S26 AerE 5740: Optimal Control

Practice Set 08: Discrete Time Finite Horizon DP

Instructor: Abhishek Halder

Problem 1. Discrete Time Finite Horizon LQR via DP

Recall the discrete time finite horizon LQR problem from Lec. 13 p. 17:

$$\begin{aligned} &\underset{\{\boldsymbol{u}_k\}_{k=0}^{T-1}}{\text{minimize}} \quad \left\{\frac{1}{2}\boldsymbol{x}_T^\top \boldsymbol{M}\boldsymbol{x}_T + \sum_{k=0}^{T-1}\left(\boldsymbol{x}_k^\top \boldsymbol{Q}_k\boldsymbol{x}_k + \boldsymbol{u}_k^\top \boldsymbol{R}_k\boldsymbol{u}_k\right)\right\} \\ &\text{subject to} \quad \boldsymbol{x}_{k+1} = \boldsymbol{A}_k\boldsymbol{x}_k + \boldsymbol{B}_k\boldsymbol{u}_k, \end{aligned}$$

where as usual, for all $k \in \{0, 1, \ldots, T-1\}$, the state vectors $\boldsymbol{x}_k \in \mathbb{R}^n$, the input $\boldsymbol{u}_k \in \mathbb{R}^m$, the cost-to-go weight matrices $\boldsymbol{Q}_k \succ \boldsymbol{0}$, $\boldsymbol{R}_k \succ \boldsymbol{0}$, and the terminal cost weight matrix $\boldsymbol{M} \succeq \boldsymbol{0}$.

(a) DP equation.

Write down the <u>DP equation</u> for this problem.

Solution for part (a):

This is a discrete time finite horizon deterministic DP. Therefore, the backward update for value function becomes:

$$V_k(\boldsymbol{x}) = \underset{\boldsymbol{u}\in\mathbb{R}^m}{\text{minimum}} \quad \left\{\frac{1}{2}\boldsymbol{x}^\top \boldsymbol{Q}_k\boldsymbol{x} + \frac{1}{2}\boldsymbol{u}^\top \boldsymbol{R}_k\boldsymbol{u} + V_{k+1}\left(\boldsymbol{A}_k\boldsymbol{x} + \boldsymbol{B}_k\boldsymbol{u}\right)\right\} \quad \forall k = T-1, T-2, \ldots, 0,$$

where $V_T(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^\top \boldsymbol{M}\boldsymbol{x}$.

(b) Solution for the DP equation.

Looking at the DP equation from part (a), it is reasonable to make a guess for value function as

$$V_k(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^\top \boldsymbol{P}_k\boldsymbol{x} \quad \forall k = T, T-1, \ldots, 0,$$

for some to-be-determined $\boldsymbol{P}_k \succeq \boldsymbol{0}$. From the known $V_T(\boldsymbol{x})$ in part (a) solution, we see that $\boldsymbol{P}_T = \boldsymbol{M}$. <u>Verify this guess for the value function, and thus solve the DP equation</u>.

Solution for part (b):

Substituting the guess for the value function in the DP equation in part (a) solution, we have

$$\frac{1}{2}\boldsymbol{x}^\top \boldsymbol{P}_{k+1}\boldsymbol{x} = \underset{\boldsymbol{u}\in\mathbb{R}^m}{\text{minimum}} \quad \left\{\frac{1}{2}\boldsymbol{x}^\top \boldsymbol{Q}_k\boldsymbol{x} + \frac{1}{2}\boldsymbol{u}^\top \boldsymbol{R}_k\boldsymbol{u} + \frac{1}{2}\left(\boldsymbol{A}_k\boldsymbol{x} + \boldsymbol{B}_k\boldsymbol{u}\right)^\top \boldsymbol{P}_{k+1}\left(\boldsymbol{A}_k\boldsymbol{x} + \boldsymbol{B}_k\boldsymbol{u}\right)\right\}. \tag{1}$$

To find the minimum value of the right-hand-side above, we set its derivative w.r.t. $\boldsymbol{u}$ to zero. This gives the minimizer $\boldsymbol{u}^{\text{opt}}$ for the above right-hand-side:

$$\left(\boldsymbol{R}_k + \boldsymbol{B}_k^\top \boldsymbol{P}_{k+1}\boldsymbol{B}_k\right)\boldsymbol{u}^{\text{opt}} = -\boldsymbol{B}_k^\top \boldsymbol{P}_{k+1}\boldsymbol{A}_k\boldsymbol{x}.$$

Since $\boldsymbol{B}_k^\top \boldsymbol{P}_{k+1}\boldsymbol{B}_k = \left(\boldsymbol{P}_{k+1}^{\frac{1}{2}}\boldsymbol{B}_k\right)^\top\left(\boldsymbol{P}_{k+1}^{\frac{1}{2}}\boldsymbol{B}_k\right) \succeq \boldsymbol{0}$, the matrix $\left(\boldsymbol{R}_k + \boldsymbol{B}_k^\top \boldsymbol{P}_{k+1}\boldsymbol{B}_k\right) \succ \boldsymbol{0}$, thus invertible. Therefore, the minimizer

$$\boldsymbol{u}^{\text{opt}} = -\left(\boldsymbol{R}_k + \boldsymbol{B}_k^\top \boldsymbol{P}_{k+1}\boldsymbol{B}_k\right)^{-1}\boldsymbol{B}_k^\top \boldsymbol{P}_{k+1}\boldsymbol{A}_k\boldsymbol{x},$$

which is exactly the time-varying state feedback we derived via PMP in Lec. 14 p. 1-2. Substituting this minimizer back in (1), the DP equation becomes

$$\frac{1}{2}\boldsymbol{x}^\top \boldsymbol{P}_k\boldsymbol{x} = \frac{1}{2}\boldsymbol{x}^\top\left[\boldsymbol{Q}_k + \boldsymbol{A}_k^\top\left(\boldsymbol{I} + \boldsymbol{P}_{k+1}\boldsymbol{B}_k\boldsymbol{R}_k^{-1}\boldsymbol{B}_k^\top\right)\boldsymbol{P}_{k+1}\boldsymbol{A}_k\right]\boldsymbol{x}. \tag{2}$$

For (2) to hold for all $\boldsymbol{x}$, we must satisfy the matrix recursion $\boldsymbol{P}_{k+1} \mapsto \boldsymbol{P}_k$ given by

$$\boldsymbol{P}_k = \boldsymbol{Q}_k + \boldsymbol{A}_k^\top\left(\boldsymbol{I} + \boldsymbol{P}_{k+1}\boldsymbol{B}_k\boldsymbol{R}_k^{-1}\boldsymbol{B}_k^\top\right)\boldsymbol{P}_{k+1}\boldsymbol{A}_k, \quad \text{subject to terminal condition } \boldsymbol{P}_T = \boldsymbol{M} \succeq \boldsymbol{0},$$

to be run backward in time from $k = T-1$ till $k = 0$. This is precisely the Riccati matrix recursion we derived in Lec. 13 p. 21-22 via PMP.