Today: Continuous Dynammic Programming

Discrete state $\rightarrow$ continuous state

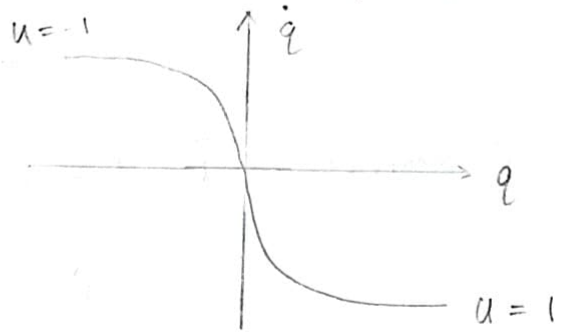
Discrete actions $\rightarrow$ continuous actions

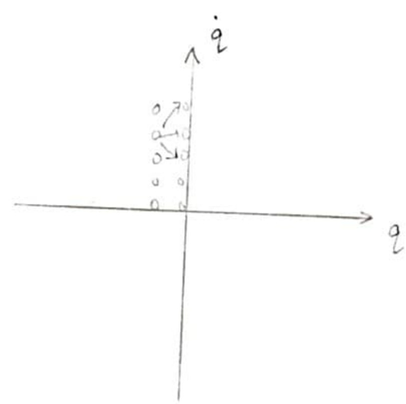
Discrete time $\rightarrow$ continuous time

HJB sufficiency theorem

Last time:

$$\ddot{q} = u \qquad |u| \leq 1$$





$$g(s,a) = \begin{cases} 1 & \text{if } s \neq s_{goal} \\ 0 & \text{if } s = s_{goal} \end{cases}$$

$$s[n+1] = f(s[n], a[n])$$

$$\forall s_i \quad J^*(s_i) = \min_a \left[g(s_i, a) + J^*(f(s_i, a)) \right]$$

$$\min_{a[\cdot]} \sum_{n=0}^{\infty} g(s[n], a[n])$$

verify $\pi^*(s_i) = \arg\min_a \left[\cancel{g(s_i,a)}\ g(s_i, a) + J^*(f(s_i, a)) \right]$

Discrete

Continuous

$$S[n+1] = f_d(s[n], a[n])$$

$$\dot{x} = f_c(x, u)$$

$$\min_{a[\cdot]} \sum_{n=0}^{\infty} g_d(s[n], a[n])$$

$$\min_{a(\cdot)} \int_0^{\infty} dt \; g_c(x, u)$$

$$\forall s \quad J^*(s) = \min_a \left[g_d(s,a) + J^*(f_d(s,a)) \right]$$

$$\forall x \quad 0 = \min_u \left[g_c(x,u) + \frac{\partial J^*}{\partial x} f(x,u) \right]$$

Hamilton - Jacob - Bellman Eq

(a. P. D. E.)

$$\boxed{\forall x \quad 0 = \min_u \left[g_c(x,u) + \frac{\partial J^*}{\partial x} f(x,u) \right]}$$

$$X[n+1] = f(x[n], u[n]).$$

$$\min_{u[\cdot]} \sum_{n=0}^{\infty} g(x[n], u[n]).$$

$$\forall x \quad J^*(x) = \min_u \left[g(x,u) + J^*(f(x,u)) \right]$$

$$x[n] \longrightarrow x(t)$$

$$x[n+1] \rightarrow x(t+dt). \quad \text{then} \quad \lim_{dt \to 0}$$

$$J^*(x) = \min_{u(\cdot)} \left[\int_0^{dt} d\tau \; g(x,u) + J^*(x(t+dt)) \right]$$

$$= \lim_{dt \to 0} \min_u \left[g(x,u)\, dt + \overset{J^*(x)}{J^*} + \frac{dJ^*}{dt} dt \right]$$

$$\Uparrow$$

$$\frac{\partial J^*}{\partial x} \dot{x} \, dt$$

$$\frac{\partial J^*}{\partial x} f(x,u) \, dt$$

$$0 = \lim_{dt \to 0} \min_u \left[g(x,u)\, dt + \frac{\partial J^*}{\partial x} f(x,u)\, dt \right]$$

$$0 = \min_u \left[g(x,u) + \frac{\partial J^*}{\partial x} f(x,u) \right]$$

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

For the optimal policy u^*,

$$-g(x,u^*) = \underbrace{\frac{\partial J^*}{\partial x} f(x,u^*)}_{\frac{dJ^*}{dt}}$$

2	→1	2
1	G	1
2	1	2
3	2	3

For an optimal trajectory,

J^* decreases with rate $g(x,u^*)$

$J^*(x)$ is a potential field

$\frac{\partial J^*}{\partial x}$ is derivative

Need some boundary cond.

$$J(x=0) = 0$$

Ex

$$x = \begin{bmatrix} q \\ \dot q \end{bmatrix}$$

$$\ddot q = u. \qquad g(x.u) = q^2 + \dot q^2 + u^2$$

$$u^* = -q - \sqrt{3}\,\dot q$$

$$J^*(x) = \sqrt{3}\,q^2 + 2q\dot q + \sqrt{3}\,\dot q^2$$

$$\frac{dJ^*}{dt} = \frac{\partial J^*}{\partial x} f(x,u) = \frac{\partial J}{\partial q}\dot q + \frac{\partial J}{\partial \dot q}\cancel{\ddot q}\ u$$

$$\frac{\partial J}{\partial q} = 2\sqrt{3}\,q + 2\dot q$$

$$\frac{\partial J}{\partial \dot q} = 2q + 2\sqrt{3}\,\dot q$$

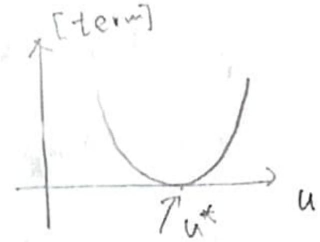
$$2\sqrt{3}\,q\dot q + 2\dot q^2 + (2q + 2\sqrt{3}\,\dot q)\,u$$

$$\frac{\partial J}{\partial x} = \begin{bmatrix} \frac{\partial J}{\partial q} & \frac{\partial J}{\partial \dot q} \end{bmatrix} \begin{bmatrix} \dot q \\ \ddot q \end{bmatrix} \qquad \frac{d}{dt}\begin{bmatrix} q \\ \dot q \end{bmatrix} = \begin{bmatrix} \dot q \\ u \end{bmatrix}$$

$$0 = \min_u \left[q^2 + \dot q^2 + u^2 + \frac{\partial J}{\partial q}\dot q + \frac{\partial J}{\partial \dot q} u \right]$$

① u^* makes the equation minimum

② when minimum, the minimum value is 0



should be 0

$$\frac{\partial[\,]}{\partial u} = 2u + \frac{\partial J}{\partial \dot{q}}$$

$$= 2u + 2q + 2\sqrt{3}\,\dot{q} = 0$$

$$u^* = -q - \sqrt{3}\,\dot{q}$$

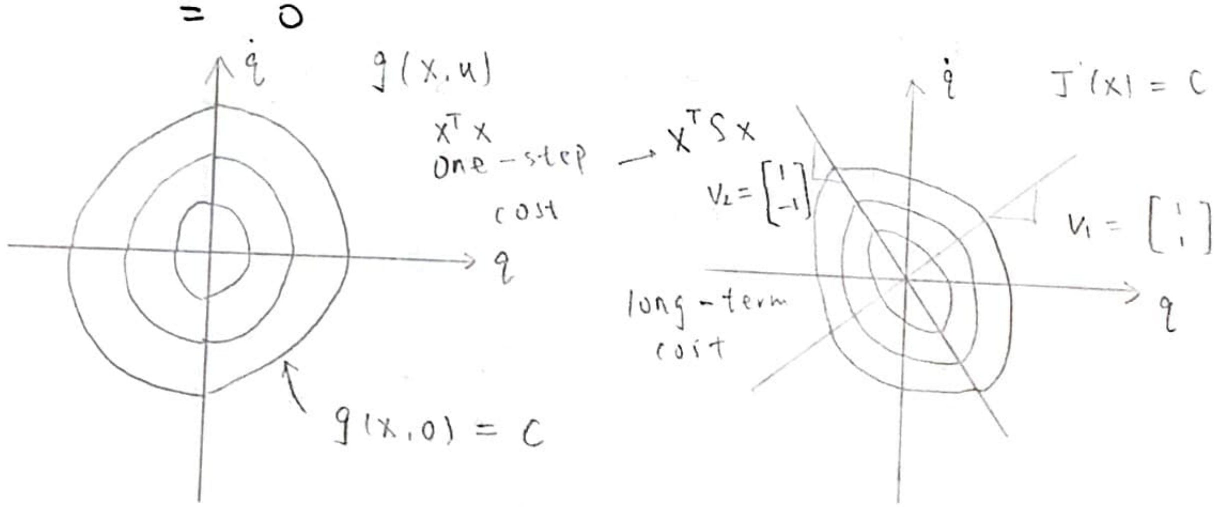
take u^* back into the equation

$$0 = q^2 + \dot{q}^2 + (-q-\sqrt{3}\,\dot{q})^2 + (2\sqrt{3}\,q + 2\dot{q})\dot{q} + (2q + 2\sqrt{3}\,\dot{q})(-q-\sqrt{3}\,\dot{q})$$

$$= 3\dot{q}^2 + 2\sqrt{3}\,q\dot{q} + q^2 - (q+\sqrt{3}\,\dot{q})^2$$

$$= 3\dot{q}^2 + 2\sqrt{3}\,q\dot{q} + q^2 - q^2 - 2\sqrt{3}\,q\dot{q} - 3\dot{q}^2$$

$$= 0$$



$$J^*(x) = x^T \begin{bmatrix} \sqrt{3} & 1 \\ 1 & \sqrt{3} \end{bmatrix} x \qquad V_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad V_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\lambda_1 = \sqrt{3}+1 \qquad \lambda_2 = \sqrt{3}-1$$

may add saturation in minimization range, or have a penalty term for large u

The theorem does not hold for not differentiable cost function J^*

Linear Quadratic Regulator

$\dot{x} = Ax + Bu$ Double - integrator

$\dot{x} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$

$g(x,u) = x^T Q x + u^T R u$

$Q = Q^T \geq 0$ $R = R^T > 0$

positive-definite

$Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ $R = 1$ previous example

$J^*(x) = x^T S x$ $S = S^T > 0$

$$\frac{\partial J}{\partial x} = 2x^T S$$

$$0 = \min_u \left[g(x,u) + \frac{\partial J}{\partial x} f(x,u) \right] =$$

$$\min_u \left[x^T Q x + u^T R u + 2x^T S (Ax + Bu) \right]$$

$$\frac{\partial [\,]}{\partial u} = 2u^T R + 2x^T S B = 0$$

$$\Rightarrow \quad -B^T S^T x = R^T u \qquad R^T = R$$

$$S^T = S$$

$$u^* = -R^{-1} B^T S x = -Kx$$

put u^* back, can find S

$$0 = Q - SBR^{-1}B^T S + SA + A^T S$$

CARE

continuous algebraic Ricatti Eq.

in programming Python / Matlab

$$[K, S] = lqr(A, B, Q, R)$$

Note:

A quadratic form is always zero, if and only if the matrix is skew-symmetric

quadratic form: $x^T A x$

A is skew-symmetric $\quad A + A^T = 0$

$$x^T A x = x^T A^T x = x^T(-A)x = -x^T A x$$

so $x^T A x = 0$ for all x

On the other hand, if $A^T + A \neq 0$, then there exists a vector x such that $x^T A x \neq 0$

~~because diagonal terms of A^T and A are the same~~

if the m^{th} diagonal entry of A is nonzero, then choosing x as the m^{th} column of the identity matrix suffices as an example of such a vector x that does not satisfy $x^T A x = 0$; indeed, $x^T A x$ would be equal to this nonzero diagonal entry. Otherwise, suppose for A, all $a_{kk} = 0$ for all k and $a_{mn} = -a_{nm}$ for some m and n. Choosing x as the sum of the m^{th} and n^{th} columns of the identity matrix suffices, $x^T A x = a_{mm} + a_{nm} + a_{mn} + a_{nn}$ is not equal to 0.

then $x^T A x = 0$ for all x $\Longleftrightarrow$ equivalent to $A + A^T = 0$.

$$u^* = - R^{-1} B^T \underbrace{S x}_{\frac{\partial J^*}{\partial x}}$$

$-\frac{\partial J^*}{\partial x}$ "go ~~d~~ downhill"

$-B^T \frac{\partial J^*}{\partial x}$ modulated by my control authority (Jacobian transpose?)

$-R^{-1} B^T \frac{\partial J^*}{\partial x}$ taking into account preference on actuators

Numerical approach to

Continuous actions

$$\min_u \left[g(x,u) + \frac{\partial J^*}{\partial x} f(x,u) \right]$$

$g(x,u) = g_1(x) + u^T R u \Leftarrow$ Quadratic in u

$f(x,u) = f_1(x) + f_2(x)\, u \Leftarrow$ Control-affine

$$\min_u \left[g_1(x) + u^T R u + \frac{\partial J}{\partial x} \left[f_1(x) + f_2(x)\, u \right] \right]$$

$$\frac{\partial [\,]}{\partial u} = 0 \qquad u^* = - R^{-1} f_2^T(x) \left(\frac{\partial J^*}{\partial x} \right)^T$$

$J(x)$ function approximation

mesh $\longleftrightarrow$ deep neural network

linear function approximators