Aalto University

The Hamilton-Jacobi-Bellman equation

CHEM-E7225 (was E7195), 2020-2021

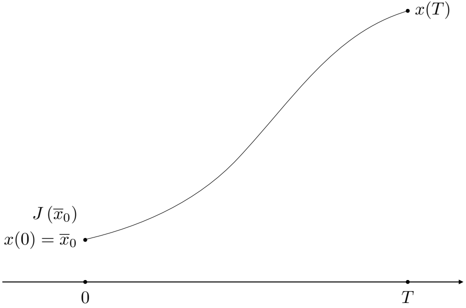
Francesco Corona (¬_¬)

Chemical and Metallurgical Engineering
School of Chemical Engineering

Principle of optimality

We discussed how to solve an optimal control problem using dynamic programming

- The principle of optimality
- Backward recursion



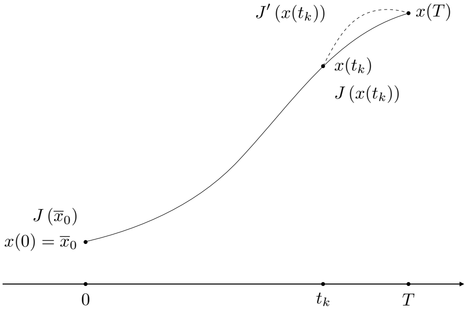
Optimal control trajectory $u^*(\cdot)$

- Transfer from $x(0)$ to $x(T)$
- Optimal cost $J(\overline{x}_0)$

The initial state is fixed, $\overline{x}_0$

Consider a point t_k in between 0 and T, 'an optimal policy has the property that whatever the previous state and control, the remaining controls must constitute an optimal policy with respect to the state resulting from the previous decision'

Principle of optimality (cont.)



Partial optimal control trajectory

$$u^*(t_k \rightsquigarrow T)$$

- Transfer from $x(t_k)$ to $x(T)$
- Optimal cost $J(x(t_k))$

Another partial control trajectory $u'(t_k \rightsquigarrow T)$ which is admissible

- Transfer from $x(t_k)$ to $x(T)$
- Cost $J(x(t_k))$

Suppose that $J(\overline{x}_0)$ is the total optimal cost to transfer the state from $x(0)$ to $x(T)$

- Then $J(x(t_k))$ must be the optimal cost from $x(t_k)$ to $x(T)$

A control trajectory such that $J'(x(t_k)) < J(x(t_k))$ would contradict the assumption

- (Equality may occur if the optimal control is not unique)

Discrete-time dynamic programming

The HJB equation

Optimal control with dynamic programming

We consider an arbitrary discrete-time dynamical system, with initial condition $x_0 = \overline{x}_0$

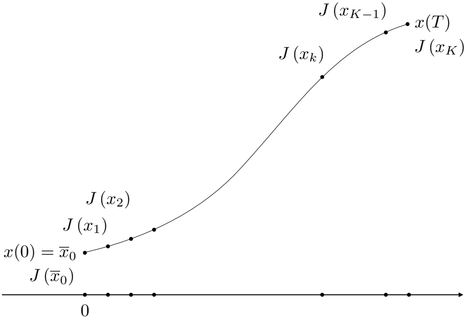
$$x_{k+1} = f\left(x_k, u_k\right), \quad k = 0, \dots, K$$

In general, the dynamics could be time-varying

Consider the cost to transfer the state from each x_k to x_K

$$J\left(x_k, u_k\right) = E\left(x_K\right) + \sum_{k=n}^{K-1} L\left(x_k, u_k\right)$$

$$n = 0, 1, \dots, K-1$$



We are interested in each of the optimal controls u_k^*, by the principle of optimality

Optimal control with dynamic programming (cont.)

At the terminal stage K, we have the optimal cost to transfer the state from x_K to x_K

$$\begin{aligned} J_K(x_K, u_K) &= E(x_K) \\ J_K^*(x_K) &= \min_{u_K} \big(J_K(x_K, u_K)\big) \\ &= J_K(x_K, u_K) \\ &= E(x_K) \end{aligned}$$

At stage $K-1$, the optimal cost to transfer the state from x_{K-1} to x_K

$$\begin{aligned} J_{K-1}(x_{K-1}, u_{K-1}) &= E(x_K) + L(x_{K-1}, u_{K-1}) \\ J_{K-1}^*(x_{K-1}) &= \min_{u_{K-1}} \big(J_{K-1}(x_{K-1}, u_{K-1})\big) \\ &= \min_{u_{K-1}} \Big[L(x_{K-1}, u_{K-1}) + \underbrace{E(x_K)}_{J_K^*(x_K)} \Big] \end{aligned}$$

Optimal control with dynamic programming (cont.)

At stage $K-2$, the optimal cost to transfer the state from x_{K-2} to x_K

$$\begin{aligned}
J_{K-2}(x_{K-2}, u_{K-2}) &= E(x_K) + L(x_{K-1}, u_{K-1}) + L(x_{K-2}, u_{K-2}) \\
J^*_{K-2}(x_{K-2}) &= \min_{u_{K-2}} \left(J_{K-2}(x_{K-2}, u_{K-2})\right) \\
&= \min_{u_{K-2}} \left[L(x_{K-2}, u_{K-2}) + \underbrace{L(x_{K-1}, u_{K-1}) + \underbrace{E(x_K)}_{J^*_K(x_K)}}_{J^*_{K-1}(x_{K-1})} \right]
\end{aligned}$$

The operation is repeated until the initial stage $K=0$, to get

$$J^*(x_0)$$

Optimal control with dynamic programming (cont.)

For an arbitrary stage k, the optimal cost to transfer the state from x_k to x_K

$$\begin{aligned} J_k^*(x_k) &= \min_{u_k} \left(J_{k-1}(x_k, u_k) \right) \\ &= \min_{u_k} \left[L(x_k, u_k) + J_{k+1}^*(x_{k+1}) \right] \end{aligned}$$

Because $x_{k+1} = f(x_k, u_k)$, we can write the dependence on the control u_k

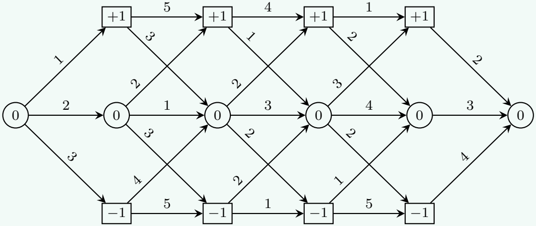
$$\begin{aligned} J_k^*(x_k) &= \min_{u_k} \left[L(x_k, u_k) + J_{k+1}^*(x_{k+1}) \right] \\ &= \min_{u_k} \left[L(x_k, u_k) + J_{k+1}^*(f(x_k, u_k)) \right] \end{aligned}$$

Optimal control with dynamic programming (cont.)

Example

In the **shortest path problem** we consider a directed graph in which the initial node is connected to the terminal node through several admissible paths consisting of arcs

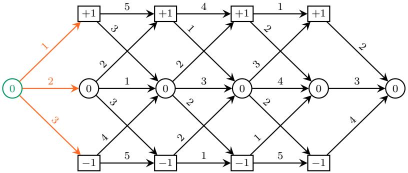
- Each path is associated to a cost, the sum of the costs of each arc



- $k \in \{0, 1, 2, 3, 4, 5\}$ denotes the stage
- $x_k = x(k) \in \{+1, 0, -1\}$ denotes the state
- $u_k = u(k) \in \{\nearrow, \rightarrow, \searrow\}$ denotes the control action

The shortest, or minimum cost, path from node (k, x_k) to the terminal node

$$J\left(k, x_k\right)$$



The shortest path from node $(0,0)$

$$J(0,0) = \min\big((0+1) + J(1,1),\ (0+2) + J(1,0),\ (0+3) + J(1,-1)\big)$$

$(0,0) \rightsquigarrow (1,+1)$, then continue along the shortest path to the terminal node

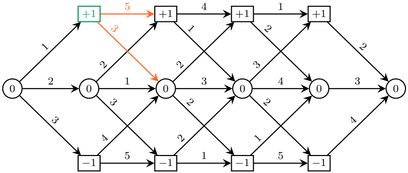
$$\underbrace{\underbrace{0+1}_{L_0(0,\nearrow)} + J(1,+1)}_{\text{path cost}}$$

$(0,0) \rightsquigarrow (1,0)$, then continue along the shortest path to the terminal node

$$\underbrace{\underbrace{0+2}_{L_0(0,\rightarrow)} + J(1,0)}_{\text{path cost}}$$

$(0,0) \rightsquigarrow (1,-1)$, then continue along the shortest path to the terminal node

$$\underbrace{\underbrace{0+3}_{L_0(0,\searrow)} + J(1,-1)}_{\text{path cost}}$$



The shortest path from node $(1,1)$

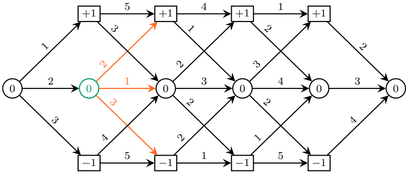
$$J(1,1) = \min(5 + J(2,+1), \; 3 + J(2,0))$$

$(1,1) \rightsquigarrow (2,+1)$, then continue along the shortest path to the terminal node

$$\underbrace{\underbrace{0+5}_{L_1(1,\rightarrow)} + J(2,+1)}_{\text{path cost}}$$

$(1,1) \rightsquigarrow (2,0)$, then continue along the shortest path to the terminal node

$$\underbrace{\underbrace{0+3}_{L1(1,\searrow)} + J(2,0)}_{\text{path cost}}$$



The shortest path from node $(1,0)$

$$J\,(1,0) = \min\big(2 + J\,(2,+1)\,,\; 1 + J\,(2,0)\,,\; 3 + J\,(2,-1)\big)$$

$(1,0) \rightsquigarrow (2,+1)$, then continue along the shortest path to the terminal node

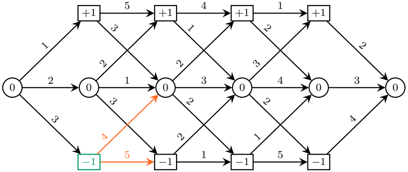
$$\underbrace{\underbrace{2}_{L_1(0,\nearrow)} + J\,(2,+1)}_{\text{path cost}}$$

$(1,0) \rightsquigarrow (2,0)$, then continue along the shortest path to the terminal node

$$\underbrace{\underbrace{1}_{L_1(0,\rightarrow)} + J\,(2,0)}_{\text{path cost}}$$

$(1,0) \rightsquigarrow (2,-1)$, then continue along the shortest path to the terminal node

$$\underbrace{\underbrace{3}_{L_1(0,\searrow)} + J\,(2,-1)}_{\text{path cost}}$$



Shortest path from node $(1, -1)$

$$J\,(1,-1) = \min\big(4 + J\,(2,0)\,,\; 5 + J\,(2,-1)\big)$$

$(1,-1) \rightsquigarrow (2,-1)$, then continue along the shortest path to the terminal node

$$\underbrace{\underbrace{1}_{L_1(-1,\nearrow)} + J\,(2,0)}_{\text{path cost}}$$

$(1,-1) \rightsquigarrow (2,0)$, then continue along the shortest path to the terminal node

$$\underbrace{\underbrace{2}_{L_1(-1,\rightarrow)} + J\,(2,1)}_{\text{path cost}}$$

The shortest path problem (cont.)

In general, the following operations are performed

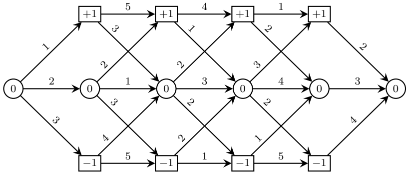
- At each stage k
    - For each state x_k
        - ⇝ Compute the optimal cost

$$J\left(k, x_k\right) = \min_{\nearrow,\rightarrow,\searrow}\left(L_k(x_k, u_k) + \underbrace{J(k+1, f\left(x_k, u_k\right))}_{\text{shortest path from } (k+1,\, x_{k+1})}\right)$$

Principle of optimality

The shortest, or minimum cost, path has the property that for any initial part of the path from the initial node to any node $(k, x_k) \in \{0, \dots, K\} \times \{1, \dots, \mathcal{N}_x\}$, the remaining path must be the shortest, or the cheapest, from node (k, x_k) to the terminal node

The shortest path problem (cont.)



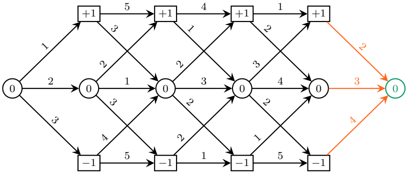
We know that the cost corresponding to the terminal part of the shortest path is given

$$J(5,0) = 0$$

We can optimise backwards and recursively compute the previous shortest path-to-go

$$J(4,x) = \begin{bmatrix} J(4,+1) \\ J(4,0) \\ J(4,-1) \end{bmatrix}$$

There is only one way to reach $(5,0)$ from each of the nodes $(4,x)$



The shortest path-to-go function

$$\underbrace{\begin{bmatrix} +\infty \\ 0 \\ +\infty \end{bmatrix}}_{J(5,x)}$$

$(4,+1) \rightsquigarrow (5,0)$, then continue along the shortest path to the terminal node

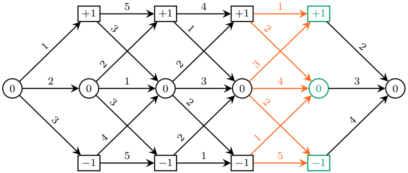
$$J\,(4,1) = \min(\underbrace{2}_{L_5(+1,\searrow)} + \underbrace{J\,(5,0)}_{=0})$$
$$= 2$$

$(4,0) \rightsquigarrow (5,0)$, then continue along the shortest path to the terminal node

$$J\,(4,0) = \min(3 + \underbrace{J\,(5,0)}_{=0})$$
$$= 3$$

$(4,-1) \rightsquigarrow (5,0)$, then continue along the shortest path to the terminal node

$$J\,(4,1) = \min(4 + \underbrace{J\,(5,0)}_{=0})$$
$$= 4$$



The shortest path-to-go function

$$\underbrace{\begin{bmatrix} 2 & +\infty \\ 3 & 0 \\ 4 & +\infty \end{bmatrix}}_{J(4\rightsquigarrow 5, x)}$$

$(3,+1) \rightsquigarrow (4,+1)(4,0)$, then along the shortest path to the terminal node

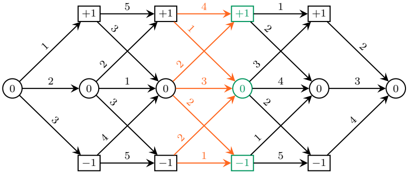
$$J(3,1) = \min(1 + \underbrace{J(4,1)}_{=2}, 3 + \underbrace{J(4,0)}_{=3})$$
$$= 3$$

$(3,0) \rightsquigarrow (4,+1)(4,0)(4,-1)$, then along the shortest path to the terminal node

$$J(3,0) = \min(3 + \underbrace{J(4,1)}_{=2}, 4 + \underbrace{J(4,0)}_{=3}, 2 + \underbrace{J(4,-1)}_{=4})$$
$$= 5$$

$(3,-1) \rightsquigarrow (4,-1)(4,0)$, then along the shortest path to the terminal node

$$J(3,-1) = \min(5 + \underbrace{J(4,-1)}_{=4}, 1 + \underbrace{J(4,0)}_{=3})$$
$$= 4$$



The shortest path-to-go function

$$\underbrace{\begin{bmatrix} 3 & 2 & +\infty \\ 5 & 3 & 0 \\ 4 & 4 & +\infty \end{bmatrix}}_{J(3 \rightsquigarrow 5, x)}$$

$(2, +1) \rightsquigarrow (3, +1)(3, 0)$, then along the shortest path to the terminal node

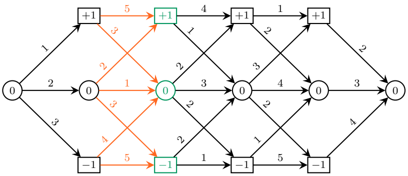
$$J(2,1) = \min(4 + \underbrace{J(3,1)}_{=3}, 1 + \underbrace{J(3,0)}_{=5})$$
$$= 6$$

$(2, 0) \rightsquigarrow (3, +1)(3, 0)(3, -1)$, then along the shortest path to the terminal node

$$J(2,0) = \min(2 + \underbrace{J(3,1)}_{=3}, 3 + \underbrace{J(3,0)}_{=5}, 2 + \underbrace{J(3,-1)}_{=4})$$
$$= 5$$

$(2, -1) \rightsquigarrow (3, -1)(3, 0)$, then along the shortest path to the terminal node

$$J(2,-1) = \min(1 + \underbrace{J(3,-1)}_{=4}, 2 + \underbrace{J(3,0)}_{=5})$$
$$= 5$$



The shortest path-to-go function

$$\underbrace{\begin{bmatrix} 6 & 3 & 2 & +\infty \\ 5 & 5 & 3 & 0 \\ 5 & 4 & 4 & +\infty \end{bmatrix}}_{J(2 \rightsquigarrow 5, x)}$$

$(1, +1) \rightsquigarrow (2, +1)(2, 0)$, then along the shortest path to the terminal node

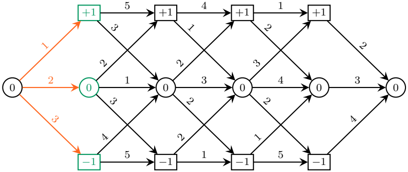
$$J\,(1,1) = \min(5 + \underbrace{J\,(2,1)}_{=6}, 3 + \underbrace{J\,(2,0)}_{=5})$$
$$= 8$$

$(1, 0) \rightsquigarrow (2, +1)(2, 0)(2, -1)$, then along the shortest path to the terminal node

$$J\,(1,0) = \min(2 + \underbrace{J\,(2,1)}_{=6}, 1 + \underbrace{J\,(2,0)}_{=5}, 3 + \underbrace{J\,(2,-1)}_{=5})$$
$$= 6$$

$(1, -1) \rightsquigarrow (2, -1)(2, 0)$, then along the shortest path to the terminal node

$$J\,(1,-1) = \min(4 + \underbrace{J\,(2,-1)}_{=5}, 5 + \underbrace{J\,(2,0)}_{=5})$$
$$= 9$$



The shortest path-to-go function

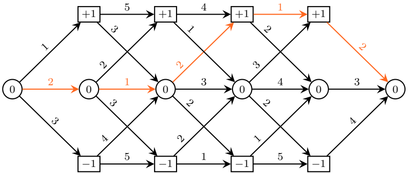
$$\underbrace{\begin{bmatrix} 8 & 6 & 3 & 2 & +\infty \\ 6 & 5 & 5 & 3 & 0 \\ 9 & 5 & 4 & 4 & +\infty \end{bmatrix}}_{J(1 \rightsquigarrow 5, x)}$$

$(0,0) \rightsquigarrow (1,+1)(1,0)(1,-1)$, then along the shortest path to the terminal node

$$J\,(0,0) = \min(1 + \underbrace{J\,(1,1)}_{=8}, 2 + \underbrace{J\,(1,0)}_{=6}, 3 + \underbrace{J\,(1,-1)}_{=9})$$

$$= 8$$

The shortest path problem (cont.)



$$\underbrace{\begin{bmatrix} & 8 & 6 & 3 & 2 & +\infty \\ 8 & 6 & 5 & 5 & 3 & 0 \\ & 9 & 5 & 4 & 4 & +\infty \end{bmatrix}}_{J(0 \rightsquigarrow 5, x)}$$

The optimal controls to the shortest path problem

$$\{u_k^*(x_k)\}_{k=0}^{5}$$

The shortest path-to-go from each node x_k

$$J(k, x)$$

□

Optimal control with dynamic programming (cont.)

Example

Consider the linear and time-invariant dynamical system from $x_0 = 8$

$$x_{k+1} = \underbrace{4x_k - 6u_k}_{f(x_k, u_k)}$$

Consider the objective function as performance index for $K = 2$

$$\underbrace{(x_2 - 20)^2}_{E(x_K)} + \frac{1}{2}\sum_{k=0}^{1} \underbrace{2x_k^2 + 4u_k^2}_{L(x_k, u_k)}$$

We are interested in the optimal controls u_0^* and u_1^*

Optimal control with dynamic programming (cont.)

We know the terminal optimal cost to transfer the state from x_K to x_K with $K = 2$

$$\begin{aligned} J_K\left(x_K, u_K\right) &= E\left(x_K\right) \\ J_K^*\left(x_K\right) &= \min_{u_K}\left(J_K\left(x_K, u_K\right)\right) \\ &= \left(x_2 - 20\right)^2 \\ &= E\left(x_K\right) \end{aligned}$$

Then, we have at stage $K-1$, cost to transfer the state from $x_K - 1$ to x_K

$$\begin{aligned} J_{K-1}\left(x_{K-1}, u_{K-1}\right) &= E\left(x_K\right) + L\left(x_{K-1}, u_{K-1}\right) \\ J_{K-1}^*\left(x_{K-1}\right) &= \min_{u_{K-1}}\left(E\left(x_K\right) + L\left(x_{K-1}, u_{K-1}\right)\right) \\ &= \min_{u_{K-1}}\left(J_K^*\left(x_K\right) + L\left(x_{K-1}, u_{K-1}\right)\right) \\ &= \min_{u_{K-1}}\left(\left(x_K - 20\right)^2 + \frac{1}{2}\left(2x_{K-1}^2 + 4u_{K-1}^2\right)\right) \\ &= \min_{u_{K-1}}\left(\left(\underbrace{\left(4x_{K-1} - 6u_{K-1}\right)}_{x_K} - 20\right)^2 + \frac{1}{2}\left(2x_{K-1}^2 + 4u_{K-1}^2\right)\right) \end{aligned}$$

Optimal control with dynamic programming (cont.)

We obtain the optimisation

$$J^*_{K-1}\left(x_{K-1}\right) = \min_{u_{K-1}} \left(\underbrace{\left(\left(4x_{K-1} - 6u_{K-1}\right) - 20\right)^2 + \frac{1}{2}\left(2x^2_{K-1} + 4u^2_{K-1}\right)}_{H_{K-1}\left(x_{K-1}, u_{K-1}\right)} \right)$$

By taking the first derivatives, we get the first-order optimality condition

$$\frac{\partial H_{K-1}\left(x_{K-1}, u_{K-1}\right)}{\partial u_{K-1}} = 0$$

We get, for a still unknown x_1, the optimal control action $u^*_{K-1}(x_{K-1})$

$$u^*_{K-1} = \frac{12x_1 - 60}{19}$$

Moreover, we have the optimal cost

$$\begin{aligned} J^*\left(x_{K-1}\right) &= \left(\left(\left(4x_{K-1} - 6u^*_{K-1}\right) - 20\right)^2 + \frac{1}{2}\left(2x^2_{K-1} + 4(u^*_{K-1})^2\right)\right) \\ &= \left(\left(4x_{K-1} - 6\frac{12x_{K-1} - 60}{19} - 20\right)^2 + \frac{1}{2}\left(2x^2_{K-1} + 4\left(\frac{12x_{K-1} - 60}{19}\right)^2\right)\right) \end{aligned}$$

Optimal control with dynamic programming (cont.)

At stage $K-2=0$, the initial stage, we have

$$
\begin{aligned}
J_{K-2}\left(x_{K-2}, u_{K-2}\right) &= E\left(x_K\right) + L\left(x_{K-1}, u_{K-1}\right) + L\left(x_{K-2}, u_{K-2}\right) \\
J^*_{K-2}\left(x_{K-2}\right) &= \min_{u_{K-2}} \left(E\left(x_K\right) + L\left(x_{K-1}, u_{K-1}\right) + L\left(x_{K-2}, u_{K-2}\right)\right) \\
&= \min_{u_{K-2}} \left(J^*_{K-1}\left(x_{K-1}\right) + L\left(x_{K-2}, u_{K-2}\right)\right) \\
&= \min_{u_{K-2}} \left(J^*\left(x_{K-1}\right) + \frac{1}{2}\left(2x^2_{K-2} + 4u^2_{K-2}\right)\right) \\
&= \min_{u_{K-2}} \left(4 \underbrace{x_{K-1}}_{4x_{K-2}-6u_{K-2}} -6 \frac{12 \overbrace{x_{K-1}}^{4x_{K-2}-6u_{K-2}} -60}{19} - 20\right)^2 \\
&\quad + \frac{1}{2}\left(2 \underbrace{x_{K-1}^2}_{4x_{K-2}-6u_{K-2}} +4 \left(\frac{12 \overbrace{x_{K-1}}^{4x_{K-2}-6u_{K-2}} -60}{19}\right)^2\right)
\end{aligned}
$$

By taking the first derivatives, we get the first-order optimality condition

$$
\frac{\partial H_{K-2}\left(x_{K-2}, u_{K-2}\right)}{\partial u_{K-2}} = 0
$$

Optimal control with dynamic programming (cont.)

$$\frac{\partial H_{K-2}(x_{K-2}, u_{K-2})}{\partial u_{K-2}} = \frac{\partial H_0(x_0, u_0)}{\partial u_0}$$
$$= 0$$

We get, for the known initial state $x_0 = \overline{x}_0$, the optimal control action $u_0^*(x_0)$,

$$u_0^* \approx 4.8$$

Given x_0 and the optimal control u_0^*, we can go forward in time and compute x_1^*

$$x_1^* = 4\overline{x}_0 - 6u_0^*$$
$$\approx 3.1$$

We can compute the next optimal control action from state $x_1 = x_1^*$,

$$u_1^* = \frac{12x_1 - 60}{16}$$
$$\approx -1.2$$

And, finally the terminal state

$$x_2^* \approx 19.6$$

□

Continuous-time dynamic programming

The HJB equation

The Hamilton-Jacobi-Bellman equation

We extend dynamic programming to the optimal control of continuous-time systems

Consider an arbitrary continuous-time system, with initial condition $x(0) = \overline{x}_0$

$$\dot{x}(t) = f\left(x(t), u(t)\right)$$

Consider the cost function, the cost of transferring the state $x(t)$ to $x(T)$ over $[t, T]$

$$J\left(x(t), u(t)\right) = E\left(x(T)\right) + \int_t^T L\left(x(\tau), u(\tau)\right) d\tau$$
$$t \in [0, T]$$

We are interested in the optimal controls $u^*(\tau)$, for all $\tau \in [t, T]$

- Principle of optimality

The Hamilton-Jacobi-Bellman equation (cont.)

$$J\left(x(t), u(t)\right) = E\left(x(T)\right) + \int_t^T L\left(x(\tau), u(\tau)\right)$$

At the boundary, we know the optimal cost of transferring the state from $x(T)$ to $x(T)$

$$\begin{aligned} J\left(x(T), u(T)\right) &= E\left(x(T)\right) \\ J^*\left(x(T)\right) &= \min_{u(T)} E\left(x(T)\right) \\ &= E\left(x(T)\right) \end{aligned}$$

The Hamilton-Jacobi-Bellman equation (cont.)

At any time t, we have the optimal cost of transferring the state from $x(t)$ to $x(T)$

$$\begin{aligned} J^*\left(x(t)\right) &= \min_{\substack{u(\tau)\\ t\leq\tau\leq T}} E\left(x(T)\right) + \int_t^T L\left(x(t), u(t)\right) d\tau \\ &= \min_{\substack{u(t)\\ t\leq\tau\leq T}} E\left(x(T)\right) + \underbrace{\int_{t+\Delta t}^T L\left(x(t), u(t)\right) d\tau + \int_t^{t+\Delta t} L\left(x(t), u(t)\right) d\tau}_{\int_t^T L(x(t),u(t))\,d\tau} \\ &= \min_{\substack{u(t)\\ t\leq\tau\leq T}} \underbrace{E\left(x(T)\right) + \int_{t+\Delta t}^T L\left(x(t), u(t)\right) d\tau}_{J^*(x(t+\Delta t))} + \int_t^{t+\Delta t} L\left(x(t), u(t)\right) d\tau \end{aligned}$$

The quantity $J^*\left(x(t+\Delta t)\right)$ is the optimal cost to transfer the state $x(t+\Delta t)$ to $x(T)$

- The minimum of $J\left(x(t+\Delta t), u(t+\Delta t)\right)$, with respect to $u(t+\Delta t)$

The Hamilton-Jacobi-Bellman equation (cont.)

$$J^*\left(x(t)\right) = \min_{\substack{u(t)\\ t\leq\tau\leq T}} \underbrace{E\left(x(T)\right) + \int_{t+\Delta t}^{T} L\left(x(t), u(t)\right) d\tau}_{J^*(x(t+\Delta t))} + \int_{t}^{t+\Delta t} L\left(x(t), u(t)\right) d\tau$$

From the principle of optimality, we have for $t \leq \tau \leq t + \Delta t$

$$J^*\left(x(t) + \Delta t\right) = \min_{\substack{u(t)\\ t\leq\tau\leq t+\Delta t}} J^*\left(x(t+\Delta t)\right) + \int_{t}^{t+\Delta t} L\left(x(t), u(t)\right) d\tau$$

Using the dynamics in integral form, we have

$$\begin{aligned} x(t+\Delta t) &= x(t) + \int_{t+\Delta t}^{T} f\left(x(\tau), u(\tau)\right) d\tau \\ &= x(t) + \Delta x(t) \end{aligned}$$

The Hamilton-Jacobi-Bellman equation (cont.)

We can expand the optimal cost $J^*\left(x(t+\Delta t)\right)$ using a Taylor's series expansion

$$J^*\left(x(t+\Delta t)\right) \approx J^*\left(x(t)\right) + \underbrace{\frac{\partial J^*\left(x(t+\Delta t)\right)}{\partial t}\Delta t}_{\text{First variation}} + \underbrace{\left[\frac{\partial J^*\left(x(t+\Delta t)\right)}{\partial x(t)}\right]^T \Delta x(t)}_{\text{Second variation}}$$

Assuming the variation in Δt to be small, we can neglet higher-order terms to get

$$J^*\left(x(t+\Delta t)\right) = \underbrace{\underbrace{J^*\left(x(t)\right)} + \underbrace{\frac{\partial J^*\left(x(t+\Delta t)\right)}{\partial t}}}_{\text{independent of } u(t)}\Delta t + \left[\frac{\partial J^*\left(x(t+\Delta t)\right)}{\partial x(t)}\right]^T \Delta x(t)$$

The Hamilton-Jacobi-Bellman equation (cont.)

Substituting and factoring out the terms that are independent of $u(t)$, we get

$$\begin{aligned} J^*\left(x(t)\right) &= \min_{\substack{u(t) \\ t \leq \tau \leq t+\Delta t}} J^*\left(x(t+\Delta t)\right) + \int_t^{t+\Delta t} L\left(x(t), u(t)\right) d\tau \\ &= J^*\left(x(t)\right) + \frac{\partial J^*\left(x(t+\Delta t)\right)}{\partial t}\Delta t \\ &\quad + \min_{\substack{u(t) \\ t \leq \tau \leq t+\Delta t}} \int_t^{t+\Delta t} L\left(x(t), u(t)\right) d\tau + \left[\frac{\partial J^*\left(x(t+\Delta t)\right)}{\partial x(t)}\right]^T \Delta x(t) \end{aligned}$$

The Hamilton-Jacobi-Bellman equation (cont.)

$$J^*\left(x(t)\right) = J^*\left(x(t)\right) + \frac{\partial J^*\left(x(t+\Delta t)\right)}{\partial t}\Delta t + \min_{\substack{u(t)\\ t\leq\tau\leq t+\Delta t}} \int_t^{t+\Delta t} L\left(x(t), u(t)\right) d\tau + \left[\frac{\partial J^*\left(x(t+\Delta t)\right)}{\partial x(t)}\right]^T \Delta x(t)$$

Simplifying $J^*\left(x(t)\right)$, dividing by Δt, and rearranging we get

$$\begin{aligned}\frac{\partial J^*\left(x(t+\Delta t)\right)}{\partial t} &= -\frac{1}{\Delta t}\min_{\substack{u(t)\\ t\leq\tau\leq t+\Delta t}} \int_t^{t+\Delta t} L\left(x(t), u(t)\right) d\tau \\ &\quad + \left[\frac{\partial J^*\left(x(t+\Delta t)\right)}{\partial x(t)}\right]^T \Delta x(t) \\ &= -\min_{\substack{u(t)\\ t\leq\tau\leq t+\Delta t}} \frac{1}{\Delta t}\int_t^{t+\Delta t} L\left(x(t), u(t)\right) d\tau \\ &\quad + \left[\frac{\partial J^*\left(x(t+\Delta t)\right)}{\partial x(t)}\right]^T \frac{\Delta x(t)}{\Delta t}\end{aligned}$$

The Hamilton-Jacobi-Bellman equation (cont.)

$$\frac{\partial J^*\left(x(t+\Delta t)\right)}{\partial t} = -\min_{\substack{u(t)\\ t\le\tau\le t+\Delta t}} \frac{1}{\Delta t}\int_t^{t+\Delta t} L\left(x(t),u(t)\right)d\tau + \left[\frac{\partial J^*\left(x(t+\Delta t)\right)}{\partial x(t)}\right]^T \frac{\Delta x(t)}{\Delta t}$$

In the limit for $\Delta t \to 0$, we get

$$\lim_{\Delta t\to 0}\left(\frac{1}{\Delta t}\int_t^{t+\Delta t} L\left(x(t),u(t)\right)d\tau + \left[\frac{\partial J^*\left(x(t+\Delta t)\right)}{\partial x(t)}\right]^T \frac{\Delta x(t)}{\Delta t}\right)$$

$$= \underbrace{\lim_{\Delta t\to 0}\left(\frac{1}{\Delta t}\int_t^{t+\Delta t} L\left(x(t),u(t)\right)d\tau\right)}_{L(x(t),u(t))} + \underbrace{\lim_{\Delta t\to 0}\left(\left[\frac{\partial J^*\left(x(t+\Delta t)\right)}{\partial x(t)}\right]^T \frac{\Delta x(t)}{\Delta t}\right)}_{\left[\frac{\partial J^*\left(x(t)\right)}{\partial x(t)}\right]^T \dot{x}(t)}$$

The Hamilton-Jacobi-Bellman equation (cont.)

After substituting the obtained quantities, we get

$$\frac{\partial J^*(x(t))}{\partial t} = -\min_{u(t)} L(x(t), u(t)) + \left[\frac{\partial J^*(x(t))}{\partial x(t)}\right]^T \underbrace{f(x(t), u(t))}_{\dot{x}(t)}$$

We define the Hamiltonian,

$$H\left(x(t), u(t), \frac{\partial J^*(x(t))}{\partial x}\right) = L(x(t), u(t)) + \left[\frac{\partial J^*(x(t))}{\partial x(t)}\right]^T \underbrace{f(x(t), u(t))}_{\dot{x}(t)}$$

As a result, we get the **Hamilton-Jacobi-Bellman equation**,

$$\begin{aligned}\frac{\partial J^*(x(t))}{\partial t} &= -\min_{u(t)} H\left(x(t), u(t), \frac{\partial J^*(x(t))}{\partial x}\right)\\ &= -H^*\left(x(t), \frac{\partial J^*(x(t))}{\partial x}\right)\end{aligned}$$

The Hamilton-Jacobi-Bellman equation (cont.)

$$\frac{\partial J^*\left(x(t)\right)}{\partial t} = -\min_{u(t)} H\left(x(t), u(t), \frac{\partial J^*\left(x(t)\right)}{\partial x}\right)$$

The Hamilton-Jacobi-Bellman requires the minimisation of the Hamiltonian

The optimal control is obtained by solving for the optimality conditions

$$\frac{\partial H\left(x(t), u(t), \frac{\partial J^*\left(x(t)\right)}{\partial x}\right)}{\partial u} = 0$$

That is, the optimal control

$$u^*(t) = u^*\left(x(t), \frac{\partial J^*\left(x(t)\right)}{\partial x}\right)$$

The Hamilton-Jacobi-Bellman equation (cont.)

$$\frac{\partial J^*\left(x(t)\right)}{\partial t} = -H^*\left(x(t), \frac{\partial J^*\left(x(t)\right)}{\partial x}\right)$$

The Hamilton-Jacobi-Bellman equation is a nonlinear partial differential equation

The boundary condition is given by the terminal cost

$$J^*\left(x^*(T)\right) = E\left(x^*(T)\right)$$

Given $J^*(x(t))$, the optimal control $u^*(t)$ is obtained by the gradient of J^*