Optimization and Control: Examples Sheet 2

LQG Models

1. [Solution] This question is meant to give practice in deriving the Riccati equation. So we start with the dynamic programming equation and make the hypothesis that $F_{s-1}(x) = \Pi_{s-1}x^2$, which is true for $s = 0$, $\Pi_0 = 1$. Then

$$\begin{aligned}
F_s(x) &= \min_u \left[Qu^2 + F_{s-1}(Ax+Bu)\right] = \min_u \left[Qu^2 + (Ax+Bu)^2\Pi_{s-1}\right] \\
&= \min_u \left[(Q + B^2\Pi_{s-1})\left(u + \frac{AB\Pi_{s-1}}{Q+B^2\Pi_{s-1}}x\right)^2 + A^2\Pi_{s-1}x^2 - \frac{(AB\Pi_{s-1})^2}{Q+B^2\Pi_{s-1}}x^2\right] \\
&= QA^2\Pi_{s-1}(Q+B^2\Pi_{s-1})^{-1}x^2,
\end{aligned}$$

where the optimal control is $u = -AB\Pi_{s-1}(Q+B^2\Pi_{s-1})^{-1}x$. This establishes the form of F and we have

$$\Pi_s^{-1} = A^{-2}\Pi_{s-1}^{-1} + B^2/QA^2.$$

This is a recurrence whose general solution is of the form $\Pi_s^{-1} = a + bA^{-2s}$ where $a = B/Q(A^2-1)$ is the particular solution, and the boundary condition $\Pi_0 = 1$ gives b. The solution is as stated.

If $A \le 1$ then $\Pi_s \to 0$ and $\Gamma_s \to A$. If $A > 1$ then $\Pi_s \to Q(A^2-1)/B^2$ and $\Gamma_s \to 1/A$.

---

2. [Solution] Suppose the cost with s attempts to go is $F_s(x) = \Pi_s x^2$, where $\Pi_0 = 1$. Then the optimality equation is

$$\begin{aligned}
F_s(x) &= \inf_u \left[EF_{s-1}(x-u+\epsilon)\right] = \inf_u \left[\Pi_{s-1}E(x-u+\epsilon)^2\right] = \Pi_{s-1}\inf_u \left[(x-u)^2 + \alpha u^2\right] \\
&= \Pi_{s-1}\inf_u \left[(1+\alpha)\left(u - \frac{x}{1+\alpha}\right)^2 - \frac{x^2}{1+\alpha} + x^2\right] = \Pi_{s-1}\left(\frac{\alpha}{1+\alpha}\right)x^2,
\end{aligned}$$

where the optimal u is $u = x/(1+\alpha)$. The minimal cost is $\Pi_s x^2$, where

$$\Pi_s = \Pi_{s-1}\left(\frac{\alpha}{1+\alpha}\right) = \left(\frac{\alpha}{1+\alpha}\right)^s.$$

---

3. [Solution] The open loop control minimizes

$$\sum_{t=0}^{h} u_t^2 + Dx_h^2 = \sum_{t=0}^{h} u_t^2 + D(x_0 + u_0 + \cdots + u_{h-1})^2.$$

Hence to be stationary with respect to u_t we require $u_t + D(x_0 + u_0 + \cdots + u_{h-1}) = 0$, implying u_t is constant, say $u_t = u$, and so $u + D(x_0 + hu) = 0$, and hence $u_t = -Dx_0/(1+hD)$.

In the closed loop case, let $F_s(x)$ be the minimal cost with time s to go in the deterministic case. Then $F_0(x) = Dx^2$ and if $F_s(x) = \pi_s x^2 + \gamma_s$ then one finds from the usual Riccati equation that

$$\pi_{s+1} = \frac{\pi_s}{1+\pi_s} \quad\Longrightarrow\quad \pi_{s+1}^{-1} = 1 + \pi_s^{-1} \quad\Longrightarrow\quad \pi_s = \frac{D}{1+sD},$$

so

$$F(x_0, 0) = F_h(x_0) = \frac{Dx_0^2}{1+hD}.$$

Also, $\gamma_{s+1} = \gamma_s + v\pi_s$. If open-loop control is used, then the control cost is unaffected by noise but x_h is changed to $x_h + \sum_{t=1}^h \epsilon_t$ giving an extra terminal cost of

$$DE\left(\sum_{t=1}^h \epsilon_t\right)^2 = hDv.$$

If closed-loop control is used then the additional cost due to noise is

$$\sum_{s=1}^h v\pi_s = \sum_{s=1}^h \frac{Dv}{1+sD} < hDv.$$

---

4. [Solution] Suppose the cost function is of the form $F_s(x) = \Pi_s x^2$. This holds at termination, with $\Pi_0(x) = 0$. The DP equation is

$$\begin{aligned}
F_s(x) &= \inf_u \left[x^2 + u^2 + EF_{s-1}(ax + \xi u)\right] \\
&= \inf_u \left[x^2 + u^2 + \Pi_{s-1}E(ax+\xi u)^2\right] \\
&= \inf_u \left[x^2 + u^2 + \Pi_{s-1}(a^2x^2 + 2abux + (b^2+\sigma^2)u^2)\right] \\
&= \inf_u \left[(1 + (b^2+\sigma^2)\Pi_{s-1})\left(u + \frac{abx\Pi_{s-1}}{1+(b^2+\sigma^2)\Pi_{s-1}}\right)^2 \right. \\
&\qquad \left. - \frac{(abx\Pi_{s-1})^2}{1+(b^2+\sigma^2)\Pi_{s-1}} + (1 + \Pi_{s-1}a^2)x^2\right]
\end{aligned}$$

Hence

$$\Pi_s = 1 + \frac{a^2\Pi_{s-1}(1+\sigma^2\Pi_{s-1})}{1+(b^2+\sigma^2)\Pi_{s-1}},$$

and the optimal control is

$$u_s = -\frac{ab\Pi_{s-1}x_s}{1+(b^2+\sigma^2)\Pi_{s-1}}.$$

This is not certainty-equivalence control because it depends on σ^2.

5. [SOLUTION] It is required to show that

$$F(x,t) = \phi(d^\top A_{h-1}\cdots A_t x_t, t)$$

for some ϕ. This is true for $t = h$. Assume true for $t+1$. Then

$$\begin{aligned} F(x,t) &= \inf_u \left[c(u,t) + F(A_t x_t + b(u,t), t+1)\right] \\ &= \inf_u \left[c(u,t) + \phi(d^\top A_{h-1}\cdots A_t x_t + d^\top A_{h-1}\cdots A_{t+1} b(u,t), t+1)\right] \end{aligned}$$

which is of the required form with

$$\phi(\xi,t) = \inf_u \left[c(u,t) + \phi(\xi + d^\top A_{h-1}\cdots A_{t+1} b(u,t), t+1)\right].$$

---

6. [SOLUTION] The entire problem can be re-written in terms of the variable $z_t = x_t + (T-t)v_t$, i.e., the value x_h would take if no further control were applied. In terms of $s = T - t$ the plant equation becomes

$$z_{t+1} = z_t + (s-1)u_t + (s-1)\epsilon_t.$$

with cost function $\sum_{t=0}^{T-1} u_t^2 + P_0 z_T^2$. Let us hypothesise that $F_{s-1}(z) = z^2\Pi_{s-1} + \gamma_{s-1}$, which is true at $s = 1$, since $F_0(z) = z^2\Pi_0$. Then

$$\begin{aligned} F_s(z) &= \inf_u \left[u^2 + EF_s(z + (s-1)u + (s-1)\epsilon)\right] \\ &= \inf_u \left[u^2 + E\left[z + (s-1)u + (s-1)\epsilon\right]^2 \Pi_{s-1} + \gamma_{s-1}\right] \\ &= \inf_u \left[u^2 + \left[(z + (s-1)u)^2 + (s-1)^2 N\right]\Pi_{s-1} + \gamma_{s-1}\right] \\ &= \inf_u \left[(1 + (s-1)^2\Pi_{s-1})\left(u + \frac{(s-1)\Pi_{s-1} z}{1 + (s-1)^2\Pi_{s-1}}\right)^2 \right. \\ &\qquad \left. - \frac{(s-1)^2\Pi_{s-1}^2 z^2}{1 + (s-1)^2\Pi_{s-1}} + \Pi_{s-1}z^2 + (s-1)^2 N\Pi_{s-1} + \gamma_{s-1}\right] \\ &= \left[-\frac{(s-1)^2\Pi_{s-1}^2 z^2}{1 + (s-1)^2\Pi_{s-1}} + \Pi_{s-1}z^2 + (s-1)^2 N\Pi_{s-1} + \gamma_{s-1}\right] \\ &= \left[\frac{\Pi_{s-1} z^2}{1 + (s-1)^2\Pi_{s-1}} + (s-1)^2 N\Pi_{s-1} + \gamma_{s-1}\right] \end{aligned}$$

Thus

$$\Pi_s = \frac{\Pi_{s-1}}{1 + (s-1)^2\Pi_{s-1}},$$

and the the optimal control is

$$u_t = -\frac{(s-1)\Pi_{s-1} z_t}{1 + (s-1)^2\Pi_{s-1}} = -(s-1)\Pi_s(x_t + s v_t).$$

By taking the reciprocal of the Riccati equation for Π_s, we have

$$\Pi_s^{-1} = \Pi_{s-1}^{-1} + (s-1) = \cdots = \Pi_0^{-1} + \sum_{i=1}^{s-1} i = \Pi_0^{-1} + \tfrac{1}{6}s(s-1)(2s-1).$$

---

7. [SOLUTION] The system can be written

$$\frac{d}{dt}\begin{bmatrix}\theta \\ \dot\theta\end{bmatrix} = \begin{bmatrix}0 & 1 \\ -\omega^2 & -2\gamma\omega\end{bmatrix}\begin{bmatrix}\theta \\ \dot\theta\end{bmatrix} + \begin{bmatrix}0 \\ \omega^2\end{bmatrix} u.$$

Hence

$$[B\ AB] = \begin{bmatrix}0 & \omega^2 \\ \omega^2 & -2\gamma\omega^3\end{bmatrix}$$

which is of rank 2 and hence the system is controllable.

---

8. [SOLUTION] For the single stick we have

$$\frac{d}{dt}\begin{bmatrix}\dot x \\ x\end{bmatrix} = \begin{bmatrix}0 & \alpha \\ 1 & 0\end{bmatrix}\begin{bmatrix}\dot x \\ x\end{bmatrix} + \begin{bmatrix}-\alpha \\ 0\end{bmatrix} u$$

Hence

$$\begin{bmatrix}B & AB\end{bmatrix} = \begin{bmatrix}-\alpha & 0 \\ 0 & -\alpha\end{bmatrix}$$

which is of full rank and so controllable. For n sticks and state variable $z = (\dot x_1, x_1, \dot x_2, x_2, \ldots, \dot x_n, x_n)^\top$ we have

$$\dot z = \begin{bmatrix} 0 & \alpha & 0 & 0 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & -\alpha & 0 & \alpha & 0 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & -\alpha & 0 & \alpha & 0 & \cdots \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \end{bmatrix} z + \begin{bmatrix} -\alpha \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \end{bmatrix} u$$

Thus

$$M = \begin{bmatrix} B & AB & A^2B & \cdots & A^{2n-2}B & A^{2n-1}B \end{bmatrix}$$

$$= \begin{bmatrix} -\alpha & 0 & -\alpha^2 & 0 & -\alpha^3 & 0 & -\alpha^4 & \cdots & -\alpha^n & 0 \\ 0 & -\alpha & 0 & -\alpha^2 & 0 & -\alpha^3 & 0 & \cdots & 0 & -\alpha^n \\ 0 & 0 & \alpha^2 & 0 & 2\alpha^3 & 0 & 3\alpha^4 & \cdots & \binom{n-1}{1}\alpha^n & 0 \\ 0 & 0 & 0 & \alpha^2 & 0 & 2\alpha^3 & 0 & \cdots & 0 & \binom{n-1}{1}\alpha^n \\ 0 & 0 & 0 & 0 & -\alpha^3 & 0 & -3\alpha^4 & \cdots & -\binom{n-1}{2}\alpha^n & 0 \\ 0 & 0 & 0 & 0 & 0 & -\alpha^3 & 0 & \cdots & 0 & -\binom{n-1}{2}\alpha^n \\ 0 & 0 & 0 & 0 & 0 & 0 & \alpha^4 & \cdots & \binom{n-1}{3}\alpha^n & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & (-1)^n\alpha^n & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & (-1)^n\alpha^n \end{bmatrix}$$

It is clear, inductively, that this matrix has rank $2n$ and so the system is controllable. Notice that since

$$z_{2n} - A^{2n}z_0 = Bu_{2n-1} + ABu_{2n-2} + \cdots + A^{2n-1}Bu_0$$

there is a control that successive brings the ends of sticks $n, n-1, \ldots$ to rest at 0. The control that minimizes $\sum_{t=0}^{2n-1} u_t^2$ has cost $z_0^\top (A^\top)^{2n}(MM^\top)^{-1}A^{2n}z_0$.

---

9. [Solution] The Riccati equation is

$$\Pi = R + A^\top \Pi A - (A^\top \Pi B + S)(Q + B^\top \Pi B)^{-1}(B^\top \Pi A + S).$$

With $A = 1$, $B = 1$, $R = 1$, $Q = 2$ and $S = 0$, this gives

$$\Pi = 1 + \Pi - \frac{\Pi^2}{\Pi + 2} \quad \Rightarrow \quad \Pi = 2$$

and $u = Kx$, with

$$\begin{aligned} K &= -(Q + B^\top \Pi B)^{-1}(B^\top \Pi A + S) \\ &= -(\Pi + 2)^{-1}\Pi = -1/2. \end{aligned}$$

The expected cost per unit time is $\mathrm{tr}(N\Pi) = 18$.

$$V = N + A^\top V A - (L + AVC^\top)(M + C^\top VC)^{-1}(L^\top + CVA^\top).$$

with $N = 9$, $L = 0$, $M = 4$, $C = 1$,

$$V = 9 + V - \frac{V^2}{4 + V} \quad \Rightarrow \quad V = 12.$$

The optimal control is $u_t = K\hat{x}_t = -(1/2)\hat{x}_t$, where $\hat{x}_t$ is the current estimate of x_t yielded by the Kalman filter

$$\hat{x}_{t+1} = A\hat{x}_t + Bu_t + H(y_{t+1} - C\hat{x}_t),$$

where $H = ((L + AVC^\top)(M + C^\top VC)^{-1} = 3/4$. So

$$\hat{x}_{t+1} = \hat{x}_t + u_t + \tfrac{3}{4}(y_{t+1} - \hat{x}_t) \quad = \quad \tfrac{3}{4}y_{t+1} + \tfrac{1}{4}\hat{x}_t + u_t.$$

---

10. [Solution] Treating t as time to go, the dynamic programming equation is

$$F_t(x) = \min_u \left\{ x^\top Rx + F_{t-1}(Ax + Bu) \right\}$$

Assuming, as given, that the optimal u is of the form $u = Kx$ and $F_t(x) = x^\top \Pi_t x$, we have

$$\Pi_t = \min_K \left\{ R + (A + BK)^\top \Pi_{t-1}(A + BK) \right\}.$$

The system is r-controllable if it is possible, from any initial x_0, to choose $u_0, \ldots, u_{r-1}$ so as to reach any prescribed value x_r at time r. Since

$$x_r = A^r x_0 + A^{r-1}Bu_0 + A^{r-2}Bu_1 + \cdots + Bu_{r-1}$$

the system is controllable if and only if $M_r = (B \mid AB \mid A^B \mid \cdots \mid A^{r-1}B)$ is of rank n. By the Caley-Hamilton theorem, if M_r is of rank n for some r then M_n is also of rank n. A system that is n-controllable is said to be controllable and this is iff M_n is of rank n.

Now $F_t(x)$ is clearly monotone increasing in t, since as t increases we simply add in more terms of cost. (It is important for this argument that there be no terminal cost). If the system is controllable it is possible to arrange that $x_n = 0$ and so no cost need be incurred after time n. Thus $F_t(x, 0)$ is bounded above uniformily in t, by the cost associated with a policy that arranges $x_n = 0$. This implies that $F_t(x) = x^\top \Pi_t x$ tends to a finite limit.

Taking x as a vector that is 0 in every component except for a 1 in the ith component, we deduce that the ith diagonal element of Π_t tends to a limit. Similarly, taking x as a vector that is 0 in every component except 1 in the ith and jth components we deduce that the (i, j) off-diagonal element of Π_t also tends to a limit. Call this limiting matrix Π. Taking limits in $\Pi_t = f(\cdots, \Pi_{t-1})$, we have $\Pi = f(\cdots, \Pi)$.

Consider

$$\hat{x}_t = A\hat{x}_{t-1} + Bu_{t-1} - H(y_t - C\hat{x}_{t-1}).$$

Taking expected values through this, we have that $\hat{x}_t$ is unbiased if $\hat{x}_{t-1}$ is unbiased. Subtracting the plant equation and substituting for y_t we have

$$\hat{x}_t - x_t = A(\hat{x}_{t-1} - x_{t-1}) - \epsilon_{t-1} - HC(x_{t-1} - \hat{x}_{t-1})$$

Thus

$$V_t = E(\hat{x}_t - x_t)(\hat{x}_t - x_t)^\top = (A + HC)V_{t-1}(A + HC)^\top + N\,.$$

By inspection this is $V_{t-1} = f(N, A^\top, C^\top, V_{t-1})$.

We have already proved that the recurrence $\Pi_t = f(R, A, B, \Pi_{t-1})$, $\Pi_0 = 0$, has a finite limit if the system is controllable, i.e., if M_n is of rank n. Therefore, simply by replacing (A, B) by $(A^\top, C^\top)$ we can deduce that $V_{t-1} = f(N, A^\top, C^\top, V_{t-1})$, $V_0 = 0$, has a finite limit if the matrix $(C^\top \mid A^\top C^\top \mid \cdots \mid (A^\top)^{n-1} C^\top)$ has rank n. The transpose of this matrix is

$$\begin{pmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{pmatrix}$$

This is the matrix that we usually use to check observability. I.e., it is of rank n iff the (noiseless) system is observable.