EXERCISE SESSION 3 - SOLUTIONS

Problem 1. (a) We use backward induction to find the optimal control sequence $u_0, \ldots, u_{N-1}$. Recall: The value function $V(x_k)$ is the minimum cost of moving the system from the state x_k to the end of the horizon.

Step $k = N$. At time step N, the value function is given by

$$V(x_N) = q_3 x_N^2.$$

Two things are important to note. Firstly, given x_N, $V(x_N)$ is fixed. Secondly, $V(x_N)$ is quadratic in x_N. Before we proceed, we rewrite $V(x_N)$ in a more convenient form

$$V(x_N) = p_N x_N^2. \tag{1}$$

Step $k = N-1$. The cost to go in $k = N-1$ is defined by

$$V(x_{N-1}) = \underset{u_{N-1}}{\text{minimize}}\{q_1 x_{N-1}^2 + q_2 u_{N-1}^2 + V(x_N)\}, \tag{2}$$

where:

- $q_1 x_{N-1}^2$ is the cost of being in x_{N-1};
- $q_2 u_{N-1}^2$ is the cost of using u_{N-1};
- $V(x_N)$ is "the cost to go".

By combining (1) and (2), we obtain

$$\begin{aligned}
V(x_{N-1}) &= \underset{u_{N-1}}{\text{minimize}}\{q_1 x_{N-1}^2 + q_2 u_{N-1}^2 + V(x_N)\}, \\
&= \underset{u_{N-1}}{\text{minimize}}\{q_1 x_{N-1}^2 + q_2 u_{N-1}^2 + p_N x_N^2\} \\
&= \underset{u_{N-1}}{\text{minimize}}\{q_1 x_{N-1}^2 + q_2 u_{N-1}^2 + p_N (a x_{N-1} + b u_{N-1})^2\} \\
&= \underset{u_{N-1}}{\text{minimize}}\{(q_1 + P_N a^2) x_{N-1}^2 + 2 p_N a b x_{N-1} u_{N-1} + (q_2 + P_N b^2) u_{N-1}^2\} \\
&= \underset{u_{N-1}}{\text{minimize}}\{c_1 u_{N-1}^2 + c_2 u_{N-1} + c_3\}
\end{aligned} \tag{3}$$

where $c_1 = q_2 + P_N b^2$, $c_2 = 2 p_N a b x_{N-1}$, and $c_3 = (q_1 + P_N a^2) x_{N-1}^2$. Note that $c_1 > 0$, since $q_2 > 0$ and $P_N = q_3 > 0$. Hence, the function $c_1 u_{N-1}^2 + c_2 u_{N-1} + c_3$ has the minimum that can be calculated by solving the equation

$$\frac{\partial (c_1 u_{N-1}^2 + c_2 u_{N-1} + c_3)}{\partial u_{N-1}} = 2 c_1 u_{N-1} + c_2 = 0. \tag{4}$$

Thus, it follows

$$u_{N-1}^* = -\frac{c_2}{2 c_1} = -\frac{a b p_N}{q_2 + P_N b^2} x_{N-1}.$$

Important observation: This is a state feedback.

By combining (3) and (4) we obtain

$$\begin{aligned}V(x_{N-1}) &= c_1(-\frac{c_2}{2c_1})^2 + c_2(-\frac{c_2}{2c_1}) + c_3 = -\frac{c_2^2}{4c_1} + c_3 = (q_1 + P_N a^2)x_{N-1}^2 - \frac{(2p_N abx_{N-1})^2}{4(q_2 + P_N b^2)} \\ &= \frac{(q_2 + P_N b^2)(q_1 + P_N a^2) - (p_N ab)^2}{q_2 + P_N b^2} x_{N-1}^2 \\ &= \frac{q_1 q_2 + q_1 P_N b^2 q_1 + q_2 P_N a^2 + P_N b^2 P_N a^2 - (p_N ab)^2}{q_2 + P_N b^2} x_{N-1}^2 \\ &= \frac{q_1 q_2 + q_1 P_N b^2 + q_2 P_N a^2}{q_2 + P_N b^2} x_{N-1}^2 = p_{N-1} x_{N-1}^2\end{aligned}$$

where

$$p_{N-1} = q_1 + \frac{q_2 P_N a^2}{q_2 + P_N b^2}.$$

Important observations:

- $V(x_{N-1})$ is again quadratic function with $p_{N-1} > 0$;
- We expressed p_{N-1} through p_N.

Hence, we can follow the same procedure for x_{N-2} and any other x_k!

Step k. To summarize, we calculate the control action at time step k as

$$u_k^* = -\frac{p_{k+1}ab}{q_2 + p_{k+1}b^2} x_k,$$

where

$$p_{k+1} = \begin{cases} q_3 & k+1 = N \\ q_1 + \frac{q_2 a^2 p_{k+2}}{q_2 + p_{k+2} b^2}, & k+1 < N. \end{cases} \tag{5}$$

(b) When $N = +\infty$, we have $V(x_k) = px_k^2$ for every k[1], which implies $p_k = p$ for every k. From (5), it follows that

$$p = q_1 + \frac{q_2 a^2 p}{q_2 + pb^2} \Rightarrow b^2 p^2 + (q_2 - q_1 b^2 - q_2 a^2)p - q_1 q_2 = 0. \tag{6}$$

By solving this quadratic equation, we obtain

$$p_{1,2} = \frac{-(q_2 - q_1 b^2 - q_2 a^2) \pm \sqrt{(q_2 - q_1 b^2 - q_2 a^2)^2 + 4b^2 q_1 q_2}}{2b^2}.$$

Some observations:

- Under the assumptions we made (the system is controllable, and q_1, q_2, and q_3 are greater than zero), we should have the unique positive solution for p (Theorem 4.1.3). Indeed, since

  $$-(q_1 b^2 + q_2 a^2 - q_2) < \sqrt{(q_2 - q_1 b^2 - q_2 a^2)^2 + 4b^2 q_1 q_2}$$

  we have

  $$p = \frac{-(q_2 - q_1 b^2 - q_2 a^2) + \sqrt{(q_2 - q_1 b^2 - q_2 a^2)^2 + 4b^2 q_1 q_2}}{2b^2},$$

  to be the only positive solution. Hence, our feedback control is given by

  $$u_k = -\frac{pab}{q_2 + pb^2} x_k.$$

[1]This is not trivial to show. See Problem 5 in the exam from 2018.

- Theorem 4.1.3 tells us that in addition to having (a, b) reachable, we need to have $(a, q_1^{1/2})$ observable to guarantee the unique positive definite solution for p. Indeed, if we do not have this condition satisfied, the unique positive solution may not exist. For example, if $q_1 = 0$ and $a < 1$, $(A, Q_1^{1/2})$ is not observable, and our solutions are

$$p_1 = \frac{-(q_2 - q_2 a^2) - \sqrt{(q_2 - q_2 a^2)^2}}{2b^2} = -\frac{q_2 - q_2 a^2}{b^2} < 0,$$
$$p_2 = \frac{-(q_2 - q_2 a^2) + \sqrt{(q_2 - q_2 a^2)^2}}{2b^2} = 0.$$

  Hence, we do not have the unique positive solution in this case.

Problem 2. See Solution to Exercise 4.5 in the compendium.

Problem 3. See Solution to Exercise 4.2 and Solution to Exercise 4.3 in the compendium.

Problem 4. See Solution to Exercise 4.8 in the compendium.