Solutions of problem set 2

Optimal control and reinforcement learning, 4SC000, TU/e

Q2, 2022-2023

---

1. Linear quadratic control

Problem 1.1

(a) For $\alpha = 1$ and $\gamma = 1$, $q = 2$, we have

$$x_{k+1} = x_k + u_k$$

and

$$\sum_{k=0}^{2} \underbrace{x_k^2 + u_k^2}_{g(x_k,u_k)} + \underbrace{0}_{J_3(x_3)}$$

which is a standard stage decision problem. Applying DP we obtain

Step 1

$$J_2(x_2) = \min_{u_2} g(x_2, u_2) + J_3(x_3) = \min_{u_2} x_2^2 + u_2^2 = x_2^2$$

where

$$u_2 = \mu(x_2) = 0 \ \ \forall_{x_2}$$

Step 2

$$J_1(x_1) = \min_{u_1} g(x_1, u_1) + J_2(x_2) = \min_{u_1} x_1^2 + u_1^2 + (x_1 + u_1)^2 = \min_{u_1} 2x_1^2 + 2u_1^2 + 2x_1u_1$$

to compute the minimum, we need to differentiate and equate to zero which leads to

$$4u_1 + 2x_1 = 0$$

where

$$u_1 = -\frac{1}{2}x_1$$

and

$$J_1(x_1) = 2x_1^2 + 2(-\frac{1}{2}x_1)^2 + 2x_1(-\frac{1}{2}x_1) = \frac{3}{2}x_1^2$$

Step 3

$$J_0(x_0) = \min_{u_0} g(x_0, u_0) + J_1(x_0) = \min_{u_0} x_0^2 + u_0^2 + \frac{3}{2}(x_0 + u_0)^2 = \min_{u_0} \frac{5}{2}x_0^2 + \frac{5}{2}u_0^2 + 3x_0u_0$$

to compute the minimum, we need to differentiate and equate to zero which leads to

$$5u_0 + 3x_0 = 0$$

where

$$u_0 = -\frac{3}{5}x_0$$

and

$$J_0(x_0) = 1.6x_0^2$$

Alternatively we can iterate the Riccati equations

$$p_k = p_{k+1} + 1 - \frac{p_{k+1}^2}{1 + p_{k+1}}$$

for $k \in \{2, 1, 0\}$ and $p_3 = 0$ and obtain the gains of the optimal policy ($u_k = K_k x_k$ with costs-to-go $x_k^2 p_k$) from

$$K_k = -\frac{p_{k+1}}{1 + p_{k+1}}.$$

We get

$$p_2 = 1, \quad K_2 = 0$$
$$p_1 = \frac{3}{2}, \quad K_1 = \frac{-1}{2}$$
$$p_0 = 1.6, \quad K_1 = \frac{-3}{5}$$

(b) Since the dynamic model is controllable and the cost matrix positive definite we know that the backward iteration

$$p_k = p_{k+1} + 1 - \frac{p_{k+1}^2}{1 + p_{k+1}}$$

will converge to a limit (say p) and thus we obtain the algebraic Riccati equation

$$p = p + 1 - \frac{p^2}{1 + p}$$

or equivalently

$$p^2 - p - 1 = 0$$

which has two roots

$$p = \frac{1 \pm \sqrt{5}}{2}.$$

We know that the algebraic Riccation equation has a unique positive definite solution, but in general can have more. The one that we are interested in is the positive definite (since $x_0^\intercal P x_0$ must be positive for $x_0 \neq 0$). Thus the optimal cost is $x_0^2 p$ where

$$p = \frac{1 + \sqrt{5}}{2} = 1.618.$$

The optimal policy is then

$$u_k = K x_k, k \in \mathbb{N}_0$$

where

$$K = -\frac{p}{1 + p} = -\frac{1 + \sqrt{5}}{3 + \sqrt{5}} = -0.618.$$

(c) $P(\gamma)$ satisfies $p = \alpha^2 p + 1 - p^2\alpha^2/(\gamma + p)$ and converges to $\frac{1}{1-\alpha^2}$ as $\gamma \to \infty$. Thus, $K(\gamma) = -\frac{\alpha^2 P(\gamma)^2}{\gamma + P(\gamma)}$ converges to zero, and thus the optimal policy is $u_k = 0$ for every $k \in \mathbb{N}_0$. The interpretation is that if the control is expensive and the system is stable $\alpha < 1$ one can run it open loop. The cost results from iterating

$$x_{k+1} = \alpha x_k$$

and computing the cost

$$\sum_{k=0}^{\infty} x_k^2 = \sum_{k=0}^{\infty} (\alpha^2)^k x_0^2 = \frac{1}{1-\alpha^2} x_0^2$$

Problem 1.2

(i) $\dot{x} = \begin{bmatrix} 0 & 1 \\ 0 & -\alpha \end{bmatrix} x + \begin{bmatrix} 0 \\ \beta \end{bmatrix} V$

$$\alpha = \frac{(b + K_t K_e/R)}{J} = 10.01$$
$$\beta = K_t/(RJ) = 1$$

(ii) Let

$$A := \begin{bmatrix} 0 & 1 \\ 0 & -\alpha \end{bmatrix}$$
$$B = \begin{bmatrix} 0 \\ \beta \end{bmatrix}$$
$$x_{k+1} = A_d x_k + B_d u_k$$
$$A_d = e^{A\tau} = \begin{bmatrix} 1.0000 & 0.0632 \\ 0 & 0.3675 \end{bmatrix}$$
$$B_d = \int_0^{\tau} e^{As} ds B = \begin{bmatrix} 0.0037 \\ 0.0632 \end{bmatrix}$$

(iii) (a) using the Matlab function *dlqr*

$$K = -\begin{bmatrix} 3.9698 & 1.3993 \end{bmatrix}$$
$$P = \begin{bmatrix} 57.4811 & 4.6278 \\ 4.6278 & 6.0073 \end{bmatrix}$$

(b) From problem 1.2 we know that the desired control inputs are given by

$$\begin{bmatrix} u_0 \\ u_1 \end{bmatrix} = \begin{bmatrix} K_0 \\ K_1 \end{bmatrix} x_0$$

$$\begin{bmatrix} K_0 \\ K_1 \end{bmatrix} = \begin{bmatrix} A_d B_d & B_d \end{bmatrix}^{-1} \times (-A_d^2)$$

and leads to $x_1 = (A_d + B_d K_0)x_0$ and $x_2 = 0$. Thus, the cost is

$$x_0^\intercal Q x_0 + x_0^\intercal K_0^\intercal R K_0 x_0 + x_0^\intercal (A_d + B_d K_0)^\intercal Q (A_d + B_d K_0) x_0 + x_0^\intercal K_1^\intercal R K_1 x_0$$
$$= x_0^\intercal P_f x_0$$

where

$$P_f = Q + K_0^\intercal R K_0 + (A_d + B_d K_0)^\intercal Q (A_d + B_d K_0) + K_1^\intercal R K_1$$

which is a qudratic function of x_0. For the numerical values given we get

$$K_0 = \begin{bmatrix} -158.2638 & -12.4309 \end{bmatrix}$$

$$K_1 = \begin{bmatrix} 58.1638 & 2.4309 \end{bmatrix}$$

$$P_f = \begin{bmatrix} 3345.4 & 231.78 \\ 231.78 & 21.92 \end{bmatrix}$$

(c) We can say that $x_0^\intercal P^* x_0 \leq x_0^\intercal P_f x_0$ for every initial condition x_0, since $P_f - P$ is a positive definite matrix.

Problem 1.3

(a)

$$\sum_{k=0}^{T} a^k = \frac{1-a}{1-a}\sum_{k=0}^{T} a^k = \frac{1}{1-a}\sum_{k=0}^{T}(a^k - a^{k+1}) = \frac{1}{1-a}(1-a^{T+1}) \tag{1}$$

(b)

$$J = \sum_{k=0}^{T} x_k^\top Q x_k = \sum_{k=0}^{T} x_0^\top (A^k)^\top Q A^k x_0 = x_0^\top [\sum_{k=0}^{T} (A^k)^\top Q A^k] x_0 \tag{2}$$

A is scalar, therefore

$$x_0^\top [\sum_{k=0}^{T} (A^k)^\top Q A^k] x_0 = x_0^2 Q \sum_{k=0}^{T} (A^2)^k \tag{3}$$

For $T \to \infty$

$$\sum_{k=0}^{T} a^k = \frac{1}{1-a} \tag{4}$$

$$x_0^2 Q \sum_{k=0}^{\infty} (A^2)^k = x_0^2 Q \frac{1}{1-A^2} \tag{5}$$

$$A^\top P A - P = -Q \Rightarrow Q = P(1-A^2) \tag{6}$$

$$J = x_0^2 P \tag{7}$$

(c) Note: We consider here only stable A.

$$\sum_{k=0}^{T} x_k^\top Q x_k = \sum_{k=0}^{T} x_0^\top (A^k)^\top Q A^k x_0 = x_0^\top [\sum_{k=0}^{T} (A^k)^\top Q A^k] x_0 \tag{8}$$

Fill in for Q

$$A^\top P A - P = -Q \tag{9}$$

And see that every time only element for k and $k+1$ are eliminated. Then

$$J = x_0^\top [\sum_{k=0}^{T} (A^k)^\top Q A^k] x_0 = x_0^\top [P - (A^T)^\top P A^T] x_0 \tag{10}$$

where the last term vanishes for $T \to \infty$.

(d) Follows by direct replacement. The explanation is that if we simply apply the control policy $u_k = Kx_k$, where K are the optimal LQR gains, to the problem

$$\min \sum_{k=0}^{\infty} x_k^\intercal Q x_k + u_k^\intercal R u_k$$

for

$$x_{k+1} = Ax_k + Bu_k$$

we get the cost

$$\sum_{k=0}^{\infty} x_k^\intercal \underbrace{(Q + K^\intercal R K)}_{:=\bar{Q}} x_k$$

where

$$x_{k+1} = \underbrace{(A + BK)}_{:=\bar{A}} x_k$$

Problem 1.4 From

$$x_{k+1} = Ax_k + Bu_k,$$

we conclude that

$$x_n = A^n x_0 + \begin{bmatrix} A^{n-1}B & A^{n-2}B & & B \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_{n-1} \end{bmatrix}$$

Then, the following control inputs achieve $x_n = \bar{x}$

$$\begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_{n-1} \end{bmatrix} = \begin{bmatrix} A^{n-1}B & A^{n-2}B & & B \end{bmatrix}^{-1} (\bar{x} - A^n x_0), \tag{11}$$

Problem 1.5

1. Running the matlab script shown in Fig. (1) we obtain

$$K_0 = \begin{bmatrix} 2.8560 & -3.3810 & -0.9790 \\ 0.0263 & 0.0605 & -1.4179 \end{bmatrix}$$

$$K_1 = \begin{bmatrix} 2.8494 & -3.3649 & -0.9777 \\ 0.0271 & 0.0589 & -1.4180 \end{bmatrix}$$

$$K_2 = \begin{bmatrix} 2.8191 & -3.2769 & -0.9749 \\ 0.0344 & 0.0450 & -1.4161 \end{bmatrix}$$

$$K_3 = \begin{bmatrix} 2.7408 & -3.0849 & -0.9967 \\ 0.0829 & -0.0448 & -1.3211 \end{bmatrix}$$

$$K_4 = \begin{bmatrix} 1.9667 & -2.6667 & -0.6667 \\ 0 & 0 & 0 \end{bmatrix}$$

The optimal policy is $u_k = K_k x_k$, $k \in \{0, 1, 2, 3, 4\}$.

```
Q = [5 0 0; 0 1 0;0 0 3];
S = [0.1 0; 0 0;0 0];
R = [1 0; 0 2];

A = [0 1 0;-3 4 1; 0 0 3];
B = [0 0;
    1 0;
    0 2];
Qh = [1 0 0;0 2 0; 0 0 0];
h = 5;
P{h+1} = Qh;
for k=h:-1:1
    K{k} = -(B'*P{k+1}*B+R)\(B'*P{k+1}*A+S');
    P{k} = A'*P{k+1}*A + Q -K{k}'*(B'*P{k+1}*B+R)*K{k};
end
[Kbar,Pbar] = dlqr(A,B,Q,R,S); Kbar = -Kbar;
```

Figure 1: Script, Problem 1.5

(ii) Considering a large h in the script ($h = 20$ was tested) we obtain $\bar{K} = K_0 = \begin{bmatrix} 2.8570 & -3.3830 & -0.9792 \\ 0.0262 & 0.0607 & -1.4179 \end{bmatrix}$. The optimal policy is $u_k = Kx_k$.

(iii) Using dlqr as shown at the bottom of Fig. (1) we obtain the same gains as in (ii).

Problem 1.6

- $B = 0$, A a matrix which is not Hurwitz (e.g. $A = 2$) Then the model boilds down to $x_{k+1} = Ax_k$ and the trajectories will blow up for any non-zero initial condition. With the parameters given above we get
$$P = 4P + Q$$
which has a unique solution $p = \frac{-Q}{3}$ which is not positive definite for any positive semi-definite Q.
- $Q = 0$, $A = 0$ -> the algebraic Riccati equation is then $P = 0$ whose solution is not positive definite.

Problem 1.7

(i) Changing the parameters A, B, Q, S,R , Q_h in the script of Question 1.5 to the ones given we obtain

$$K_0 = \begin{bmatrix} -1.1328 & -3.2656 \end{bmatrix}$$
$$K_1 = \begin{bmatrix} -1.1328 & -3.2655 \end{bmatrix}$$
$$K_2 = \begin{bmatrix} -1.1324 & -3.2647 \end{bmatrix}$$
$$K_3 = \begin{bmatrix} -1.1250 & -3.2500 \end{bmatrix}$$
$$K_4 = \begin{bmatrix} -1 - 3 \end{bmatrix}$$

The optimal policy is $u_k = K_k x_k$, $k \in \{0, 1, 2, 3, 4\}$.

(ii) Considering a large h in the script ($h = 20$ was tested) we obtain $\bar{K} = K_0 = \begin{bmatrix} -1.1328 - 3.2656 \end{bmatrix}$. The optimal policy is $u_k = Kx_k$

(iii) dlqr, only works if R is positive definite, but as (ii) shows this is not needed for the algebraic Riccati equation to have a solution.

Problem 1.8 This exercise was taken from the sample exam 2 of 2015/2016, question 3, and the solution is already provided.

2. Optimal estimation and output feedback

Problem 2.1 (a)(i) The policy is the same as in 1.1. (certainty equivalence property in the context of linear quadratic control), that is,

$$\begin{aligned}
u_2 &= \mu_2(x_2) = 0\\
u_1 &= \mu_1(x_1) = -\frac{1}{2}x_1\\
u_0 &= \mu_0(x_0) = -\frac{3}{5}x_0
\end{aligned}$$

Since $u_2 = 0$ we only need to compute L_0 and L_1 to obtain $\hat{x}_{0|0}$ and $\hat{x}_{1|1}$ and compute the optimal policy. We get (a)(ii) Let $I_0 = \{y_0\}$, $I_1 = \{u_0, y_0, y_1\}$, $I_2 = \{u_0, y_0, y_1, u_1, y_2\}$. Using the separation principle in the context of output feedback linear quadratic control we have that the optimal policy is obtained by replacing the state by a state estimate in the optimal state feedback policy

$$\begin{aligned}
u_2 &= \mu_2(I_2) = 0\\
u_1 &= \mu_1(I_1) = -\frac{1}{2}\hat{x}_{1|1}\\
u_0 &= \mu_0(I_0) = -\frac{3}{5}\hat{x}_{0|0}
\end{aligned}$$

where we can use the Kalman filter to obtain $\hat{x}_{0|0}$ and $\hat{x}_{1|1}$ since the measurement noise and state disturbances are Gaussian.

To compute the Kalman filter we start by noticing that the initial state follows a Gaussian distribution with mean $\hat{x}_{0|-1} = \bar{x}_0 = 0$ and covariance $\Phi_{0|-1} = \mathbb{E}[(x_0 - \bar{x}_0)^2] = 10$. Then we can obtain the error covariances $\Phi_{0|0}$, $\Phi_{1|0}$, $\Phi_{1|1}$, $\Phi_{2|1}$, … a priori by iterating

$$\begin{aligned}
\Phi_{k+1|k} &= A\Phi_{k|k}A^\intercal + W\\
\Phi_{k|k} &= \Phi_{k|k-1} - \Phi_{k|k-1}C^\intercal(C\Phi_{k|k-1}C^\intercal + V)^{-1}C\Phi_{k|k-1}
\end{aligned}$$

where $W = 1$ and $V = \frac{1}{9}$ are the disturbances and noise covariances, respectively, $A = B = C = 1$, and the Kalman gains from

$$L_k = \Phi_{k|k-1}C^\intercal(C\Phi_{k|k-1}C^\intercal + V)^{-1}.$$

The Kalman filter equations are given by

$$\begin{aligned}
\hat{x}_{k+1|k} &= A\hat{x}_{k|k} + Bu_k\\
\hat{x}_{k|k} &= \hat{x}_{k|k-1} + L_k(y_k - C\hat{x}_{k|k-1})
\end{aligned}$$

$$\begin{aligned}
\Phi_{0|0} &= 10 - \frac{10^2}{10+\frac{1}{9}} = \frac{10}{91}, \quad L_0 = \frac{10}{10+\frac{1}{9}} = \frac{90}{91}\\
\Phi_{1|0} &= \frac{10}{91} + 1 = \frac{101}{91}, L_1 = \frac{\frac{101}{91}}{\frac{101}{91}+\frac{1}{9}} = 0.909
\end{aligned}$$

Thus,

$$\begin{aligned}
\hat{x}_{0|0} &= \frac{90}{91}y_0\\
\hat{x}_{1|0} &= \frac{90}{91}y_0 + u_0\\
\hat{x}_{1|1} &= \hat{x}_{1|0} + L_1(y_1 - C\hat{x}_{1|0}) = \frac{90}{91}y_0 + u_0 + 0.909(y_1 - (\frac{90}{91}y_0 + u_0))
\end{aligned}$$

(b)(i) For the average cost problem in the context of linear quadratic control the policy coincides with the policy for the deterministic infinite horizon policy computed in 1.1, i.e.,

$$u_k = -0.618x_k.$$

(b)(ii) Again using the separation principle the optimal policy for an output feedback average cost problem coincides with the optimal policy derived in (b)(i) by replacing the state x_k by the state estimate $\hat{x}_{k|k}$, i.e., the optimal policy is

$$u_k = -0.618\hat{x}_{k|k}.$$

Since the noise and disturbances in the model are Gaussian this state estimate coincides with the Kalman filter estimate and since we are interest in an average cost (for which the initial transient of the state and state estimates is irrelevant) we can consider the stationary solution of the Kalman filter

$$\begin{aligned}\hat{x}_{k+1|k} &= \hat{x}_{k|k} + u_k \\ \hat{x}_{k|k} &= \hat{x}_{k|k-1} + L(y_k - \hat{x}_{k|k-1})\end{aligned}$$

for $k \in \{0, 1, 2, \dots\}$ where

$$L = \frac{\Phi}{\Phi + \frac{1}{9}},$$

and Φ is the unique positive definite solution to the algebraic Riccati equation

$$\Phi = \Phi - \frac{\Phi^2}{(\Phi + \frac{1}{9}))} + 1 \leftrightarrow \Phi^2 = (\Phi + \frac{1}{9}).$$

This quadratic equation has two solutions 1.1009 and -0.1009 and we take the positive definite one. The corresponding stationary Kalman gains are

$$L = 0.9083.$$

Problem 2.2 (a) We have $x_{k+1} = x_k$, $y_k = x_k + n_k$ which means that the covariance matrix of the disturbances is $W = 0$. Then from the standard Kalman filter equations

$$\begin{aligned}\hat{x}_{k+1|k} &= \hat{x}_{k|k} \\ \hat{x}_{k|k} &= \hat{x}_{k|k-1} + L_k(y_k - \hat{x}_{k|k-1})\end{aligned}$$

where

$$L_k = \frac{\phi_{k|k-1}}{(\phi_{k|k-1} + \sigma_n^2)}$$

and

$$\begin{aligned}\phi_{k+1|k} &= \phi_{k|k} \\ \phi_{k|k} &= \phi_{k|k-1} - \frac{\phi_{k|k-1}^2}{(\phi_{k|k-1} + \sigma_n^2)} = \frac{\sigma_n^2 \phi_{k|k-1}}{(\phi_{k|k-1} + \sigma_n^2)}\end{aligned}$$

To simplify the notation, let $\hat{x}_k := \hat{x}_{k|k}$, $\phi_k = \phi_{k|k-1}$. Note that $\phi_k = \phi_{k|k-1} = \phi_{k|k} = \mathbb{E}[(x_k - \hat{x}_k)^2 | I_k]$. Then,

$$\hat{x}_0 = \bar{x}_0 + \frac{\phi_0}{\phi_0 + \sigma_n^2}(y_0 - \bar{x}_0) = \bar{x}_0 \frac{\sigma_n^2}{\phi_0 + \sigma_n^2} + y_0 \frac{1}{\phi_0 + \sigma_n^2}$$

Note that this corresponds to a weighted average (the weights depend on the ratio between the initial error covariance ϕ_0 and the noise covariance σ_n^2).

$$\begin{aligned}\hat{x}_1 &= \hat{x}_0 + \frac{\phi_1}{\phi_1 + \sigma_n^2}(y_1 - \hat{x}_0) \\ \hat{x}_2 &= \hat{x}_1 + \frac{\phi_2}{\phi_2 + \sigma_n^2}(y_2 - \hat{x}_1)\end{aligned}$$

where

$$\phi_1 = \frac{\sigma_n^2 \phi_0}{(\phi_0 + \sigma_n^2)}$$
$$\phi_2 = \frac{\sigma_n^2 \phi_1}{(\phi_1 + \sigma_n^2)}$$

Note that $\hat{x}_1$ is a linear combination of y_1, y_0 and $\hat{x}_0$ and $\hat{x}_2$ is a linear combination of y_2 y_1, y_0 and $\hat{x}_0$.
(b) $\lim_{k\to\infty} \mathbb{E}[(x_k - \hat{x}_k)^2 | I_k] = \lim_{k\to\infty} \phi_k$ Since $\phi_k > 0$ and $\phi_{k+1} < \phi_k$ we conclude that the limit must be zero. This can be explained as follows: as k increases we have more and more measurements about the same initial value, and therefore more and more certainty about its value which is obtained by computing a linear combination of the measurements (averaging out the measurements).

Problem 2.3 A similar problem is solved in the appendix of Slides II-2 and therefore the solution is omitted.

Problem 2.4

(a)

There are several ways of solving the problem, for example: (i) we can write the cost in terms of u_0, u_1 and find the optimal solution of the unconstrained problem; (ii) we can solve a constrained optimization problem (the constraints are $x_1 = Ax_0 + Bu_0$, $x_2 = Ax_1 + Bu_1$) by writing the Lagrangian; (iii) we can note that this is a standard linear control problem and thus we can find an optimal policy with DP (using Riccati equations) and use this policy to compute an optimal path. Here, we follow (iii) since this will be useful for problem (b). Therefore we have

$$u_0 = K_0 x_0$$
$$u_1 = K_1 x_1$$
$$u_2 = K_2 x_2$$

where,

$$K_k = -(B^\intercal P_{k+1} B + R)^{-1} B^\intercal P_{k+1} A$$

and

$$P_k = A^\intercal P_{k+1} A + Q - A^\intercal P_{k+1} B (B^\intercal P_{k+1} B + R)^{-1} B^\intercal P_{k+1} A$$

for $k \in \{0, 1, 2\}$ and $P_3 = Q_3$, where $Q_3 = 0$ since the terminal cost (absent in the cost function) is $x_3^\intercal Q_3 x_3$ for $Q_3 = 0$. Then,

$$P_3 = 0$$
$$P_2 = I$$
$$P_1 = \begin{bmatrix} 2 & 0.2 \\ 0.2 & 1.1309 \end{bmatrix}$$

and

$$K_2 = 0$$
$$K_1 = \begin{bmatrix} 0 & -\frac{1}{1.1} \end{bmatrix}$$
$$K_0 = \begin{bmatrix} -0.1625 & -0.9513 \end{bmatrix}$$

To find the optimal path we iterate the dynamics and compute the corresponding control inputs with the optimal policy:

$$u_0 = K_0 x_0 = \begin{bmatrix} -0.1625 & -0.9513 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = -1.1137$$
$$x_1 = Ax_0 + Bu_0 = \begin{bmatrix} 1 & \frac{1}{5} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1.2 \\ -0.1137 \end{bmatrix}$$
$$u_1 = K_1 x_1 = 0.1034$$
$$u_2 = 0$$

(b) Due to the certainty equivalence property in the context of linear quadratic control we have that the optimal policy is the same as the one computed in (a) where the disturbances were considered to be identically zero. That is the optimal policy is

$$\begin{aligned} u_0 &= K_0 x_0 \\ u_1 &= K_1 x_1 \\ u_2 &= K_2 x_2 \end{aligned}$$

where

$$\begin{aligned} K_2 &= 0 \\ K_1 &= \begin{bmatrix} 0 & -\frac{1}{1.1} \end{bmatrix} \\ K_0 &= \begin{bmatrix} -0.1625 & -0.9513 \end{bmatrix} \end{aligned}$$

Note that since x_0 is deterministic u_0 is constant and therefore known a priori but x_1 is a random variable which depends on the random variable w_0, and therefore u_1 is not known a priori.

(c) (i) Due to the separation principle in the context of output-feedback linear quadratic control we have

$$\begin{aligned} u_0 &= K_0 \hat{x}_{0|0} \\ u_1 &= K_1 \hat{x}_{1|1} \\ u_2 &= K_2 \hat{x}_{2|2} \end{aligned}$$

where the gains coincide with the ones obtained for the optimal state feedback problem

$$\begin{aligned} K_2 &= 0 \\ K_1 &= \begin{bmatrix} 0 & -\frac{1}{1.1} \end{bmatrix} \\ K_0 &= \begin{bmatrix} -0.1625 & -0.9513 \end{bmatrix} \end{aligned}$$

and the state estimates $\hat{x}_{k|k}$ coincide with the Kalman filter state estimates $\mathbb{E}[x_k|I_k]$ (since the noise and disturbance variables are Gaussian). Here, $\mathbb{E}[x_k|I_k]$ denotes the expected value of the distribution $\mathrm{P}_{x_k|I_k}$.

To compute the Kalman filter we start by noticing that the initial state follows a Gaussian distribution with mean $\hat{x}_{0|-1} = \bar{x}_0 = \begin{bmatrix} 1 & 1 \end{bmatrix}^\intercal$ and zero covariance $\Phi_{0|-1} = \mathbb{E}[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^\intercal] = 0$. Then we can obtain the error covariances $\Phi_{0|0}$, $\Phi_{1|0}$, $\Phi_{1|1}$, $\Phi_{2|1}$, $\ldots$ a priori by iterating

$$\begin{aligned} \Phi_{k+1|k} &= A\Phi_{k|k}A^\intercal + W \\ \Phi_{k|k} &= \Phi_{k|k-1} - \Phi_{k|k-1}C^\intercal(C\Phi_{k|k-1}C^\intercal + V)^{-1}C\Phi_{k|k-1} \end{aligned}$$

where $W = \frac{1}{2}I$ and $V = \frac{1}{4}$ are the disturbances and noise covariances, respectively, and the Kalman gains from

$$L_k = \Phi_{k|k-1}C^\intercal(C\Phi_{k|k-1}C^\intercal + V)^{-1}.$$

The Kalman filter equations are given by

$$\begin{aligned} \hat{x}_{k+1|k} &= A\hat{x}_{k|k} + Bu_k \\ \hat{x}_{k|k} &= \hat{x}_{k|k-1} + L_k(y_k - C\hat{x}_{k|k-1}) \end{aligned}$$

Since $K_2 = 0$ then $u_2 = 0$ and we only need to compute L_0 and L_1 to obtain $\hat{x}_{0|0}$ and $\hat{x}_{1|1}$ and compute the optimal policy. We get

$$\begin{aligned} &\Phi_{0|0} = 0, \quad L_0 = 0 \\ &\Phi_{1|0} = W = \frac{1}{2}I, L_1 = \frac{1}{2}IC^\intercal(C\frac{1}{2}IC^\intercal + \frac{1}{4})^{-1} = \frac{2}{3}\begin{bmatrix} 1 \\ 0 \end{bmatrix} \end{aligned}$$

Since $L_0 = 0$,

$$\hat{x}_{0|0} = \hat{x}_{0|-1} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Moreover,

$$\hat{x}_{1|0} = A\hat{x}_{0|0} + Bu_0 = (A + BK_0)x_0 = \begin{bmatrix} 1 & \frac{1}{5} \\ -0.1625 & 0.0487 \end{bmatrix} x_0 = \begin{bmatrix} 1.2 \\ -0.1137 \end{bmatrix}$$

and

$$\hat{x}_{1|1} = \hat{x}_{1|0} + L_1(y_1 - C\hat{x}_{1|0}) = \begin{bmatrix} 1.2 \\ -0.1137 \end{bmatrix} + \frac{2}{3}\begin{bmatrix} 1 \\ 0 \end{bmatrix}(y_1 - 1.2) = \begin{bmatrix} 1.2 + \frac{2}{3}(y_1 - 1.2) \\ -0.1137 \end{bmatrix}$$

Then

$$\begin{aligned} u_0 &= K_0\hat{x}_{0|0} = \begin{bmatrix} -0.1625 & -0.9513 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = -1.1137 \\ u_1 &= K_1\hat{x}_{1|1} = \begin{bmatrix} 0 & -\frac{1}{1.1} \end{bmatrix}(\begin{bmatrix} 1.2 + \frac{2}{3}(y_1 - 1.2) \\ -0.1137 \end{bmatrix}) = 0.1034 \\ u_2 &= 0 \end{aligned}$$

It is interesting to notice that in this case the policy is to apply a deterministic value for u_0, u_1, u_2 no matter what the measurements y_0, y_1, y_2 are (actually these deterministic values for the *policy* coincide with the ones obtained in (a) for the optimal path without disturbances). This follows from the facts that: (i) the initial state x_0 is deterministic and therefore y_0 is irrelevant and u_0 is constant; (ii) $K_1L_1 = 0$ and therefore u_1 does not depend on y_1, (iii) u_2 is zero since the terminal cost is zero.

(c)(ii) Again due to the separation principle in the context of linear quadratic control, we have that the optimal policy for this average cost optimal control problem is:

$$u_k = K\hat{x}_{k|k}$$

where

$$K = -(B^\intercal PB + R)^{-1}B^\intercal PA$$

and P is the unique positive definite solution to the algebraic Riccati equation

$$P = A^\intercal PA + Q - A^\intercal PB(B^\intercal PB + R)^{-1}B^\intercal PA;$$

and the state estimate $\hat{x}_{k|k}$ is obtained by iterating the stationary Kalman filter

$$\begin{aligned} \hat{x}_{k+1|k} &= A\hat{x}_{k|k} + Bu_k \\ \hat{x}_{k|k} &= \hat{x}_{k|k-1} + L(y_k - C\hat{x}_{k|k-1}) \end{aligned}$$

for $k \in \{0, 1, 2, \dots\}$ where

$$L = \Phi C^\intercal(C\Phi C^\intercal + V)^{-1},$$

and Φ is the unique positive definite solution to the algebraic Riccati equation

$$\Phi = A\Phi A^\intercal - A\Phi C^\intercal(C\Phi C^\intercal + V)^{-1}C\Phi A^\intercal + W.$$

Note that indeed $\hat{x}_{k|k}$ is a function of I_k since it depends only on current and previous outputs y_ℓ, $\ell \in \{0, 1, \dots, k\}$ and previous inputs u_ℓ, $\ell \in \{0, 1, \dots, k-1\}$. With the help of Matlab function dlqr.m and kalman.m we can obtain the solution to these algebraic Riccati equations:

$$\begin{aligned} P &= \begin{bmatrix} 6.6069 & 1.2044 \\ 1.2044 & 1.3506 \end{bmatrix} \\ \Phi &= \begin{bmatrix} 0.8781 & 0.7510 \\ 0.7510 & 3.4230 \end{bmatrix} \\ L &= \begin{bmatrix} 0.7784 \\ 0.6658 \end{bmatrix} \\ K &= \begin{bmatrix} -0.8303 & -1.0971 \end{bmatrix} \end{aligned}$$

(d)(i) We wish to minimize

$$\mathbb{E}[x_0^\intercal Q x_0 + u_0^\intercal R u_0 + x_1^\intercal Q x_1 + u_1^\intercal R u_1 + x_2^\intercal Q x_2 + u_2^\intercal R u_2]$$

when $u_1 = 0$. Applying DP at stages 0 and 2, it is clear that the value of u_2 (for a given realization of I_2) that minimizes the cost to go $\mathbb{E}[u_2^\intercal R u_2 | I_2]$ is $u_2 = 0$. Taking this into account and writing $x_1 = Ax_0 + Bu_0 + w_0$ and $x_2 = Ax_1 + w_1 = A^2 x_0 + ABu_0 + Aw_0 + w_1$ in terms of the only decision variable u_0, the cost to be minimized is

$$\begin{aligned}
&\mathbb{E}[x_0^\intercal Q x_0 + u_0^\intercal R u_0 + (Ax_0 + Bu_0 + w_0)^\intercal Q (Ax_0 + Bu_0 + w_0) \\
&+ (A^2 x_0 + ABu_0 + Aw_0 + w_1)^\intercal Q (A^2 x_0 + ABu_0 + Aw_0 + w_1)] = \\
&\mathbb{E}[x_0^\intercal \bar{Q} x_0 + u_0^\intercal \bar{R} u_0 + x_0^\intercal \bar{S} u_0 + (\bar{A} x_0 + \bar{B} u_0 + \bar{w}_0)^\intercal Q (\bar{A} x_0 + \bar{B} u_0 + \bar{w}_0)] + c
\end{aligned}$$

where c is a given constant and

$$\begin{aligned}
\bar{A} &= A^2 \\
\bar{B} &= AB \\
\bar{Q} &= Q + A^\intercal Q A \\
\bar{R} &= R + B^\intercal Q B \\
\bar{S} &= A^\intercal Q B \\
\bar{w}_0 &= Aw_0 + w_1
\end{aligned}$$

Solving this optimal control problem to find a policy for u_0 is a standard finite-horizon linear quadratic optimal control problem (with disturbances) and the optimal policy is

$$u_0 = \underline{K}_0 \hat{x}_{0|0}$$

where

$$\underline{K}_0 = -(\bar{B}^\intercal Q \bar{B} + \bar{R})^{-1} (\bar{B}^\intercal Q \bar{A} + \bar{S}^\intercal) = \begin{bmatrix} -0.0935 - 0.9720 \end{bmatrix}$$

and $\hat{x}_{0|0} = x_0 = [1\,1]^\intercal$ (as discussed before). Therefore we obtain $u_0 = -1.0654$, i.e., a policy at time $k = 0$ which is independent of the value of y_0.

(d)(ii) Using a similar reasoning to the one used in (d)(i) we have that if we let $\bar{x}_\ell = x_{2\ell}$ and $\bar{u}_\ell = u_{2\ell}$ then

$$\bar{x}_{\ell+1} = \bar{A}\bar{x}_\ell + \bar{B}\bar{u}_\ell + \bar{w}_\ell,$$

and the cost is given by

$$\lim_{\bar{h}\to\infty} \frac{1}{\bar{h}} \mathbb{E}[\sum_{\ell=0}^{\bar{h}} \bar{x}_\ell^\intercal Q \bar{x}_\ell + \bar{x}_\ell^\intercal \bar{S} \bar{u}_\ell + \bar{u}_\ell^\intercal R \bar{u}_\ell]$$

for $\hat{h} = \frac{h}{2}$ for even values of h. The matrices $\bar{A}$, $\bar{B}$, $\bar{Q}$, $\bar{S}$, $\bar{R}$ are as defined in (d)(i) and

$$\bar{w}_\ell = Aw_{2\ell} + w_{2\ell+1}$$

are independent and identically distributed random variables (which are Gaussian since the sum of Gaussian random variables is a Gaussian). We have

$$\mathbb{E}[\bar{w}_\ell \bar{w}_\ell^\intercal] = A\mathbb{E}[w_{2\ell} w_{2\ell}^\intercal] A^\intercal + \underbrace{\mathbb{E}[w_{2\ell} w_{2\ell+1}^\intercal]}_{=0} + \mathbb{E}[w_{2\ell+1} w_{2\ell+1}^\intercal] = AWA^T + W$$

Moreover, let

$$\bar{y}_\ell = \begin{bmatrix} y_{2\ell} \\ y_{2\ell-1} \end{bmatrix}$$

and note that

$$\begin{aligned} y_{2\ell} &= Cx_{2\ell} = C\bar{x}_\ell \\ y_{2\ell-1} &= Cx_{2\ell-1} = CAx_{2(\ell-1)} + CBu_{2(\ell-1)} = CA\bar{x}_{\ell-1} + CB\bar{u}_{\ell-1} \end{aligned}.$$

Then this is a standard average cost optimal control problem with Gaussian noise and disturbances. In particular, appealing again to the separation principle, the solution is

$$u_{2\ell} = \bar{u}_\ell = \bar{K}\hat{\bar{x}}_{\ell|\ell} = \bar{K}\hat{x}_{2\ell|2\ell}$$

where

$$\bar{K} = -(\bar{B}^\intercal \bar{P}\bar{B} + \bar{R})^{-1}(\bar{B}^\intercal \bar{P}\bar{A} + \bar{S}^\intercal)$$

and $\bar{P}$ is the unique positive definite solution to the algebraic Riccati equation

$$\bar{P} = \bar{A}^\intercal \bar{P}_{k+1}\bar{A} + \bar{Q} - (\bar{S} + \bar{A}^\intercal \bar{P}\bar{B})(\bar{B}^\intercal \bar{P}\bar{B} + \bar{R})^{-1}(\bar{B}^\intercal \bar{P}\bar{A} + \bar{S}^\intercal);$$

and the state estimate $\hat{x}_{2\ell|2\ell} = \mathbb{E}[x_{2\ell}|I_{2\ell}]$ is obtained by iterating the stationary Kalman filter

$$\begin{aligned} \hat{x}_{k+1|k} &= A\hat{x}_{k|k} + Bu_k \\ \hat{x}_{k|k} &= \hat{x}_{k|k-1} + L(y_k - C\hat{x}_{k|k-1}) \end{aligned}$$

for $k \in \{0, 1, 2, \dots\}$ where

$$L = \Phi C^\intercal (C\Phi C^\intercal + V)^{-1},$$

and Φ is the unique positive definite solution to the algebraic Riccati equation

$$\Phi = A\Phi A^\intercal - A\Phi C^\intercal (C\Phi C^\intercal + V)^{-1} C\Phi A^\intercal + W.$$

Note that indeed $\hat{x}_{k|k}$ is a function of I_k since it depends only on current and previous outputs y_r, $r \in \{0, 1, \dots, k\}$ and previous inputs u_r, $r \in \{0, 1, \dots, k-1\}$. Note that this Kalman filter has the same structure of the Kalman filter in (c).(ii). However, $u_k = 0$ if k is odd and the state estimate is only used at even times. With the help of Matlab function dlqr.m and kalman.m we can obtain the solution to these algebraic Riccati equations:

$$\begin{aligned} P &= \begin{bmatrix} 6.6274 & 1.2045 \\ 1.2045 & 1.3536 \end{bmatrix} \\ \Phi &= \begin{bmatrix} 0.8781 & 0.7510 \\ 0.7510 & 3.4230 \end{bmatrix} \\ L &= \begin{bmatrix} 0.7784 \\ 0.6658 \end{bmatrix} \\ K &= \begin{bmatrix} -0.7905 & -1.1269 \end{bmatrix} \end{aligned}$$

where L, Φ is the same as in (c).(ii).

Problem 2.5 This exercise was taken from the sample exam 1 of 2015/2016, the solution of this sample exam is already provided.

Problem 2.6 We start by noticing that, if we define

$$J_0(x_0) = (x_0 - \bar{x}_0)^\intercal \bar{\Theta}_0^{-1}(x_0 - \bar{x}_0) + (y_0 - Cx_0)^\intercal V^{-1}(y_0 - Cx_0) \tag{12}$$

we have that

$$J_0(x_0) = (x_0 - \hat{x}_{0|0})^\intercal \Theta_{0|0}^{-1}(x_0 - \hat{x}_{0|0}) + c_0 \tag{13}$$

where c_0 is a constant (in the sense that it does not depend on x_0) and

$$\hat{x}_{0|0} = \bar{x}_0 + L_0(y_0 - C\bar{x}_0),$$

where

$$L_0 = \bar{\Theta}_0 C^\intercal (C\bar{\Theta}_0 C^\intercal + V)^{-1}$$

and

$$\Theta_{0|0} = \bar{\Theta}_0 - \bar{\Theta}_0 C^\intercal (C\bar{\Theta}_0 C^\intercal + V)^{-1} C\bar{\Theta}_0 \tag{14}$$

To prove this, we will need the following identities:

$$\begin{aligned} \Theta_{0|0}^{-1} &= \bar{\Theta}_0^{-1} + C^\intercal V^{-1} C & (15)\\ (I - L_0 C)^\intercal \Theta_{0|0}^{-1} &= \bar{\Theta}_0^{-1} & (16)\\ L_0^\intercal \Theta_{0|0}^{-1} &= V^{-1} C & (17) \end{aligned}$$

The first one can be obtained from the matrix inversion lemma[1]

$$\begin{aligned} \Theta_{0|0}^{-1} &= (\bar{\Theta}_0 - \bar{\Theta}_0 C^\intercal (C\bar{\Theta}_0 C^\intercal + V)^{-1} C\bar{\Theta}_0)^{-1} \\ &= \bar{\Theta}_0^{-1} + C^\intercal (C\bar{\Theta}_0 C^\intercal + V^{-1} - C\bar{\Theta}_0 C^\intercal) C \\ &= \bar{\Theta}_0^{-1} + C^\intercal V^{-1} C \end{aligned} \tag{18}$$

the second one is equivalent to $\bar{\Theta}_0 (I - L_0 C)^\intercal = \Theta_{0|0}$, which can be concluded directly from (14), and the third one follows from

$$\begin{aligned} V^{-1} C\Theta_{0|0} &= V^{-1} C(\bar{\Theta}_0 - \bar{\Theta}_0 C^\intercal (C\bar{\Theta}_0 C^\intercal + V)^{-1} C\bar{\Theta}_0) \\ &= (V^{-1} - V^{-1}\bar{\Theta}_0 C^\intercal (C\bar{\Theta}_0 C^\intercal + V)^{-1}) C\bar{\Theta}_0 \\ &= \underbrace{(V^{-1} - V^{-1}\bar{\Theta}_0 C^\intercal (C\bar{\Theta}_0 C^\intercal + V)^{-1})(C\bar{\Theta}_0 C^\intercal + V)}_{=I} (C\bar{\Theta}_0 C^\intercal + V)^{-1} C\bar{\Theta}_0 = L_0^\intercal \end{aligned}$$

Then, using these identities we obtain

$$\begin{aligned} (x_0 - \hat{x}_{0|0})^\intercal \Theta_{0|0}^{-1} (x_0 - \hat{x}_{0|0}) &= (x_0 - ((I - L_0 C)\bar{x}_0 + L_0 y_0))^\intercal \Theta_{0|0}^{-1} (x_0 - ((I - L_0 C)\bar{x}_0 + L_0 y_0)) \\ &= x_0^\intercal \underbrace{\Theta_{0|0}^{-1}}_{(\bar{\Theta}_0^{-1} + C^\intercal V^{-1} C)} x_0 - 2\bar{x}_0^\intercal \underbrace{(I - L_0 C)^\intercal \Theta_{0|0}^{-1}}_{\bar{\Theta}_0^{-1}} x_0 - 2y_0^\intercal \underbrace{L_0^\intercal \Theta_{0|0}^{-1}}_{V^{-1} C} x_0 + \bar{c}_0 \\ &= (x_0 - \bar{x}_0)^\intercal \bar{\Theta}_0^{-1} (x_0 - \bar{x}_0) + (y_0 - C x_0)^\intercal V^{-1} (y_0 - C x_0) - c_0 \end{aligned}$$

where c_0 and $\bar{c}_0$ do not depend on x_0.

If we define

$$\tilde{x}_k = x_{h-k} \quad \tilde{u}_k = \hat{w}_{h-k} \quad \tilde{A} = A^{-1} \quad \tilde{B} = -A^{-1} \quad \tilde{R} = W^{-1} \quad \tilde{Q} = V^{-1} \quad \tilde{Q}_h = \Theta_{0|0} \quad \tilde{r}_k = y_{h-k}$$

we obtain a standard tracking control problem

$$\min_{\tilde{u}_0, \ldots, \tilde{u}_{h-1}} \sum_{k=0}^{h-1} (\tilde{r}_k - C\tilde{x}_k)^\intercal \tilde{Q} (\tilde{r}_k - C\tilde{x}_k) + (x_{0|0} - \tilde{x}_h)^\intercal Q_h (x_{0|0} - \tilde{x}_h)$$

subject to

$$\tilde{x}_{k+1} = \tilde{A}\tilde{x}_k + \tilde{B}\tilde{u}_k$$

However, we will work with the original coordinates, but this observation justifies that we can solve the problem by running the cost-to-come equations

$$J_{k+1}(x_{k+1}) = \min_{\bar{w}_{k+1}} (y_{k+1} - C x_{k+1})^\intercal V^{-1} (y_{k+1} - C x_{k+1}) + \bar{w}_{k+1}^\intercal W^{-1} \bar{w}_{k+1} + J_k(A^{-1} x_{k+1} - A^{-1} \bar{w}_{k+1}) \tag{19}$$

[1] Matrix inversion lemma: $(A + UCV)^{-1} = A^{-1} - A^{-1} U (C^{-1} + V A^{-1} U)^{-1} V A^{-1}$

for $k \in \{0, 1, \ldots, h-1\}$.

Given the measurements $y_0,\ y_1,\ \ldots$ and the initial filter estimate x_0 we can define the Kalman filter estimates, for $k \in \{0, 1, \ldots, h\}$,

$$\begin{aligned}\hat{x}_{k+1|k} &= A\hat{x}_{k|k}\\ \hat{x}_{k|k} &= \hat{x}_{k|k-1} + L_k(y_k - C\hat{x}_{k|k-1})\end{aligned}$$

where

$$\begin{aligned}L_k &= \Theta_{k|k-1}C^\intercal(C\Theta_{k|k-1}C^\intercal + V)^{-1}\\ \Theta_{k+1|k} &= A\Theta_{k|k}A^\intercal + W\\ \Theta_{k|k} &= \Theta_{k|k-1} - \Theta_{k|k-1}C^\intercal(C\Theta_{k|k-1}C^\intercal + V)^{-1}C\Theta_{k|k-1},\end{aligned}$$

which will be important to prove the result.

Now assuming that, for a given $k \in \{0, \ldots, h\}$

$$J_k(x_k) = (x_k - \hat{x}_{k|k})^\intercal\Theta_{k|k}^{-1}(x_k - \hat{x}_{k|k}) + c_k \tag{20}$$

where c_k is a constant, if we manage to prove that

$$J_{k+1}(x_{k+1}) = (x_{k+1} - \hat{x}_{k+1|k+1})^\intercal\Theta_{k+1|k+1}^{-1}(x_{k+1} - \hat{x}_{k+1|k+1}) + c_{k+1} \tag{21}$$

we will obtain by induction that

$$J_h(x_h) = (x_h - \hat{x}_{h|h})^\intercal\Theta_{h|h}^{-1}(x_h - \hat{x}_{h|h}) + c_h$$

from which the desired result follows

$$\bar{x}_h = \text{argmin} J_h(x_h) = \hat{x}_{h|h}$$

Plugging (20) into (19) we obtain

$$\begin{aligned}&\min_{\bar{w}_{k+1}}\quad (y_{k+1} - Cx_{k+1})^\intercal V^{-1}(y_{k+1} - Cx_{k+1}) + \bar{w}_{k+1}^\intercal W^{-1}\bar{w}_{k+1}\\ &\qquad\qquad + (A^{-1}x_{k+1} - A^{-1}\bar{w}_{k+1} - \hat{x}_{k|k})^\intercal\Theta_{k|k}^{-1}(A^{-1}x_{k+1} - A^{-1}\bar{w}_{k+1} - \hat{x}_{k|k}))\\ &= (y_{k+1} - Cx_{k+1})^\intercal V^{-1}(y_{k+1} - Cx_{k+1}) +\\ &\min_{\bar{w}_{k+1}} \bar{w}_{k+1}^\intercal W^{-1}\bar{w}_{k+1} + (A^{-1}(x_{k+1} - \hat{x}_{k+1|k}) - A^{-1}\bar{w}_{k+1})^\intercal\Theta_{k|k}^{-1}(A^{-1}(x_{k+1} - \hat{x}_{k+1|k}) - A^{-1}\bar{w}_{k+1})\\ &= \min_{\bar{w}_{k+1}} \bar{w}_{k+1}^\intercal(W^{-1} + A^{-\intercal}\Theta_{k|k}^{-1}A^{-1})\bar{w}_{k+1} - 2\bar{w}_{k+1}^\intercal A^{-\intercal}\Theta_{k|k}^{-1}A^{-1}(x_{k+1} - \hat{x}_{k+1|k}) + d\end{aligned}$$

where $A^{-\intercal} := (A^{-1})^\intercal$ and

$$d := (y_{k+1} - Cx_{k+1})^\intercal V^{-1}(y_{k+1} - Cx_{k+1}) + (x_{k+1} - \hat{x}_{k+1|k})^\intercal A^{-\intercal}\Theta_{k|k}^{-1}A^{-1}(x_{k+1} - \hat{x}_{k+1|k})$$

The minimizer is given by

$$\bar{w}_{k+1} = (W^{-1} + A^{-\intercal}\Theta_{k|k}^{-1}A^{-1})^{-1}A^{-\intercal}\Theta_{k|k}^{-1}A^{-1}(x_{k+1} - \hat{x}_{k+1|k})$$

leading to

$$\begin{aligned}&J_{k+1}(x_{k+1}) = (y_{k+1} - Cx_{k+1})^\intercal V^{-1}(y_{k+1} - Cx_{k+1})\\ &\quad + (x_{k+1} - \hat{x}_{k+1|k})^\intercal(A^{-\intercal}\Theta_{k|k}^{-1}A^{-1} - A^{-\intercal}\Theta_{k|k}^{-1}A^{-1}(W^{-1} + A^{-\intercal}\Theta_{k|k}^{-1}A^{-1})^{-1}A^{-\intercal}\Theta_{k|k}^{-1}A^{-1})(x_{k+1} - \hat{x}_{k+1|k})\end{aligned}$$

The second term is given by

$$(x_{k+1} - \hat{x}_{k+1|k})^\intercal\Theta_{k+1|k}^{-1}(x_{k+1} - \hat{x}_{k+1|k})$$

which can be concluded from

$$\begin{aligned}
\Theta_{k+1|k}^{-1} &= (A\Theta_{k|k}A^\intercal + W)^{-1} \\
&= A^{-\intercal}(\Theta_{k|k} + A^{-1}WA^{-\intercal})^{-1}A^{-1} \\
&= A^{-\intercal}(\Theta_{k|k}^{-1} - \Theta_{k|k}^{-1}A^{-1}(W^{-1} + A^{-\intercal}\Theta_{k|k}^{-1}A^{-1})^{-1}A^{-\intercal}\Theta_{k|k}^{-1})A^{-1} \\
&= (A^{-\intercal}\Theta_{k|k}^{-1}A^{-1} - A^{-\intercal}\Theta_{k|k}^{-1}A^{-1}(W^{-1} + A^{-\intercal}\Theta_{k|k}^{-1}A^{-1})^{-1}A^{-\intercal}\Theta_{k|k}^{-1}A^{-1})
\end{aligned}$$

where in the third equality we used the matrix inversion lemma. Then

$$\begin{aligned}
J_{k+1}(x_{k+1}) = (y_{k+1} - Cx_{k+1})^\intercal V^{-1}(y_{k+1} - Cx_{k+1}) \\
+ (x_{k+1} - \hat{x}_{k+1|k})^\intercal \Theta_{k+1|k}^{-1}(x_{k+1} - \hat{x}_{k+1|k})
\end{aligned}$$

Using similar arguments to the ones used to prove that (12) equals (12) we can conclude that

$$J_{k+1}(x_{k+1}) = (x_{k+1} - \hat{x}_{k+1|k+1})^\intercal \Theta_{k+1|k+1}^{-1}(x_{k+1} - \hat{x}_{k+1|k+1}) + c_{k+1}$$

which is (21) and concludes the proof.

3. Discretization and static optimization

Problem 3.1

(i) $A_d = 1.1052$, $B_d = 0.2103$.

(ii) $A_d = \begin{bmatrix} 1.0513 & 0.3076 \\ 0 & 1.0000 \end{bmatrix}, \quad B_d = \begin{bmatrix} 0.3097 \\ 0.3500 \end{bmatrix}$

(iii) $A_d = \begin{bmatrix} 0.6967 & -0.7174 \\ 0.7174 & 0.6967 \end{bmatrix}, \quad B_d = \begin{bmatrix} -0.0758 \\ 0.1793 \end{bmatrix}$

(iv) $A_d = \begin{bmatrix} 7.384 & 7.384 \\ 0 & 7.384 \end{bmatrix}, \quad B_d = \begin{bmatrix} 2.0973 \\ 3.1945 \end{bmatrix}$

(v) $A_d = \begin{bmatrix} 2.7183 & 0 & 0 \\ 0 & 0.3679 & 0 \\ 0 & 0 & 1.0000 \end{bmatrix}, \quad B_d = \begin{bmatrix} 0.8591 \\ 0.6321 \\ 0.5000 \end{bmatrix}$

Problem 3.2

(i) The optimal policy obtained after discretization is depicted in Figure 2.

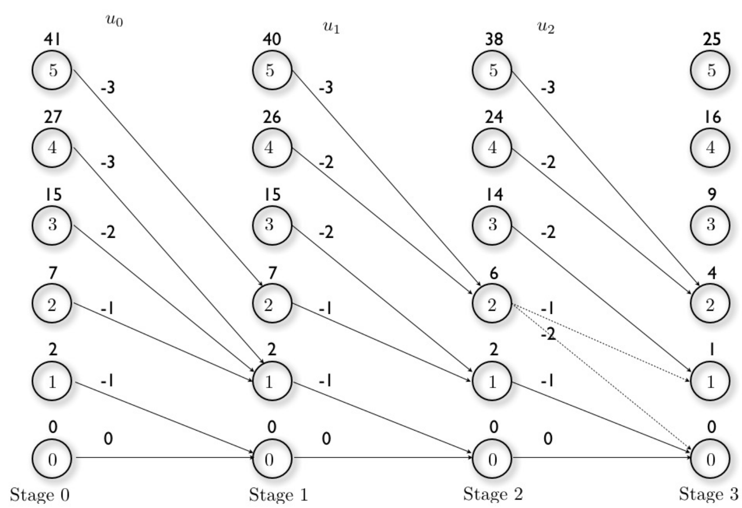

Figure 2: Optimal policy, problem 2.1

(ii) The costs for each possible initial condition coincide with the cost to go at stage 0.

4. Static optimization formulation of optimal control problems

Problem 4.1

(i)

$$L(x,\lambda) = x^\top X x + \lambda(c^\top x - 1) \tag{22}$$

$$0 = \frac{\partial}{\partial x} L = 2Xx + \lambda c \rightarrow x = -\frac{\lambda}{2} X^{-1} c \tag{23}$$

$$0 = \frac{\partial}{\partial \lambda} L = c^\top x - 1 \rightarrow 0 = -\frac{\lambda}{2} c^\top X^{-1} c - 1 \rightarrow \lambda = -2 \frac{1}{c^\top X^{-1} c} \tag{24}$$

$$\rightarrow x = \frac{1}{c^\top X^{-1} c} X^{-1} c \tag{25}$$

For $n = 2$ the cost is a parabola and the constraint is a line.

$x = [0.2973 \; 0.3514]^\top$ and $J = 0.2162$

(ii)

$$L(x,\lambda) = c^\top x + \lambda(x^\top X x - 1) \tag{26}$$

$$0 = \frac{\partial}{\partial x} L = 2\lambda X x + c \rightarrow x = -\frac{1}{2\lambda} X^{-1} c \tag{27}$$

$$0 = \frac{\partial}{\partial \lambda} L = x^\top X x - 1 \rightarrow 0 = (\frac{1}{2\lambda} X^{-1} c)^\top X \frac{1}{2\lambda} X^{-1} c - 1 \rightarrow \lambda^2 = \frac{1}{4} c^\top X^{-1} c \tag{28}$$

$$\rightarrow x = \pm \frac{1}{\sqrt{c^\top X^{-1} c}} X^{-1} c \tag{29}$$

For $n = 2$ the cost is a line and the constraint is an ellipsoid around the origin.

$x = [-0.2973 \;\; -0.3514]^\top$ and $J = -1$

(iii)

$$L(x,\lambda) = x^\top X x + \lambda(x^\top x - 1) \tag{30}$$

$$0 = \frac{\partial}{\partial x} L = 2Xx + \lambda 2x \rightarrow (X + \lambda I)x = 0 \tag{31}$$

$$\rightarrow \text{thus the } x \text{ that satisfy this condition are the eigenvectors } v \text{ of X, and } \lambda \text{ are the eigenvalues} \tag{32}$$

$$0 = \frac{\partial}{\partial \lambda} L = x^\top x - 1 \quad \text{the normalized eigenvectors satisfy this condition} \tag{33}$$

$$\rightarrow x^* = \frac{v}{|v|} \tag{34}$$

For $n = 2$ the cost is a parabola and the constraint is the unit circle.

$x = [-0.7071 \;\; -0.7071]^\top$ and $J = 1$

Problem 4.2

(i)

$$\left(\sum_{k=0}^{1} x_k^2 + u_k^2\right) + x_2^2 = x_0^2 + u_0^2 + x_1^2 + u_1^2 + \left(\frac{1}{2}(x_1 + u_1)\right)^2 \tag{35}$$

$$= x_0^2 + u_0^2 + \left(\frac{1}{2}(x_0 + u_0)\right)^2 + u_1^2 + \left(\frac{1}{2}\left(\frac{1}{2}(x_0 + u_0) + u_1\right)\right)^2 \tag{36}$$

$$= \left(1 + \frac{1}{4} + \frac{1}{16}\right)x_0^2 + \left(\frac{1}{2} + \frac{1}{8}\right)x_0u_0 + \left(1 + \frac{1}{4} + \frac{1}{16}\right)u_0^2 + \left(1 + \frac{1}{4}\right)u_1^2 + \left(\frac{1}{4}\right)x_0u_1 + \left(\frac{1}{4}\right)u_0u_1 \tag{37}$$

$$= [u_0\ u_1]A[u_0\ u_1]^\top + B[u_0\ u_1]^\top + c \tag{38}$$

with $A = [21/16\ 1/8; 1/8\ 5/4]$, $B = [5/8\ 1/4]$, $c = 21/16$, optimum $[u_0^*\ u_1^*]$

$$2A[u_0^*\ u_1^*]^\top + B^\top = 0 \tag{39}$$

which gives two linear equations for two unknowns, which can be solved easily.
$u^* = [-3/13\ \ -1/13]^\top$.

(ii)

$$\left(\sum_{k=0}^{1} x_k^2 + u_k^2\right) + x_2^2 = x_0^2 + u_0^2 + x_1^2 + u_1^2 + x_2^2 \tag{40}$$

$$L = x_0^2 + u_0^2 + x_1^2 + u_1^2 + x_2^2 + \lambda_1\left(x_2 - \frac{1}{2}(x_1 + u_1)\right) + \lambda_2\left(x_1 - \frac{1}{2}(x_0 + u_0)\right) \tag{41}$$

$$0 = \frac{\partial}{\partial x_2}L = 2x_2 + \lambda_1 \tag{42}$$

$$0 = \frac{\partial}{\partial x_1}L = 2x_1 - \frac{1}{2}\lambda_1 + \lambda_2 \tag{43}$$

$$0 = \frac{\partial}{\partial \lambda_2}L = x_1 - \frac{1}{2}(x_0 + u_0) \tag{44}$$

$$0 = \frac{\partial}{\partial \lambda_1}L = x_2 - \frac{1}{2}(x_1 + u_1) \tag{45}$$

$$0 = \frac{\partial}{\partial u_1}L = 2u_1 - \frac{1}{2}\lambda_1 \tag{46}$$

$$0 = \frac{\partial}{\partial u_0}L = 2u_0 - \frac{1}{2}\lambda_2 \tag{47}$$

which gives 6 linear equations for 6 unknowns (x_0 is given), which can be solved easily.
$u^* = [-3/13\ \ -1/13]^\top$.

Problem 4.3 (i)

$$\begin{aligned}
x_1 &= \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} x_0 + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u_0 \\
x_2 &= \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} x_1 + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u_1 \\
&= \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} x_0 + \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} u_0 + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u_1 \\
&= \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u_0 + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u_1
\end{aligned}$$

Cost function

$$u_0^2 + u_1^2 + \|x_2\|^2$$

$$\begin{aligned} x_2^\intercal x_2 &= (\begin{bmatrix}1\\2\end{bmatrix} + \begin{bmatrix}0\\1\end{bmatrix} u_0 + \begin{bmatrix}1\\0\end{bmatrix} u_1)^\intercal (\begin{bmatrix}1\\2\end{bmatrix} + \begin{bmatrix}0\\1\end{bmatrix} u_0 + \begin{bmatrix}1\\0\end{bmatrix} u_1) \\ &= 5 + 2u_0 + u_1 + 2u_0 + u_0^2 + u_1 + u_1^2 \\ &= 5 + \begin{bmatrix}4 & 2\end{bmatrix}\begin{bmatrix}u_0\\u_1\end{bmatrix} + \begin{bmatrix}u_0 & u_1\end{bmatrix}\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}\begin{bmatrix}u_0\\u_1\end{bmatrix} \end{aligned}$$

Therefore

$$u_0^2 + u_1^2 + \|x_2\|^2 = \underbrace{5}_{z} + \underbrace{\begin{bmatrix}4 & 2\end{bmatrix}}_{Y}\begin{bmatrix}u_0\\u_1\end{bmatrix} + \begin{bmatrix}u_0 & u_1\end{bmatrix}\underbrace{\begin{bmatrix}2 & 0\\0 & 2\end{bmatrix}}_{X}\begin{bmatrix}u_0\\u_1\end{bmatrix}$$

(ii)

$$f(u) = u^\intercal X u + Y^\intercal u + z$$

$$\begin{aligned} \nabla f(u) = 0 &\equiv 2Xu + Y = 0 \\ &\equiv u = -\frac{1}{2}\begin{bmatrix}2 & 0\\0 & 2\end{bmatrix}^{-1}\begin{bmatrix}4\\2\end{bmatrix} = \begin{bmatrix}-1\\-\frac{1}{2}\end{bmatrix} \end{aligned}$$

(it is a minimum because $X > 0$).

Problem 4.4

(i) For J the given quadratic cost, the Lagrangian is given by

$$L = J + \sum_{k=0}^{h-1} \lambda_{k+1}^\top ((Ax_k + Bu_k) - x_{k+1}) \tag{48}$$

$$0 = \frac{\partial}{\partial x_h} L = Q_h x_h - \lambda_h \tag{49}$$

$$0 = \frac{\partial}{\partial x_k} L = Qx_k + A^\top \lambda_k - \lambda_{k+1} \tag{50}$$

$$0 = \frac{\partial}{\partial \lambda_k} L = (Ax_k + Bu_k) - x_{k+1} \tag{51}$$

$$0 = \frac{\partial}{\partial u_k} L = Ru_k + B^\top \lambda_{k+1} \tag{52}$$

Note that λ_k has multiple elements.

(ii)

$$P_h x_h = Q_h x_h \to P_h = Q_h \tag{53}$$

$$Ru_k = -B^\top P_{k+1} x_{k+1} = -B^\top P_{k+1}(Ax_k + Bu_k) \to Ru_k + B^\top P_{k+1} B u_k = -B^\top P_{k+1} A x_k \tag{54}$$

$$\to u_k = -(R + B^\top P_{k+1} B)^{-1} B^\top P_{k+1} A x_k \tag{55}$$

$$P_k x_k = A^\top P_{k+1}(Ax_k + Bu_k) + Qx_k = (A^\top P_{k+1} A + Q)x_k - A^\top P_{k+1} B (R + B^\top P_{k+1} B)^{-1} B^\top P_{k+1} A x_k \tag{56}$$

$$\to P_k = (A^\top P_{k+1} A + Q) - A^\top P_{k+1} B (R + B^\top P_{k+1} B)^{-1} B^\top P_{k+1} A \tag{57}$$

5. Approximate dynamic programming

Problem 5.1 (i) Optimal policy.

This is a standard stage-decision problem with $g_0, g_1 = 0$, $g_2(x_2)$, $x_2 = A_{\sigma_1}x_1$, and $x_1 = A_{\sigma_1}x_0$. Applying DP we have

$$\begin{aligned}
J_2(x_2) &= x_2^\intercal x_2\\
J_1(x_1) &= \min_{\sigma_1\in\{0,1\}} g_1(x_1,\sigma_1) + J_2(x_2) = J_2(A_{\sigma_1}x_1)\\
&= \min_{\sigma_1\in\{0,1\}} x_1^\intercal A_{\sigma_1}^\intercal A_{\sigma_1} x_1\\
&= \min\{x_1^\intercal \begin{bmatrix}0.8 & 0.2\\ 0.1 & 0.8\end{bmatrix}\begin{bmatrix}0.8 & 0.1\\ 0.2 & 0.8\end{bmatrix}x_1, x_1^\intercal\begin{bmatrix}1 & 0\\ 0.1 & 0.5\end{bmatrix}\begin{bmatrix}1 & 0.1\\ 0 & 0.5\end{bmatrix}x_1\}\\
&= \min\{x_1^\intercal\begin{bmatrix}0.68 & 0.24\\ 0.24 & 0.65\end{bmatrix}x_1, x_1^\intercal\begin{bmatrix}1 & 0.1\\ 0.1 & 0.26\end{bmatrix}x_1\}
\end{aligned}$$

Optimal policy at stage $k = 1$

$$\sigma_1 = \begin{cases} 0 \text{ if } x_1^\intercal\begin{bmatrix}0.68 & 0.24\\ 0.24 & 0.65\end{bmatrix}x_1 < x_1^\intercal\begin{bmatrix}1 & 0.1\\ 0.1 & 0.26\end{bmatrix}x_1\\ 1 \text{ otherwise}\end{cases}$$

Note that the optimal policy is not unique -in particular we can replace $<$ by $\leq$ in the expression above. Note also that this leads to a conic state partition

$$\sigma_1 = \begin{cases} 0 \text{ if } x_1^\intercal \underbrace{(\begin{bmatrix}0.68 & 0.24\\ 0.24 & 0.65\end{bmatrix} - \begin{bmatrix}1 & 0.1\\ 0.1 & 0.26\end{bmatrix})}_{\begin{bmatrix}-0.32 & 0.14\\ 0.14 & 0.39\end{bmatrix}} x_1 < 0\\ 1 \text{ otherwise}\end{cases}$$

In fact, making $x_1 = (y,z)$ we have that the states for which $\sigma_0 = 0$ are described by

$$\{(y,z)\in\mathbb{R}^2 : -0.32y^2 + 0.28zy + 0.39z^2 < 0\}$$

To find the boundary of this region we set $y = \alpha z$ and obtain $-0.32\alpha^2 + 0.28\alpha + 0.39 = 0 \to \alpha = 1.625$, $\alpha = -0.75$. The policy can then be graphically represented as in Figure 3.

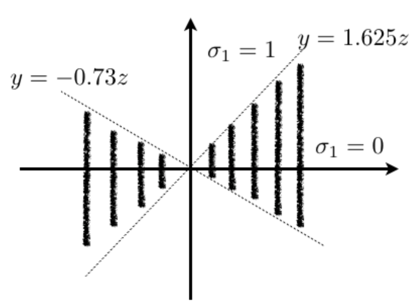

Figure 3: Policy

Optimal policy at stage $k = 0$

$$J_1(x_1) = \min\{\|A_0x_1\|^2, \|A_1x_1\|^2\}$$

$$\begin{aligned}
J_0(x_0) &= \min_{\sigma_0\in\{0,1\}} \{\|A_0A_{\sigma_0}x_0\|^2, \|A_1A_{\sigma_0}x_0\|^2\} \\
&= \min\{\|A_0A_0x_0\|^2, \|A_0A_1x_0\|^2, \|A_1A_0x_0\|^2, \|A_1A_1x_0\|^2 \\
&= \min\{x_0^\intercal Z_1x_0, x_0^\intercal Z_2x_0, x_0^\intercal Z_3x_0, x_0^\intercal Z_4x_0\}
\end{aligned}$$

$$\begin{aligned}
Z_1 &= (A_0A_0)^\intercal A_0A_0 \\
Z_2 &= (A_0A_1)^\intercal A_0A_1 \\
Z_3 &= (A_1A_0)^\intercal A_1A_0 \\
Z_4 &= (A_1A_1)^\intercal A_1A_1
\end{aligned}$$

Optimal policy for σ_0

$$\sigma_0 = \begin{cases} 0 \text{ if } x_0^\intercal Z_1x_0 < x_0^\intercal Z_ix_0 \text{ for } i\in\{2,3,4\} \text{ of if } x_0^\intercal Z_3x_0 < x_0^\intercal Z_ix_0 \text{ for } i\in\{1,2,4\} \\ 1 \text{ otherwise.} \end{cases}$$

(ii) Rollout policy with horizon $H = 1$
stage 0

$$\min_{\sigma_k:k\in\{0,\dots,H-1\}} \sum_{k=0}^{H-1} g_k(x_k,\sigma_k) + \sum_{k=H}^{h} g(x_k, \underbrace{1}_{basepolicy}) + g_h(x_h)$$

$$x_k = A_1A_1A_1\dots A_1A_{\sigma_{H-1}}\dots A_{\sigma_0}x_0, \quad k > H-1.$$

For $H = 1$ and cost function $g_3(x_3) = \|x_3\|^2$ we have

$$x_3 = A_1A_1A_{x_0}x_0$$

and the approximate cost-to-go (to be minimize by the ADP policy) is

$$\min_{\sigma_0\in\{0,1\}} \|A_1A_1A_{\sigma_0}x_0\|^2$$

Thus,

$$\sigma_0 = \begin{cases} 0 \text{ if } \|A_1A_1A_0x_0\|^2 < \|A_1A_1A_1x_0\|^2 \\ 1 \text{ otherwise.} \end{cases}$$

Using a similar reasoning

$$\sigma_1 = \begin{cases} 0 \text{ if } \|A_1A_0x_1\|^2 < \|A_1A_1x_1\|^2 \\ 1 \text{ otherwise.} \end{cases}$$

$$\sigma_2 = \begin{cases} 0 \text{ if } \|A_0x_2\|^2 < \|A_1x_2\|^2 \\ 1 \text{ otherwise.} \end{cases}$$

Problem 5.2 (i) MPC policy. At time ℓ,

$$\min_{\theta_\ell,\theta_{\ell+1}} \|x_\ell\|_Q^2 + \|x_{\ell+1}\|_Q^2$$

$$\min_{\theta_\ell} \|x_\ell\|_Q^2 + \|x_\ell + \tau \begin{bmatrix}\cos(\theta_\ell)\\ \sin(\theta_\ell)\end{bmatrix} \|$$

which is equivalent to minimizing

$$\min_{\theta_\ell} \|x_\ell + \tau \begin{bmatrix} \cos(\theta_\ell) \\ \sin(\theta_\ell) \end{bmatrix} \|$$

or

$$\min_{v_1, v_2} \|x_\ell + \tau \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \|$$

where

$$\{(v_1, v_2) = (\cos(\theta), \sin(\theta)), \theta \in [-\pi, \pi)\} \equiv \{(v_1, v_2) \in \mathbb{R}^2 | \|v\|^2 = 1\}.$$

Thus, the optimization problem is equivalent to

$$\min_{v \in \mathbb{R}^2} \|x_\ell + \tau v\|^2$$

such that $\|v\|^2 = 1$ and $\cos(\theta) = v_1$, $\sin(\theta) = v_2$. This implies that

$$\theta = \begin{cases} \arctan \frac{v_2}{v_1} \text{ if } v_1 \neq 0 \\ \frac{\pi}{2} \text{ if } v_1 = 0 \& v_2 > 0 \\ -\frac{\pi}{2} \text{ if } v_1 = 0 \& v_2 < 0 \end{cases}.$$

(ii) Need to solve the constrained optimization problem

$$\min_{v \in \mathbb{R}^2} \|x_0 + \tau v\|_Q^2$$
$$\text{s.t.} \|v\|^2 = 1$$

when

$$\tau = 0.1, Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad x_0 = \begin{bmatrix} 1 & 1 \end{bmatrix}^\intercal$$

$$\min_{v \in \mathbb{R}^2} \| \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 0.1 \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \|^2 = (1 + 0.1v_1)^2 + (1 + 0.1v_2)^2$$

s.t. $v_1^2 + v_2^2 = 1$

Lagrangian

$$L(v, \lambda) = (1 + 0.1v_1)^2 + (1 + 0.1v_2)^2 + \lambda(v_1^2 + v_2^2 - 1)$$

$$\frac{\partial L}{\partial v_1} = 0 \to 2(1 + 0.1v_1) + 2\lambda v_1 = 0 \to v_1 = \frac{-1}{\lambda + 0.1}$$
$$\frac{\partial L}{\partial v_2} = 0 \to 2(1 + 0.1v_2) + 2\lambda v_2 = 0 \to v_2 = \frac{-1}{\lambda + 0.1}$$

$$v_1^2 + v_2^1 = 1 \to (\frac{-1}{\lambda + 0.1})^2 + (\frac{-1}{\lambda + 0.1})^2 - 1 = 0 \to \lambda = -0.1 \pm \sqrt{2}$$

Therefore, we have

$$(v_1, v_2) = (-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}) \to \theta = \frac{5\pi}{4} (minimum)$$
$$(v_1, v_2) = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}) \to \theta = \frac{\pi}{4} (maximum)$$

Problem 5.3

(i) (a) From the figures, we can see that by taking $\sigma = 1$ if $yz < 0$ and $\sigma = 2$ otherwise, the system converges.

(b) The same switching condition as in the CT case can be applied.

(ii) (a)

$$\sigma_k = \text{argmin} x_k^\intercal (I + A_{\sigma_k}^\intercal A_{\sigma_k}) x_k$$

or $\sigma_k = 1$ if $x_k^\intercal (A_1^\intercal A_1 - A_2^\intercal A_2) x_k < 0$. Plot the lines that separate the regions and see from that whether the MPC law is stabilizing. MPC does not guarantee stability of the system in general.

(b) Use the result from Problem 1.3 with $Q = I$ on x_1:

$$\sum_{k=0}^{\infty} x_k^\intercal Q x_k = x_0^\intercal Q x_0 + \sum_{k=1}^{\infty} x_k^\intercal Q x_k = x_0^\intercal Q x_0 + x_1^\intercal P x_1$$

where P is the solution to the infinite horizon Lyapunov equation $P = \text{dlyap}(A_1^\intercal, Q)$;

$$\sigma_0 = \text{argmin} x_0^\intercal Q x_0 + x_1^\intercal P x_1 = \text{argmin} x_0^\intercal Q x_0 + x_0^\intercal A_{\sigma_0}^\intercal P A_{\sigma_0} x_0.$$

For $\beta = 1$ there is no positive definite solution for P, hence the base policy is not stabilizing and cannot be applied.

Problem 5.4 Solution not provided.