Dynamic Programming & Optimal Control

Lecture 9
Deterministic Continuous-Time Optimal Control and the HJB Equation

Fall 2023

Prof. Raffaello D'Andrea

ETH Zurich

Learning Objectives

Topic: Deterministic Continuous-Time Optimal Control and the HJB Equation

Objectives

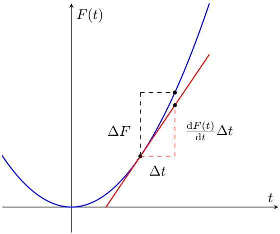
- You know the *continuous-time optimal control* setup.
- You know the *Hamilton-Jacobi-Bellman (HJB)* equation and its properties and limitations.
- You know the connection between *Dynamic Programming* and the HJB equation.

Outline

Calculus refresher (1/8)

The derivative of a function $F : \mathbb{R} \rightarrow \mathbb{R}$ is:

$$\frac{\mathrm{d}F(t)}{\mathrm{d}t} = \lim_{h \to 0} \frac{F(t+h) - F(t)}{h}.$$



Example: Some functions and their derivatives.

- $F(t) = t^2 \Rightarrow \dfrac{\mathrm{d}F(t)}{\mathrm{d}t} = 2t$.

  When evaluated at $\bar{t} = 10$, $\left.\dfrac{\mathrm{d}F(t)}{\mathrm{d}t}\right|_{t=\bar{t}} = 20$.

- $F(t) = \sin(t) \Rightarrow \dfrac{\mathrm{d}F(t)}{\mathrm{d}t} = \cos(t)$.

  When evaluated at $\bar{t} = \pi$, $\left.\dfrac{\mathrm{d}F(t)}{\mathrm{d}t}\right|_{t=\bar{t}} = -1$.

Calculus refresher (2/8)

Given two functions $F : \mathbb{R} \to \mathbb{R}$, $G : \mathbb{R} \to \mathbb{R}$, the derivative of the composition $F \circ G : \mathbb{R} \to \mathbb{R}$ is given by the chain rule:

$$\frac{\mathrm{d}(F \circ G)(t)}{\mathrm{d}t} = \frac{\mathrm{d}F(G(t))}{\mathrm{d}t} = \left.\frac{\mathrm{d}F(\mathrm{x})}{\mathrm{d}\mathrm{x}}\right|_{\mathrm{x}=G(t)} \frac{\mathrm{d}G(t)}{\mathrm{d}t}.$$

Example: Chain rule.

Let $F(\mathrm{x}) = \mathrm{x}^2$ and $G(t) = \sin(t)$. Then:

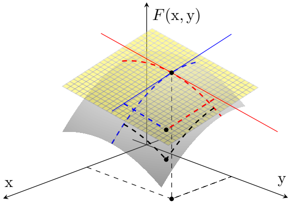
$$\begin{aligned}\frac{\mathrm{d}F(G(t))}{\mathrm{d}t} &= \left.\frac{\mathrm{d}F(\mathrm{x})}{\mathrm{d}\mathrm{x}}\right|_{\mathrm{x}=G(t)} \frac{\mathrm{d}G(t)}{\mathrm{d}t} \\ &= \left. 2\mathrm{x}\right|_{\mathrm{x}=G(t)} \cos(t) \\ &= 2\sin(t)\cos(t).\end{aligned}$$

Calculus refresher (3/8)

Consider now a function of two variables, $F(\mathrm{x},\mathrm{y})$. The partial derivative of $F(\cdot,\cdot)$ with respect to its first argument, x, is

$$\frac{\partial F(\mathrm{x},\mathrm{y})}{\partial \mathrm{x}} = \lim_{h\to 0} \frac{F(\mathrm{x}+h,\mathrm{y}) - F(\mathrm{x},\mathrm{y})}{h}.$$

Analogously we define the partial derivative with respect to y.



Example: Partial derivatives.

Consider $F(\mathrm{x},\mathrm{y}) = \mathrm{x}^2\mathrm{y}$. Then:

$$\frac{\partial F(\mathrm{x},\mathrm{y})}{\partial \mathrm{x}} = 2\mathrm{x}\mathrm{y} \quad \text{and} \quad \frac{\partial F(\mathrm{x},\mathrm{y})}{\partial \mathrm{y}} = \mathrm{x}^2.$$

Calculus refresher (4/8)

The partial derivative of $F(\mathrm{x},\mathrm{y})$ with respect to x approximates the variation of the value of the function due to a small variation in x only. Thus,

$$\Delta F \approx \frac{\partial F(\mathrm{x},\mathrm{y})}{\partial \mathrm{x}} \Delta \mathrm{x} + \frac{\partial F(\mathrm{x},\mathrm{y})}{\partial \mathrm{y}} \Delta \mathrm{y}.$$

When there is a dependency between x and y, and it is captured via a third variable t as $\mathrm{x} = x(t)$ and $\mathrm{y} = y(t)$, we expect $\Delta \mathrm{x} \approx \frac{\mathrm{d}x(t)}{\mathrm{d}t}\Delta t$ and $\Delta \mathrm{y} \approx \frac{\mathrm{d}y(t)}{\mathrm{d}t}\Delta t$. Thus:

$$\Delta F \approx \frac{\partial F(\mathrm{x},\mathrm{y})}{\partial \mathrm{x}} \frac{\mathrm{d}x(t)}{\mathrm{d}t} \Delta t + \frac{\partial F(\mathrm{x},\mathrm{y})}{\partial \mathrm{y}} \frac{\mathrm{d}y(t)}{\mathrm{d}t} \Delta t.$$

As we will see shortly, this is an application of a more general chain rule.

Calculus refresher (5/8)

The Jacobian of $F(\mathrm{x}_1,\ldots,\mathrm{x}_n) = \begin{bmatrix} F_1(\mathrm{x}_1,\ldots,\mathrm{x}_n) & \ldots & F_p(\mathrm{x}_1,\ldots,\mathrm{x}_n) \end{bmatrix}^\top$ is:

$$\frac{\partial(F_1,\ldots,F_p)(\mathrm{x}_1,\ldots,\mathrm{x}_n)}{\partial(\mathrm{x}_1,\ldots,\mathrm{x}_n)} = \begin{bmatrix} \frac{\partial F_1(\mathrm{x}_1,\ldots,\mathrm{x}_n)}{\partial \mathrm{x}_1} & \ldots & \frac{\partial F_1(\mathrm{x}_1,\ldots,\mathrm{x}_n)}{\partial \mathrm{x}_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_p(\mathrm{x}_1,\ldots,\mathrm{x}_n)}{\partial \mathrm{x}_1} & \ldots & \frac{\partial F_p(\mathrm{x}_1,\ldots,\mathrm{x}_n)}{\partial \mathrm{x}_n} \end{bmatrix}.$$

Example: Simple Jacobian.

Consider $F(\mathrm{x}_1,\mathrm{x}_2,\mathrm{x}_3) = \begin{bmatrix} F_1(\mathrm{x}_1,\mathrm{x}_2,\mathrm{x}_3) & F_2(\mathrm{x}_1,\mathrm{x}_2,\mathrm{x}_3) \end{bmatrix}^\top = \begin{bmatrix} \mathrm{x}_1\mathrm{x}_2^2 & \mathrm{x}_2\mathrm{x}_3 \end{bmatrix}^\top$.
Then:

$$\frac{\partial(F_1,F_2)(\mathrm{x}_1,\mathrm{x}_2,\mathrm{x}_3)}{\partial(\mathrm{x}_1,\mathrm{x}_2,\mathrm{x}_3)} = \begin{bmatrix} \mathrm{x}_2^2 & 2\mathrm{x}_1\mathrm{x}_2 & 0 \\ 0 & \mathrm{x}_3 & \mathrm{x}_2 \end{bmatrix}.$$

Calculus refresher (6/8)

The Jacobian is the derivative for vector-valued, multiple-variable functions:

$$\frac{\mathrm{d}F(\mathrm{x})}{\mathrm{dx}} = \frac{\partial(F_1,\ldots,F_p)(\mathrm{x}_1,\ldots,\mathrm{x}_n)}{\partial(\mathrm{x}_1,\ldots,\mathrm{x}_n)}, \quad \mathrm{x} = \begin{bmatrix}\mathrm{x}_1 & \ldots & \mathrm{x}_n\end{bmatrix}^\top .$$

It captures the variation of the value of $F(\cdot)$ due to variations of all the arguments:

$$\begin{aligned}\Delta F = \begin{bmatrix}\Delta F_1 \\ \vdots \\ \Delta F_p\end{bmatrix} &\approx \underbrace{\frac{\partial(F_1,\ldots,F_p)(\mathrm{x}_1,\ldots,\mathrm{x}_n)}{\partial(\mathrm{x}_1,\ldots,\mathrm{x}_n)}}_{=\frac{\mathrm{d}F(\mathrm{x})}{\mathrm{dx}}} \underbrace{\begin{bmatrix}\Delta\mathrm{x}_1 \\ \vdots \\ \Delta\mathrm{x}_n\end{bmatrix}}_{=\Delta\mathrm{x}} \\ &= \begin{bmatrix}\sum_{j=1}^n \frac{\partial F_1(\mathrm{x}_1,\ldots,\mathrm{x}_n)}{\partial\mathrm{x}_j}\Delta\mathrm{x}_j \\ \vdots \\ \sum_{j=1}^n \frac{\partial F_p(\mathrm{x}_1,\ldots,\mathrm{x}_n)}{\partial\mathrm{x}_j}\Delta\mathrm{x}_j\end{bmatrix}.\end{aligned}$$

Calculus refresher (7/8)

Given $F : \mathbb{R}^n \to \mathbb{R}^p$ and $G : \mathbb{R}^k \to \mathbb{R}^n$, the chain rule reads:

$$\frac{\mathrm{d}F(G(\mathrm{x}))}{\mathrm{dx}} = \left.\frac{\mathrm{d}F(\mathrm{y})}{\mathrm{dy}}\right|_{\mathrm{y}=G(\mathrm{x})} \frac{\mathrm{d}G(\mathrm{x})}{\mathrm{dx}}.$$

Example: Chain rule.

Consider $\mathrm{y} = \begin{bmatrix}\mathrm{y}_1 & \mathrm{y}_2\end{bmatrix}^\top$, $\mathrm{x} = \begin{bmatrix}\mathrm{x}_1 & \mathrm{x}_2\end{bmatrix}^\top$, $F(\mathrm{y}) = \mathrm{y}_1 + \mathrm{y}_2^2$ and $G(\mathrm{x}) = \begin{bmatrix}\mathrm{x}_1\mathrm{x}_2^2 & \mathrm{x}_2\end{bmatrix}^\top$. Then:

$$\begin{aligned}\frac{\mathrm{d}F(G(\mathrm{x}))}{\mathrm{dx}} &= \left.\frac{\mathrm{d}F(\mathrm{y})}{\mathrm{dy}}\right|_{\mathrm{y}=G(\mathrm{x})} \frac{\mathrm{d}G(\mathrm{x})}{\mathrm{dx}} \\ &= \left.\begin{bmatrix}1 & 2\mathrm{y}_2\end{bmatrix}\right|_{\mathrm{y}=G(\mathrm{x})} \begin{bmatrix}\mathrm{x}_2^2 & 2\mathrm{x}_1\mathrm{x}_2 \\ 0 & 1\end{bmatrix} \\ &= \begin{bmatrix}1 & 2\mathrm{x}_2\end{bmatrix} \begin{bmatrix}\mathrm{x}_2^2 & 2\mathrm{x}_1\mathrm{x}_2 \\ 0 & 1\end{bmatrix} \\ &= \begin{bmatrix}\mathrm{x}_2^2 & 2\mathrm{x}_2(\mathrm{x}_1+1)\end{bmatrix}.\end{aligned}$$

Calculus refresher (8/8)

For $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$, $G : \mathbb{R}^k \times \mathbb{R}^l \to \mathbb{R}^n$ the partial derivative of $F(\cdot,\cdot)$ with respect to the first argument (which is a vector!) is:

$$\begin{aligned}\frac{\partial F(\mathrm{z},\mathrm{y})}{\partial \mathrm{z}} &= \frac{\partial(F_1,\dots,F_p)(\mathrm{z}_1,\dots,\mathrm{z}_n,\mathrm{y}_1,\dots,\mathrm{y}_m)}{\partial(\mathrm{z}_1,\dots,\mathrm{z}_n)} \\ &= \begin{bmatrix} \frac{\partial F_1(\mathrm{z}_1,\dots,\mathrm{z}_n,\mathrm{y}_1,\dots,\mathrm{y}_m)}{\partial \mathrm{z}_1} & \cdots & \frac{\partial F_1(\mathrm{z}_1,\dots,\mathrm{z}_n,\mathrm{y}_1,\dots,\mathrm{y}_m)}{\partial \mathrm{z}_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_p(\mathrm{z}_1,\dots,\mathrm{z}_n,\mathrm{y}_1,\dots,\mathrm{y}_m)}{\partial \mathrm{z}_1} & \cdots & \frac{\partial F_p(\mathrm{z}_1,\dots,\mathrm{z}_n,\mathrm{y}_1,\dots,\mathrm{y}_m)}{\partial \mathrm{z}_n} \end{bmatrix},\end{aligned}$$

and the chain rule reads:

$$\frac{\partial F(G(\mathrm{x},\mathrm{u}),\mathrm{y})}{\partial \mathrm{x}} = \left.\frac{\partial F(\mathrm{z},\mathrm{y})}{\partial \mathrm{z}}\right|_{\mathrm{z}=G(\mathrm{x},\mathrm{u})} \frac{\partial G(\mathrm{x},\mathrm{u})}{\partial \mathrm{x}}.$$

Calculus refresher (recap)

In the remaining lectures:

- when a function has a single variable (vector or scalar), $\dfrac{\mathrm{d}F(\mathrm{x})}{\mathrm{dx}} = \dfrac{\partial F(\mathrm{x})}{\partial \mathrm{x}}$;
- the partial derivative of $F(t, \mathrm{x}, \mathrm{u})$ with respect to t, when subject to $\mathrm{u} = \mu(t, \mathrm{x})$, is

$$\frac{\partial F(t,\mathrm{x},\mu(t,\mathrm{x}))}{\partial t} = \left.\frac{\partial F(t,\mathrm{x},\mathrm{u})}{\partial t}\right|_{\mathrm{u}=\mu(t,\mathrm{x})} + \left.\frac{\partial F(t,\mathrm{x},\mathrm{u})}{\partial \mathrm{u}}\right|_{\mathrm{u}=\mu(t,\mathrm{x})} \frac{\partial \mu(t,\mathrm{x})}{\partial t};$$

- the derivative of $V(t, \mathrm{x})$ when subject to $\mathrm{x} = x(t)$ is

$$\frac{\mathrm{d}V(t,x(t))}{\mathrm{d}t} = \left.\frac{\partial V(t,\mathrm{x})}{\partial t}\right|_{\mathrm{x}=x(t)} \underbrace{\frac{\mathrm{d}t}{\mathrm{d}t}}_{=1} + \left.\frac{\partial V(t,\mathrm{x})}{\partial \mathrm{x}}\right|_{\mathrm{x}=x(t)} \frac{\mathrm{d}x(t)}{\mathrm{d}t};$$

- we make the approximation argument formal via Taylor expansion:

$$\Delta V = V(t+\Delta t, x(t+\Delta t)) - V(t, x(t)) = \frac{\mathrm{d}V(t,x(t))}{\mathrm{d}t}\Delta t + o(\Delta t),$$

where $\lim_{\Delta t \to 0}(o(\Delta t)/\Delta t) = 0$.

Outline

Problem Setup (1/4)

Dynamics

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

where

- time $t \in \mathbb{R}_{\geq 0}$ and T is the terminal time;
- state $x(t) \in \mathcal{S} := \mathbb{R}^n, \ \forall t \in [0, T]$;
- control input $u(t) \in \mathcal{U} \subseteq \mathbb{R}^m, \ \forall t \in [0, T]$, $\mathcal{U}$ is the input space;
- $f(\cdot, \cdot)$ is the function capturing the system evolution.

Problem Setup (2/4)

Feedback control law

Let $\mu(\cdot, \cdot)$ be an admissible control law that maps state $\mathrm{x} \in \mathcal{S}$ at time t to the control input:

$$\mu(t, \mathrm{x}) \in \mathcal{U}, \ \forall t \in [0, T], \ \forall \mathrm{x} \in \mathcal{S}.$$

Let Π denote the set of all admissible control laws.

Problem Setup (3/4)

Cost

We consider the following scalar-valued cost function:

$$h(x(T)) + \int_0^T g(x(\tau), u(\tau))\mathrm{d}\tau,$$

where $u(\tau) \in \mathcal{U}$ for all $\tau \in [0, T]$.

For an initial time t and state $\mathrm{x} \in \mathcal{S}$, the closed loop cost associated with feedback control law $\mu(\cdot, \cdot) \in \Pi$ is:

$$J_\mu(t, \mathrm{x}) := h(x(T)) + \int_t^T g(x(\tau), \mu(\tau, x(\tau)))\mathrm{d}\tau,$$

subject to

$$\begin{aligned} \dot{x}(\tau) &= f(x(\tau), \mu(\tau, x(\tau))), \qquad t \leq \tau \leq T, \\ x(t) &= \mathrm{x}\,. \end{aligned}$$

Problem Setup (4/4)

Objective

Construct an optimal feedback control law $\mu^* \in \Pi$ such that:

$$J_{\mu^*}(0, \mathrm{x}) \leq J_\mu(0, \mathrm{x})\,, \quad \forall \mu \in \Pi,\, \forall \mathrm{x} \in \mathcal{S}.$$

The associated closed loop cost $J^*(t, \mathrm{x}) := J_{\mu^*}(t, \mathrm{x})$ is called the optimal cost-to-go at state x and time t, and $J^*(\cdot, \cdot)$ is the optimal cost-to-go function or optimal value function.

Assumption 9.1: Existence and uniqueness
For any admissible control input $u(\cdot)$, initial time $t \in [0, T]$ and initial condition $x(t) \in \mathcal{S}$, there exists a unique state trajectory $x(\tau)$ that satisfies: $$\dot{x}(\tau) = f(x(\tau), u(\tau)), \qquad t \leq \tau \leq T.$$

Assumption 9.1 is required for the problem to be well defined. Ensuring that it is satisfied for a particular problem requires tools from the theory of differential equations, and is beyond the scope of this class.

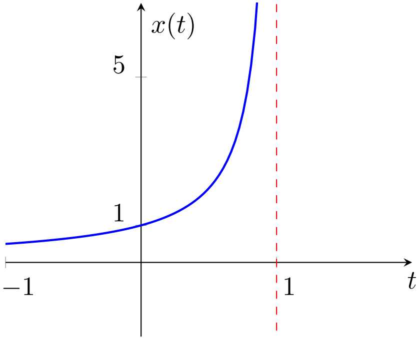
Example: Existence

Consider $\dot{x}(t) = x(t)^2$, $x(0) = 1$.
The solution to the ODE is

$$x(t) = \frac{1}{1-t}$$

$\Rightarrow$ it has a finite escape time: $x(t) \to \infty$ as $t \to 1$

$\Rightarrow$ a solution does not exist for $T \geq 1$.



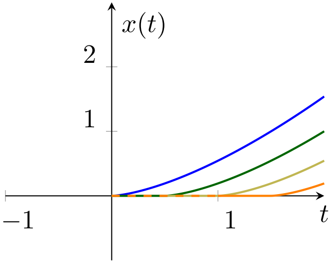
Example: Uniqueness

Consider $\dot{x}(t) = x(t)^{\frac{1}{3}}$, $x(0) = 0$.
Infinitely many solutions exist:

$$x(t) = 0 \;\; \forall t \geq 0 \tag{1}$$

or, for any $\tau \geq 0$,

$$x(t) = \begin{cases} 0 & \text{for } 0 \leq t \leq \tau \\ \left(\frac{2}{3}(t-\tau)\right)^{3/2} & \text{for } t > \tau \end{cases} \tag{2}$$



Outline

Deterministic Continuous-Time Optimal Control and the HJB Equation

Preliminaries

Problem Setup

The Hamilton-Jacobi-Bellman (HJB) Equation

Additional reading material

The HJB (1/4)

We now derive a partial differential equation (PDE) which is satisfied by the optimal cost-to-go function under certain assumptions that we will see later.

This PDE is the continuous-time analog of the DPA, and will be motivated applying the DPA to a discrete-time approximation of the continuous-time optimal control problem.

First, divide the time horizon $[0, T]$ into N pieces and define $\delta := \frac{T}{N}$.

Let $x_k := x(k\delta)$ for $k = 0, 1, ..., N$, $u_k := u(k\delta)$ for $k = 0, 1, ..., N-1$, and approximate the differential equation $\dot{x}(k\delta) = f(x(k\delta),\, u(k\delta))$ by:

$$\frac{x_{k+1} - x_k}{\delta} = f(x_k, u_k), \quad k = 0, 1, ..., N-1,$$

which is equivalent to

$$x_{k+1} = x_k + f(x_k, u_k)\delta, \quad k = 0, 1, ..., N-1.$$

The HJB (2/4)

Similarly, the cost function is approximated by

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

The state space and the control space remain unchanged.

Let $J_k(\mathrm{x})$ be the cost-to-go at stage k and state x for the auxiliary problem and apply the DPA:

$$J_N(\mathrm{x}) = h(\mathrm{x}), \quad \forall \mathrm{x} \in \mathcal{S},$$

$$J_k(\mathrm{x}) = \min_{\mathrm{u} \in \mathcal{U}} \left(\underbrace{g(\mathrm{x}, \mathrm{u})\delta}_{\text{stage cost}} + \underbrace{J_{k+1}(\mathrm{x} + f(\mathrm{x}, \mathrm{u})\delta)}_{\text{cost-to-go at stage } k+1} \right), \quad \forall \mathrm{x} \in \mathcal{S},\ k = N-1, ..., 0.$$

The HJB (3/4)

Assume $J_k(\mathrm{x}) = J^*(k\delta, \mathrm{x}) + o(\delta)$ for all $\mathrm{x} \in \mathcal{S}$, $k = 0, \ldots, N-1$, where $\lim_{\delta \to 0} (o(\delta)/\delta) = 0$.

Thus, for any $\mathrm{u} \in \mathcal{U}$,

$$J_{k+1}(\mathrm{x} + f(\mathrm{x}, \mathrm{u})\delta) = J^*((k+1)\delta,\ \mathrm{x} + f(\mathrm{x}, \mathrm{u})\delta) + o(\delta).$$

Assuming the optimal cost-to-go function $J^*(\cdot, \cdot)$ of the continuous-time formulation is differentiable with respect to t and x, we can express it using a first order Taylor series around $(k\delta, \mathrm{x})$, and evaluate it at $((k+1)\delta, \mathrm{x} + f(\mathrm{x}, \mathrm{u})\delta)$:

$$\begin{aligned} &J^*((k+1)\delta, \mathrm{x} + f(\mathrm{x}, \mathrm{u})\delta) \\ &\qquad = J^*(k\delta, \mathrm{x}) + \frac{\partial J^*(k\delta, \mathrm{x})}{\partial t}\delta + \frac{\partial J^*(k\delta, \mathrm{x})}{\partial \mathrm{x}} f(\mathrm{x}, \mathrm{u})\delta + o(\delta). \end{aligned}$$

The HJB (4/4)

Thus, with $t = k\delta$,

$$
\begin{aligned}
J^*(t, \mathrm{x}) &= \min_{\mathrm{u} \in \mathcal{U}} \left(g(\mathrm{x}, \mathrm{u})\delta + J^*(t, \mathrm{x}) + \frac{\partial J^*(t, \mathrm{x})}{\partial t}\delta + \frac{\partial J^*(t, \mathrm{x})}{\partial \mathrm{x}} f(\mathrm{x}, \mathrm{u})\delta + o(\delta) \right) \\
\Leftrightarrow 0 &= \min_{\mathrm{u} \in \mathcal{U}} \left(g(\mathrm{x}, \mathrm{u})\delta + \frac{\partial J^*(t, \mathrm{x})}{\partial t}\delta + \frac{\partial J^*(t, \mathrm{x})}{\partial \mathrm{x}} f(\mathrm{x}, \mathrm{u})\delta + o(\delta) \right) \\
\Leftrightarrow 0 &= \min_{\mathrm{u} \in \mathcal{U}} \left(g(\mathrm{x}, \mathrm{u}) + \frac{\partial J^*(t, \mathrm{x})}{\partial t} + \frac{\partial J^*(t, \mathrm{x})}{\partial \mathrm{x}} f(\mathrm{x}, \mathrm{u}) + \frac{o(\delta)}{\delta} \right).
\end{aligned}
$$

Taking the limit of the above equation as $N \to \infty$ (i.e., $\delta \to 0$):

$$
0 = \min_{\mathrm{u} \in \mathcal{U}} \left(g(\mathrm{x}, \mathrm{u}) + \frac{\partial J^*(t, \mathrm{x})}{\partial t} + \frac{\partial J^*(t, \mathrm{x})}{\partial \mathrm{x}} f(\mathrm{x}, \mathrm{u}) \right), \quad \forall t \in [0, T], \, \forall \mathrm{x} \in \mathcal{S},
$$

subject to the terminal condition $J^*(T, \mathrm{x}) = h(\mathrm{x})$ for all $\mathrm{x} \in \mathcal{S}$. This is known as the **Hamilton-Jacobi-Bellman (HJB)** equation.

Example 1 (1/4)

Consider

$$\dot{x}(t) = u(t), \quad |u(t)| \leq 1,\ 0 \leq t \leq 1,$$

with cost function

$$\frac{1}{2}x(1)^2,$$

that is $h(x(1)) = \frac{1}{2}x(1)^2$ and $g(\mathrm{x}, \mathrm{u}) = 0$ for all $\mathrm{x} \in \mathcal{S}$ and $\mathrm{u} \in \mathcal{U}$.
Since we only care about the square of the terminal state, we can construct a candidate optimal policy that drives the state towards 0 as quickly as possible, and maintaining it at 0 once it is at 0:

$$\begin{aligned} \mu(t, \mathrm{x}) &= \begin{cases} -1 & \text{if } \mathrm{x} > 0 \\ 0 & \text{if } \mathrm{x} = 0 \\ 1 & \text{if } \mathrm{x} < 0 \end{cases} \\ &= -\operatorname{sgn}(\mathrm{x}). \end{aligned}$$

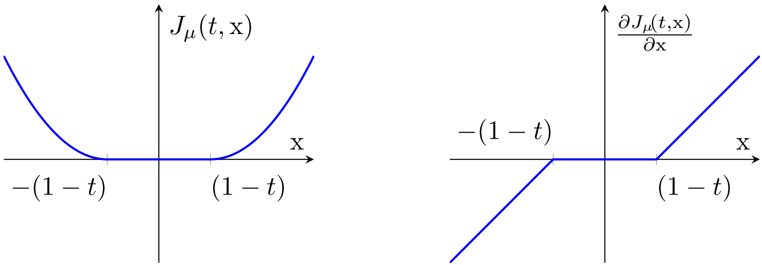
Example 1 (2/4)

For a given initial time t and initial state x, the cost $J_\mu(t, \mathrm{x})$ associated with this policy is:

$$J_\mu(t, \mathrm{x}) = \frac{1}{2}\left(\max\{0, |\mathrm{x}| - (1-t)\}\right)^2 .$$

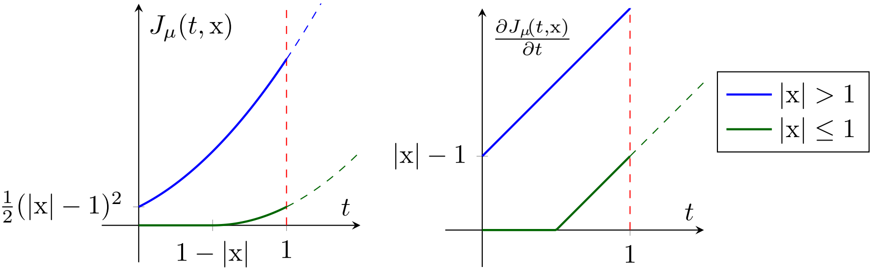
We will now verify that this cost function satisfies the HJB equation. In theorem (9.1) we will prove that this implies that J_μ is indeed the optimal cost-to-go function.

For a fixed t, $\frac{\partial J_\mu(t,\mathrm{x})}{\partial \mathrm{x}} = \mathrm{sgn}(\mathrm{x}) \max\{0, |\mathrm{x}| - (1-t)\}$.



Example 1 (3/4)

For a fixed x, $\frac{\partial J_\mu(t,\mathrm{x})}{\partial t} = \max\{0, |\mathrm{x}| - (1-t)\}$.



Example 1 (4/4)

Note that $J_\mu(1,\mathrm{x}) = \frac{1}{2}\mathrm{x}^2$, so the boundary condition is satisfied. Furthermore,

$$\begin{aligned}\min_{|\mathrm{u}|\leq 1}\left(\frac{\partial J_\mu(t,\mathrm{x})}{\partial t} + \frac{\partial J_\mu(t,\mathrm{x})}{\partial \mathrm{x}}\mathrm{u}\right) &= \min_{|\mathrm{u}|\leq 1}\Big((1+\operatorname{sgn}(\mathrm{x})\mathrm{u})(\max\{0, |\mathrm{x}| - (1-t)\})\Big)\\ &= 0,\end{aligned}$$

where the minimum is attained by $\mathrm{u} = -\operatorname{sgn}(\mathrm{x})$.

Thus $J_\mu(t,\mathrm{x}) = J^*(t,\mathrm{x})$, and $\mu^*(t,\mathrm{x}) = -\operatorname{sgn}(\mathrm{x})$ is an optimal policy.

This was a simple example: In general solving the HJB equation is nontrivial!

Example 2 (1/2)

We will now look at a problem for which the optimal cost-to-go may not be smooth. The HJB equation as we introduced it entails differentiability of the cost-to-go. We will elaborate on non-smooth solutions at the end of the lecture.

Consider the system

$$\dot{x}(t) = x(t)u(t), \quad |u(t)| \leq 1, \ 0 \leq t \leq 1,$$

with cost $h(x(1)) = x(1)$ and $g(\mathrm{x}, \mathrm{u}) = 0$ for all $\mathrm{x} \in \mathcal{S}$ and $\mathrm{u} \in \mathcal{U}$.

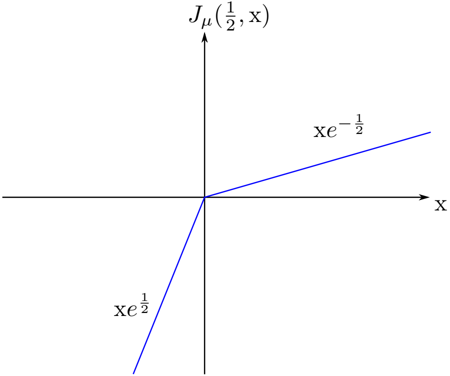
One can show that an optimal policy is the following:

$$\mu(t, \mathrm{x}) = \begin{cases} -1 & \text{if } \mathrm{x} > 0 \\ 0 & \text{if } \mathrm{x} = 0 \\ 1 & \text{if } \mathrm{x} < 0 \end{cases}.$$

Its associated cost-to-go function is:

$$J_\mu(t, \mathrm{x}) = \begin{cases} e^{-1+t}\mathrm{x} & \mathrm{x} > 0 \\ e^{1-t}\mathrm{x} & \mathrm{x} < 0 \\ 0 & \mathrm{x} = 0 \end{cases}.$$

Example 2 (2/2)



The associated cost-to-go function is not differentiable with respect to x at $\mathrm{x} = 0$ and $t = \frac{1}{2}$, thereby not satisfying the HJB equation as we introduced it. This illustrates the fact that (this version of) the HJB equation is in general not a necessary condition for optimality, but only sufficient.

Theorem 9.1: Sufficiency of the HJB

Suppose $V(t,\mathrm{x})$ is continuously differentiable in t and x and solves the HJB equation, that is:

$$\min_{\mathrm{u}\in\mathcal{U}}\left[g(\mathrm{x},\mathrm{u})+\frac{\partial V(t,\mathrm{x})}{\partial t}+\frac{\partial V(t,\mathrm{x})}{\partial \mathrm{x}}f(\mathrm{x},\mathrm{u})\right]=0 \quad ,\forall \mathrm{x}\in\mathcal{S},\, 0\le t\le T,$$

$$\text{s.t.} \quad V(T,\mathrm{x})=h(\mathrm{x}), \quad \forall \mathrm{x}\in\mathcal{S}.$$

If Assumption 9.1 holds, $V(t,\mathrm{x})$ is equal to the optimal cost-to-go function:

$$V(t,\mathrm{x})=J^*(t,\mathrm{x}), \quad \forall \mathrm{x}\in\mathcal{S},\, 0\le t\le T.$$

Furthermore, the mapping $\mu(t,\mathrm{x})$ minimizing the HJB equation is an optimal feedback law.

Proof of Theorem 9.1 (1/2)

For any initial time $t \in [0, T]$ and any initial condition $x(t) = \mathrm{x}$, $\mathrm{x} \in \mathcal{S}$, let $\hat{u}(\tau) \in \mathcal{U}$ for all $\tau \in [t, T]$ be any admissible control trajectory, $\hat{x}(\tau)$ be the corresponding state trajectory where $\hat{x}(\tau)$ is the unique solution to the ODE $\dot{x}(\tau) = f(x(\tau), \hat{u}(\tau))$.

From the HJB we have for all $\tau \in [0, T]$,

$$
\begin{aligned}
0 &\leq g(\hat{x}(\tau), \hat{u}(\tau)) + \left.\frac{\partial V(\tau, \mathrm{x})}{\partial \tau}\right|_{\mathrm{x}=\hat{x}(\tau)} + \left.\frac{\partial V(\tau, \mathrm{x})}{\partial \mathrm{x}}\right|_{\mathrm{x}=\hat{x}(\tau)} f(\hat{x}(\tau), \hat{u}(\tau)) \\
&= g(\hat{x}(\tau), \hat{u}(\tau)) + \frac{\mathrm{d}V(\tau, \hat{x}(\tau))}{\mathrm{d}\tau}.
\end{aligned}
$$

Integrating the above inequality over $\tau \in [t, T]$ yields:

$$
\begin{aligned}
0 &\leq \int_t^T g(\hat{x}(\tau), \hat{u}(\tau))\mathrm{d}\tau + V(T, \hat{x}(T)) - V(t, \mathrm{x}) \\
\Leftrightarrow V(t, \mathrm{x}) &\leq h(\hat{x}(T)) + \int_t^T g(\hat{x}(\tau), \hat{u}(\tau))\mathrm{d}\tau.
\end{aligned}
$$

Proof of Theorem 9.1 (2/2)

In the previous slide we obtained

$$V(t, \mathrm{x}) \leq h(\hat{x}(T)) + \int_t^T g(\hat{x}(\tau), \hat{u}(\tau)) \mathrm{d}\tau.$$

The inequality becomes equality for the minimizing $\mu(\tau, x(\tau))$ of the HJB equation:

$$V(t, \mathrm{x}) = h(x(T)) + \int_t^T g(x(\tau), \mu(\tau, x(\tau))) \mathrm{d}\tau,$$

where $x(\tau)$ is the unique solution to the ODE $\dot{x}(\tau) = f(x(\tau), \mu(\tau, x(\tau)))$ with $x(t) = \mathrm{x}$. Thus $V(t, \mathrm{x})$ is the optimal cost-to-go at state x at time t,

$$V(t, \mathrm{x}) = J^*(t, \mathrm{x}), \quad \forall \mathrm{x} \in \mathcal{S}, \forall t \in [0, T],$$

and $\mu(\cdot, \cdot)$ is an optimal feedback law.

The HJB equation – Summary

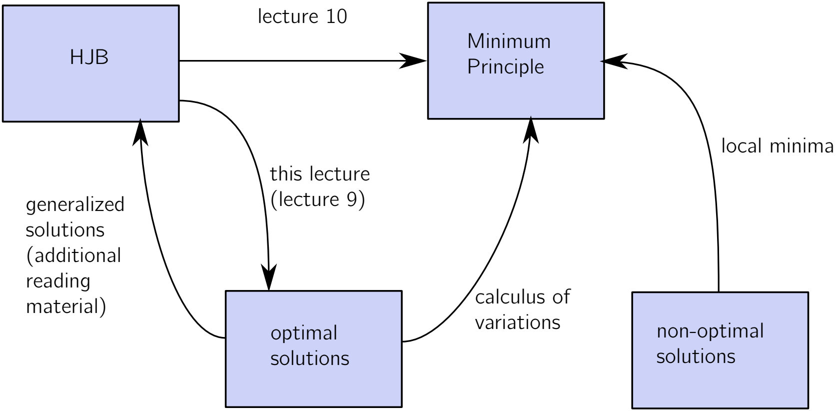


Figure: Optimal solutions and their relation to the HJB equation and the Minimum Principle (lecture 10).

Outline

Deterministic Continuous-Time Optimal Control and the HJB Equation

Preliminaries

Problem Setup

The Hamilton-Jacobi-Bellman (HJB) Equation

Additional reading material

Additional reading material

The HJB equation we introduced in this lecture entails differentiability of the solution, but we have seen that there may be solutions to the optimal control problem that are not differentiable. As a result, the HJB equation we introduced is only a sufficient condition for optimality. However, it is possible to derive another version of the HJB equation that is also a necessary condition for optimality resorting to the so-called viscosity solutions:

- A tutorial on viscosity solutions for the HJB equation: `https://ethz.ch/content/dam/ethz/special-interest/mavt/dynamic-systems-n-control/idsc-dam/Lectures/Optimal-Control/Additional%20Material/tutorial.pdf`

- Variational Analysis, by R. Tyrrell Rockafellar and Roger J. B. Wets.