

Hamilton-Jacobi-Bellman Equation

Feb 25, 2008

What is it?

The Hamilton-Jacobi-Bellman (HJB) equation is the continuous-time analog to the discrete deterministic dynamic programming algorithm

Discrete VS Continuous

$$x_{k+1} = f(x_k, u_k)$$

$$k \in 0, \dots, N$$

$$g_N(x_N) + \sum_{k=0}^{N-1} g_k(x_k, u_k)$$

$$J_N(x_N) = g_N(x_N)$$

$$J_k(x_k) = \min_{u_k \in U_k} \{ g_k(x_k, u_k) + J_{k+1}(x_k, u_k) \}$$

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

$$0 \leq t \leq T$$

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

$$V(T, x) = h(x)$$

$$0 = \min_{u \in U} \{ g(x, u) + \nabla_t V(t, x) + \nabla_x V(t, x)' f(x, u) \}$$

HJB Equation

- Extension of Hamilton-Jacobi equation (classical mechanics)
- Solution is the optimal cost-to-go function
- Applications
 - path planning
 - medical
 - financial

Derivation

- Start with continuous time interval $t \in [0, T]$
- Discretize into N pieces so that $\delta = \frac{T}{N}$
- Denote $x_k = x(k\delta), \quad k = 0, \dots, N$
 $u_k = u(k\delta), \quad k = 0, \dots, N$

Derivation

- Remember

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

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

- Approximate the continuous time by

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

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

Derivation

$J^*(t, x)$: Optimal cost-to-go function for
continuous time problem

$\tilde{J}^*(t, x)$: Optimal cost-to-go function for
discrete time approximation

Derivation

- From discrete time DP

$$J_N(x_N) = g_N(x_N)$$

$$J_k(x_k) = \min_{u_k \in U_k} \{ g_k(x_k, u_k) + J_{k+1}(x_k, u_k) \}$$

- For the discrete time approximation

$$\tilde{J}^*(N\delta, x) = h(x)$$

$$\tilde{J}^*(k\delta, x) = \min_{u_k \in U_k} \{ \delta g(x, u) + \tilde{J}^*(\delta(k+1), x + \delta f(x, u)) \}$$

Derivation

- Reminder: Taylor series expansion for $f(x, y)$

$$f(x + \Delta x, y + \Delta y) = \sum_{i=0}^{\infty} \left\{ \frac{1}{i!} [\Delta x \nabla_x + \Delta y \nabla_y]^i f(x, y) \right\}$$

- Assume that the Taylor series expansion exists

$$\tilde{J}^*(k\delta + \delta, x + \delta f(x, u)) = \sum_{i=0}^{\infty} \left\{ \frac{1}{i!} [\delta \nabla_t + \delta f(x, u) \nabla_x]^i \tilde{J}^*(k\delta, x) \right\}$$

- Ignoring all higher order terms $o(\delta)$

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

Derivation

- Combine

$$\tilde{J}^*(k\delta, x) = \min_{u_k \in U_k} \{ \delta g(x, u) + \tilde{J}^*(\delta(k+1), x + \delta f(x, u)) \}$$

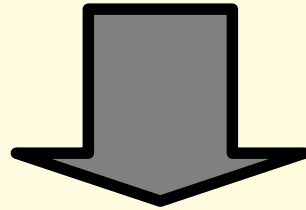
$$\begin{aligned} \tilde{J}^*(k\delta + \delta, x + \delta f(x, u)) &= \tilde{J}^*(k\delta, x) + \delta \nabla_t \tilde{J}^*(k\delta, x) \\ &\quad + \delta f(x, u) \nabla_x \tilde{J}^*(k\delta, x)' + o(\delta) \end{aligned}$$

- We get

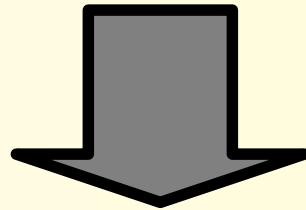
$$\begin{aligned} \tilde{J}^*(k\delta, x) &= \min_{u \in U} \{ \delta g(x, u) + \tilde{J}^*(k\delta, x) + \delta \nabla_t \tilde{J}^*(k\delta, x) + \\ &\quad \delta f(x, u) \nabla_x \tilde{J}^*(k\delta, x)' + o(\delta) \} \end{aligned}$$

Derivation

$$\cancel{\tilde{J}^*(k\delta, x)} = \min_{u \in U} \{ \cancel{\delta g(x, u) + \tilde{J}^*(k\delta, x)} + \delta \nabla_t \tilde{J}^*(k\delta, x) + \delta f(x, u) \nabla_x \tilde{J}^*(k\delta, x)' + o(\delta) \}$$



$$0 = \min_{u \in U} \{ \cancel{\delta} g(x, u) + \cancel{\delta} \nabla_t \tilde{J}^*(k\delta, x) + \cancel{\delta} f(x, u) \nabla_x \tilde{J}^*(k\delta, x)' + o(\delta) \}$$



$$0 = \min_{u \in U} \{ g(x, u) + \nabla_t \tilde{J}^*(k\delta, x) + f(x, u) \nabla_x \tilde{J}^*(k\delta, x)' + o(\delta) \}$$

Derivation

$$0 = \min_{u \in U} \{ g(x, u) + \nabla_t \tilde{J}^*(k\delta, x) + f(x, u) \nabla_x \tilde{J}^*(k\delta, x)' + o(\delta) \}$$

- Take the limit $\delta \rightarrow 0 \quad k \rightarrow \infty \quad k\delta = t$
- Assume that $\lim_{\delta \rightarrow 0, k \rightarrow \infty, k\delta = t} \tilde{J}^*(k\delta, x) = J^*(t, x)$

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

$$J^*(T, x) = h(x)$$

Claim

If $V(t,x)$ is a solution to the HJB equation, $V(t,x)$ equals the optimal cost-to-go function for all t and x

Proof

- From $0 = \min_{u \in U} \{ g(x, u) + \nabla_t V(t, x) + f(x, u) \nabla_x V(t, x)' \}$

- With any control and state trajectory

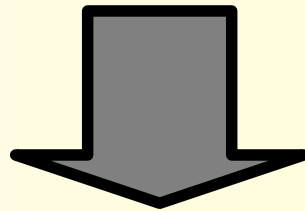
$$\begin{aligned} & \{ \hat{u}(t) | t \in [0, T] \} \\ & \{ \hat{x}(t) | t \in [0, T] \} \end{aligned}$$

$$0 \leq g(\hat{x}(t), \hat{u}(t)) + \nabla_t V(t, \hat{x}(t)) + f(\hat{x}(t), \hat{u}(t)) \nabla_x V(t, \hat{x}(t))'$$

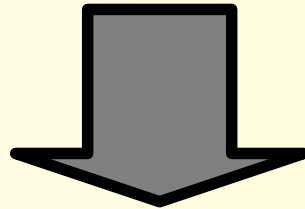
Proof

- Substitute in $\dot{\hat{x}}(t) = f(\hat{x}(t), \hat{u}(t))$

$$0 \leq g(\hat{x}(t), \hat{u}(t)) + \nabla_t V(t, \hat{x}(t)) + \dot{\hat{x}}(t) \nabla_x V(t, \hat{x}(t))'$$



$$0 \leq g(\hat{x}(t), \hat{u}(t)) + \frac{dV(t, \hat{x}(t))}{dt}$$



$$0 \leq \int_0^T g(\hat{x}(t), \hat{u}(t)) dt + \int_0^T \frac{dV(t, \hat{x}(t))}{dt} dt$$

Proof

- Evaluate $0 \leq \int_0^T g(\hat{x}(t), \hat{u}(t)) dt + \int_0^T \frac{dV(t, \hat{x}(t))}{dt} dt$

$$0 \leq \int_0^T g(\hat{x}(t), \hat{u}(t)) dt + V(T, \hat{x}(T)) - V(0, \hat{x}(0))$$

- For any state and control trajectory

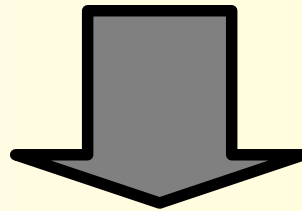
$$V(0, x(0)) \leq h(\hat{x}(T)) + \int_0^T g(\hat{x}(t), \hat{u}(t)) dt$$

Proof

- For optimal state and control trajectory

$$\begin{aligned} & \{u^*(t) | t \in [0, T]\} \\ & \{x^*(t) | t \in [0, T]\} \end{aligned}$$

$$V(0, x(0)) = h(x^*(T)) + \int_0^T g(x^*(t), u^*(t)) dt = J^*(0, x(0))$$



$$V(0, x(0)) = J^*(0, x(0))$$

HJB Example

- Consider the simple scalar system

$$\dot{x}(t) = u(t) \quad |u(t)| \leq 1 \quad \forall t \in [0, T]$$

- The terminal cost $V(T, x) = \frac{1}{2} x^2$

- HJB Equation $0 = \min_{|u| \leq 1} \{ \nabla_t V(t, x) + u \nabla_x V(t, x) \}$

HJB Example

- Candidate control policy

$$u^*(t, x) = -\text{sgn}(x)$$

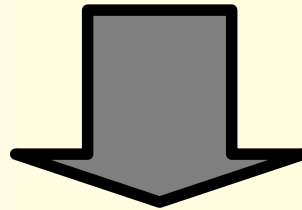
- Optimal cost-to-go function

$$J^*(t, x) = \frac{1}{2} (\max \{ 0, |x| - (T - t) \})^2$$

- Check to see if it solves the HJB equation

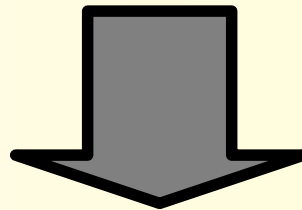
HJB Example

$$J^*(t, x) = \frac{1}{2} (\max \{ 0, |x| - (T - t) \})^2$$



$$\nabla_t J^*(t, x) = \max \{ 0, |x| - (T - t) \}$$

$$\nabla_x J^*(t, x) = \text{sgn}(x) \max \{ 0, |x| - (T - t) \}$$



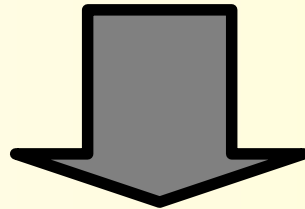
$$0 = \min_{|u| \leq 1} \{ \nabla_t V(t, x) + u \nabla_x V(t, x) \} = \min_{|u| \leq 1} \{ 1 + u \text{sgn}(x) \} \max \{ 0, |x| - (T - t) \}$$

LQR and HJB

- Remember the continuous time LQR

$$\dot{x}(t) = Ax(t) + Bu(t)$$

$$J = x(T)' Q_T x(T) + \int_0^T (x(t)' Q x(t) + u(t)' R u(t)) dt$$



$$g(x, u) = x' Q x + u' R u$$

$$h(x) = x' Q_T x$$

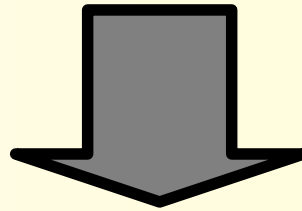
LQR and HJB

- The HJB Equation

$$0 = \min_u \{ x' Q x + u' R u + \nabla_t V(t, x) + (Ax + Bu) \nabla_x V(t, x)' \}$$

$$V(T, x) = x' Q_T x$$

- Try a solution of the form $V(t, x) = x' K(t) x$



$$\nabla_x V(t, x) = 2K(t)x$$

$$\nabla_t V(t, x) = x' \dot{K}(t)x$$

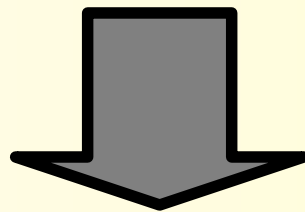
LQR and HJB

- Substitute in the HJB equation

$$0 = \min_u \{ x' Q x + u' R u + x' \dot{K}(t) x + 2x' K(t) A x + 2x' K(t) B u \}$$

- Differentiate to find minimum

$$0 = \frac{\partial}{\partial u} \{ x' Q x + u' R u + x' \dot{K}(t) x + 2x' K(t) A x + 2x' K(t) B u \}$$

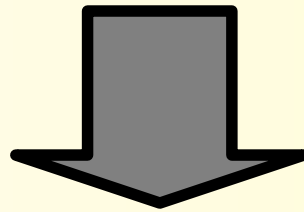


$$2B' K(t) x + 2R u = 0 \quad \Rightarrow \quad u = -R^{-1} B' K(t) x$$

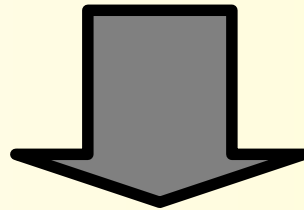
LQR and HJB

- Therefore, at minimum u

$$0 = x' (\dot{K}(t) + K(t)A + A'K(t) - K(t)BR^{-1}B'K(t) + Q)x$$



$$0 = \dot{K}(t) + K(t)A + A'K(t) - K(t)BR^{-1}B'K(t) + Q$$



$$\dot{K}(t) = -K(t)A - A'K(t) + K(t)BR^{-1}B'K(t) - Q$$

Continuous-time Riccati Equation

LQR and HJB

- Given that $K(t)$ satisfies the Riccati equation

$$J^*(t, x) = V(t, x) = x' K(t) x$$

- And the optimal control policy is

$$u^*(t, x) = -R^{-1} B' K(t) x$$