4F3: An optimisation based approach to control CRIB

Version GV/4

1 The motion of an inkjet printer head is to be optimised. The printer head moves accordingly to the law

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

where x_k is the current position, x_{k+1} is the next position, and u_k is the offset.

(a) Let r be the desired target position and consider the finite horizon cost

$$J_1 = (x_T - r)^2 + \sum_{k=0}^{T-1} (x_k - r)^2 \qquad \text{for } T = 2 .$$

(i) Using dynamic programming, find the optimal cost J_1^*, optimal trajectory x_k^*, and optimal (unconstrained) input sequence u_k, from the generic initial state $x_0^* = x_0$. [30%]

System dynamics and cost function lead to the dynamic programming equation

$$\boldsymbol{V(x,k) = \min_{u \in \mathbb{R}} (x-r)^2 + V(x+u, k+1) \qquad k = 0, 1}$$

with final condition $V(x,2) = (x-r)^2$. We have

$$\boldsymbol{V(x,1) = \min_{u \in \mathbb{R}} (x-r)^2 + (x+u-r)^2 = (x-r)^2 \qquad \text{for } u^* = r - x ,}$$
$$\boldsymbol{V(x,0) = \min_{u \in \mathbb{R}} (x-r)^2 + (x+u-r)^2 = (x-r)^2 \qquad \text{for } u^* = r - x .}$$

For the generic initial condition x_0 the optimal cost is given by

$$\boldsymbol{J_1^* = V(x_0, 0) = (x_0 - r)^2 .}$$

The optimal trajectory is $x_{k+1}^* = x_k^* + (r - x_k^*) = r$, thus

$$\boldsymbol{x_0^* = x_0 , \quad x_1^* = r , \quad x_2^* = r .}$$

The optimal input sequence is $u_k = r - x_k^*$, thus

$$\boldsymbol{u_0^* = r - x_0 , \quad u_1^* = 0 .}$$

The printer head moves to the target r in one step, driven by the offset $u_0 = r - x_0$. No additional control action is then required ($u_1^* = 0$).
Alternative solution: define the error $e = x - r$ and derive the error dynamics $e_{k+1} = e_k + u_k$. J_1 becomes $\sum_{k=0}^{T} e_k^2 = \sum_{k=0}^{T-1} e_k^2 + e_T^2$. Then solve the finite-horizon linear quadratic regulator problem.

(ii) Explain why the modified cost

$$J_2 = (x_T - r)^2 + \sum_{k=0}^{T-1} \left((x_k - r)^2 + u_k^2\right) \qquad \text{for } T = 2$$

gives a smoother optimal motion x_k^* than J_1. Does J_2 guarantee that $x_T^* = r$? [10%]

J_2 penalizes large target errors $(x_k - r)^2$ and large control offsets u_k^2 at the same time. The combination of these two objectives will result in a optimal trajectory that moves in the direction of the target r driven by "small" offsets u_k. In comparison to J_1 the head will not necessarily move in one step to the target but will use the whole horizon to get closer to the target through small offsets u_k. Indeed, J_2 does not guarantee $x_2^* = r$.

(b) Consider the cost

$$J_3 = (x_T - r)^2 \qquad \text{for } T = 2\ .$$

Find the optimal input sequence constrained to $-1 \le u_k^* \le 1$, optimal trajectory x_k^*, and optimal cost J_3^*, for the initial condition $x_0 = 0$ and target $r = 2$. [30%]

The dynamic programming equation for J_3 reads

$$V(x, k) = \min_{-1 \le u \le 1} V(x + u, k + 1) \qquad k = 0, 1$$

with final condition $V(x, 2) = (x - r)^2$. To solve the problem, define the saturation function

$$\text{sat}(z) := \begin{cases} -1 & \text{if } z \le -1 \\ z & \text{if } |z| \le 1 \\ 1 & \text{if } z \ge 1 \end{cases} \qquad z \in \mathbb{R}\ .$$

Then,

$$V(x, 1) = \min_{-1 \le u \le 1} (x + u - r)^2 = (x - r - \text{sat}(x - r))^2 \quad \text{for } u^* = -\text{sat}(x - r)$$

and

$$V(x, 0) = \min_{-1 \le u \le 1} (x + u - r - \text{sat}(x + u - r))^2\ .$$

For $r = 2$ and $x_0 = 0$ the optimal cost is

$$J_3^* = V(0, 0) = \min_{-1 \le u \le 1} (u - 2 - \text{sat}(u - 2))^2 = (1 - 2 - \text{sat}(1 - 2))^2 = 0 \quad \text{for } u^* = 1$$

For $x_0^* = 0$, the optimal input is thus $u_0^* = 1$, which brings the state to $x_1^* = x_0^* + u_1^* = 1$. Then, the next optimal input $u_1^* = -\text{sat}(x_1^* - r) = 1$ brings the state to the target $x_2^* = 2$.

(c) A simple physical model of the printer head is given by the dynamics of a frictionless mass of unit weight, with position p, and velocity v

$$\dot{p}(t) = v(t) \qquad \dot{v}(t) = u(t)$$

where u is the external driving force. An appropriate cost is

$$J_4 = \int_0^\infty (p(t) - r)^2 + u(t)^2 \, dt \, .$$

By formulating the problem of minimising this cost as a *linear quadratic regulator* (LQR) problem, find the optimal (unconstrained) control $u(t)$. [30%]

Define the error $e = [\, e_p \, , \, e_v \,]^T$ where $e_p = p - r$ and $e_v = v$. Then, the error dynamics read

$$\dot{e} = Ae + Bu \qquad A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \; B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

and the cost reads

$$J_4 = \int_0^\infty e(t)^T Q e(t) + u(t)^2 \, dt \qquad Q = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},$$

This is a standard continuous-time infinite-horizon linear quadratic regulator problem. Optimal costs and optimal control input are, respectively,

$$J_4^*(e_0) = e_0^T X e_0 \qquad\qquad u^*(t) = -B^T X e(t)$$

where $X = X^T > 0$ solves the Control Algebraic Riccati Equation

$$0 = Q + XA + A^T X - XBB^T X \, .$$

For instance,

$$0 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} X_1 & X_3 \\ X_3 & X_2 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} X_1 & X_3 \\ X_3 & X_2 \end{bmatrix} - \begin{bmatrix} X_1 & X_3 \\ X_3 & X_2 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} X_1 & X_3 \\ X_3 & X_2 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & X_1 \\ 0 & X_3 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ X_1 & X_3 \end{bmatrix} - \begin{bmatrix} X_3^2 & X_3 X_2 \\ X_3 X_2 & X_2^2 \end{bmatrix}$$

which gives

$$X_3^2 = 1 \quad \rightarrow \quad X_3 = 1$$
$$X_2^2 = 2X_3 \quad \rightarrow \quad X_2 = \sqrt{2}$$
$$X_1 = X_3 X_2 \quad \rightarrow \quad X_1 = \sqrt{2}$$

Note that $X_3 = \pm 1$ but the restriction to $X_3 = 1$ guarantees feasibility of the second equation. $X_2 = \pm\sqrt{2}$ but the restriction to $X_2 = \sqrt{2}$ is needed for $X > 0$. Thus,

$$u^*(t) = -B^T X e(t) = -\begin{bmatrix} 0 & 1 \end{bmatrix}\begin{bmatrix} \sqrt{2} & 1 \\ 1 & \sqrt{2} \end{bmatrix} e(t) = -p(t) + r - \sqrt{2} v(t) \, .$$

2 (a) Consider the first order system with transfer function

$$G(s) = \frac{1}{s+1}\ .$$

(i) Describe the two methods of computing the $\mathcal{H}_\infty$ norm of a stable system, one in the frequency domain and the other in the time domain / state space. [15%]

In the frequency domain, the $\mathcal{H}_\infty$ norm of G is given by

$$\| G \|_\infty = \sup_\omega \bar{\sigma}(G(j\omega))$$

where $\bar{\sigma}$ is the largest singular value of $G(j\omega)$. The $\mathcal{H}_\infty$ norm of G can be estimated numerically by gridding over frequency.

The $\mathcal{H}_\infty$ norm of G has also the interpretation of the the largest amplification factor among $\mathcal{L}_{2,[0,\infty)}$ signals

$$\| G \|_\infty = \sup_{u \in \mathcal{L}_{2,[0,\infty)}} \frac{\| y \|_2}{\| u \|_2} \quad \text{where } \bar{y}(s) = G(s)\bar{u}(s)$$

This interpretation provides an alternative way to compute the $\mathcal{H}_\infty$ norm, which uses the state-space realization fo G given by

$$\dot{x} = Ax + Bu\ , \qquad y = Cx\ .$$

It can be shown that $\frac{\|y\|_2}{\|u\|_2} \le \gamma$ if and only if the Riccati equation

$$A^T X + XA + C^T C + \frac{1}{\gamma^2} XBB^T X = 0$$

has a positive solution $X = X^T > 0$. A bisection algorithm can then be used to find the smallest γ for which the Riccati equation has a solution.

(ii) Compute the $\mathcal{H}_\infty$ norm of $G(s)$. [15%]

Using the first method,

$$\sup_\omega |G(j\omega)| = \sup_\omega \left(G(j\omega)^* G(j\omega)\right)^{1/2} = \sup_\omega \left(\frac{1}{\omega^2+1}\right)^{1/2} = 1 \qquad \text{at } \omega = 0\ .$$

Using the second method, for $A = -1$, $B = 1$, and $C = 1$ the Riccati equation

$$0 = -2X + 1 + \frac{X^2}{\gamma^2} \quad \Longrightarrow \quad X = \frac{\gamma^2}{2}\left(2 \pm \sqrt{4 - \frac{4}{\gamma^2}}\right)$$

one of which is real and positive if $\gamma \ge 1$.

(b) Consider the associated generalised plant P_1 with realisation

$$\dot{x} = -x + u + w_1, \quad y = x + w_2, \quad z = \begin{bmatrix} x \\ u \end{bmatrix}$$

where w_1 and w_2 represent additive noise on state x and measured output y, respectively. z is the performance output and u is the control input.

(i) Define the $\mathcal{H}_\infty$ optimal control problem for the generalised plant P_1. [10%]

Find a controller $\bar{y}_c(s) = K(s)\bar{u}_c(s)$ such that the closed loop of controller K and generalized plant P_1 given by $u = y_c, u_c = y$ is stable and $\| z \|_2 \leq \gamma \| w \|_2$ for all $w \in \mathcal{L}_{2,[0,\infty)}$.

Using linear fractional transformations, the $\mathcal{H}_\infty$ control problem is about finding a stabilizing controller $K(s)$ such that $\| \mathcal{F}_l(P_1(s), K(s)) \|_\infty \leq \gamma$.

(ii) Derive state-feedback and output-feedback $\mathcal{H}_\infty$ controllers which guarantee that the closed loop transfer function from w to z has $\mathcal{H}_\infty$ norm smaller than or equal to 1. [40%]

The generalized plant can be written in the form

$$\begin{bmatrix} \dot{x} \\ \hline z \\ y \end{bmatrix} = \left[\begin{array}{c|cc} A & \begin{bmatrix} B_1 & 0 \end{bmatrix} & B_2 \\ \hline \begin{bmatrix} C_1 \\ 0 \end{bmatrix} & 0 & \begin{bmatrix} 0 \\ I \end{bmatrix} \\ C_2 & \begin{bmatrix} 0 & I \end{bmatrix} & 0 \end{array}\right] \begin{bmatrix} x \\ \hline w \\ u \end{bmatrix}$$

where $A = -1$, $B_1 = B_2 = C_1 = C_2 = 1$. The two identity matrices have dimension 1×1. Controllability and observability conditions are all fulfilled.

Using the 4F3 datasheet, the state feedback controller is given by

$$y_c = -B_2^T X x = -Xx$$

where X is the solution to

$$\begin{aligned} 0 &= XA + A^T X + C_1^T C_1 - X(B_2 B_2^T - \gamma^{-2} B_1 B_1^T)X \\ &= -2X + 1 - \left(1 - \frac{1}{\gamma^2}\right) X^2 = -2X + 1 \quad \Longrightarrow \quad X = 1/2 . \end{aligned}$$

Note that $A - B_2 B_2^T X = -1.5$ and $A - B_2 B_2^T X + \frac{1}{\gamma^2} B_1 B_1^T X = -1$ are both stable.

From the 4F3 datasheet, the output feedback controller is given by

$$\begin{bmatrix} \dot{x}_k \\ \hline u \end{bmatrix} = \left[\begin{array}{c|c} A + \frac{1}{\gamma^2} B_1 B_1^T X - B_2 F - H C_2 & -H \\ \hline F & 0 \end{array}\right] \begin{bmatrix} x_k \\ \hline y \end{bmatrix} = \left[\begin{array}{c|c} -1 - H & -H \\ \hline F & 0 \end{array}\right] \begin{bmatrix} x_k \\ \hline y \end{bmatrix}$$

for $F = B_2^T X = 1/2$ and $H = YC_2^T = Y$, where Y satisfies

$$0 = B_1B_1^T + Y(A + \frac{1}{\gamma^2}B_1B_1^T X)^T + (A + \frac{1}{\gamma^2}B_1B_1^T X)Y - Y(C_2^T C_2 - \gamma^{-2}F^T F)Y$$

$$= 1 - Y - \left(1 - \frac{1}{4\gamma^2}\right)Y^2 = 1 - Y - \frac{3Y^2}{4} \implies Y \in \{-2, 2/3\}\,.$$

The second is a stabilizing solution, thus the controller reads

$$\left[\begin{array}{c} \dot{x}_k \\ \hline u \end{array}\right] = \left[\begin{array}{c|c} -5/3 & -2/3 \\ \hline 1/2 & 0 \end{array}\right]\left[\begin{array}{c} x_k \\ \hline y \end{array}\right].$$

(c) Consider the simplified generalised plant P_2 with realisation

$$\dot{x} = -x + u + w, \quad z = y = x$$

where w represents additive noise. Using linear matrix inequalities, describe how to find the state-feedback $\mathcal{H}_\infty$ controller $u = Kx$ which minimises the $\mathcal{H}_\infty$ norm of the closed loop transfer function from w to z. [20%]

The generalized plant has the form

$$\dot{x} = Ax + Bu + Bw\,,\ z = y = Cx \qquad \text{for } A = -1,\ B = C = 1\,.$$

We look for a controller $u = Kx$ that guarantees

$$\frac{d}{dt}V \le -z^2 + \gamma^2 w^2 \qquad \text{for } V = x^T X x$$

which is equivalent to solve the linear matrix inequality

$$\begin{bmatrix} X(A+BK) + (A+BK)^T X + C^T C & PB \\ B^T P & -\gamma^2 I \end{bmatrix} \le 0$$

For $Y = X^{-1}$ we get the equivalent formulation

$$\begin{bmatrix} (A+BK)Y + Y(A+BK)^T + YC^T CY & B \\ B^T & -\gamma^2 I \end{bmatrix} \le 0$$

and using $Z = KY$ we get

$$\begin{bmatrix} AY + BZ + YA^T + Z^T B^T + YC^T CY & B \\ B^T & -\gamma^2 I \end{bmatrix} \le 0\,.$$

By Schur complement,

$$\begin{bmatrix} AY + BZ + YA^T + Z^T B^T & B & YC^T \\ B^T & -\gamma^2 I & 0 \\ CY & 0 & -I \end{bmatrix} \le 0$$

which is a linear matrix inequality in the unknowns $Y = Y^T > 0$, Z, and $\gamma > 0$. Minimizing over γ returns Y and Z from which the minimizing $\mathcal{H}_\infty$ state-feedback controller reads

$$u = Kx = ZY^{-1}x \ .$$

3 (a) List some advantages and disadvantages of predictive control, referring to one or more different application areas in your answer. [25%]

Advantages: takes constraints into account, particularly useful for operating near constraints. Intutitive for system operators.

Disadvantages: needs an accurate model and full state feedback. Computationally expensive online computation. Lack of transparency and guarantees.

(b) Write down the standard form of a *quadratic programming* (QP) optimisation problem. [10%]

Given matrices $Q > 0$ and A and vectors c and b, a QP is the optimisation problem:

$$\min_{\theta} \frac{1}{2}\theta^T Q\theta + c^T\theta$$

$$\text{subject to: } A\theta \le b$$

(c) Consider a linear discrete time dynamical system

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

Predictive control is to be applied to this system with a receding horizon cost function

$$\sum_{i=k}^{k+N-1} \left(x_{i+1}^T Q x_{i+1} + u_i^T R u_i \right)$$

and constraints

$$|u_k| \le U$$

for some U and all k. Assuming that the full state vector x_k can be measured at each step, show how the problem of choosing the control input at each step can be written as a QP problem. [40%]

Since

$$x_1 = Ax_0 + Bu_0$$
$$x_2 = Ax_1 + Bu_1$$

etc, we can write

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} \triangleq \begin{bmatrix} A \\ A^2 \\ \vdots \\ A^N \end{bmatrix} x_0 + \begin{bmatrix} B & 0 & \cdots & 0 \\ AB & B & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A^{N-1}B & A^{N-2}B & \cdots & B \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_{N-1} \end{bmatrix}.$$

This can be written as x = $\Phi x_0 + \Gamma$u, where $x_0 \triangleq x = x(k)$ The cost function can be rewritten as:

$$V(x,\mathrm{u}) = x_0^T Q x_0 + \begin{bmatrix} x_1 \\ x_2 \\ \vdots \end{bmatrix}^T \begin{bmatrix} Q & & \\ & Q & \\ & & \ddots \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \end{bmatrix} + \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_{N-1} \end{bmatrix}^T \begin{bmatrix} R & & & \\ & R & & \\ & & \ddots & \\ & & & R \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_{N-1} \end{bmatrix}$$

$$= x^T Q x + \mathrm{x}^T \Omega \mathrm{x} + \mathrm{u}^T \Psi \mathrm{u}.$$

Note that:

$$Q \geq 0 \implies \Omega \geq 0$$
$$R > 0 \implies \Psi > 0.$$

Putting this together:

$$\begin{aligned} V(x,\mathrm{u}) &= x^T Q x + (\Phi x + \Gamma \mathrm{u})^T \Omega (\Phi x + \Gamma \mathrm{u}) + \mathrm{u}^T \Psi \mathrm{u} \\ &= \mathrm{u}^T (\Gamma^T \Omega \Gamma + \Psi)\mathrm{u} + 2\mathrm{u}^T \Gamma^T \Omega \Phi x + x^T (Q + \Phi^T \Omega \Phi) x \\ &= \frac{1}{2} \mathrm{u}^T G \mathrm{u} + \mathrm{u}^T F x + x^T (Q + \Phi^T \Omega \Phi) x \end{aligned}$$

for

$$G \triangleq 2(\Psi + \Gamma^T \Omega \Gamma)$$
$$F \triangleq 2\Gamma^T \Omega \Phi.$$

which is in the correct form, since x is a constant and the final term can be removed for calculating the optimal u. We write the constraints as

$$\begin{bmatrix} -I \\ I \end{bmatrix} \mathrm{u} \leq U$$

This is now in the standard form for a QP.

(d) Explain how, and why, the cost function and constraints should be modified to ensure stability and feasibility. Equations are not required. [25%]

(i) design a stabilizing state f/b K.

(ii) Find (by solving a Ricatti equation) a P such that $x_N^T P X_N$ is a control Lyapunov function for the unconstrained sub-optimal f/b law $u = Kx$.

(iii) add the terminal cost $x_N^T P X_N$ to the cost function, so that the value function of the unconstrained optimal control is a control lyapunov function for the receding horizon control law.

(iv) Find a terminal constraint $M_N x_N \leq b_N$ which is both invariant for the control law $u = Kx$ and constraint admissible and add this constraint to the QP. This ensures that feasible solutions remain feasible.

4 Consider the following three control problems. For each problem choose a suitable algorithm from the areas of Predictive Control, Reinforcement Learning or Optimal Control (a different algorithm for each problem). In each case describe the algorithm and its application to the problem in detail, explaining why it is appropriate to the problem. Included in your answers should be definitions of the terms *value function*, *policy iteration*, *value iteration* and the action-value function, Q. You may make reasonable assumptions.

(a) Control Problem 1: Determining the minimum fuel required, and the corresponding throttle settings to land a spacecraft on the surface of the moon from orbit. An accurate model of the spacecraft is available. You may assume that the spacecraft moves in one orbital plane (i.e. that the desired landing point is directly below the spacecraft at one point in its orbit) and that the orientation of the spacecraft can be controlled instantaneously. [33%]

The state-space is of low dimension, and the problem is deterministic, so it is feasible to solve this directly by dynamic programming. The value function $V(s)$ would be the minimum fuel required to land the spacecraft from state s (∞ if it's not possible). Choose a suitable time step, let $V(s_{landed}) = 0$ and work backwards. A complete answer would include a description of both value and policy iteration and an argument about which is to be preferred here (which would depend on assumptions that the student makes)

(b) Control Problem 2: Maintaining straight and level flight for a damaged aircraft, where several control surfaces are damaged or unavailable but an accurate linearised mathematical model is available for both the aircraft and the available controls and it is known which controls are not available. [33%]

Predictive control is suitable here, and a reasonable alternative to designing individual controllers for each failure condition. A complete answer would include a definition of the value function for the finite horizon problem and a description of the receding horizon formulation. Constraints should be placed on the deviation of the aircraft from straight and level flight and on the magnitude of the control inputs.

(c) Control Problem 3: An offline study into the possibility of controlling a heavily damaged aircraft, including being able to execute turns, where several control surfaces and one or more engines are unavailable, and parts of the wing are missing. An accurate simulation environment is available, incorporating important nonlinear characteristics from a detailed fluid dynamics model, and it is expected that an unconventional nonlinear control strategy would be required. [34%]

Some form of RL would be appropriate here, with Q-learning being the expected

answer. The action-value function should be defined. The reward should be a combination of a reward for landing safely plus a reward for maintaining control of the aircraft. The state-space is of too high a dimension for a tabular approach, so some way of approximating Q is required, eg a deep neural net.

END OF PAPER

THIS PAGE IS BLANK

Assessor's report, Module 4F3

Q1: Linear quadratic control

Popular question, good answers by most of the students. Part A: most of the students answered correctly but some adopted the LQR approach, which is not correct (not well posed, does not meet the condition R¿0). Part B: well addressed by most of the students. Part C: good average. Main issues were about LQR formulation. A few blank answers.

Q2: H_∞ norm and control

Popular question. Part A: good answers in general. A few students did not remember the approach based on Riccati's equation. Part B: well addressed by most students. Part C: just a few students did it correctly.

Q3: Predictive Control

Popular question on model predictive control, attempted by all candidates. Routine and very well answered on the whole.

Q4: Reinforcement Learning

This question included parts on the new material in the course, reinforcement learning. Attempted by approximately half the candidates. The average mark was OK, but this is because there were a lot of mediocre answers. The question was descriptive, but few candidates took seriously the requirement for *detailed* descriptions of the algorithms chosen.