Version GV/2

EGT3
ENGINEERING TRIPOS PART IIB

Tuesday 09 May 2023   2 to 3.40

Module 4F3

AN OPTIMISATION BASED APPROACH TO CONTROL

*Answer not more than **three** questions.*

All questions carry the same number of marks.

*The **approximate** percentage of marks allocated to each part of a question is indicated in the right margin.*

*Write your candidate number **not** your name on the cover sheet.*

STATIONERY REQUIREMENTS
Single-sided script paper

SPECIAL REQUIREMENTS TO BE SUPPLIED FOR THIS EXAM
CUED approved calculator allowed
Attachment: 4F3 data sheet (two pages).
Engineering Data Book

10 minutes reading time is allowed for this paper at the start of the exam.

You may not start to read the questions printed on the subsequent pages of this question paper until instructed to do so.

You may not remove any stationery from the Examination Room.

1 Consider a system with transfer function representation

$$\begin{bmatrix} \bar{z}(s) \\ \bar{y}(s) \end{bmatrix} = \begin{bmatrix} P_{11}(s) & P_{12}(s) \\ P_{21}(s) & P_{22}(s) \end{bmatrix} \begin{bmatrix} \bar{w}(s) \\ \bar{u}(s) \end{bmatrix} \tag{1}$$

and control policy $\bar{u}(s) = K(s)\bar{y}(s)$.

(a) Explain what is meant by the lower Linear Fractional Transformation (LFT) $\mathcal{F}_l\,(P(s), K(s))$ for the system above, and derive an expression for it. [20%]

(b) The system in (1) has the following state space realization

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} w_1 + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

$$z = \begin{bmatrix} x_1 + x_2 \\ u \end{bmatrix}$$

$$y = \begin{bmatrix} 1 & 0 \end{bmatrix} x + w_2$$

where x is the state vector and w_1, w_2 are external disturbances.
Describe how you would solve the $\mathcal{H}_2$ optimal control problem

$$\min_{K(s)\ \text{stabilising}} \|\mathcal{F}_l\,(P(s), K(s))\|_2$$

Furthermore, if the stabilizing solutions to the CARE and FARE equations associated with this problem are, respectively,

$$X = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \alpha, \qquad Y = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \alpha$$

where $\alpha = 2 + \sqrt{5}$, derive a state space realization for the optimal control policy, expressing your answer in terms of α. [35%]

(c) Consider now the system

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

$$z = \begin{bmatrix} x_1 + x_2 \\ u \end{bmatrix}$$

$$y = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

with state feedback policy $u = Kx$, where K is a constant matrix.

Consider the optimal control problem

$$\min_{K \text{ stabilising}} 2\pi \int_0^{\infty} z^T z dt$$

Find the matrix K that solves this problem, and find also the minimum cost when the initial condition is $x_1(0) = x_2(0) = 1$. [30%]

(d) Discuss the connection and differences between the optimal control problems considered in parts (b) and (c). [15%]

2 Consider the system

$$x_{k+1} = x_k + u_k \tag{2}$$

where $x_k \in \mathbb{R}$, $u_k \in \mathbb{R}$ for $k = 0, \ldots, h-1$, and $x_h \in \mathbb{R}$. Consider also the cost function

$$J(x_0, u_0, \ldots, u_{h-1}) = x_h^2 + \sum_{k=0}^{h-1} (x_k^2 + u_k^2)$$

and the optimal control problem

$$\min_{u_0, \ldots, u_{h-1}} J(x_0, u_0, \ldots, u_{h-1})$$

for a given initial condition x_0.

(a) Describe what is meant by the value function for the optimal control problem described above. Write also the dynamic programming equation for this problem. [15%]

(b) Show that the value function at step k is of the form $g(k)x_k^2$ and use this to derive an iteration for $g(k)$. [35%]

(c) Find the optimal value of J when $x_0 = 1$, $h = 3$. [15%]

(d) Consider the following modified cost function for the system in (2):

$$\tilde{J}(x_0, u_0, \ldots, u_{h-1}) = \alpha^{2h} x_h^2 + \sum_{k=0}^{h-1} (\alpha^{2k} x_k^2 + \alpha^{2k} u_k^2)$$

where $\alpha \in \mathbb{R}$ is a constant. Using the transformation $\tilde{x}_k = \alpha^k x_k$, $\tilde{u}_k = \alpha^k u_k$, derive an iteration that gives the optimal control policy when the optimal control problem described at the beginning of the question is solved, but with the cost function J replaced with the cost function $\tilde{J}$. [35%]

3 (a) Describe briefly the receding horizon principle and explain why model predictive control with a receding horizon may not result in a stable feedback system. [10%]

(b) Consider the system

$$x(l+1) = 2x(l) + u(l)$$

with a model predictive controller using the receding horizon cost function

$$J(x) = \sum_{k=0}^{1} qx_k^2 + u_k^2 \quad \text{(subject to } x_0 = x,\ x_1 = 2x_0 + u_0\text{)}$$

Find explicitly the resulting feedback law, and hence show that the resulting feedback system is stable if, and only if, $q > 1$. [40%]

(c) Assume $q < 1$. For the system in part (b), the controller that minimizes the infinite horizon cost

$$\sum_{l=0}^{\infty} qx(l)^2 + u(l)^2 \tag{3}$$

is given by $u(l) = -K/(1+K)x(l)$, where $K > 0$ solves $K = K/(1+K) + q$, with an optimal cost of $Kx(0)^2$.

Consider now the modified system

$$x(l+1) = 2x(l) + \operatorname{sat} u(l)$$

where

$$\operatorname{sat} u = \begin{cases} 1, & u > 1 \\ u, & -1 \le u \le 1 \\ -1, & u < -1 \end{cases}$$

(i) Explain why it is not possible to stabilize the system for all possible values of $x(0)$. [10%]

(ii) How should the receding horizon cost function of part (b) be modified in order to ensure that the new feedback system is stable, and that (3) is minimized, for as wide a range of $x(0)$ as possible? Justify, and explain carefully, your answer. [40%]

4 (a) Define the action-value function Q and write down its update equation for the Q-learning algorithm. [15%]

(b) Consider the system

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

with the initial condition $x_0 = N$, for N integer, and the objective of minimizing one of the following control objectives

(i) $\sum_k u_k^2 + \sum_k x_k^2$

(ii) $\max\{\max_k\{u_k\gamma^k\}, \max_k\{x_k\gamma^k\}\}$, for $\gamma > 1$.

The action u is also required to be integer. These are both reasonable objectives, but only one of these problems can be solved using Q-learning. Which one, and why?. Explain your reasoning carefully. [25%]

(c) For the same system as in part (b) consider now the problem of minimising

$$\sum_k |u_k| + \sum_k |x_k|$$

for

$$0 \leq x \leq 4 \text{ and } |u| \leq 2$$

Choices of u that would result in x going outside that range are not admissible, e.g. when $x = 1$ then u must be chosen from $\{-1, 0, 1, 2\}$ etc.

(i) Find an optimal solution by using four iterations of Q-learning, exploring the whole state-action space at each iteration and starting with $Q_0(x, u) = 0$ for all x and u, so the first iteration, Q_1, is the result of updating the Q-value of each state and action pair once etc. [40%]
[You might find it convenient to tabulate the Q values on a sequence of 5x5 grids.]

(ii) Explain, using one state action pair as an example, how you would verify that the Q-values have converged to their optimal value. [10%]

(iii) What is the optimal trajectory starting from $x(0) = 2$? [10%]

END OF PAPER

Module 4F3: An optimisation based approach to control

(available in the exam)

1. (a) For the dynamical system satisfying, $\dot{x} = f(x, u),\ x(0) = x_0$, and the cost function

$$J(x_0, u(\cdot)) = \int_0^T c(x(t), u(t))\, dt \ + J_T(x(T))$$

then under suitable assumptions the value function, $V(x, t)$, satisfies the Hamilton-Jacobi-Bellman PDE,

$$-\frac{\partial V(x,t)}{\partial t} = \min_{u \in U} \left(c(x, u) + \frac{\partial V(x,t)}{\partial x} f(x, u) \right), \quad V(x, T) = J_T(x).$$

(b) For $f(x, u) = Ax + Bu$, $c(x, u) = x^T Q x + u^T R u$, and $J_T(x) = x^T X_T x$, if $X(t)$ satisfies the Riccati ODE,

$$-\dot{X} = Q + XA + A^T X - XBR^{-1}B^T X, \quad X(T) = X_T,$$

then $J_{opt} = x_0^T X(0) x_0$ and $u_{opt}(t) = -R^{-1} B^T X(t) x(t)$.

2. For the discrete-time system satisfying $x_{k+1} = Ax_k + Bu_k$ with x_0 given and cost function,

$$J(x_0, u_0, u_1, \ldots, u_{h-1}) = \sum_{k=0}^{h-1} \left(x_k^T Q x_k + u_k^T R u_k \right) + x_h^T X_h x_h,$$

if X_k satisfies the backward difference equation,

$$X_{k-1} = Q + A^T X_k A - A^T X_k B (R + B^T X_k B)^{-1} B^T X_k A,$$

then $J_{opt} = x_0^T X_0 x_0$ and optimal control signal, $u_k = -(R + B^T X_{k+1} B)^{-1} B^T X_{k+1} A x_k$.

3. For the system satisfying,

$$\begin{bmatrix} \dot{x} \\ \hline z \\ y \end{bmatrix} = \left[\begin{array}{c|cc} A & \begin{bmatrix} B_1 & 0 \end{bmatrix} & B_2 \\ \hline \begin{bmatrix} C_1 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} & \begin{bmatrix} 0 \\ I \end{bmatrix} \\ C_2 & \begin{bmatrix} 0 & I \end{bmatrix} & 0 \end{array}\right] \begin{bmatrix} x \\ \hline w \\ u \end{bmatrix} \quad \text{where} \quad \begin{cases} (A, B_2) & \text{controllable} \\ (A, C_1) & \text{observable} \\ (A, B_1) & \text{controllable} \\ (A, C_2) & \text{observable} \end{cases}$$

(a) The optimal $\mathcal{H}_2$ controller is given by,

$$\begin{bmatrix} \dot{x}_k \\ \hline u \end{bmatrix} = \left[\begin{array}{c|c} A - B_2F - HC_2 & -H \\ \hline F & 0 \end{array}\right] \begin{bmatrix} x_k \\ \hline y \end{bmatrix}$$

where $F = B_2^T X$, $H = YC_2^T$, and X and Y are stabilising solutions to

$$0 = XA + A^TX + C_1^TC_1 - XB_2B_2^TX \quad \text{(CARE)}$$

and

$$0 = YA^T + AY + B_1B_1^T - YC_2^TC_2Y \quad \text{(FARE)}$$

(b) The controller given by,

$$\begin{bmatrix} \dot{x}_k \\ \hline u \end{bmatrix} = \left[\begin{array}{c|c} \hat{A} - B_2F - HC_2 & -H \\ \hline F & 0 \end{array}\right] \begin{bmatrix} x_k \\ \hline y \end{bmatrix}$$

where $F = B_2^T X$, $H = YC_2^T$, $\hat{A} = A + \frac{1}{\gamma^2}B_1B_1^TX$, and X and Y are stabilising solutions to,

$$XA + A^TX + C_1^TC_1 - X(B_2B_2^T - \gamma^{-2}B_1B_1^T)X = 0$$

and

$$Y\hat{A}^T + \hat{A}Y + B_1B_1^T - Y\big(C_2^TC_2 - \gamma^{-2}F^TF\big)Y = 0,$$

satisfies$\|T_{w\to z}\|_\infty \leq \gamma$.