Problem set 2

Stage decision problems

Optimal control and reinforcement learning, TU/e, 2022-2023

Outline

Linear quadratic control

Optimal estimation and output feedback

Discretization

Static optimization formulation of optimal control problems

Approximate dynamic programming

Linear quadratic control

Problem 1.1 Consider a first order discrete-time linear system

$$x_{k+1} = \alpha x_k + u_k, \quad k \in \mathbb{N}_0,$$

and the cost

$$\sum_{k=0}^{q} x_k^2 + \gamma u_k^2.$$

(a) Considering $\alpha = 1$, $\gamma = 1$ and $q = 2$, determine the policy for u_0, u_1, and u_2 that minimizes the cost.

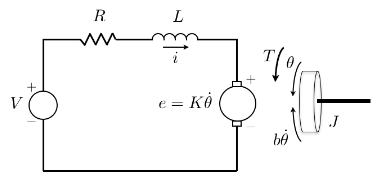
(b) Considering $\alpha = 1$, $\gamma = 1$ and $q = \infty$, determine the policy for $\{u_k, k \in \mathbb{N}_0\}$ that minimizes the cost.

(c) Suppose now that $|\alpha| < 1$ and $q = \infty$. Provide and interpret the policy that minimizes the cost when $\gamma \to \infty$.

Problem 1.2 A schematic model of a DC motor is depicted in the figure. The torque generated by the motor is assumed to be proportional to the armature current $T = K_t i$ and the back emf, e, is assumed to be proportional to the angular velocity of the shaft by a constant factor $e = K_e \dot{\theta}$. From the figure we can derive the following equations based on Newton's law and Kirchhoff's voltage law

$$J\ddot{\theta} + b\dot{\theta} = K_t i$$

$$L\frac{di}{dt} + Ri = V - K_e\dot{\theta}$$



Assuming that the inductance is very small $L \approx 0$, we can write

$$J\ddot{\theta} + (b + \frac{K_t K_e}{R})\dot{\theta} = \frac{K_t}{R}V.$$

We consider the following parameters $J = 0.01$, $b = 0.1$, $K_e = 0.01$, $K_t = 0.01$, $R = 1$, $L = 0$.

(i) Find a state-space representation for the evolution of $x(t) := \begin{bmatrix}\theta(t) & \dot{\theta}(t))\end{bmatrix}^\intercal$.

(ii) Discretize this model finding a dynamic system describing the evolution of $x_k := x(\tau k)$ for $\tau = 0.1$ and $k \in \mathbb{N}_0$. Confirm your answer with the Matlab function *c2d*.

(iii) Suppose that we are interested in regulation the state to zero while minimizing a quadratic cost

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

where

$$Q = \begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix}, \quad R = 0.1.$$

(a) Compute the matrix K^* of the optimal policy $u_k = K^* x_k$ and the matrix P^* of the optimal cost $x_0^\intercal P^* x_0$ by using the Matlab function *dlqr*.

(b) Compute the control input that regulates the state to zero in the minimum number of steps as a function of the initial condition x_0. Show that the resulting cost takes the form $x_0^\intercal P_f x_0$, for some matrix P_f.[Hint: Solve Problem 1.4].

(c) Compare the costs obtained in the previous questions (a) and (b).

Problem 1.3 In this problem we wish to prove that the cost

$$\sum_{k=0}^{T} x_k^\intercal Q x_k, \qquad T = \infty, \quad Q = Q^\intercal,$$

for an asymptotically stable discrete-time system

$$x_{k+1} = Ax_k, \quad k \in \mathbb{N}_0,$$

is given by $x_0^\intercal P x_0$ where P is the solution to the following (Lyapunov) equation

$$A^\intercal P A - P = -Q. \tag{1}$$

(a) Start by proving that

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

if $a \in (-1,1)$ by multiplying the left-hand side by $\frac{1-a}{1-a}$.

(b) Use (a) to prove the desired result for the case $A \in \mathbb{R}$, and $|A| < 1$.

(c) Inspired by (a) and (b), prove the desired result for the general case $A \in \mathbb{R}^{n\times n}$, where the eigenvalues of A lie inside the unit circle.

(d) Confirm that if A and Q are replaced by $(A+BK)$ and $Q + K^\intercal R K$, respectively, for $K = -(B^\intercal P B + R)^{-1} B^\intercal P A$, the Lyapunov equation coincides with the Riccati equation

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

Provide an explanation for this fact.

Problem 1.4 Consider the following discrete-time linear time-invariant system

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

where $x_k \in \mathbb{R}^n$, $u_k \in \mathbb{R}$, $k \in \mathbb{N}_0$, and the initial condition x_0 at time $k=0$ is arbitrary. Suppose that we wish to drive the state to a desired final value $\bar{x}$. Show that the following control values u_k for $k \in \{0,1,\dots,n-1\}$ result in $x_n = \bar{x}$:

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

provided that the matrix on the right-hand side is invertible (in which case the system is said to be controllable).

Problem 1.5 Consider the following standard linear quadratic problem

$$\min \sum_{k=0}^{h-1} x_k^\intercal Q x_k + u_k^\intercal R u_k + 2x_k^\intercal S u_k + g_h(x_h) \tag{3}$$

$$x_{k+1} = Ax_k + Bu_k \tag{4}$$

for

$$Q = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix} \quad S = \begin{bmatrix} 0.1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \quad R = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix},$$

$$A = \begin{bmatrix} 0 & 1 & 0 \\ -3 & 4 & 1 \\ 0 & 0 & 3 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 2 \end{bmatrix} \tag{5}$$

and

$$g_h(x_h) = x_h^\intercal Q_h x_h, \quad Q_h = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{6}$$

(i) Using Matlab provide the gains K_k of the optimal policy

$$u_k = K_k x_k, \quad k \in \{0, 1, 2, 3, 4\}$$

when $h = 5$ by iterating the Riccati equations.

(ii) Using the script developed in (i) obtain the optimal policy

$$u_k = K x_k$$

when $g_h(x_h) = 0$ and $h = \infty$.

(iii) Confirm the results of (ii) using the Matlab function dlqr.m

Problem 1.6 The goal of this exercise is to provide examples of linear quadratic regulation problems

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

where

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

for which the optimal control policy cannot be obtained by solving the algebraic Riccati equation

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

(i) Provide an example of a pair (A, B) for which there does not exist a positive definite solution to the algebraic Riccati equation for any Q and R.

(ii) Provide an example of a pair (A, Q) for which there does not exist a positive definite solution to the algebraic Riccati equation for any B and R.

Problem 1.7 The goal of this exercise is to show that we do not need R to be positive definite to obtain the optimal policy for the standard linear regular problem with the algebraic Riccati equation. Consider (3), (4) and the following parameters

$$A = \begin{bmatrix} 1/2 & 1 \\ 1 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad R = 0, \quad S = 0, \quad Q = I, \quad Qh = I.$$

(i) Using Matlab provide the gains K_k of the optimal policy

$$u_k = K_k x_k, \quad k \in \{0, 1, 2, 3, 4\}$$

when $h = 5$ by iterating the Riccati equations.

(ii) Using the script developed in (i) obtain the optimal policy

$$u_k = Kx_k$$

when $g_h(x_h) = 0$ and $h = \infty$.

(iii) Try to use Matlab function dlqr.m and see that it does not work.

Problem 1.8 Consider the following optimal control problem

$$\min \frac{1}{T+1}\mathbb{E}[\sum_{k=0}^{T} x_k^\intercal \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} x_k + u_k^\intercal \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} u_k]$$

for

$$x_{k+1} = \begin{bmatrix} \frac{3}{2} & 0 \\ 0 & 1 \end{bmatrix} x_k + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} u_k + w_k, \quad k \in \mathbb{N}_{\geq 0}$$

where $x_k \in \mathbb{R}^2$, $u_k \in \mathbb{R}^2$, are the state and the control input at time $k \in \mathbb{N}_{\geq 0}$ and $w_k \in \mathbb{R}^2$, $k \in \mathbb{N}_{\geq 0}$, are independent and identically distributed random vectors.

(i) Find an optimal policy for u_0 and u_1 when $T = 2$.

(ii) Compute the value of the expected optimal cost if $x_0 = \begin{bmatrix} 1 & 2 \end{bmatrix}^\intercal$ and $\mathbb{E}[w_k w_k^\intercal] = 3I$ for every k.

(iii) Find the optimal control policy for u_k, $k \in \mathbb{N}_{\geq 0}$ when $T \to \infty$.

Optimal estimation and output feedback

Problem 2.1 The goal of this exercise is to repeat Problem 1.1 but considering state disturbances and noisy measurements. To this effect consider a first order discrete-time linear system

$$x_{k+1} = \alpha x_k + u_k + w_k, \quad k \in \mathbb{N}_0,$$

where $\{w_k | k \in \mathbb{N}_0\}$ are Gaussian independent and identically distributed random variables with variance $\mathbb{E}[w_k^2] = \sigma_w^2$ and consider the cost

$$\mathbb{E}[\sum_{k=0}^{q} x_k^2 + \gamma u_k^2].$$

(a) Considering $\alpha = 1$, $\gamma = 1$ and $q = 2$, $\sigma_w = 1$, determine the policy for u_0, u_1, and u_2 that minimizes the cost when:

- (i) the value of state x_k is known at every stage k.
- (ii) the value of state x_k is measured at every stage k by a noisy sensor

$$y_k = x_k + n_k$$

where $\{n_k | k \in \mathbb{N}_0\}$ are Gaussian independent and identically distributed random variables with variance $\mathbb{E}[n_k^2] = \sigma_n^2$, for $\sigma_n = \frac{1}{3}$. Assume that the initial state x_0 follows a Gaussian distribution with mean $\bar{x}_0 = 0$ and variance $\sigma_0^2 = 10$.

(b) Considering $\alpha = 1$, $\gamma = 1$ and that the cost is now given by

$$\lim_{q \to \infty} \frac{1}{q+1} \mathbb{E}[\sum_{k=0}^{q} x_k^2 + \gamma u_k^2].$$

Determine the policy for $\{u_k | k \in \mathbb{N}_0\}$ that minimizes the cost when

- (i) the value of state x_k is known at every stage k.
- (ii) the value of state x_k is measured at every stage k by a noisy sensor as in (a).(ii). Assume the same distribution for x_0 as in (a).(ii).

Problem 2.2 Consider the problem of estimating the value of a constant x_0 from noisy measurements

$$y_k = x_0 + n_k, \quad k \in \mathbb{N}_0,$$

where n_k are zero mean independent and identically distributed Gaussian random variables with variance $\mathbb{E}[n_k^2] = \sigma_n^2$. The initial state of x_0 is also Gaussian with mean $\bar{x}_0$ and variance $\phi_0 = \mathbb{E}[(x_0 - \bar{x}_0)^2]$.

(i) Compute the Kalman filter estimate $\hat{x}_k$[1] of x_0 at times $k \in \{0, 1, 2\}$, i.e., based on $I_0 = \{y_0\}$, $I_1 = \{y_0, y_1\}$, $I_2 = \{y_0, y_1, y_2\}$ and the associated error variances $\mathbb{E}[(x_k - \hat{x}_k)^2 | \mathcal{I}_k]$.

[Suggestion: Note that $y_k = x_k + n_k$, for $x_{k+1} = x_k$, $k \in \mathbb{N}_0$]

(ii) Compute the asymptotic error variance $\lim_{k\to\infty} \mathbb{E}[(x_k - \hat{x}_k)^2 | \mathcal{I}_k]$ and provide an explanation for the value obtained.

Problem 2.3 Consider a mass-spring-damper system for which the mass displacement $z(t)$ is described by

$$m\ddot{z}(t) = -c_f\dot{z}(t) - k_s z(t) + f_d(t), \quad t \in \mathbb{R}_{\geq 0}$$

where $m = 1$, $k_s = 5$, $c_f = 1$ and $f_d(t)$ models force disturbances.

(a) Find a state-space representation for the system considering $x(t) = \begin{bmatrix} z(t) & \dot{z}(t) \end{bmatrix}^\intercal$.

(b) Find a discrete-time system describing how $x(\tau h)$ evolves, for a sampling period $\tau = 0.1$ and for $k \in \mathbb{N}_0$ and considering the approximation for the disturbances $f_d(t) \approx f_d(k\tau)$ where $f_d(k\tau)$ are zero mean independent and identically distributed Gaussian random variables with $\mathbb{E}[f_d(k\tau)^2] = 0.1$.

(c) Suppose that the following output is available to estimate the state

$$y(k\tau) = Cx(k\tau) + n_k, \quad k \in \mathbb{N}_0,$$

where n_k are zero mean independent and identically distributed Gaussian random variables with $\mathbb{E}[n_k^2] = \sigma_n^2$.

(i) Provide the Kalman filter giving an estimate of $x(\tau h)$ for $k \in \{0, 1, 2\}$ assuming there the initial distribution of the state is Gaussian with covariance matrix

$$\mathbb{E}[(x(0) - \mathbb{E}[x(0)])(x(0) - \mathbb{E}[x(0)])^\intercal] = 2I,$$

where I denotes the identity matrix.

(ii) Using Matlab function kalman.m provide the stationary Kalman filter and the asymptotic error covariance matrix.

Problem 2.4 Consider the following discrete-time linear system

$$x_{k+1} = Ax_k + Bu_k + \omega_k, \quad k \in \mathbb{N}_0$$

where x_k, u_k, and ω_k are the state, control input, and state disturbance at time k,

$$A = \begin{bmatrix} 1 & \frac{1}{5} \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

and $x_0 = \begin{bmatrix} 1 & 1 \end{bmatrix}^\intercal$.

[1] The Kalman filter estimate $\hat{x}_k$ is the one that minimizes the error variance $\mathbb{E}[(x_k - \hat{x}_k)^2]$

(a) Suppose that $\omega_k = 0$ for every $k \in \mathbb{N}_0$. Compute the optimal control inputs u_0, u_1, u_2 minimizing the cost

$$J(x_0, u_0, u_1, \dots, u_{h-1}) = \sum_{k=0}^{h-1} x_k^\intercal Q x_k + u_k^\intercal R u_k \tag{7}$$

for $h = 3$,

$$Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad R = 0.1.$$

(b) Suppose that ω_k are zero mean independent and identically distributed Gaussian random variables with $\mathbb{E}[\omega_k \omega_k^\intercal] = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$ and that the full state x_k is available for feedback. Compute the optimal policy $u_k = \mu_k(x_k)$ for $k \in \{0, 1, 2\}$ minimizing the cost

$$\mathbb{E}[J(x_0, \mu_0(x_0), \mu_1(x_1), \mu_2(x_2))],$$

where J is described in (7).

(c) Suppose that ω_k are as in (b) but only the following output is available for feedback

$$y_k = C x_k + n_k$$

where $C = \begin{bmatrix} 1 & 0 \end{bmatrix}$ and n_k are zero mean independent and identically distributed Gaussian random variables with $\mathbb{E}[n_k^2] = \frac{1}{4}$.

(i) Compute the optimal policy $u_k = \mu_k(\mathcal{I}_k)$ for $k \in \{0, 1, 2\}$, where $\mathcal{I}_0 = (y_0)$ and $\mathcal{I}_k = (y_0, \dots, y_k, u_0, \dots, u_{k-1})$ are the information sets, minimizing the cost

$$\mathbb{E}[J(x_0, \mu_0(\mathcal{I}_0), \mu_1(\mathcal{I}_1), \mu_2(\mathcal{I}_2))], \tag{8}$$

where J is described in (7).

(ii) Compute the optimal policy $u_k = \mu_k(\mathcal{I}_k)$ for $k \in \mathbb{N}_0$ minimizing the cost

$$\lim_{h \to \infty} \frac{1}{h} \mathbb{E}[J(x_0, \mu_0(\mathcal{I}_0), \mu_1(\mathcal{I}_1), \mu_2(\mathcal{I}_2), \dots, \mu_{h-1}(\mathcal{I}_{h-1}))]. \tag{9}$$

(d) Suppose now that the actuators can only be updated at even times $k = 2\ell$, $\ell \in \mathbb{N}_0$ but the output is still available at every time step $k \in \mathbb{N}_0$ (multi-rate control problem). At odd times $k = 2\ell + 1$ we have $u_k = 0$.

(i) Compute the optimal policy $u_k = \mu_k(\mathcal{I}_k)$ for $k \in \{0, 2\}$ minimizing the cost (8).

(ii) Compute the optimal policy $u_k = \mu_k(\mathcal{I}_k)$ for $k = 2\ell$, $\ell \in \mathbb{N}_0$ minimizing the cost (9).

Problem 2.5 Consider the following system

$$x_{k+1} = \frac{5}{6} x_k + u_k + w_k$$

where the control policy is an open-loop policy with $u_k = 1$ for every $k \in \mathbb{N}_{\geq 0}$ and w_k, $k \in \mathbb{N}_{\geq 0}$, are zero-mean independent and identically distributed Gaussian random variables with $\mathbb{E}[w_k^2] = 1$. We wish to estimate the state of the system x_k based on the measurements

$$y_\ell = 2x_\ell + \frac{1}{2}v_\ell, \quad \ell = \{0, 1, \dots, k\}$$

where v_ℓ, $\ell \in \mathbb{N}_{\geq 0}$ are zero-mean independent and identically distributed Gaussian random variables with $\mathbb{E}[v_k^2] = 2$. The initial state x_0 is uncertain and modeled as a Gaussian random variable with mean $\bar{x}_0 = \frac{1}{2}$ and covariance $\bar{\Phi}_0 = 1$.

(i) Show that the Kalman filter estimate $\hat{x}_{2|2} = \mathbb{E}[x_2|y_0, y_1, y_2, u_0 = 1, u_1 = 1]$ of x_2 can be written as

$$\hat{x}_{2|2} = a_0 y_0 + a_1 y_1 + a_2 y_2 + c,$$

and provide the numerical values of a_0, a_1, a_2 and c.

(ii) Show that the estimate $\hat{x}_{k|k} = \mathbb{E}[x_k|y_0, \dots, y_k, u_0 = 1, \dots, u_{k-1} = 1]$ obtained by the stationary Kalman filter can be written as

$$\hat{x}_{k|k} = \sum_{\ell=0}^{k} h_{k-\ell} y_\ell + d_k$$

and provide an expression for h_k, $k \in \mathbb{N}_{\geq 0}$, and for the scalars d_k.

Problem 2.6 Suppose that we wish to estimate the state, at a given time $k = h$, of a linear system:

$$x_{k+1} = Ax_k + w_k, \quad k \in \mathbb{N}_0 := \{0, 1, 2, \dots h-1\},$$

where each w_k denotes the state disturbance at time k. Assuming that A is invertible, we have:

$$x_k = A^{-1}x_{k+1} - A^{-1}\bar{w}_{k+1} \tag{10}$$

where $\bar{w}_{k+1} = w_k$. The following output measurements

$$y_k = Cx_k + v_k$$

are available at times $k \in \{0, 1, 2, \dots, h\}$, where v_k denotes the output noise at time k. The initial state is unknown but an initial estimate, denoted by $\bar{x}_0$, is available. In order to obtain the state estimate at time h, denoted by $\bar{x}_h$, one can: (i) first solve the optimal control problem

$$\begin{aligned} J_h(x_h) = \min_{\bar{w}_1, \dots, \bar{w}_h} & (x_0 - \bar{x}_0)^\intercal \bar{\Theta}_0^{-1}(x_0 - \bar{x}_0) + (y_0 - Cx_0)^\intercal V^{-1}(y_0 - Cx_0) \\ & + \sum_{k=1}^{h} \bar{w}_k^\intercal W^{-1}\bar{w}_k + (y_k - Cx_k)^\intercal V^{-1}(y_k - Cx_k) \end{aligned}$$

subject to (10), for given positive definite weighting matrices W, V, $\bar{\Theta}$; and then (ii) find

$$\bar{x}_h = \operatorname{argmin} J_h(x_h).$$

This can be interpreted as follows: estimate the states $x_0, \dots, x_h$ by taking those which correspond to ('can be explained by') small values of state disturbances, output noise, and deviation from the initial condition estimate according to the penalties in this cost function and take the last state estimate of x_h.

The goal of this exercise is to solve the given optimal control problem and show that $\bar{x}_h$ coincides with the solution of the Kalman filter, $\hat{x}_{h|h}$ obtained by iterating, for $k \in \{0, 1, \dots, h\}$,

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

with initial condition $\hat{x}_{0|-1}$, where, for $k \in \{0, 1, \dots, h\}$,

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

with initial condition $\Theta_{0|-1} = \bar{\Theta}_0$ (in this sense this control problem is the dual of the estimation problem addressed in class).

Suggestion: Consider the DP algorithm going forward in time (instead of backwards) with dynamics (10) and the initial 'cost-to-come' (instead of the cost-to-go)

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

and use (multiple times!) the matrix inversion lemma

$$
(A + UCV)^{-1} = A^{-1} - A^{-1}U(C^{-1} + VA^{-1}U)^{-1}VA^{-1}
$$

The following identities may also be useful

$$
\begin{aligned}
\Theta_{0|0}^{-1} &= \bar{\Theta}_0^{-1} + C^\intercal V^{-1}C \\
(I - L_0C)^\intercal\Theta_{0|0}^{-1} &= \bar{\Theta}_0^{-1} \\
L_0^\intercal\Theta_{0|0}^{-1} &= V^{-1}C
\end{aligned}
$$

Discretization

Problem 3.1 Consider a linear system

$$\dot{x}(t) = Ax(t) + Bu(t)$$

and suppose that $u(t) = u_k$ when $t \in [k\tau, (k+1)\tau)$. Let $x_k := x(k\tau)$. Obtain the matrices A_d and B_d of the discrete-time system

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

when

(i) $A = 1$, $B = 2$, $\tau = 0.1$.

(ii) $A = \begin{bmatrix} 1 & 6 \\ 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 5 \\ 7 \end{bmatrix}$, $\tau = 0.05$.

(iii) $A = \begin{bmatrix} 0 & -4 \\ 4 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ $\tau = 0.2$.

(iv) $A = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$, $B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\tau = 1$.

(v) $A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$, $\tau = 0.5$.

Problem 3.2 Consider a first order discrete-time linear system

$$x_{k+1} = x_k + u_k, \quad k \in \mathbb{N}_0,$$

and the cost

$$\sum_{k=0}^{h-1} x_k^2 + u_k^2 + x_h^2,$$

where $h = 3$. Suppose that $x_0 \in X$, where $X := \{0, 1, 2, 3, 4, 5\}$, and that the input is limited to the following values $u_k \in U$, $U := \{-5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5\}$.

(i) Specify an optimal policy $u_0 = \mu_0(x_0)$, $u_1 = \mu_1(x_1)$, $u_2 = \mu_2(x_2)$ for the following values of the state $x_0 \in X$, $x_1 \in X$, $x_2 \in X$.

(ii) Determine the optimal cost for each possible initial condition.

Static optimization formulation of optimal control problems

Problem 4.1 Solve each of the following optimization problems assuming that X is a positive definite matrix.

(i) $$\min_{x\in\mathbb{R}^n} x^\intercal X x \quad \text{s.t. } c^\intercal x = 1.$$

(ii) $$\min_{x\in\mathbb{R}^n} c^\intercal x \quad \text{s.t. } x^\intercal X x = 1.$$

(iii) $$\min_{x\in\mathbb{R}^n} x^\intercal X x \quad \text{s.t. } x^\intercal x = 1.$$

Suppose that $n = 2$. Can you provide a geometric condition to find the optimal solution for each of the problems? Illustrate this using

$$X = \begin{bmatrix} 2.5 & -1.5 \\ -1.5 & 2.5 \end{bmatrix}, c = [1 \ \ 2].$$

Problem 4.2 Consider the following stage decision problem, with cost function

$$\Big(\sum_{k=0}^{1} x_k^2 + u_k^2\Big) + x_2^2$$

and dynamic model

$$x_{k+1} = \frac{1}{2}(x_k + u_k), \quad k \in \{0, 1\}$$

for an initial condition $x_0 = 1$. Find the control inputs u_0, u_1 minimizing the cost using the following method:

(i) write the cost function in terms of u_0 and u_1 and solve an unconstrained problem.

(ii) use the method of the Lagrange multipliers.

Problem 4.3 Consider the following dynamic model

$$x_{k+1} = Ax_k + Bu_k, \quad k \in \mathbb{N}_0$$

where $x_0 = \begin{bmatrix} 0 & 1 \end{bmatrix}^\intercal$ and

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

and the cost function

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

(i) Find matrices X, Y, and a scalar z such that the cost function takes the form

$$u^\intercal X u + Y^\intercal u + z$$

where

$$u = \begin{bmatrix} u_0 \\ u_1 \end{bmatrix}$$

(ii) Find the control inputs u_0 and u_1 that minimize the cost function.

Problem 4.4 Consider a stage decision problem with linear model

$$x_{k+1} = Ax_k + Bu_k, \quad k \in 0, \dots, h-1$$

for a given initial condition $x_0 = v$ and quadratic cost

$$\frac{1}{2}\Big(\sum_{k=0}^{h-1} x_k^\intercal Q x_k + u_k^\intercal R u_k + x_h^\intercal Q_h x_h\Big),$$

where $Q > 0$, $R > 0$ and $Q_h > 0$, and $x_k \in \mathbb{R}^n$, $u_k \in \mathbb{R}^m$.

(i) Use the method of the Lagrange multipliers to establish that the control input minimizing the cost function can be obtained from the solution $x_k, k \in \{1, \dots, h\}$, $u_k, k \in \{0, \dots, h-1\}$, $\lambda_k, k \in \{1, \dots, h\}$, to the following set of linear equations

$$\begin{aligned} \lambda_k &= A^\intercal \lambda_{k+1} + Qx_k, \quad k \in \{1, \dots, h-1\}, \\ x_{k+1} &= Ax_k + Bu_k, \quad k \in \{0, \dots, h-1\}, \quad x_0 = v \\ Ru_k &= -B^\intercal \lambda_{k+1} \\ \lambda_h &= Q_h x_h \end{aligned}$$

(ii) Show that if we assume that

$$\lambda_k = P_k x_k, k \in \{1, \dots, h\}$$

then P_k satisfies the Riccati equations for $k \in \{1, \dots, h\}$.

Approximate dynamic programming

Problem 5.1 Consider the following switched linear system

$$x_{k+1} = A_{\sigma_k} x_k, \quad k \in \mathbb{N}_0,$$

where $\sigma_k \in \{0, 1\}$ is the control input, and

$$A_0 = \begin{bmatrix} 0.8 & 0.1 \\ 0.2 & 0.8 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 1 & 0.1 \\ 0 & 0.5 \end{bmatrix}.$$

Suppose that we wish to minimize

$$\|x_h\|^2,$$

for some $h \in \mathbb{N}$.

(i) Provide the optimal policy for σ_0 and σ_1 when $h = 2$.

(ii) Consider the base policy $\sigma_k = 1, \forall_{k \in \{0,\ldots,h-1\}}$. Considering this base policy, derive the rollout policy with horizon $H = 1$ for σ_0, σ_1 and σ_2 when $h = 3$.

Problem 5.2 Consider the following dynamic system

$$\begin{bmatrix} \dot{y}(t) \\ \dot{z}(t) \end{bmatrix} = \begin{bmatrix} \cos(\theta(t)) \\ \sin(\theta(t)) \end{bmatrix}$$

where $\theta(t), \in [-\pi, \pi)$, is the control input, and $x(t) = \begin{bmatrix} y(t) & z(t) \end{bmatrix}^\intercal$ is the state. Suppose that the control input is implemented digitally and satisfies

$$\theta(t) = \theta_k, \quad t \in [k\tau, (k+1)\tau],$$

where τ is the sampling period, and consider the following Euler discretization

$$x_{k+1} = x_k + \tau \begin{bmatrix} \cos(\theta_k) \\ \sin(\theta_k) \end{bmatrix}$$

where $x_k := x(k\tau)$.The goal is to minimize

$$\sum_{k=0}^{\infty} \|x_k\|_Q^2$$

where $\|x\|_Q^2 := x^\intercal Q x$ for $x \in \mathbb{R}^2$, and Q is a positive definite matrix.

(i) Show that the model predictive control (MPC) policy with horizon $H = 2$, i.e., taking into account the cost $\sum_{k=\ell}^{\ell+H-1} \|x_k\|_Q^2$, is

$$\theta_k(x_k) = \mu(x_k), \quad k \in \mathbb{N}_0$$

where

$$\mu(x) = \begin{cases} \text{arctang}(\dfrac{v_2(x)}{v_1(x)}), \text{ if } v_1(x) \neq 0 \\ \dfrac{\pi}{2}, \text{ if } v_2(x) \geq 0 \text{ and } v_1(x) = 0 \\ -\dfrac{\pi}{2}, \text{ if } v_2(x) < 0 \text{ and } v_1(x) = 0 \end{cases},$$

and $v(x) := (v_1(x), v_2(x))$, for $x \in \mathbb{R}^2$ is the solution to the problem

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

(ii) Provide the optimal control input at stage 0 when $\tau = 0.1$,

$$Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

and $x_0 = \begin{bmatrix} 1 & 1 \end{bmatrix}^\intercal$.

Problem 5.3 Consider the following continuous-time switched linear system

$$\dot{x}(t) = B_{\sigma(t)} x(t), \quad t \in \mathbb{R}_{\geq 0}$$

where $x(t) = \begin{bmatrix} y(t) & z(t) \end{bmatrix}^\intercal$

$$B_1 = \begin{bmatrix} -(1-\beta) & 1 \\ -9 & -(1-\beta) \end{bmatrix}, \quad B_2 = \begin{bmatrix} -(1-\beta) & 9 \\ -1 & -(1-\beta) \end{bmatrix}.$$

The scheduling variable $\sigma(t)$ should be seen as a control input. Suppose that the system is controlled digitally with a sampling period of $\tau = 0.1$. Then

$$x_{k+1} = A_{\sigma_k} x_k$$

where $x_k = x(k\tau)$, $\sigma(t) = \sigma_k$, for $t \in [k\tau, (k+1)\tau)$, $A_1 = e^{B_1 \tau}$ and $A_2 = e^{B_2 \tau}$. The figure below depicts two trajectories for $x(t)$ and $x_k := x(kh)$ when $\beta = 1$ and: (a) $\sigma(t) = 1$ for every t and $x_0 = [0 \;\; 1]^\intercal$; (b) $\sigma(t) = 2$ for every t and $x_0 = [1 \;\; 0]^\intercal$.

(i) Suppose that $\beta = 1$.

   (a) Provide a policy for the control input $\sigma(t)$ which stabilizes the continuous-time system, i.e., is such that $\lim_{t \to \infty} \|x(t)\| = 0$.

   (b) Provide a policy for the control input σ_k which stabilizes the discrete-time system, i.e., is such that $\lim_{k \to \infty} \|x_k\| = 0$.

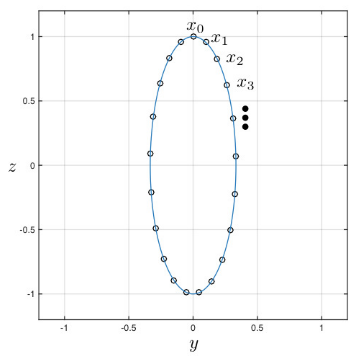
(ii) Consider now a cost

$$\sum_{k=0}^{\infty} x_k^\intercal x_k.$$

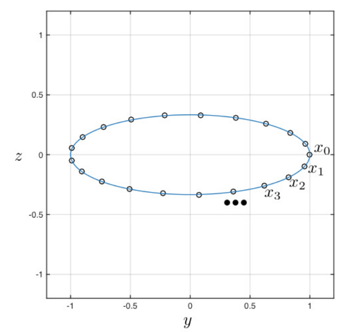
(a) Derive the policy for the control input σ_k obtained using model predictive control with horizon $H = 2$, i.e., based on the minimization of $x_k^\intercal x_k + x_{k+1}^\intercal x_{k+1}$ at each time k. Provide the region in the state-space where $\sigma_k = 1$ and where $\sigma_k = 2$, when $\beta = 1$ and when $\beta = 0.9$. Do these policies stabilize the discrete-time system?

(b) Considering $\beta = 0.9$ derive a rollout policy with horizon $H = 1$ assuming that the base policy is $\sigma_k = 1$ for every $k \geq 0$. Is it possible to use this base policy when $\beta = 1$?

[Hint: Compute first the cost of the based policy using Problem 2.2. and then derive the rollout policy.]



(a)



(b)

Problem 5.4 Consider the following discrete-time system

$$x_{k+1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} x_k + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u_k,$$

where $u_k \in \mathbb{R}$ and $x_k \in \mathbb{R}^2$. Two stabilizing controllers are available

$$u_k = \begin{bmatrix} -\frac{1}{2} & -1 \end{bmatrix} x_k, \quad u_k = \begin{bmatrix} -\frac{1}{3} & -\frac{6}{5} \end{bmatrix} x_k.$$

The goal is to switch between the two controllers in order to minimize

$$\sum_{k=0}^{\infty} \|x_k\|^2.$$

(i) Formulate the problem as the optimal control of a switched linear system

$$x_{k+1} = A_{\sigma_k} x_k$$

for $\sigma_k \in \{1, 2\}$, $k \in \mathbb{N}_0$, with cost function

$$\sum_{k=0}^{\infty} x_k^\intercal Q x_k,$$

for some positive definite matrix Q. Provide the matrices A_1, A_2, and Q.

(ii) Using the Lyapunov equation, show that the cost of the base policy $\sigma_k = 1$, $\forall k \in \mathbb{N}_0$, takes the form

$$x_0^\intercal \bar{P} x_0.$$

Provide the matrix $\bar{P}$.

(iii) Considering the same base policy $\sigma_k = 1$, $\forall k \in \mathbb{N}_0$, derive the rollout policy with horizon $H = 1$ for $\{\sigma_k\}_{k=0}^{\infty}$.

(iv) Let $\hat{J}(x_0)$ denote the cost when the rollout policy with horizon $H = 1$ is applied for a given initial condition x_0. It is possible to prove that

$$\hat{J}(x_0) \leq x_0^\intercal \bar{P} x_0, \quad \text{for every } x_0. \tag{11}$$

Use (11) to establish that the rollout policy with horizon $H = 1$ asymptotically stabilizes the system, in the sense that $\|x_k\| \to 0$ as $k \to \infty$.