EGT3
ENGINEERING TRIPOS PART IIB

Tuesday 06 May 2025 9.30 to 11.10

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 the system

$$x_{k+1} = f(x_k, u_k) \qquad (1)$$

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

$$\min_{u_0,\ldots,u_{h-1}} J(x_0, u_0, \ldots, u_{h-1}) \qquad (2)$$

where

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

and x_0 is a given initial condition.

(a) Write down the dynamic programming equation for this problem and explain what is meant by the value function $V(x, k)$. [10%]

(b) Explain the significance of $V(x, 0)$ and $V(x, h)$. Also discuss advantages and disadvantages of dynamic programming. [15%]

(c) Consider now the system

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

associated with the optimal control problem (1)-(2) with cost function

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

where $\alpha \in \mathbb{R}$, $\lambda \in \mathbb{R}$ are constants.

(i) Derive, using the dynamic programming equation, an iteration that can be used to find the value function for each k. [25%]

(ii) Find an expression for the value the optimal cost will tend to as $h \to \infty$. [10%]

(iii) Find the optimal cost when $h = 1$, $\alpha = \lambda = 1/2$ with initial condition $x_0 = 1/2$. [10%]

(d) Consider the optimal control problem in part (c) but with the cost now given by

$$J(x_0, u_0, u_1 \ldots) = \sum_{k=0}^{\infty} \left((x_k - r)^2 + (u_k - r(1-\alpha))^2\right) \qquad (3)$$

where $r \in \mathbb{R}$ is a constant. Find the optimal cost and the optimal control input u_k for this problem. Discuss also when the cost function in (3) is a relevant one to consider. [30%]

2 (a) Explain what is meant by the $\mathcal{H}_2$ norm and the $\mathcal{H}_\infty$ norm of a transfer function $G(s)$. [10%]

(b) Consider the system

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= w_1 + u \\ z_1 &= 2x_1, \quad z_2 = u \\ y &= x_1 \end{aligned}$$

where $x_1(t) \in \mathbb{R}$, $x_2(t) \in \mathbb{R}$ are the states of the system, $w_1(t)$ is an external disturbance, and $u(t)$ is the control input. Consider also a controller with transfer function $K(s)$ such that $\bar{u}(s) = K(s)\bar{y}(s)$, where $\bar{u}(s), \bar{y}(s)$ are the Laplace transforms of the signals u, y respectively.

Find a controller K that will minimize the $\mathcal{H}_2$ norm of the closed-loop transfer function from w_1 to z, where $z = [\begin{matrix} z_1 & z_2 \end{matrix}]^T$. [40%]

(c) Consider the system in part (b) but with $y = x_1$ replaced by

$$y_1 = x_1, \quad y_2 = x_2$$

and the controller $K(s)$ satisfies $\bar{u}(s) = K(s)\bar{y}(s)$ where $y = [\begin{matrix} y_1 & y_2 \end{matrix}]^T$.

(i) Find a controller K that will lead to an $\mathcal{H}_\infty$ norm that is less than $\sqrt{2}$, for the closed-loop transfer function from w_1 to z . [30%]

(ii) Describe a procedure that can be used to minimize the $\mathcal{H}_\infty$ norm considered in part (c)(i). [10%]

(iii) Explain if there exists a controller $K(s)$ such that the $\mathcal{H}_\infty$ norm considered in part (c)(i) is less than 1. [10%]

3 Consider a discrete-time system given by

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) \end{aligned} \tag{4}$$

where $x(k)$, $u(k)$ and $y(k)$ are vector-valued states, inputs and outputs of the system. We would like to find a predictive receding horizon control law for $u(k)$ to drive the output $y(k)$ towards a (nonzero) setpoint y_s. At each time index $k > 0$, we must solve the finite horizon optimal control problem given by

$$\min_{u_0,\ldots,u_{N-1}} (y_N - y_s)^\top P(y_N - y_s) + \sum_{i=0}^{N-1} \Big((y_i - y_s)^\top Q(y_i - y_s) + \Delta u_i^\top \Delta u_i\Big) \tag{5}$$

subject to

$$\begin{aligned} -1 \le \Delta u_i &\le 1, \quad i = 0,\ldots,N-1 \\ x_0 &= x(k) \\ u_{-1} &= u(k-1) \end{aligned} \tag{6}$$

where $\Delta u_i = u_i - u_{i-1}$. Recall x_i, u_i and y_i denote predicted variables at time index i within a prediction window.

(a) Considering the objective of the problem, explain why it is desirable to penalize $\Delta u_i^\top \Delta u_i$, rather than $u_i^\top u_i$, in the cost function (5). [10%]

(b) Let $\theta_x = [x_1^\top, \ldots, x_N^\top]^\top$ and $\theta_u = [u_0^\top, \ldots, u_{N-1}^\top]^\top$. Find matrices M_x and M_u and vectors f_x and f_u such that the expression being minimized in (5) is equal to

$$\theta_x^\top M_x \theta_x + \theta_u^\top M_u \theta_u + f_x^\top \theta_x + f_u^\top \theta_u + \text{constant} \tag{7}$$

The M and f variables in (7) should be written in terms of the system matrices in (4), the setpoint y_s, and the data $x(k)$, $u(k-1)$. For simplicity, use $N = 2$. [25%]

(c) Hence show that the problem (5)-(6) can be formulated equivalently as a standard Quadratic Programming (QP) problem in the decision variable vector $\theta = [\theta_x^\top, \theta_u^\top]^\top$. [25%]

(d) State conditions for the QP problem in part (c) to be strictly convex. What does that imply regarding the uniqueness of the optimal receding horizon control law? [15%]

(e) Show that it is possible to weaken the conditions in part (d) by using the model (4) to eliminate θ_x from the decision variables of the QP problem. [25%]

4 The motion of a mobile robot on the plane is described by the kinematic model

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} \sin\theta & 0 \\ \cos\theta & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v \\ \omega \end{bmatrix}$$

where x and y are the Cartesian coordinates of the robot on the plane, θ is the orientation of the robot with respect to the x-axis, v is the magnitude of the robot's forward velocity (i.e., along the θ direction) in m/s, and ω is the robot's angular velocity in rad$/s$. In this model, $s = [x, y, \theta]^\top$ is the state vector, and $a = [v, \omega]^\top$ is the input vector. The robot has a mass of M, and a moment of inertia around its axis of rotation of J.

Consider the case where $s(0) = [0, 0, 0]^\top$, and the objective is to move the robot towards the target coordinates $(x_f, y_f) = (5, 4)$ while minimizing the robot's kinetic energy.

(a) Propose a cost functional for the objective above that results in an episodic optimal control problem. Explain why the problem is episodic, and state its stopping set. [20%]

(b) Explain the main difficulties in using Reinforcement Learning to solve the episodic problem above. [10%]

(c) We now assume that the input $a(t)$ is constructed for $t \geq 0$ by repeatedly applying any combination of the following actions:

$$\begin{aligned} &a_0 : \text{apply } v = 0,\ \omega = 0 \text{ for 1 second;} \\ &a_1 : \text{apply } v = 0,\ \omega = \pi/4 \text{ for 1 second;} \\ &a_2 : \text{apply } v = 0,\ \omega = -\pi/4 \text{ for 1 second;} \\ &a_3 : \begin{cases} \text{if } \theta = n\pi/2 \text{ for some } n \in \mathbb{Z}, & \text{apply } v = 1,\ \omega = 0 \text{ for 1 second;} \\ \text{otherwise,} & \text{apply } v = \mu,\ \omega = 0 \text{ for 1 second;} \end{cases} \end{aligned} \tag{8}$$

where $\mu > 0$ is a constant value. By discretizing time with a sampling period of 1 second, and assuming $s(0) = [0, 0, 0]^\top$, find a value for μ in Eq. (8) that facilitates the application of Reinforcement Learning to the problem. Describe how $s(k + 1)$ is obtained from $s(k)$ and $a(k)$, and propose a cost for the discrete-time problem. [35%]

(d) Explain how the discrete-time problem of part (c) can be solved using Q-learning with the ϵ-greedy policy. For a step size $\alpha = 0.5$, write down the first three iterations of the algorithm assuming that $s(0) = [0, 0, 0]^\top$, $a(0) = a_1$, $a(1) = a_2$ and $a(2) = a_1$ are selected and the action-value is initialized as $Q_0(s, a) = 0$ for all state-action pairs (s, a). [35%]

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

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

and

$$0 = YA^T + AY + B_1B_1^T - YC_2^T C_2 Y \quad \text{(FARE)}$$

(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 $F = B_2^T X$, $H = YC_2^T$, $\hat{A} = A + \frac{1}{\gamma^2} B_1 B_1^T X$, and X and Y are stabilising solutions to,

$$XA + A^T X + C_1^T C_1 - X(B_2B_2^T - \gamma^{-2} B_1 B_1^T)X = 0$$

and

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

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