Control for Robotics: From Optimal Control to Reinforcement Learning

Lingjie Zhang

2024 SS

Assignment 1: Optimal Control and Dynamic Programming

Contents

Problem 1.1 Finite Horizon Dynamic Programming

(a)

The policy found using the dynamic programming algorithm is expected to behave in a way that minimizes the total cost over the given time horizon.

Different values of q and r will change the robot's behavior as follows:

- Increasing q will increase penalty of position errors in the cost function, which leading to a policy that prioritizes minimizing position errors.
- Increasion r will increase penalty of control inputs in the cost function, which leading to a policy that prioritizes smoother control inputs to save energy.

(b)

· **Dynamics**

$$x_{k+1} = x_k + u_k + w_k \tag{1.1}$$

· **Cost**

$$x_2^2 + \sum_{k=0}^{1}(qx_k^2 + ru_k^2) \tag{1.2}$$

Let $q = 5/2$ and $r = 1$ and assume that $w_k = 0$ for all k:

· **Initialization**

$$J_2(x_2) = x_2^2 \tag{1.3}$$

· **Recursion**

$\triangleright$ *Step* $k = 1$

$$\begin{aligned} J_1(x_1) &= \min_{u_1}(\frac{5}{2}xx_1^2 + u_1^2 + J_2(x_2)) && (1.4) \\ &= \min_{u_1}(\frac{5}{2}x_1^2 + u_1^2 + (x_1 + u_1)^2) \text{ (sub dynamics)} && (1.5) \\ &= \min_{u_1}(\frac{7}{2}x_1^2 + 2u_1x_1 + 2u_1^2) && (1.6) \end{aligned}$$

Solve for optimal u_1 by differentiating the cost and setting it to zero:

$$u_1 = -\frac{x_1}{2} \tag{1.7}$$

Substitute u_1 back to $J_1(x_1)$:

$$J_1(x_1) = 3x_1^2 \tag{1.8}$$

▷ *Step* $k = 0$

$$J_0(x_0) = \min_{u_0}(\frac{5}{2}x_0^2 + u_0^2 + J_1(x_1)) \tag{1.9}$$

$$= \min_{u_0}(\frac{11}{2}x_0^2 + 4u_0^2 + 6x_0u_0) \tag{1.10}$$

Solve for optimal u_0 by differentiating the cost and setting it to zero:

$$u_0 = -\frac{3x_0}{4} \tag{1.11}$$

Substitute u_0 back to $J_0(x_0)$:

$$J_0(x_0) = \frac{13}{4}x_0^2 \tag{1.12}$$