) \cdot \delta \tag{3}$$

and the cost function (2) is

$$h(x_N, N\delta) + \sum_{k=0}^{N-1} g(x_k, u_k, k\delta) \cdot \delta \tag{4}$$

Hamilton-Jacobi-Bellman equation

$J^*(t, x)$: optimal cost-to-go at time t and state x for the continuous-time problem

$\tilde{J}^*(t, x)$: optimal cost-to-go at time t and state x for the discrete-time problem

- DP-algorithm:

$$
\begin{aligned}
\tilde{J}^*(N\delta, x) &= h(x), \\
\tilde{J}^*(k\delta, x) &= \min_{u \in U} [g(x, u, k\delta)\delta + \tilde{J}^*((k+1)\delta, x + f(x, u, k\delta)\delta)]
\end{aligned}
$$

for all $k = 0, 1, ..., N-1$; note: $t = k\delta$

Hamilton-Jacobi-Bellman equation

- Let us expand $\tilde{J}^*$ by the first-order Taylor series at $(k\delta, x)$:

$$\begin{aligned}\tilde{J}^*((k+1)\delta, x + f(x,u,k\delta)\delta) \quad = \quad & \tilde{J}^*(k\delta, x) + \nabla_t \tilde{J}^*(k\delta, x)\delta \\ & + \nabla_x \tilde{J}^*(k\delta, x)^T f(x,u,k\delta)\delta \\ & + o(\delta)\end{aligned}$$

  where the higher order tems vanish: $\lim_{\delta\to 0} o(\delta)/\delta = 0$
- Let us substitute the expansion to the previous DP algorithm:

$$\begin{aligned}\tilde{J}^*(k\delta, x) \quad = \quad & \min_{u\in U}\big[g(x,u,k\delta)\delta + \tilde{J}^*(k\delta, x) + \nabla_t \tilde{J}^*(k\delta, x)\delta \\ & \qquad + \nabla_x \tilde{J}^*(k\delta, x)^T f(x,u,k\delta)\delta + o(\delta)\big]\end{aligned}$$

Hamilton-Jacobi-Bellman equation

- Let us divide by δ and assume that

$$\lim_{k\to\infty,\delta\to 0} \tilde{J}^*(k\delta, x) = J^*(t, x), \qquad \forall t, x \tag{5}$$

- The equation for the cost-to-go in the continuous-time problem

$$0 = \min_{u\in U}[g(x,u,t)+\nabla_t J^*(t,x)+\nabla_x J^*(t,x)^T f(x,u,t)], \quad \forall t, x, \tag{6}$$

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

Hamilton-Jacobi-Bellman equation

- Equation (6) is the **HJB equation**
- HJB is a *partial differential equation* (PDE)
- It is solved by the optimal cost-to-go $J^*(t,x)$
- For each optimal cost-to-go, we attach the corresponding optimal control law:

$$u^*(t) = \arg\min_{u \in U} \left[g(x^*, u, t) + \nabla_x J^*(t, x^*)^T f(x^*, u, t) \right] \quad (7)$$

- Optimal control law is not determined for all $x(t)$, but only for one trajectory $x(t) = x^*(t)$; the optimal state trajectory is determined from the optimal cost-to-go $J^*(t, x^*)$

Hamilton-Jacobi-Bellman equation

- Let $J^*(t, x)$ for all t, x be the optimal cost-to-go. Then it satisfies the HJB equation (6).
- Sufficiency. Let $V(t, x)$ be any function satisfying the HJB equation. Then it is an optimal cost-to-go.

- The following formulation simplifies solving HJB
- The **Hamiltonian** is defined

$$H(x, u, p, t) = g(x, u, t) + p^T f(x, u, t) \tag{8}$$

  where $p(t, x) = \nabla_x J^*(t, x)$ is (the Lagrange) *costate variable*
- Now, HJB is

$$0 = \nabla_t J^*(t, x) + H(x, u^*(x, p^*, t), p^*, t) \tag{9}$$

- Now, if we know the control u^* that minimizes the Hamiltonian, we can solve HJB (9)

Notation

- Let us shorten the notation by dropping out the time dependency: $x \equiv x(t)$, $u \equiv u(t)$, $p \equiv p(t)$
- The partial differentials are denoted by subscripts: $F_t \equiv \partial F(t)/\partial t$, also sometimes $\partial_t F \equiv \partial F(t)/\partial t$
- Note the order of differentiation:

$$G_{tx} \equiv \frac{\partial}{\partial x}\left[\frac{\partial}{\partial t} G(t, x(t))\right] = \partial_{xt} G$$

- Note that even though x is time dependent, $x \equiv x(t)$, the partial differential takes x as *independent* variable:

$$G(t, x) = t^2 - xt, \quad G_t = 2t - x, \quad G_x = -t, \quad G_{tx} = -1$$

Hamilton-Jacobi-Bellman equation
Example

- System:

$$\dot{x} = x + u$$

- Cost:

$$\frac{1}{4}x^2(T) + \int_0^T \frac{1}{4}u^2$$

- Hamiltonian:

$$H(x, u, p) = \frac{1}{4}u^2 + px + pu$$

- What u minimizes the Hamiltonian? The first-order condition:

$$\frac{\partial H(x,u,p)}{\partial u} = \frac{1}{2}u + p = 0 \quad \Rightarrow u^* = -2p \tag{10}$$

- HJB, where $p = J_x^*$:

$$\begin{aligned} 0 &= J_t^* + \tfrac{1}{4}(-2p)^2 + px - 2p^2 \\ &= J_t^* - p^2 + px \\ &= J_t^* - (J_x^*)^2 + xJ_x^* \end{aligned} \tag{11}$$

- Let us try to find a solution of the form, where $K \equiv K(t)$:

$$J = \frac{1}{2}Kx^2, \quad J_x = Kx, \quad J_t = \frac{1}{2}\dot{K}x^2 \tag{12}$$

- Final condition $J^*(T,x) = \frac{1}{4}x^2(T)$ gives $K(T) = \frac{1}{2}$
- Let us substitute the trial (12) into HJB:

$$0 = \frac{1}{2}\dot{K}x^2 - K^2x^2 + Kx^2 \tag{13}$$

- This equation must be satisfied $\forall x$:

$$\frac{1}{2}\dot{K} - K^2 + K = 0 \tag{14}$$

- By separation and using the final condition:

$$K = \frac{e^{T-t}}{e^{T-t} + e^{t-T}} \tag{15}$$

- Now, we have

$$J_x^* = \frac{e^{T-t}x}{e^{T-t} + e^{t-T}}, \qquad u^* = -\frac{2e^{T-t}x}{e^{T-t} + e^{t-T}} \tag{16}$$

that are functions of time-state pairs (t, x)

- Note: the solutions to ordinary differential equations usually contain arbitrary constants; partial differential equations usually contain arbitrary *functions* (like $K(t)$ here)

Hamilton-Jacobi-Bellman equation

- The necessary condition for optimality: J^* has to satisfy HJB equation
- HJB is also a sufficient condition
- Terminology: $J^* = V(t, x)$, where the solution V to HJB is called the *value function*
- For example, HJBs in finance give infinite horizon stochastic PDEs; HJB is also important for modern macro economics
- HJB is usually solved by numeric integration

Summary

- Continuous-time problem
- Discretizing this problem, DP algorithm and derivation of HJB equation
- Hamiltonian