Version GV/2

EGT3
ENGINEERING TRIPOS PART IIB

Tuesday 4 May 2021 9 to 10.40

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 and at the top of each answer sheet.*

STATIONERY REQUIREMENTS
Write on single-sided paper.

SPECIAL REQUIREMENTS TO BE SUPPLIED FOR THIS EXAM
CUED approved calculator allowed
Attachment: 4F3 data sheet (two pages).
You are allowed access to the electronic version of the Engineering Data Books.

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

The time taken for scanning/uploading answers is 15 minutes.

Your script is to be uploaded as a single consolidated pdf containing all answers.

1 The structure of a data network is represented by the graph in Fig. 1. Nodes represent networking devices, typically routers. Edges represent bidirectional connections between routers. Each edge number indicates the time needed for a packet to travel between routers. We want to find the minimal-time routing path from any initial node to node 7.

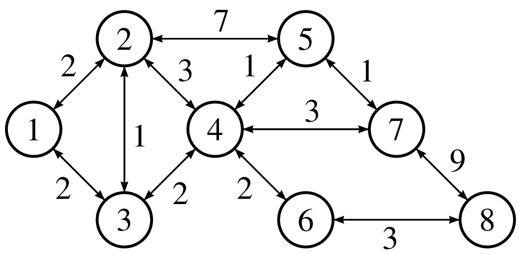


Fig. 1

(a) Discuss Bellman's principle of optimality as it applies to the specific context of minimal-time routing problems. [15%]

(b) Define the cost of a path as the sum of the transmission times along it. Using dynamic programming, show how to find the minimal-time routing path to send packets from each node to node 7. [25%]

(c) Suppose there is an uncertainty of ± 1 time units on each communication between routers, i.e. in the delay associated with each edge. For the worst case scenario (largest delay on each edge), will the minimal-time routing paths to node 7 change? Explain your answer. [10%]

(d) Explain why an edge with *negative* time weight would make the optimisation problem ill-conditioned. [10%]

(e) For each edge from node i to node j, denote by w_{ij} their communication time and consider the indicator variable $x_{ij} \in \{0, 1\}$.

(i) Define the mathematical optimisation problem for the minimal-time routing problem from node 1 to node 7. [30%]

(ii) Assume now that each variable x_{ij} is continuous and non-negative. What kind of optimisation problem is that? (least-squares / linear programming / convex optimisation). Explain your answer. [10%]

2 A two-compartment system models drug diffusion between body organs. Using x_1 and x_2 to denote drug concentrations in the two compartments, the model is given by

$$\dot{x}_1 = -(k_{12}+d)x_1 + k_{21}x_2 + u \qquad \dot{x}_2 = k_{12}x_1 - (k_{21}+d)x_2 + w_1 \qquad y = x_2 + w_2\,,$$

where $k_{12} = 10$, $k_{21} = 20$ are the flow rates between compartments, and $d = 1$ is the drug's degradation rate. The input u models drug injection. The inputs w_1 and w_2 model disturbances affecting concentration in the second compartment and measurement noise, respectively. The output y represents the measured concentration in the second compartment.

(a) Write the system in matrix form and compute the transfer function $T_{w_1\to y}(s)$ from input w_1 to output y. [20%]

(b) Define the norms $\|T_{w_1\to y}\|_2$ and $\|T_{w_1\to y}\|_\infty$ and discuss their meaning in terms of the performance of this system. [15%]

(c) Compute $\|T_{w_1\to y}\|_2$. Explain the steps of your computation. [25%]

(d) It is required to find the controller $K(s)$, from y to u, which solves the $\mathcal{H}_2$ optimal control problem from the disturbance $w = [\,w_1 \;\; w_2\,]^T$ to the performance output $z = [\,x_2 \;\; u\,]^T$. To this end:

(i) Formalise the $\mathcal{H}_2$ optimal control problem using the notion of generalised plant $P(s)$ and linear fractional transformations. [15%]

(ii) Write the state-space realisation of the generalised plant P that is needed to solve the $\mathcal{H}_2$ optimal control problem. Define numerically each matrix you use. [10%]

(iii) Using

$$X = \begin{bmatrix} 0.0491 & 0.0542 \\ 0.0542 & 0.0753 \end{bmatrix} \qquad \text{and} \qquad Y = \begin{bmatrix} 0.1944 & 0.1072 \\ 0.1072 & 0.0747 \end{bmatrix}$$

as solutions to the CARE and FARE equations, respectively, derive the state matrices of the optimal $\mathcal{H}_2$ controller $\dot{x}_k = A_k x_k + B_k y$, $u = C_k x_k$. [15%]

3 It is desired to find an *approximate* solution to the problem of minimising

$$J(x_0) = \sum_{k=0}^{\infty} x_k^T Q x_k + u_k^T R u_k$$

for the system

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

for a given initial condition x_0, subject to the constraint $||u_k||_\infty \leq U$ for all k (i.e., all elements of the vector u_k satisfy $|u_k(\cdot)| \leq U$).

(a) Explain in detail how predictive control, with a horizon length of $N = 2$, can be used to find a controller which ensures that $J(x_0)$ is minimised for sufficiently small x_0. Give the precise form of any Riccati equations or any Quadratic Programs (QPs) that must be solved (it is not necessary to solve them). QPs should be written in the standard form. [60%]

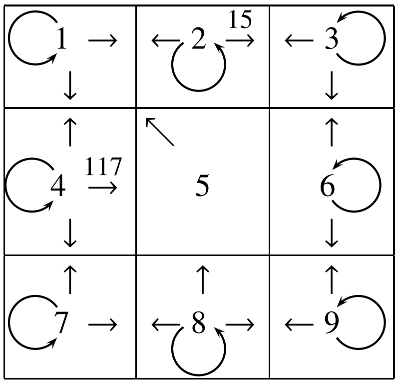
(b) Explain why you have chosen the QP you did in your answer to (a). What are its advantages and disadvantages compared to other equivalent QPs? [10%]

(c) What would be the advantage of increasing the horizon length N? [15%]

(d) Explain why, in practice, a horizon length of $N = 2$ is unlikely to be sufficient. [15%]

4 (a) Define the state-action value function $Q(s, a)$. When can it be more useful than the value function $V(s)$? [10%]

(b) A gridworld is labelled with states $x \in \{1, \ldots, 9\}$ as shown



For all states except 5, an action involves a move from a state to an adjacent state, or staying in the same place, with all moves allowed except from 2 or 6 to 5. The only action from state 5 is a return to the starting state 1. Each arrow in the figure represents a state action pair.

The cost of visiting each state is given by the following array, where the position in this array corresponds to the position in the gridworld:

1	5	5
4	100	5
4	4	0

(So, for example, the cost of visiting state 5 is 100, and that of visiting states 2, 3, 6 is 5.)

(i) Two of the arrows in the top figure have been labelled with the corresponding values of $Q(s, a)$ for that state and action. What is the value of Q for the others? [30%]

(ii) What is the minimal episodic cost when starting at state 1, and what is the optimal path. [20%]

(iii) If Q-learning with an ϵ-greedy action selection method is used, with $\epsilon = 0.1$ and a discount factor $\lambda = 0.9$, estimate a lower bound on the average episodic cost after a large number of trials. [20%]

(iv) Explain, briefly, the SARSA algorithm. Given the same discount factor and ϵ-greedy action selection method what path might you expect it to find? (You are not required to do any calculations for this part.) [20%]

END OF PAPER

THIS PAGE IS BLANK

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,

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

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

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

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,

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

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

K Glover, 2013
I. Lestas (no change), 2021