Version GV/2

EGT3
ENGINEERING TRIPOS PART IIB

Monday 6 May 2024 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 the discrete-time system

$$x_{k+1} = f(x_k) + g(u_k)$$

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

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

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) Write down the dynamic programming equation for this problem and discuss its significance. [25%]

(b) Consider the continuous-time system

$$\dot{x} = f(x) + g(u)$$

where $x(t) \in \mathbb{R}, u(t) \in \mathbb{R}$, with initial condition $x(0) = x_0$. Consider the cost function

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

By considering the discretization of the continuous time system

$$x(t + \delta t) = x(t) + f(x(t))\delta t + g(u(t))\delta t + O\big((\delta t)^2\big)$$

and the corresponding dynamic programming equation, derive the Hamilton-Jacobi-Bellman PDE satisfied by the optimal control input $u(t)$. [35%]

(c) Consider the optimal control problem in part (b) with $T \to \infty$, $f(x) = -x$, $g(u) = u$ and

$$c(x(t), u(t)) = e^{2\alpha t}(x^2(t) + u^2(t)) \quad \text{where } \alpha < 0.$$

(i) Use the transformation $\tilde{x}(t) = e^{\alpha t}x(t)$, $\tilde{u}(t) = e^{\alpha t}u(t)$ to show that this optimal control problem can be transformed to an infinite horizon Linear Quadratic Regulator problem. [20%]

(ii) Find the optimal control input $u(t)$. [20%]

2 Consider the system with state space realization

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 2 \\ 2 \end{bmatrix} w_1 + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

$$z = \begin{bmatrix} 2x_1 + 2x_2 \\ u \end{bmatrix}$$

$$y = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + w_2$$

where $x_1(t) \in \mathbb{R}$, $x_2(t) \in \mathbb{R}$ are the states of the system, $w_1(t), w_2(t)$ are external disturbances, and $u(t)$ is the control input.

Consider also a control policy $\bar{u}(s) = K(s)\bar{y}(s)$, and the $\mathcal{H}_2$ optimal control problem

$$\min_{K(s) \text{ stabilising}} \|T_{w\to z}\|_2 \qquad (1)$$

where $T_{w\to z}$ denotes the transfer function from w to z.

(a) The CARE equation associated with this optimal control problem has two solutions of the form

$$X = \begin{bmatrix} \alpha & \beta \\ \beta & \beta \end{bmatrix}.$$

Find these two solutions and show that $\alpha = 2\beta$ in these solutions. [30%]

(b) The FARE equation associated with this optimal control problem has two solutions of the form

$$Y = \begin{bmatrix} \gamma & \gamma \\ \gamma & \delta \end{bmatrix}.$$

Find these two solutions and show that $\delta = 2\gamma = 2\beta$ in these solutions. [20%]

(c) Find the state space realization of the optimal controller and its transfer function $K(s)$, expressing your answers in terms of β. [30%]

(d) If in equation (1) the objective is to minimize the $\mathcal{H}_\infty$ norm of $T_{w\to z}$, instead of its $\mathcal{H}_2$ norm, describe how you would find the optimal controller. [20%]

3 (a) What is meant by

(i) a convex function;

(ii) a convex set?

Sketch examples of a function and a set which are not convex, illustrating why this is the case. [20%]

When minimising the value of a function $f(x)$ over a set $x \in X$ why is it helpful for both the function and the set to be convex? [10%]

(b) A plant with state x_k and input u_k at time k is described by the discrete-time state-space model $x_{k+1} = Ax_k + Bu_k$. A predictive controller is required to minimise the cost function

$$V(x_0, u_0, u_1, ..., u_{N-1}) = x_N^T P x_N + \sum_{k=0}^{N-1} (x_k^T Q x_k + u_k^T R u_k)$$

subject to the constraints $Mx_k + Eu_k \leq b$ for $k = 0, 1, ..., N-1$ and the given dynamics.

(i) Under what conditions can this be written as a convex optimisation problem in the decision variables $u_0, u_1, \ldots, u_{N-1}$ and $x_1, x_2, \ldots, x_N$? Verify that, under these conditions, the function to be minimised and the constraints to be satisfied are convex. [30%]

(ii) Show that this problem can be written as a standard quadratic programming problem of the form

$$\text{minimise } \theta^T H \theta \text{ subject to } F\theta - f = 0 \text{ and } G\theta - g \leq 0$$

for suitable matrices F, G, H and suitable vectors f, g, with the vector θ containing the decision variables $u_0, u_1, \ldots, u_{N-1}$ and $x_1, x_2, \ldots, x_N$. [15%]

(c) If the aim is to approximate the solution to minimising $\sum_{k=0}^{\infty}(x_k^T Q x_k + u_k^T R u_k)$ subject to the same constraints then how should P be chosen, what further constraint should be added to the formulation and how should N be chosen? [25%]

4 Figure 1 shows a Markov decision process with the rewards for each state/action pair labelled (e.g. the reward for taking action 2 in state 2 is -1).

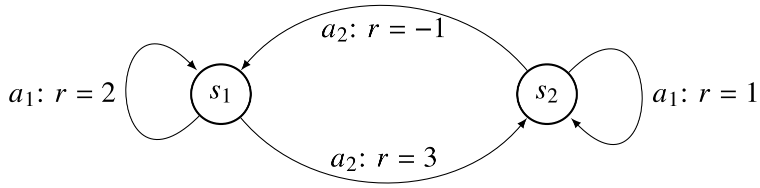

Fig. 1

(a) Describe the Monte Carlo Exploring Starts (MCES) and Q-learning algorithms as they apply to finding the optimal discounted reward for this process. What would be an appropriate learning rate? [20%]

(b) Assuming a discount factor $\lambda = 0.8$ write down the value function and the associated optimal control by inspection. Verify that this value function satisfies the Bellman equation at each state. [20%]

(c) Write down the sequence of states, actions and rewards that follow from starting in state 1 and applying action 1, and subsequently following the policy of taking action 2 in state 1 and action 1 in state 2. [15%]

(d) Find the resulting action-values that would result from applying the following algorithms to the data in Part (c), using a learning rate of 1 and assigning an initial value of 0 to all action-values:

(i) Q-learning, using all the available data sequentially; [10%]

(ii) First-visit MCES. [10%]

(e) For each set of tentative action-values derived in Part (d) write down the sequence of states, actions and rewards that would follow from starting in state 1 and following the greedy policy. [15%]

(f) Which algorithm converges quickest for this problem, and when might the other be preferred? [10%]

END OF PAPER

THIS PAGE IS BLANK

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,

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

(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 $F=B_2^TX$, $H=YC_2^T$, 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 $F=B_2^TX$, $H=YC_2^T$, $\hat{A}=A+\frac{1}{\gamma^2}B_1B_1^TX$, 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 $\|T_{w\to z}\|_\infty\le\gamma$.