Solution for homework 3

Exercise 5.2

a) We have a linear-quadratic problem with imperfect state information. Thus the optimal control law is:

$$\mu_k^*(I_k) = L_k\, E\{x_k \,|\, I_k\},$$

where L_k is a gain matrix given by the Riccati formula. Since the system and cost matrices A_k, B_k, Q_k, R_k are all equal to 1:

$$\begin{aligned} L_k &= -(R_k + B_k{}' K_{k+1} B_k)^{-1} B_k' K_{k+1} A_k \\ &= \frac{-K_{k+1}}{1 + K_{k+1}} \end{aligned}$$

and with $K_N = 1$,

$$\begin{aligned} K_k &= A_k{}'[K_{k+1} - K_{k+1} B_k (R_k + B_k{}' K_{k+1} B_k)^{-1} B_k' K_{k+1}] A_k + Q_k \\ &= \frac{1 + 2K_{k+1}}{1 + K_{k+1}}. \end{aligned}$$

For this particular problem, $E\{x_k \,|\, I_k\}$ can be calculated easily and *is equal to the exact value of the state* x_k. To see this note that given x_k and I_{k+1}:

$$z_{k+1} = x_{k+1} + v_{k+1} = x_k + u_k + w_k + v_{k+1}.$$

So

$$z_{k+1} - u_k - x_k = w_k + v_{k+1}.$$

Now $w_k + v_{k+1}$ can take on four possible values: $\pm 1 \pm \frac{1}{4}$. If at time $k+1$ the known value $z_{k+1} - u_k - x_k$ comes out to be $1 \pm \frac{1}{4}$ then we know that $w_k = 1$ and $x_{k+1} = x_k + u_k + 1$ becomes known. If $z_{k+1} - u_k - x_k$ comes out to be $-1 \pm \frac{1}{4}$ then we know that $w_k = -1$ and $x_{k+1} = x_k + u_k - 1$ becomes known. Also note that, given z_0, we can compute the exact value of x_0. Thus the estimator for $E\{x_k \,|\, I_k\}$ is given by:

$$E\{x_0 \,|\, I_0\} = \begin{cases} 2, & \text{if } z_0 = 2 \pm \frac{1}{4} \\ -2, & \text{if } z_0 = -2 \pm \frac{1}{4} \end{cases}$$

$$E\{x_{k+1} \,|\, I_{k+1}\} = \begin{cases} E\{x_k \,|\, I_k\} + u_k + 1, & \text{if } z_{k+1} - E\{x_k \,|\, I_k\} - u_k = 1 \pm \frac{1}{4} \\ E\{x_k \,|\, I_k\} + u_k - 1, & \text{if } z_{k+1} - E\{x_k \,|\, I_k\} - u_k = -1 \pm \frac{1}{4} \end{cases}.$$

An alternative approach to compute $E\{x_k \,|\, I_k\}$ is based on the fact that:

$$x_k - \sum_{i=0}^{k-1} u_i = x_0 + \sum_{i=0}^{k-1} w_i \quad \in \mathit{Integers}$$

since x_0 and w_k take on integer values. So we have:

$$z_k - \sum_{i=0}^{k-1} u_i = x_k - \sum_{i=0}^{k-1} u_i + v_k \quad \in \mathit{Integer}\ \pm \frac{1}{4}.$$

Thus the estimator will be the true value of x_k which is the nearest integer to $z_k - \sum u_i$ plus $\sum u_i$.

Exercise 5.7

a) We have

$$
\begin{aligned}
p_{k+1}^j &= P(x_{k+1} = j \mid z_0, \ldots, z_{k+1}, u_0, \ldots, u_k) \\
&= P(x_{k+1} = j \mid I_{k+1}) \\
&= \frac{P(x_{k+1} = j, z_{k+1} \mid I_k, u_k)}{P(z_{k+1} \mid I_k, u_k)} \\
&= \frac{\sum_{i=1}^n P(x_k = i) P(x_{k+1} = j \mid x_k = i, u_k) P(z_{k+1} \mid u_k, x_{k+1} = j)}{\sum_{s=1}^n \sum_{i=1}^n P(x_k = i) P(x_{k+1} = s \mid x_k = i, u_k) P(z_{k+1} \mid u_k, x_{k+1} = s)} \\
&= \frac{\sum_{i=1}^n p_k^i p_{ij}(u_k) r_j(u_k, z_{k+1})}{\sum_{s=1}^n \sum_{i=1}^n p_k^i p_{is}(u_k) r_s(u_k, z_{k+1})}.
\end{aligned}
$$

Rewriting p_{k+1}^j in vector form, we have

$$p_{k+1}^j = \frac{r_j(u_k, z_{k+1})[P(u_k)'P_k]_j}{\sum_{s=1}^n r_s(u_k, z_{k+1})[P(u_k)'P_k]_s}, \qquad j = 1, \ldots, n.$$

Therefore,

$$P_{k+1} = \frac{[r(u_k, z_{k+1})] * [P(u_k)'P_k]}{r(u_k, z_{k+1})'P(u_k)'P_k}.$$

b) The DP algorithm for this system is:

$$\begin{aligned}
\bar{J}_{N-1}(P_{N-1}) &= \min_u \left\{ \sum_{i=1}^n p_{N-1}^i \sum_{j=1}^n p_{ij}(u) g_{N-1}(i, u, j) \right\} \\
&= \min_u \left\{ \sum_{i=1}^n p_{N-1}^i \big[G_{N-1}(u)\big]_i \right\} \\
&= \min_u \big\{ P_{N-1}' G_{N-1}(u) \big\}
\end{aligned}$$

$$\begin{aligned}
\bar{J}_k(P_k) &= \min_u \left\{ \sum_{i=1}^n p_k^i \sum_{j=1}^n p_{ij}(u) g_k(i, u, j) + \sum_{i=1}^n p_k^i \sum_{j=1}^n p_{ij}(u) \sum_{\theta=1}^q r_j(u, \theta) \bar{J}_{k+1}(P_{k+1} \mid P_k, u, \theta) \right\} \\
&= \min_u \left\{ P_k' G_k(u) + \sum_{\theta=1}^q r(u, \theta)' P(u)' P_k \bar{J}_{k+1} \left[\frac{[r(u, \theta)] * [P(u)'P_k]}{r(u, \theta)'P(u)'P_k} \right] \right\}.
\end{aligned}$$

c) For $k = N - 1$,

$$\begin{aligned}
\bar{J}_{N-1}(\lambda P_{N-1}') &= \min_u \big\{ \lambda P_{N-1}' G_{N-1}(u) \big\} \\
&= \min_u \big\{ \sum_{i=1}^n \lambda p_{N-1}^i \big[G_{N-1}(u)\big]_i \big\} \\
&= \min_u \{ \lambda \sum_{i=1}^n p_{N-1}^i [G_{N-1}(u)]_i \} \\
&= \lambda \min_u \big\{ \sum_{i=1}^n p_{N-1}^i [G_{N-1}(u)]_i \big\} \\
&= \lambda \min_u \big\{ \sum_{i=1}^n p_{N-1}^i [G_{N-1}(u)]_i \big\} \\
&= \lambda \bar{J}_{N-1}(P_{N-1}).
\end{aligned}$$

Now assume $\bar{J}_k(\lambda P_k) = \lambda \bar{J}_k(P_k)$. Then,

$$\begin{aligned}
\bar{J}_{k-1}(\lambda P_{k-1}') &= \min_u \left\{ \lambda P_{k-1}' G_{k-1}(u) + \sum_{\theta=1}^q r(u, \theta)' P(u)' \lambda P_{k-1} \bar{J}_k(P_k | P_{k-1}, u, \theta) \right\} \\
&= \min_u \left\{ \lambda P_{k-1}' G_{k-1}(u) + \lambda \sum_{\theta=1}^q r(u, \theta)' P(u)' P_{k-1} \bar{J}_k(P_k | P_{k-1}, u, \theta) \right\} \\
&= \lambda \min_u \left\{ P_{k-1}' G_{k-1}(u) + \sum_{\theta=1}^q r(u, \theta)' P(u)' P_{k-1} \bar{J}_k(P_k | P_{k-1}, u, \theta) \right\} \\
&= \lambda \bar{J}_{k-1}(P_{k-1}). \qquad \text{Q.E.D.}
\end{aligned}$$

12

For any u, $r(u,\theta)'P(u)'P_k$ is a scalar. Therefore, letting $\lambda = r(u,\theta)'P(u)'P_k$, we have

$$\bar{J}_k(P_k) = \min_u \left\{ P_k' G_k(u) + \sum_{\theta=1}^{q} r(u,\theta)'P(u)'P_k \bar{J}_{k+1}\left[\frac{[r(u,\theta)] * [P(u)'P_k]}{r(u,\theta)'P(u)'P_k}\right]\right\}$$
$$= \min_u \left[P_k' G_k(u) + \sum_{\theta=1}^{q} \bar{J}_{k+1}\big([r(u,\theta)] * [P(u)'P_k]\big)\right].$$

d) For $k = N-1$, we have $\bar{J}_{N-1}(P_{N-1}) = \min_u [P_{N-1}' G_{N-1}(u)]$, and so $\bar{J}_{N-1}(P_{N-1})$ has the desired form

$$\bar{J}_{N-1}(P_{N-1}) = \min\big[P_{N-1}'\alpha_{N-1}^1, \ldots, P_{N-1}'\alpha_{N-1}^m\big],$$

where $\alpha_{N-1}^j = G_{N-1}(u^j)$ and u^j is the jth element of the control constraint set.

Assume that

$$\bar{J}_{k+1}(P_{k+1}) = \min\big[P_{k+1}'\alpha_{k+1}^1, \ldots, P_{k+1}'\alpha_{k+1}^{m_{k+1}}\big].$$

Then, using the expression from part (c) for $\bar{J}_k(P_k)$,

$$\bar{J}_k(P_k) = \min_u \left[P_k' G_k(u) + \sum_{\theta=1}^{q} \bar{J}_{k+1}\big([r(u,\theta)] * [P(u)'P_k]\big)\right]$$
$$= \min_u \left[P_k' G_k(u) + \sum_{\theta=1}^{q} \min_{m=1,\ldots,m_{k+1}} \left[\big\{[r(u,\theta)] * [P(u)'P_k]\big\}'\alpha_{k+1}^m\right]\right]$$
$$= \min_u \left[P_k' G_k(u) + \sum_{\theta=1}^{q} \min_{m=1,\ldots,m_{k+1}} \left[P_k' P(u) r(u,\theta)'\alpha_{k+1}^m\right]\right]$$
$$= \min_u \left[P_k' \left\{ G_k(u) + \sum_{\theta=1}^{q} \min_{m=1,\ldots,m_{k+1}} \left[P(u) r(u,\theta)'\alpha_{k+1}^m\right]\right\}\right]$$
$$= \min\left[P_k'\alpha_k^1, \ldots, P_k'\alpha_k^{m_k}\right],$$

where $\alpha_k^1, \ldots, \alpha_k^{m_k}$ are all possible vectors of the form

$$G_k(u) + \sum_{\theta=1}^{q} P(u) r(u,\theta)'\alpha_{k+1}^{m_{u,\theta}},$$

as u ranges over the finite set of controls, θ ranges over the set of observation vector indexes $\{1, \ldots, q\}$, and $m_{u,\theta}$ ranges over the set of indexes $\{1, \ldots, m_{k+1}\}$. The induction is thus complete.

For a quick way to understand the preceding proof, based on polyhedral concavity notions, note that the conclusion is equivalent to asserting that $\bar{J}_k(P_k)$ is a positively homogeneous, concave polyhedral function. The preceding induction argument amounts to showing that the DP formula of part (c) preserves the positively homogeneous, concave polyhedral property of $\bar{J}_{k+1}(P_{k+1})$. This is indeed evident from the formula, since taking minima and nonnegative weighted sums of positively homogeneous, concave polyhedral functions results in a positively homogeneous, concave polyhedral function.

Exercise 5.14

a) The state is (x_k, d_k), where x_k is the current offer under consideration and d_k takes the value 1 or 2 depending on whether the common distribution of the system disturbance, w_k, is F_1 or F_2. The variable d_k stays constant (i.e., satisfies $d_{k+1} = d_k$ for all k), but is not observed perfectly. Instead, the sample offer values $w_0, w_1, \ldots$ are observed ($w_k = x_{k+1}$), and provide information regarding the value of d_k. In particular, given the a priori probability q and the demand values $w_0, \ldots, w_{k-1}$, we can calculate the conditional probability that w_k will be generated according to F_1.

b) A suitable sufficient statistic is (x_k, q_k), where

$$q_k = P(d_k = 1 \mid w_0, \ldots, w_{k-1}).$$

The conditional probability q_k evolves according to

$$q_{k+1} = \frac{q_k F_1(w_k)}{q_k F_1(w_k) + (1 - q_k) F_2(w_k)}, \qquad q_0 = q,$$

where $F_i(w_k)$ denotes probability under the distribution F_i, and assuming that w_k can take a finite number of values under the distributions F_1 and F_2. Let $w^1, w^2, \ldots, w^n$ be the possible values w_k can take under either distribution.

We have the following DP algorithm:

$$\begin{aligned}
J_N(x_N, q_N) &= x_N \\
J_k(x_k, q_k) &= \max\Big[(1+r)^{N-k} x_k, E\big\{J_{k+1}(x_{k+1}, q_{k+1})\big\}\Big] \\
&= \max\Bigg[(1+r)^{N-k} x_k, \\
&\sum_{i=1}^{n} \big(q_k F_1(w^i) + (1-q_k) F_2(w^i)\big) J_{k+1}\left(w^i, \frac{q_k F_1(w^i)}{q_k F_1(w^i) + (1-q_k) F_2(w^i)}\right)\Bigg]
\end{aligned}$$

As in the text, we renormalize the cost-to go so that each stage has the same cost function for stopping. Let

$$V_k(x_k, q_k) = \frac{J_k(x_k, q_k)}{(1+r)^{N-k}}.$$

Then we have

$$V_N(x_N, q_N) = x_N,$$

$$V_k(x_k, q_k) = \max\Big[x_k, \alpha_k(q_k)\Big],$$

where

$$\alpha_k(q_k) = (1+r)^{-1} \sum_{i=1}^{n} \big(q_k F_1(w^i) + (1-q_k) F_2(w^i)\big) V_{k+1}\left(w^i, \frac{q_k F_1(w^i)}{q_k F_1(w^i) + (1-q_k) F_2(w^i)}\right),$$

which is independent of x_k. Each stopping set therefore has a threshold format, $T_k = \{x \mid x \geq \alpha_k(q_k)\}$, where the threshold depends on q_k.

Because $V_{N-1}(x, q) \geq V_N(x, q)$ for all x, q, we have by the monotonicity property for stationary problems that $V_k(x, q) \geq V_{k+1}(x, q)$ for all x, q, k, which implies $\alpha_k(q) \geq \alpha_{k+1}(q)$ for all q, k.