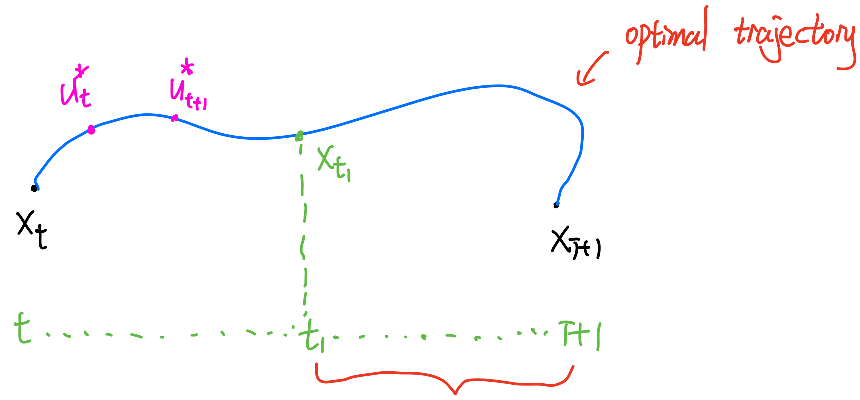
Dynamic Programming – Discrete time OC problem

$$\min_{u_{t:T}} J(x_t, u_{t:T}, t) = \sum_{s=t}^{T} L(x_s, u_s) + \ell(x_{T+1})$$

$$\text{s.t.} \quad x_{s+1} = f_D(x_s, u_s) \quad \forall s \in \{t, \dots, T\}$$

Principle of Optimality

In an optimal sequence of decisions, each subsequence must also be optimal. Thus, if we take any state along the optimal state trajectory, then the remaining subtrajectory is also optimal.



This trajectory must minimize $J(x_{t_1}, u_{t_1:T}, t_1)$

$$\min_{u_{t:T}} J(x_t, u_{t:T}, t) = \sum_{s=t}^{T} L(x_s, u_s) + \ell(x_{T+1})$$

$$= \sum_{s=t}^{t_1} L(x_s, u_s) + \underbrace{\sum_{s=t_1}^{T} L(x_s, u_s) + \ell(x_{T+1})}_{J(x_{t_1}, u_{t_1:T}, t)}$$

for all $t_1 > t$

$$V(x_t,t) = \min_{u_{t:T}} J(x_t, u_{t:T}, t)$$ (optimal cost) value function

$$= \min_{u_{t:T}} \{ L(x_t,u_t) + J(x_{t+1}, u_{t+1:T}, t) \}$$

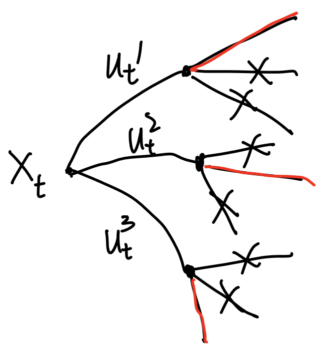
$$= \min_{u_{t:T}} \{ L(x_t,u_t) + V(x_{t+1}, t+1) \}$$

(Bellman Equation)

(Bellman backup)

$$= \min_{u_t} \{ L(x_t,u_t) + V(x_{t+1}, t+1) \}$$

(Value Iteration)



reduces the search from sequence to "pointwise" control.

$$V(x_t,t) = \min_{u_t} \{ L(x_t,u_t) + V(x_{t+1}, t+1) \}$$

$V(x_T, T)$

$V(x_{T-1}, T-1)$

$$V(x_{T+1}, T+1) = \ell(x_{T+1})$$

↑ terminal condition

Deriving the optimal control

$$u_t^*(x_t) = \arg\min_{u_t} \{ L(x_t,u_t) + V(x_{t+1}, t+1) \}$$

↑ feedback control

State-action value: $Q(x_t, u_t) = L(x_t,u_t) + V(x_{t+1}, t+1)$

$$u_t^*(x_t) = \arg\min_{u_t} Q(x_t, u_t)$$

Linear Quadratic Regulator Problem

$$J(x_1, u_{1:H}) = \sum_{t=1}^{H} (q x_t^2 + r u_t^2) + q_f x_{H+1}^2 \qquad q, r, q_f > 0$$

$$\text{s.t.} \quad x_{t+1} = a x_t + b u_t$$

@ $t = H+1$

$$V(x, H+1) = q_f x^2 \quad \leftarrow \text{terminal condition}$$

$$V(x, H) = \min_{u_H} \{ (q x^2 + r u_H^2) + V(x_+, H+1) \}$$

$$= \min_{u_H} \{ (q x^2 + r u_H^2) + q_f x_+^2 \} \qquad x_+ = a x + b u_H \text{ (dynamics)}$$

$$= \min_{u_H} \{ (q x^2 + r u_H^2) + q_f \underbrace{(a x + b u_H)^2}_{= a^2 x^2 + b^2 u_H^2 + 2ab\, x u_H} \}$$

$$= \min_{u_H} \{ \underbrace{(q + q_f a^2) x^2 + (r + q_f b^2) u_H^2 + 2ab q_f x u_H}_{= g(u_H)} \}$$

$\frac{\partial g}{\partial u_H} = 0$ at optimal u_H

$$\Rightarrow \quad 2(r + q_f b^2) u_H + 2ab\, q_f x = 0$$

$$\Rightarrow \quad u_H = \frac{-2ab\, q_f}{r + q_f b^2} x$$

$u_H = -K_H x$ ← linear feedback control

← simple yet optimal

$$V(x, H) = q x^2 + r u_H^2 + q_f (a x + b u_H^*)^2$$
$$= P_H x^2$$

(+) globally optimal control laws

(+) state feedback (w/o any re-optimization)

(+) really easy to implement

(−) limited scope

(−) simple dynamics

(−) no constraints

(−) quadratic cost function