Optimal Control for Linear Quadratic Regulators

Recap: Dynamic Programming

$$\pi^\star = \{\pi_0^\star, \pi_1^\star, \ldots, \pi_{H-1}^\star\}$$

We use Dynamic Programming, and do DP backward in time; start at $H - 1$

$$Q_{H-1}^\star(s, a) = r(s, a) \qquad \pi_{H-1}^\star(s) = \arg\max_a Q_{H-1}^\star(s, a)$$

$$V_{H-1}^\star(s) = \max_a Q_{H-1}^\star(s, a) = Q_{H-1}^\star(s, \pi_{H-1}^\star(s))$$

Recap: Dynamic Programming

$$\pi^\star = \{\pi_0^\star, \pi_1^\star, \ldots, \pi_{H-1}^\star\}$$

We use Dynamic Programming, and do DP backward in time; start at $H-1$

$$Q_{H-1}^\star(s,a) = r(s,a) \qquad \pi_{H-1}^\star(s) = \arg\max_a Q_{H-1}^\star(s,a)$$

$$V_{H-1}^\star(s) = \max_a Q_{H-1}^\star(s,a) = Q_{H-1}^\star(s, \pi_{H-1}^\star(s))$$

Now assume that we have already computed $V_{h+1}^\star$, $h \leq H-2$

(i.e., we know how to perform optimally at $h+1$)

$$Q_h^\star(s,a) = r(s,a) + \mathbb{E}_{s' \sim P(\cdot|s,a)} V_{h+1}^\star(s')$$

$$\pi_h^\star(s) = \arg\max_a Q_h^\star(s,a)$$

Recap: The Linear Quadratic Regulator (LQR)

$$\min_{\pi_0,\ldots,\pi_{H-1}} \quad E\left[x_H^\top Q x_H + \sum_{h=0}^{H-1} (x_h^\top Q x_h + u_h^\top R u_h)\right]$$

such that $\quad x_{h+1} = Ax_h + Bu_h + w_h, \;\; u_h = \pi_h(x_h) \quad x_0 \sim \mu_0, \; w_h \sim N(0, \sigma^2 I)$,

Here, $x_h \in \mathbb{R}^d, u_h \in \mathbb{R}^k$,

the disturbance $w_t \in \mathbb{R}^d$ is Gaussian noise
$A \in \mathbb{R}^{d\times d}$ and $B \in \mathbb{R}^{d\times k}$;
$Q \in \mathbb{R}^{d\times d}$ and $R \in \mathbb{R}^{k\times k}$ are pd matrices

V/Q functions:

- Value function $V_h^\pi : \mathbb{R}^d \to \mathbb{R}$ as

$$V_h^\pi(x) = \mathbb{E}\left[x_H^\top Q x_H + \sum_{t=h}^{H-1} (x_t^\top Q x_t + u_t^\top R u_t) \,\middle|\, \pi, x_h = x\right],$$

- And $Q_h^\pi : \mathbb{R}^d \times \mathbb{R}^k \to \mathbb{R}$ as:

$$Q_h^\pi(x, u) = \mathbb{E}\left[x_H^\top Q x_H + \sum_{t=h}^{H-1} (x_t^\top Q x_t + u_t^\top R u_t) \,\middle|\, \pi, x_h = x, u_h = u\right],$$

Optimal Value functions:

$$V_h^\star(x) = \min_{\pi_h, \pi_{h+1}, \ldots, \pi_{H-1}} \mathbb{E}\left[x_H^T Q x_H + \sum_{t=h}^{H-1} x_t^\top Q x_t + u_t^\top R u_t \,\middle|\, u_t = \pi_t(x_t), x_h = x \right]$$

Optimal Value functions:

$$V_h^\star(x) = \min_{\pi_h, \pi_{h+1}, \ldots, \pi_{H-1}} \mathbb{E}\left[x_H^T Q x_H + \sum_{t=h}^{H-1} x_t^\top Q x_t + u_t^\top R u_t \middle| u_t = \pi_t(x_t), x_h = x \right]$$

Theorem:

$V_h^\star$ is a quadratic function, i.e., $V_h^\star(x) = x^\top P_h x + p_h$,

and optimal policy is linear:

$\pi_h^\star(x) = -K_h^\star x$,

and P_h & $K_h^\star$ can be computed exactly

Optimal Value functions:

$$V_h^\star(x) = \min_{\pi_h, \pi_{h+1}, \ldots, \pi_{H-1}} \mathbb{E}\left[x_H^T Q x_H + \sum_{t=h}^{H-1} x_t^\top Q x_t + u_t^\top R u_t \,\middle|\, u_t = \pi_t(x_t), x_h = x \right]$$

Theorem:

$V_h^\star$ is a quadratic function, i.e., $V_h^\star(x) = x^\top P_h x + p_h$,

and optimal policy is linear:

$\pi_h^\star(x) = - K_h^\star x$,

and P_h & $K_h^\star$ can be computed exactly

Today: we prove the above theorem and derive optimal policies

Key Steps to Deriving Optimal Control

Again, we will do dynamic programming **backward in time**, i.e., from H to 0

Key Steps to Deriving Optimal Control

Again, we will do dynamic programming **backward in time**, i.e., from H to 0

1. **Base case:**
Show that $V_H^\star(x)$ is quadratic for all $x \in \mathbb{R}^d$

Key Steps to Deriving Optimal Control

Again, we will do dynamic programming **backward in time**, i.e., from H to 0

1. **Base case:**
Show that $V^\star_H(x)$ is quadratic for all $x \in \mathbb{R}^d$

2. **Inductive hypothesis:** Assume $V^\star_{h+1}(x)$ is quadratic $\forall x$:

- show that $Q^\star_h(x, u)$ is quadratic in both (x, u)
- Derive the optimal policy $\pi^\star_h(x) = \arg\min_u Q^\star_h(x, u)$, and show that it's linear

Key Steps to Deriving Optimal Control

Again, we will do dynamic programming **backward in time**, i.e., from H to 0

1. **Base case:**
Show that $V_H^\star(x)$ is quadratic for all $x \in \mathbb{R}^d$

2. **Inductive hypothesis:** Assume $V_{h+1}^\star(x)$ is quadratic $\forall x$:

- show that $Q_h^\star(x,u)$ is quadratic in both (x,u)
- Derive the optimal policy $\pi_h^\star(x) = \arg\min_u Q_h^\star(x,u)$, and show that it's linear

3. **Conclusion:**
show $V_h^\star(x)$ is quadratic for all x;

Base case at H

Recall our cost functions:

$$\min_{\pi_0,\ldots,\pi_{H-1}} \quad \mathbb{E}\left[x_H^\top Q x_H + \sum_{h=0}^{H-1} (x_h^\top Q x_h + u_h^\top R u_h)\right]$$

Base case at H

Recall our cost functions:

$$\min_{\pi_0,\ldots,\pi_{H-1}} \quad \mathbb{E}\left[x_H^\top Q x_H + \sum_{h=0}^{H-1} (x_h^\top Q x_h + u_h^\top R u_h) \right]$$

So, at time step H, given x, the cost-to-go is $x^\top Q x$ regardless..

$$V_H^\star(x) = x^\top Q x, \ \forall x \in \mathbb{R}^d$$

Base case at H

Recall our cost functions:

$$\min_{\pi_0,\ldots,\pi_{H-1}} \quad \mathbb{E}\left[x_H^\top Q x_H + \sum_{h=0}^{H-1} (x_h^\top Q x_h + u_h^\top R u_h)\right]$$

So, at time step H, given x, the cost-to-go is $x^\top Q x$ regardless..

$$V_H^\star(x) = x^\top Q x, \ \forall x \in \mathbb{R}^d$$

Denote $P_H := Q, p_H = 0,$

we write $V_H^\star(x) = x^\top P_H x + p_H$

(Goal: derive recursive formulation of P_h, & p_h)

Induction Step:

$$\min_{\pi_0,\ldots,\pi_{H-1}} \quad E\left[x_H^\top Q x_H + \sum_{h=0}^{H-1} (x_h^\top Q x_h + u_h^\top R u_h) \right]$$

such that $\quad x_{h+1} = Ax_h + Bu_h + w_h, \; u_h = \pi_h(x_h) \quad x_0 \sim \mu_0, \, w_h \sim N(0, \sigma^2 I)$,

Assume $V^\star_{h+1}(x) = x^\top P_{h+1} x + p_{h+1}$, for all x, where $P_{h+1} \in \mathbb{R}^{d\times d}, p_{h+1} \in \mathbb{R}$

Induction Step:

$$\min_{\pi_0,\ldots,\pi_{H-1}} \quad E\left[x_H^\top Q x_H + \sum_{h=0}^{H-1}(x_h^\top Q x_h + u_h^\top R u_h)\right]$$

such that $\quad x_{h+1} = Ax_h + Bu_h + w_h,\ u_h = \pi_h(x_h) \quad x_0 \sim \mu_0,\ w_h \sim N(0,\sigma^2 I)$,

Assume $V^\star_{h+1}(x) = x^\top P_{h+1} x + p_{h+1}$, for all x, where $P_{h+1} \in \mathbb{R}^{d\times d}, p_{h+1} \in \mathbb{R}$

$$Q^\star_h(x,u) = c(x,u) + \mathbb{E}_{x'\sim P(x,u)} V^\star_{h+1}(x')$$

Induction Step:

$$\min_{\pi_0,\ldots,\pi_{H-1}} \quad E\left[x_H^\top Q x_H + \sum_{h=0}^{H-1} (x_h^\top Q x_h + u_h^\top R u_h)\right]$$

such that $\quad x_{h+1} = Ax_h + Bu_h + w_h,\ u_h = \pi_h(x_h) \quad x_0 \sim \mu_0,\ w_h \sim N(0,\sigma^2 I)$,

Assume $V^\star_{h+1}(x) = x^\top P_{h+1} x + p_{h+1}$, for all x, where $P_{h+1} \in \mathbb{R}^{d\times d}, p_{h+1} \in \mathbb{R}$

$$Q^\star_h(x,u) = c(x,u) + \mathbb{E}_{x'\sim P(x,u)} V^\star_{h+1}(x')$$

$$= x^\top Q x + u^\top R u + \mathbb{E}_{x'\sim P(x,u)} V^\star_{h+1}(x')$$

Induction Step:

$$\min_{\pi_0,\ldots,\pi_{H-1}} \quad E\left[x_H^\top Q x_H + \sum_{h=0}^{H-1} (x_h^\top Q x_h + u_h^\top R u_h) \right]$$

such that $\quad x_{h+1} = Ax_h + Bu_h + w_h,\ u_h = \pi_h(x_h) \quad x_0 \sim \mu_0,\ w_h \sim N(0,\sigma^2 I)$,

Assume $V^\star_{h+1}(x) = x^\top P_{h+1} x + p_{h+1}$, for all x, where $P_{h+1} \in \mathbb{R}^{d\times d}, p_{h+1} \in \mathbb{R}$

$$Q^\star_h(x,u) = c(x,u) + \mathbb{E}_{x'\sim P(x,u)} V^\star_{h+1}(x')$$

$$= x^\top Q x + u^\top R u + \mathbb{E}_{x'\sim P(x,u)} V^\star_{h+1}(x')$$

$$= x^\top Q x + u^\top R u + \mathbb{E}_{w\sim\mathcal{N}(0,\sigma^2 I)} \left[V^\star_{h+1}\left(Ax + Bu + w\right) \right]$$

Induction Step:

$$\min_{\pi_0,\ldots,\pi_{H-1}} \quad E\left[x_H^\top Q x_H + \sum_{h=0}^{H-1} (x_h^\top Q x_h + u_h^\top R u_h)\right]$$

such that $\quad x_{h+1} = Ax_h + Bu_h + w_h,\ u_h = \pi_h(x_h) \quad x_0 \sim \mu_0,\ w_h \sim N(0,\sigma^2 I)$,

Assume $V^\star_{h+1}(x) = x^\top P_{h+1} x + p_{h+1}$, for all x, where $P_{h+1} \in \mathbb{R}^{d\times d}, p_{h+1} \in \mathbb{R}$

$$Q^\star_h(x,u) = c(x,u) + \mathbb{E}_{x'\sim P(x,u)} V^\star_{h+1}(x')$$

$$= x^\top Q x + u^\top R u + \mathbb{E}_{x'\sim P(x,u)} V^\star_{h+1}(x')$$

$$= x^\top Q x + u^\top R u + \mathbb{E}_{w\sim\mathcal{N}(0,\sigma^2 I)}\left[V^\star_{h+1}(Ax + Bu + w)\right]$$

$$= x^\top Q x + u^\top R u + \mathbb{E}_{w\sim\mathcal{N}(0,\sigma^2 I)}\left[(Ax + Bu + w)^\top P_{h+1}(Ax + Bu + w) + p_{h+1}\right]$$

Induction Step:

$$\min_{\pi_0,\ldots,\pi_{H-1}} \quad E\left[x_H^\top Q x_H + \sum_{h=0}^{H-1}(x_h^\top Q x_h + u_h^\top R u_h)\right]$$

such that $\quad x_{h+1} = Ax_h + Bu_h + w_h,\ u_h = \pi_h(x_h) \quad x_0 \sim \mu_0,\ w_h \sim N(0,\sigma^2 I)$,

Assume $V^\star_{h+1}(x) = x^\top P_{h+1} x + p_{h+1}$, for all x, where $P_{h+1} \in \mathbb{R}^{d\times d}, p_{h+1} \in \mathbb{R}$

$$Q^\star_h(x,u) = c(x,u) + \mathbb{E}_{x'\sim P(x,u)} V^\star_{h+1}(x')$$

$$= x^\top Q x + u^\top R u + \mathbb{E}_{x'\sim P(x,u)} V^\star_{h+1}(x')$$

$$= x^\top Q x + u^\top R u + \mathbb{E}_{w\sim\mathcal{N}(0,\sigma^2 I)}\left[V^\star_{h+1}(Ax+Bu+w)\right]$$

$$= x^\top Q x + u^\top R u + \mathbb{E}_{w\sim\mathcal{N}(0,\sigma^2 I)}\left[(Ax+Bu+w)^\top P_{h+1}(Ax+Bu+w) + p_{h+1}\right]$$

$$= x^\top\left(Q + A^\top P_{h+1} A\right)x + u^\top\left(R + B^\top P_{h+1} B\right)u + 2x^\top A^\top P_{h+1} B u + \mathbb{E}_{w\sim\mathcal{N}(0,\sigma^2 I)}\left[w^\top P_{h+1} w\right] + p_{h+1}$$

Induction Step:

$$\min_{\pi_0,\ldots,\pi_{H-1}} \quad E\left[x_H^\top Q x_H + \sum_{h=0}^{H-1} (x_h^\top Q x_h + u_h^\top R u_h)\right]$$

such that $\quad x_{h+1} = Ax_h + Bu_h + w_h,\ u_h = \pi_h(x_h) \quad x_0 \sim \mu_0,\ w_h \sim N(0,\sigma^2 I)$,

Assume $V^\star_{h+1}(x) = x^\top P_{h+1} x + p_{h+1}$, for all x, where $P_{h+1} \in \mathbb{R}^{d\times d}, p_{h+1} \in \mathbb{R}$

$$Q_h^\star(x,u) = c(x,u) + \mathbb{E}_{x'\sim P(x,u)} V^\star_{h+1}(x')$$

$$= x^\top Q x + u^\top R u + \mathbb{E}_{x'\sim P(x,u)} V^\star_{h+1}(x')$$

$$= x^\top Q x + u^\top R u + \mathbb{E}_{w\sim\mathcal{N}(0,\sigma^2 I)}\left[V^\star_{h+1}\left(Ax + Bu + w\right)\right]$$

$$= x^\top Q x + u^\top R u + \mathbb{E}_{w\sim\mathcal{N}(0,\sigma^2 I)}\left[(Ax + Bu + w)^\top P_{h+1}(Ax + Bu + w) + p_{h+1}\right]$$

$$= x^\top\left(Q + A^\top P_{h+1} A\right)x + u^\top\left(R + B^\top P_{h+1} B\right)u + 2x^\top A^\top P_{h+1} B u + \mathbb{E}_{w\sim\mathcal{N}(0,\sigma^2 I)}\left[w^\top P_{h+1} w\right] + p_{h+1}$$

$$= x^\top\left(Q + A^\top P_{h+1} A\right)x + u^\top\left(R + B^\top P_{h+1} B\right)u + 2x^\top A^\top P_{h+1} B u + \mathrm{tr}\left(\sigma^2 P_{h+1}\right) + p_{h+1}$$

Induction Step (continue)

$$Q_h^\star(x,u) = c(x,u) + \mathbb{E}_{x' \sim P(x,u)}\left[V_{h+1}^\star(x')\right]$$

$$= x^\top\left(Q + A^\top P_{h+1} A\right)x + u^\top\left(R + B^\top P_{h+1} B\right)u + 2x^\top A^\top P_{h+1} B u + \mathsf{tr}\left(\sigma^2 P_{h+1}\right) + p_{h+1}$$

Induction Step (continue)

$$Q_h^\star(x,u) = c(x,u) + \mathbb{E}_{x' \sim P(x,u)}\left[V_{h+1}^\star(x')\right]$$

$$= x^\top \left(Q + A^\top P_{h+1} A\right) x + u^\top \left(R + B^\top P_{h+1} B\right) u + 2x^\top A^\top P_{h+1} B u + \mathrm{tr}\left(\sigma^2 P_{h+1}\right) + p_{h+1}$$

$$\pi_h^\star(x) = \arg\min_u Q_h^\star(x,u)$$

Induction Step (continue)

$$Q_h^\star(x,u) = c(x,u) + \mathbb{E}_{x' \sim P(x,u)}\left[V_{h+1}^\star(x')\right]$$

$$= x^\top\left(Q + A^\top P_{h+1} A\right)x + u^\top\left(R + B^\top P_{h+1} B\right)u + 2x^\top A^\top P_{h+1} B u + \mathrm{tr}\left(\sigma^2 P_{h+1}\right) + p_{h+1}$$

$$\pi_h^\star(x) = \arg\min_u Q_h^\star(x,u)$$

Set $\nabla_u Q_h^\star(x,u) = 0$, and solve for u:

Induction Step (continue)

$$Q_h^\star(x,u) = c(x,u) + \mathbb{E}_{x' \sim P(x,u)}\left[V_{h+1}^\star(x')\right]$$

$$= x^\top\left(Q + A^\top P_{h+1} A\right)x + u^\top\left(R + B^\top P_{h+1} B\right)u + 2x^\top A^\top P_{h+1} B u + \mathrm{tr}\left(\sigma^2 P_{h+1}\right) + p_{h+1}$$

$$\pi_h^\star(x) = \arg\min_u Q_h^\star(x,u)$$

Set $\nabla_u Q_h^\star(x,u) = 0$, and solve for u:

$$\pi_h^\star(x) = -\underbrace{(R + B^\top P_{h+1} B)^{-1} B^\top P_{h+1} A}_{:= K_h^\star}\, x$$

Induction Step (continue)

$$Q_h^\star(x,u) = c(x,u) + \mathbb{E}_{x' \sim P(x,u)}\left[V_{h+1}^\star(x')\right]$$

$$= x^\top\left(Q + A^\top P_{h+1} A\right)x + u^\top\left(R + B^\top P_{h+1} B\right)u + 2x^\top A^\top P_{h+1} B u + \text{tr}\left(\sigma^2 P_{h+1}\right) + p_{h+1}$$

$$\pi_h^\star(x) = \arg\min_u Q_h^\star(x,u)$$

Set $\nabla_u Q_h^\star(x,u) = 0$, and solve for u:

$$\pi_h^\star(x) = -\underbrace{(R + B^\top P_{h+1} B)^{-1} B^\top P_{h+1} A}_{:= K_h^\star}\, x$$

$$:= -K_h^\star x$$

Concluding the Induction step:

$$Q_h^\star(x,u) = x^\top \left(Q + A^\top P_{h+1} A\right) x + u^\top \left(R + B^\top P_{h+1} B\right) u + 2x^\top A^\top P_{h+1} B u + \mathrm{tr}\left(\sigma^2 P_{h+1}\right) + p_{h+1}$$

$$\pi_h^\star(x) = -\underbrace{(R + B^\top P_{h+1} B)^{-1} B^\top P_{h+1} A}_{:= K_h^\star}\, x$$

Concluding the Induction step:

$$Q_h^\star(x,u) = x^\top\left(Q + A^\top P_{h+1} A\right)x + u^\top\left(R + B^\top P_{h+1} B\right)u + 2x^\top A^\top P_{h+1} B u + \mathrm{tr}\left(\sigma^2 P_{h+1}\right) + p_{h+1}$$

$$\pi_h^\star(x) = \underbrace{-(R + B^\top P_{h+1} B)^{-1} B^\top P_{h+1} A}_{:= K_h^\star}\, x$$

$$V_h^\star(x) = Q_h^\star(x, \pi_h^\star(x))$$

Concluding the Induction step:

$$Q_h^\star(x,u) = x^\top \left(Q + A^\top P_{h+1} A\right) x + u^\top \left(R + B^\top P_{h+1} B\right) u + 2x^\top A^\top P_{h+1} B u + \mathrm{tr}\left(\sigma^2 P_{h+1}\right) + p_{h+1}$$

$$\pi_h^\star(x) = -\underbrace{(R + B^\top P_{h+1} B)^{-1} B^\top P_{h+1} A}_{:= K_h^\star}\, x$$

$$V_h^\star(x) = Q_h^\star(x, \pi_h^\star(x))$$

We can express $V_h^\star(x)$ as $V_h^\star(x) = x^\top P_h x + p_h,$ where:

Concluding the Induction step:

$$Q_h^\star(x,u) = x^\top\left(Q + A^\top P_{h+1}A\right)x + u^\top\left(R + B^\top P_{h+1}B\right)u + 2x^\top A^\top P_{h+1}Bu + \mathrm{tr}\left(\sigma^2 P_{h+1}\right) + p_{h+1}$$

$$\pi_h^\star(x) = \underbrace{-\,(R + B^\top P_{h+1}B)^{-1}B^\top P_{h+1}A}_{:=K_h^\star}\,x$$

$$V_h^\star(x) = Q_h^\star(x, \pi_h^\star(x))$$

We can express $V_h^\star(x)$ as $V_h^\star(x) = x^\top P_h x + p_h,$ where:

$$P_h = Q + A^\top P_{h+1}A - A^\top P_{h+1}B(R + B^\top P_{h+1}B)^{-1}B^\top P_{h+1}A,$$
$$p_h = \mathrm{tr}\left(\sigma^2 P_{h+1}\right) + p_{h+1}$$

Concluding the Induction step:

$$Q_h^\star(x,u) = x^\top\left(Q + A^\top P_{h+1} A\right)x + u^\top\left(R + B^\top P_{h+1} B\right)u + 2x^\top A^\top P_{h+1} B u + \mathrm{tr}\left(\sigma^2 P_{h+1}\right) + p_{h+1}$$

$$\pi_h^\star(x) = -\underbrace{(R + B^\top P_{h+1} B)^{-1} B^\top P_{h+1} A}_{:=K_h^\star}\, x$$

$$V_h^\star(x) = Q_h^\star(x, \pi_h^\star(x))$$

We can express $V_h^\star(x)$ as $V_h^\star(x) = x^\top P_h x + p_h$, where:

$$P_h = Q + A^\top P_{h+1} A - A^\top P_{h+1} B (R + B^\top P_{h+1} B)^{-1} B^\top P_{h+1} A,$$ **Ricatti Equation**

$$p_h = \mathrm{tr}\left(\sigma^2 P_{h+1}\right) + p_{h+1}$$

Summary:

Summary:

$V_H^\star(x) = x^\top Q x$, define $P_H = Q, p_H = 0,$

Summary:

$V_H^\star(x) = x^\top Q x$, define $P_H = Q, p_H = 0$,

We have shown that $V_h^\star(x) = x^\top P_h x + p_h$, where:

$$P_h = Q + A^\top P_{h+1} A - A^\top P_{h+1} B (R + B^\top P_{h+1} B)^{-1} B^\top P_{h+1} A,$$

$$p_h = \mathrm{tr}\left(\sigma^2 P_{h+1}\right) + p_{h+1}$$

Summary:

$$V_H^\star(x) = x^\top Q x, \text{ define } P_H = Q, p_H = 0,$$

We have shown that $V_h^\star(x) = x^\top P_h x + p_h$, where:

$$P_h = Q + A^\top P_{h+1} A - A^\top P_{h+1} B (R + B^\top P_{h+1} B)^{-1} B^\top P_{h+1} A,$$

$$p_h = \mathrm{tr}\left(\sigma^2 P_{h+1}\right) + p_{h+1}$$

Along the way, we also have shown that $\pi_h^\star(x) = -K_h^\star x$ where:

$$\pi_h^\star(x) = -\underbrace{(R + B^\top P_{h+1} B)^{-1} B^\top P_{h+1} A}_{:= K_h^\star}\, x$$

Summary:

$$V_H^\star(x) = x^\top Q x, \text{ define } P_H = Q, p_H = 0,$$

We have shown that $V_h^\star(x) = x^\top P_h x + p_h,$ where:

$$P_h = Q + A^\top P_{h+1} A - A^\top P_{h+1} B (R + B^\top P_{h+1} B)^{-1} B^\top P_{h+1} A,$$

$$p_h = \text{tr}\left(\sigma^2 P_{h+1}\right) + p_{h+1}$$

Along the way, we also have shown that $\pi_h^\star(x) = -K_h^\star x$ where:

$$\pi_h^\star(x) = -\underbrace{(R + B^\top P_{h+1} B)^{-1} B^\top P_{h+1} A}_{:= K_h^\star}\, x$$

Optimal control has nothing to do with initial distribution, and the noise!

Some Basic Extension:

Time dependent costs and transitions:

$$\min_{\pi_0,\ldots,\pi_{H-1}} \quad E\left[x_H^\top Q_H x_H + \sum_{h=0}^{H-1} (x_h^\top Q_h x_h + u_h^\top R_h u_h) \right]$$

such that $\quad x_{h+1} = A_h x_h + B_h u_h + w_h, \; u_h = \pi_h(x_h) \quad x_0 \sim \mu_0, \, w_h \sim N(0, \sigma^2 I)\,,$

Some Basic Extension:

Time dependent costs and transitions:

$$\min_{\pi_0,\ldots,\pi_{H-1}} \quad E\left[x_H^\top Q_H x_H + \sum_{h=0}^{H-1} (x_h^\top Q_h x_h + u_h^\top R_h u_h)\right]$$

such that $\quad x_{h+1} = A_h x_h + B_h u_h + w_h,\ u_h = \pi_h(x_h) \quad x_0 \sim \mu_0,\ w_h \sim N(0,\sigma^2 I)\,,$

Same derivation, we will have the following Ricatti Equation:

$$P_h = Q_h + A_h^\top P_{h+1} A_h - A_h^\top P_{h+1} B_h (R_h + B_h^\top P_{h+1} B_h)^{-1} B_h^\top P_{h+1} A_h,$$

$$p_h = \mathrm{tr}\left(\sigma^2 P_{h+1}\right) + p_{h+1}$$

Some Basic Extension:

More generally…

$$\min_{\pi_0,\ldots,\pi_{H-1}} \quad E\left[x_H^\top Q_H x_H + x_H^\top q_H + c_H + \sum_{h=0}^{H-1} (x_h^\top Q_h x_h + u_h^\top R_h u_h + u_h^\top M_h x_h + x_h^\top q_h + u_h^\top r_h + c_h)\right]$$

such that $\quad x_{h+1} = A_h x_h + B_h u_h + v_h + w_h, \;\; u_h = \pi_h(x_h) \quad x_0 \sim \mu_0, \; w_h \sim N(0,\sigma^2 I)\,,$

Some Basic Extension:

More generally…

$$\min_{\pi_0,\ldots,\pi_{H-1}} \quad E\left[x_H^\top Q_H x_H + x_H^\top q_H + c_H + \sum_{h=0}^{H-1} (x_h^\top Q_h x_h + u_h^\top R_h u_h + u_h^\top M_h x_h + x_h^\top q_h + u_h^\top r_h + c_h) \right]$$

such that $\quad x_{h+1} = A_h x_h + B_h u_h + v_h + w_h, \;\; u_h = \pi_h(x_h) \quad x_0 \sim \mu_0, \; w_h \sim N(0, \sigma^2 I)\,,$

Same DP idea and similar derivation
(HW1 question)

Some Basic Extension:

Tracking a pre-defined trajectory:

$$\min_{\pi_0,\ldots,\pi_{H-1}} \quad E\left[(x_H - x_H^\star)^\top Q_H (x_H - x_H^\star) + \sum_{h=0}^{H-1} (x_h - x_h^\star)^\top Q_h (x_h - x_h^\star) + (u_h - u_h^\star)^\top R_h (u_h - u_h^\star)\right]$$

such that $\quad x_{h+1} = A_h x_h + B_h u_h + w_h,\ \ u_h = \pi_h(x_h) \quad x_0 \sim \mu_0,\ w_h \sim N(0,\sigma^2 I)\,,$

Some Basic Extension:

Tracking a pre-defined trajectory:

$$\min_{\pi_0,\ldots,\pi_{H-1}} \quad E\left[(x_H - x_H^\star)^\top Q_H (x_H - x_H^\star) + \sum_{h=0}^{H-1} (x_h - x_h^\star)^\top Q_h (x_h - x_h^\star) + (u_h - u_h^\star)^\top R_h (u_h - u_h^\star)\right]$$

such that $\quad x_{h+1} = A_h x_h + B_h u_h + w_h, \;\; u_h = \pi_h(x_h) \quad x_0 \sim \mu_0, \; w_h \sim N(0,\sigma^2 I)\,,$

We can simply complete the square
and we reduce back to the setting in the previous slide!

Known transition versus black-box access

So far, we studied Policy Evaluation, Policy Iteration, Value Iteration, and DP-based approach,

we have assumed that transition is **known**,
i.e., $P(s' \mid s, a), \forall s, a, s'$ is known, or $A, B, \mathcal{N}(0, \sigma^2 I)$ are known

Known transition versus black-box access

So far, we studied Policy Evaluation, Policy Iteration, Value Iteration, and DP-based approach,

we have assumed that transition is **known**,
i.e., $P(s' \mid s, a), \forall s, a, s'$ is known, or $A, B, \mathcal{N}(0, \sigma^2 I)$ are known

Starting from this Thursday, we start considering **unknown** transition:

Known transition versus black-box access

So far, we studied Policy Evaluation, Policy Iteration, Value Iteration, and DP-based approach,

we have assumed that transition is **known**,
i.e., $P(s' \mid s, a), \forall s, a, s'$ is known, or $A, B, \mathcal{N}(0, \sigma^2 I)$ are known

Starting from this Thursday, we start considering **unknown** transition:

We start w/ black-box access to P, or $f(x, u, w)$:

Known transition versus black-box access

So far, we studied Policy Evaluation, Policy Iteration, Value Iteration, and DP-based approach,

we have assumed that transition is **known**,
i.e., $P(s' | s, a), \forall s, a, s'$ is known, or $A, B, \mathcal{N}(0, \sigma^2 I)$ are known

Starting from this Thursday, we start considering **unknown** transition:

We start w/ black-box access to P, or $f(x, u, w)$:

We can **reset the system to any** (s, a), and observe $s' \sim P(\cdot | s, a)$,

Known transition versus black-box access

So far, we studied Policy Evaluation, Policy Iteration, Value Iteration, and DP-based approach,

we have assumed that transition is **known**,
i.e., $P(s' | s, a), \forall s, a, s'$ is known, or $A, B, \mathcal{N}(0, \sigma^2 I)$ are known

Starting from this Thursday, we start considering **unknown** transition:

We start w/ black-box access to P, or $f(x, u, w)$:

We can **reset the system to any** (s, a), and observe $s' \sim P(\cdot | s, a)$,

Or we can reset to any (x, u), and observe $x' = f(x, u, w)$
(w being some unknown noisy disturbance)

Summary today:

1. We use DP to derive the optimal control for LQR (Ricatti equation)!

2. Never try to remember the exact form!
Only need to understand the way we derive it (again DP!)

Next Lecture:

Control for Nonlinear system w/ black-box access to $f(x, u)$
(In general, very hard, we will study approximate algorithm
and only aim for locally optimal solutions)