Version GV/3

EGT3
ENGINEERING TRIPOS PART IIB

Tuesday 03 May 2022 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 We want to solve a path planning problem for a robot moving on the chessboard represented in Fig. 1a. The robot always occupies a single box and moves **u** - up, **d** - down, **l** - left, **r** - right, **n** - none. Assume that each move takes approximately one second. We want the robot to move from the starting position, S, to the end position, E, in minimum time. Shaded boxes represent forbidden regions, where the robot should not enter.

(a) Formulate and solve the path planning problem using dynamic programming. Use the state space $X = \{(i, j) \mid i \text{ is the box row, } j \text{ is the box column}\}$ and the input space $U = \{\mathbf{u}, \mathbf{d}, \mathbf{l}, \mathbf{r}, \mathbf{n}\}$.

(i) Define stage cost and terminal cost to solve the path planning problem via dynamic programming. Explain your choices. [20%]

(ii) Using dynamic programming, show how to find the value function or cost to go $V : X \to \mathbb{R}$. At each update, show the value of V at each chessboard box. [20%]

(iii) Find the optimal input sequence and the optimal trajectory from S to E. How would these change if we extended the chessboard with additional rows at the top and at the bottom? Justify your answers. [15%]

(b) Now that a path has been found, the robot has to navigate it optimally. Consider the abstract path in Fig. 1b, where the variable z denotes the distance along the path, from z_S to z_E. The robot motion satisfies the equation $\dot{z} = u$, where the input u represents the robot's speed. The energy cost is measured by $\int_0^6 u(t)^2 dt$.

(i) Show how to find the minimum energy to drive the robot from z_S to z_E using Riccati equations. [20%]

(ii) Compute the optimal cost and optimal control u^*. [25%]

(a) (b)

Fig. 1

2 Figure 2 represents a shock absorber with unit mass $M = 1$, spring stiffness k, and damping coefficient c. The displacement x_1 and the velocity x_2 of the shock absorber satisfy

$$\dot{x} = Ax + Bu + Bw \qquad A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{1}$$

where $x = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^T$, $u = -kx_1 - cx_2$ represents the internal force that spring and damper exert on the mass M, and w is an external force acting on the system. The measured output is $y = x_1$.

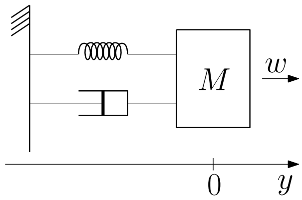

Fig. 2

(a) $\mathcal{H}_2$ analysis and control:

(i) Compute the $\mathcal{H}_2$ norm of the transfer function from w to y, $||T_{w\to y}||_2$, for nominal parameters $k = 1$ and $c = 1$. Explain its significance in terms of the impulse response of the system and in terms of the system transients measured by $||y||_\infty$. [20%]

(ii) Compute the stiffness and damping parameters, k and c, that guarantee optimal $\mathcal{H}_2$ norm $||T_{w\to z}||_2$ from the input w to the performance output $z = \begin{bmatrix} y & u \end{bmatrix}^T$. [25%]

(iii) Prove that $||T_{w\to y}||_2 \leq ||T_{w\to z}||_2$, which illustrates how optimising the latter is a way to improve the former (while moderating the control action). Then, in comparison to your results in part (a) (i), show that k and c computed in part (a) (ii) make $||T_{w\to y}||_2$ smaller. [20%]

(b) $\mathcal{H}_\infty$ analysis and control:

(i) Show that $||T_{w\to y}||_\infty = 1$ if $k = 1$ and $c \geq 2$. [15%]

(ii) Show how to find the values of k and c that minimise the $\mathcal{H}_\infty$ norm from w to $z = \begin{bmatrix} y & u \end{bmatrix}^T$ using linear matrix inequalities. [20%]

3 (a) Outline how an algorithm that can minimise some norm of a vector $\underline{z}$, subject to constraints on $\underline{z}$, can be used as the basis of a model predictive control scheme. Details are not required, and in particular you are not required to construct any matrices. [10%]

(b) (i) Show, using sketches or otherwise, that for $\alpha > 0$

$$\min_u |z-u| + \alpha|u| = \begin{cases} |z|, & \text{if } \alpha \geq 1 \\ \alpha|z|, & \text{if } \alpha \leq 1 \end{cases}$$

(Hint: Assume first that $z > 0$, in which case the minimum occurs for u somewhere in the range $[0, z]$, and then generalise.) [10%]

(ii) Show, using sketches or otherwise, that

$$\min_u \max\{|z-u|, \alpha|u|\} = \frac{\alpha}{1+\alpha}|z|$$

(Hint: The hint in part (b)(i) also applies here.) [10%]

(c) Consider the problem of minimising each of the following three cost functions, subject to

$$x_0 = x, \quad x_{k+1} = 2x_k + u_k : k = 0, 1$$

(i) $J_1(x) = |x_1| + |x_2| + \alpha(|u_0| + |u_1|)$

(ii) $J_2(x) = x_1^2 + x_2^2 + \alpha^2(u_0^2 + u_1^2)$

(iii) $J_\infty(x) = \max(|x_1|, |x_2|, \alpha|u_0|, \alpha|u_1|)$

Assume $\alpha > 0$. In each case find $\min_{u_0,u_1} J_\bullet$ and the minimising u_0, and describe how they vary with α. [30%]

(Hint: you may find it easiest in each case to minimise over u_1 for fixed u_0 and then minimise the resulting answer over u_0.)

(d) Explain how your answers to part (c) can be used to construct a receding horizon controller for the system

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

For what range of α is the closed loop system stable in each case? [30%]

(e) Comment briefly on the results. [10%]

4 Figure 3 shows a gridworld in which an agent starts in square 4 and proceeds by moving horizontally or vertically (but not diagonally) to an adjacent square. Each visit to a square incurs a cost of 1, with the exception of the lava filled square 7 for which the cost is 100. If the agent moves to square 7 then it is immediately transported back to the start square 4 after incurring the cost of 100. The episode ends when the agent reaches square 6. The solution which minimises the overall cost is $4 \to 5 \to 6$ with a cost of 2.

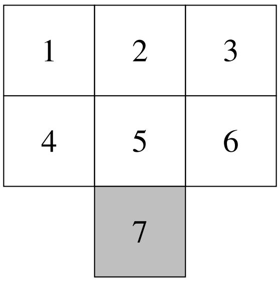

Fig. 3

(a) Explain briefly how the SARSA algorithm, with ϵ-greedy action selection, may be applied to the problem of completing this task with a small total cost. [10%]

(b) Consider the following collection of five episodes

$$1)\ 4 \to 5 \to 6, \quad 2)\ 4 \to 1 \to 2 \to 5 \to 6, \quad 3)\ 4 \to 5 (\to 7) \to 4 \to 5 \to 6$$

$$4)\ 4 \to 5 (\to 7) \to 4 \to 5 \to 6, \quad 5)\ 4 \to 5 (\to 7) \to 4 \to 5 \to 6$$

(where the brackets around the $\to 7$ are to show that this is not the resulting next state, since the next state after applying the action $\to 7$ is actually 4.)

(i) If the value of $Q(s,a)$ is initially set to ∞ for all state-action pairs then what are the revised values of $Q(s,a)$ after each episode? (You may find it helpful to show these using numbers and arrows on one of more copies you have made of Fig. 3. Note that after the first move from 5 to 6 then the Q value for that action is set to 1 since the episode terminates.) [30%]

(ii) Explain how it is possible that these might be the first five episodes that occur. [10%]

(iii) The algorithm will find eventually different solutions for large and small ϵ. What are they and what approximately is the critical value of ϵ that separates them? [30%]

(c) If Q-learning, again with with ϵ-greedy action selection, is applied to this problem then what is the solution found and what is the average episodic cost? [20%]

END OF PAPER

Module 4F3: Optimal and Predictive Control

Data Sheet (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$\|T_{w \to z}\|_\infty \leq \gamma$.

K Glover, 2013
F. Forni (no change), 2019