6.3. Partial observations and the separation principle

We now consider control problems associated with the full state-space model

$$x_{k+1} = A(k)x_k + B(k)u_k + C(k)w_k \tag{6.3.1}$$

$$y_k = H(k)x_k + G(k)w_k. \tag{6.3.2}$$

As before, the initial state x_0 has mean and covariance m_0, P_0 and is uncorrelated with w_k. In this case the state x_k cannot be measured directly, but 'noisy observations' $y^k = (y_0, y_1, \ldots, y_k)$ are available at time k. Thus the control u_k will be a feedback function of the form

$$u_k = u_k(y^k). \tag{6.3.3}$$

This is the 'full LQG problem'. The difficulty here is, of course, that knowledge of y^k does not (except in special cases) determine x_k exactly, and the current state x_k is just what is needed for controlling the system at time k. We deal with this by replacing the state-space model (6.3.1), (6.3.2) by the corresponding *innovations representation.* As discussed in Section 3.4, this provides an equivalent model in the form

$$\hat{x}_{k+1|k} = A(k)\hat{x}_{k|k-1} + B(k)u_k + K(k)v_k \tag{6.3.4}$$

where the innovations process v_k is given by

$$v_k = y_k - H(k)\hat{x}_{k|k-1} \tag{6.3.5}$$

so that y_k satisfies

$$y_k = H(k)\hat{x}_{k|k-1} + v_k. \tag{6.3.6}$$

The Kalman gain $K(k)$ is given by (3.3.5). The new 'state' of the system is $\hat{x}_{k|k-1}$ and this *is* determined exactly by y^{k-1}. We thus reduce the situation to one in which the state is known, and can then apply the results of the previous section to determine optimal control policies. First, however, the status of the innovations representation (6.3.4), (6.3.6) must be clarified. We do this before continuing with our discussion of optimal control problems in Section 6.3.2 below.

6.3.1 The Kalman filter for systems with feedback control

In the derivation of the Kalman filtering formulae in Section 3.3 it was assumed that $\{w_k\}$ was a weak-sense white noise (w_k and w_l uncorrelated for $k \neq l$) and that $\{u_k\}$ was a *deterministic sequence.* Under these conditions $\hat{x}_{k|k-1}$ given by (6.3.4) is the best linear (more precisely, affine) estimator of x_k given y^{k-1}, and the input/output properties of the model (6.3.4), (6.3.6) are identical to those of the original model (6.3.1), (6.3.2). Now, however, we wish to consider controls u_k which are not deterministic but which are feedback functions as in (6.3.3). Further, there is no reason why $u_k(y^k)$ should be a linear function of y^k. Suppose in fact that this function is nonlinear. Combining (6.3.3)–(6.3.5), we see that $\hat{x}_{k|k-1}$ satisfies

$$\hat{x}_{k+1|k} = A(k)\hat{x}_{k|k-1} + B(k)u_k(y^k) + K(k)(y_k - H(x)\hat{x}_{k|k-1}). \quad (6.3.7)$$

Given the sequence $y^j = (y_0, y_1, \ldots, y_j)$, one can use this equation for $k = 0, 1, \ldots, j$ to compute $\hat{x}_{j+1|j}$. Thus $\hat{x}_{j+1|j}$ is a function of y^j, say

$$\hat{x}_{j+1|j} = g_j(y^j).$$

Now g_j is a nonlinear function, due to the nonlinearity of u_k in (6.3.7). So $\hat{x}_{j+1|j}$ cannot possibly be the best *linear* estimator of x_{j+1} given y^j, as it would be were u_k deterministic. To get round this apparently awkward fact, we use the alternative interpretation of the Kalman filter, namely that if the w_k are independent *normal* random vectors and x_0 is normal, then $\hat{x}_{j+1|j}$ is the *conditional expectation* of x_{j+1} given y^j. The advantage of this formulation is that there is no requirement that a conditional expectation should be a linear function of the conditioning random variables.

Theorem 6.3.1

Suppose that, in the model (6.3.1), (6.3.2), $x_0, w_0, w_1, \ldots$ are normally distributed and that u_k is a feedback control as in (6.3.3). Let $\hat{x}_{k|k-1}$ be generated by the Kalman filter equation of Theorem (3.3.1). Then

$$\hat{x}_{k|k-1} = E[x_k | y^{k-1}]. \quad (6.3.8)$$

The innovations process (6.3.5) is a *normal* white-noise sequence.

PROOF The proof relies on Proposition 1.1.6 which shows that

$$E[x_{j+1}|y^j] = E[x_{j+1}|\bar{y}^j]$$

if $y^j, \bar{y}^j$ are random vectors which are related to each other in a one-to-one way, i.e. there are functions h_j, h_j^{-1} such that

$$\bar{y}^j = h_j(y^j), \quad y^j = h_j^{-1}(\bar{y}^j).$$

As in Section 3.4, let us write the state x_k in (6.3.1) as $x_k = \bar{x}_k + x_k^*$, and correspondingly $y_k = \bar{y}_k + y_k^*$, where $\bar{x}_k$, x_k^*, $\bar{y}_k$, y_k^* satisfy:

$$\left.\begin{aligned} \bar{x}_{k+1} &= A(k)\bar{x}_k + C(k)w_k, \quad \bar{x}_0 = x_0 - m_0 \\ \bar{y}_k &= H(k)\bar{x}_k + G(k)w_k \end{aligned}\right\} \tag{6.3.9}$$

$$\left.\begin{aligned} x_{k+1}^* &= A(k)x_k^* + B(k)u_k(y^k), \quad x_0^* = m_0 \\ y_k^* &= H(k)x_k^*. \end{aligned}\right\} \tag{6.3.10}$$

Equations (6.3.9) are linear, so that $\bar{x}_{k+1}$, $\bar{y}_k$ are zero-mean normal random vectors for all k. x_{k+1}^* and y_k^* are random vectors which depend on $\bar{y}^k$ since $u_k(y^k) = u_k(\bar{y}^k + y^{*k})$. Applying the standard Kalman filter results from Section 3.3 we see that $\hat{\bar{x}}_{k+1|k} := E[\bar{x}_{k+1}|\bar{y}^k]$ satisfies

$$\hat{\bar{x}}_{k+1|k} = A(k)\hat{\bar{x}}_{k|k-1} + K(k)(\bar{y}_k - H(k)\hat{\bar{x}}_{k|k-1}) \tag{6.3.11}$$

where $K(k)$ is given by (3.3.5). We cannot obtain (6.3.4) immediately by adding (6.3.11) to (6.3.10) because the conditioning random variable is $\bar{y}^k$ and not y^k as required. However, $\bar{y}^k$ and y^k are equivalent in the sense mentioned earlier. Indeed, plainly from (6.3.10), x_k^*, and hence y_k^*, is determined by $y^{k-1} = (y_0, y_1, \ldots, y_{k-1})$. Thus

$$\bar{y}_k = y_k - y_k^* =: h_k(y^k).$$

Conversely, suppose $\bar{y}^k = (\bar{y}_0, \bar{y}_1, \ldots, \bar{y}_k)$ is given; then y_k is determined. We show this by induction. Suppose that for $j = 0, 1, \ldots, k$ there are functions f_j such that

$$y_j = f_j(\bar{y}^j). \tag{6.3.12}$$

Then given $\bar{y}^k$ we can calculate y_j, $0 \le j \le k$, and hence y_{k+1}^*, using (6.3.10). But now

$$y_{k+1} = \bar{y}_{k+1} + y_{k+1}^* =: f_{k+1}(\bar{y}^{k+1})$$

Thus (6.3.12) holds for $j = k + 1$. At time zero,

$$y_0^* = H(0)x_0^* = H(0)m_0$$

and m_0 is known, so that

$$y_0 = H(0)m_0 + \bar{y}_0 =: f_0(\bar{y}_0).$$

Thus (6.3.12) holds for all j, and $f_j = h_j^{-1}$.

This argument shows that $\bar{y}^k$ and y^k are obtained from each other in a one-to-one fashion, and hence that

$$\hat{\bar{x}}_{k+1|k} = E[\bar{x}_{k+1}|\bar{y}^k] = E[\bar{x}_{k+1}|y^k].$$

Now x^*_{k+1} is a function of y^k, so that

$$E[x^*_{k+1} | y^k] = x^*_{k+1}.$$

Combining these relations, we obtain

$$\begin{aligned} E[x_{k+1} | y^k] &= E[x^*_{k+1} + \bar{x}_{k+1} | y^k] \\ &= x^*_{k+1} + \hat{\bar{x}}_{k+1|k}. \end{aligned}$$

Adding the equations (6.3.10) and (6.3.11) shows that $\hat{x}_{k+1|k} := E[x_{k+1} | y^k]$ satisfies (6.3.4). Thus (6.3.4) is indeed the Kalman filter when u_k is a feedback control, as long as the disturbance process w_k is a normal white-noise process. As regards the innovations process v_k, note that

$$\begin{aligned} v_k &= y_k - H(k)\hat{x}_{k|k-1} \\ &= \bar{y}_k + y^*_k - H(k)(x^*_k + \hat{\bar{x}}_{k|k-1}) \\ &= \bar{y}_k - H(k)\hat{\bar{x}}_{k|k-1}. \end{aligned}$$

Thus v_k coincides with the innovations process corresponding to the control-free system (6.3.9). It is therefore a normal white-noise process with covariance

$$E[v_k v_k^{\mathrm{T}}] = H(k)P(k)H^{\mathrm{T}}(k) + G(k)G^{\mathrm{T}}(k) \tag{6.3.13}$$

as in Section 3.4. □

It is perhaps worth pointing out that, even if w_k is a normal white-noise process, the state process x_k is not necessarily normal, since (6.3.1), (6.3.2) determine x_k as a possibly nonlinear function of w^{k-1}. However, the *conditional distribution* of x_k given y^{k-1} is normal, since x_k has the representation

$$\begin{aligned} x_k &= x^*_k + \hat{\bar{x}}_{k|k-1} + \tilde{\bar{x}}_{k|k-1} \\ &= \hat{x}_{k|k-1} + \tilde{\bar{x}}_{k|k-1} \end{aligned}$$

where $\tilde{\bar{x}}_{k|k-1} = \bar{x}_k - \hat{\bar{x}}_{k|k-1}$ is a normal random vector with mean 0 and covariance $P(k)$ given by (3.3.6). Thus the conditional distribution of x_k given y^{k-1} is $N(\hat{x}_{k|k-1}, P(k))$.

6.3.2 The linear regulator problem

Let us now return to the control problem of choosing u_k to minimize the cost

$$C_N(u) = E\left(\sum_{k=0}^{N-1} \|D(k)x_k + F(k)u_k\|^2 + x_N^{\mathrm{T}} Q x_N \right). \tag{6.3.14}$$

This is the same form of cost as in Section 6.2 but a different class of controls is involved. In this section we shall consider feedback controls of the form

$$u_k = u_k(y^{k-1}) \tag{6.3.15}$$

rather than $u_k(y^k)$ as discussed above. Controls (6.3.15) are of course a sub-class of those previously considered – we are now insisting that the control u_k should depend on the observations y_j for times j up to, but not including k, whereas previously dependence on y_k also was allowed. This restriction is introduced for two reasons. Practically, it means that 'instant' data processing is then not required: at time k we record the new observation y_k, and apply the control $u_k(y^{k-1})$ which can be computed somewhat in advance since it does not depend on y_k. Mathematically, controls (6.3.15) are related, as will be seen below, to our formulation of the Kalman filter as a predictor, giving the best estimate $\hat{x}_{k|k-1}$ of x_k given y^{k-1}. Analogous results can be obtained for controls $u_k(y^k)$, but these involve the Kalman filter in the form which computes the *current* state estimate $\hat{x}_{k|k}$, and this is somewhat more complicated.

The cost $C_N(u)$ in (6.3.14) is expressed in terms involving the state variables x_k; we wish, however, to use the innovations representation (6.3.4) in which the state variable is $\hat{x}_{k|k-1}$. The first task is therefore to re-express $C_N(u)$ in a way which involves $\hat{x}_{k|k-1}$ rather than x_k, and this is done by introducing conditional expectations as follows:

$$C_N(u) = E\left(\sum_{k=0}^{N-1} E[\|D(k)x_k + F(k)u_k\|^2 | y^{k-1}] + E[x_N^{\mathrm{T}} Q x_N | y^{N-1}]\right). \tag{6.3.16}$$

Now x_k can be expressed in the form

$$x_k = \hat{x}_{k|k-1} + \tilde{x}_{k|k-1}$$

where $\hat{x}_{k|k-1}$ is a function of y^{k-1} and the estimation error $\tilde{x}_{k|k-1}$ is independent of y^{k-1} with distribution $N(0, P(k))$. We can simplify the terms in (6.3.16) using this fact and properties of conditional expectations. The last term is:

$$\begin{aligned} E[x_N^{\mathrm{T}} Q x_N | y^{N-1}] &= E[(\hat{x}_{N|N-1} + \tilde{x}_{N|N-1})^{\mathrm{T}} Q(\hat{x}_{N|N-1} + \tilde{x}_{N|N-1}) | y^{N-1}] \\ &= \hat{x}_{N|N-1}^{\mathrm{T}} Q \hat{x}_{N|N-1} + E[\tilde{x}_{N|N-1}^{\mathrm{T}} Q \tilde{x}_{N|N-1} | y^{N-1}] \\ &= \hat{x}_{N|N-1}^{\mathrm{T}} Q \hat{x}_{N|N-1} + \mathrm{tr}[P(N)Q]. \end{aligned}$$

Similarly the kth term in the sum becomes

$$\begin{aligned}&E[(D(k)\hat{x}_{k|k-1} + F(k)u_k + D(k)\tilde{x}_{k|k-1})^{\mathrm{T}}\\&\quad\cdot(D(k)\hat{x}_{k|k-1} + F(k)u_k + D(k)\tilde{x}_{k|k-1})|y^{k-1}]\\&= \|D(k)\hat{x}_{k|k-1} + F(k)u_k\|^2 + \mathrm{tr}[P(k)D^{\mathrm{T}}(k)D(k)]\end{aligned}$$

where we have used the fact that u_k is a function of y^{k-1}. Thus

$$\begin{aligned}C_N(u) = E\left(\sum_{k=0}^{N-1} \|D(k)\hat{x}_{k|k-1} + F(k)u_k\|^2 + \hat{x}^{\mathrm{T}}_{N|N-1}Q\hat{x}_{N|N-1}\right)\\ + \sum_{k=0}^{N-1} \mathrm{tr}[D(k)P(k)D^{\mathrm{T}}(k)] + \mathrm{tr}[P(N)Q]. \qquad (6.3.17)\end{aligned}$$

This expresses $C_N(u)$ in a way which involves the state $\hat{x}_{k|k-1}$ of the innovations representation. The important thing to notice about this expression is that the first term is identical to the original expression (6.3.14) with x_k replaced by $\hat{x}_{k|k-1}$, and that the remaining two terms are constants which do not depend in any way on the choice of u_k. Thus minimizing $C_N(u)$ is equivalent to minimizing

$$E\left(\sum_{k=0}^{N-1} \|D(k)\hat{x}_{k|k-1} + F(k)u_k\|^2 + \hat{x}^{\mathrm{T}}_{N|N-1}Q\hat{x}_{N|N-1}\right) \qquad (6.3.18)$$

where the dynamics of $\hat{x}_{k|k-1}$ are given by (6.3.4), namely

$$\hat{x}_{k|k-1} = A(k)\hat{x}_{k|k-1} + B(k)u_k + K(k)v_k. \qquad (6.3.19)$$

Since the innovations process v_k is a sequence of independent normal random variables, the problem (6.3.18)–(6.3.19) is the standard 'completely observable' regulator problem considered in the previous section. All coefficients are as before except for the 'noise' term $K(k)v_k$ in (6.3.19). However, it was noted in Section 6.2 that the optimal control for the linear regulator does not depend on the noise covariance. Therefore the optimal control coefficients are the same as in the completely observable case. We have obtained the following result:

Theorem 6.3.2

The optimal control for the noisy observations problem (6.3.1), (6.3.2), (6.3.14) is

$$\hat{u}^1_k = -M(k)\hat{x}_{k|k-1} \qquad (6.3.20)$$

where $M(k)$ is given as before by (6.1.16). The cost of this policy is

$$C_N(\hat{u}^1) = m_0^{\mathrm{T}} S(0) m_0 + \operatorname{tr}[P(N)Q] + \sum_{k=0}^{N-1} \operatorname{tr}[D(k)P(k)D^{\mathrm{T}}(k) + G(k)G^{\mathrm{T}}(k))K^{\mathrm{T}}(k)S(k+1)]. + K(k)(H(k)P(k)H^{\mathrm{T}}(k) \tag{6.3.21}$$

PROOF Only the expression (6.3.21) for the optimal cost remains to be verified. We use the expression (6.2.9) for the completely observable case. First, note that the initial condition for (6.3.19) is deterministic: $\hat{x}_{0|-1} = 0$. Next, consider the contribution of the 'noise' term $K(k)v_k$. Define

$$\tilde{v}_k = [H(k)P(k)H^{\mathrm{T}}(k) + G(k)G^{\mathrm{T}}(k)]^{-1/2} v_k$$

(the inverse exists since by our standing assumptions $G(k)G^{\mathrm{T}}(k) > 0$). From (6.3.13) we see that $E[\tilde{v}_k \tilde{v}_k^{\mathrm{T}}] = I$, so that $\tilde{v}_k$ is a normalized white-noise process, and (6.3.19) can be written

$$\hat{x}_{k+1|k} = A(k)\hat{x}_{k|k-1} + B(k)u_k + K(k)[H(k)P(k)H^{\mathrm{T}}(k) + G(k)G^{\mathrm{T}}(k)]^{1/2}\tilde{v}_k.$$

This is now in the standard form of (6.2.1) with a new 'C-matrix' $K[HPH^{\mathrm{T}} + GG^{\mathrm{T}}]^{1/2}$ and we can read off the optimal cost from (6.2.9). Remembering that the two constant terms from (6.3.17) must also be included, we obtain (6.3.21). □

Let us summarize the computations needed in order to implement the control policy described in Theorem 6.3.2. They are as follows:

(a) Solve the matrix Riccati equation of dynamic programming backwards from the terminal time to give matrices $S(N), \ldots, S(0)$:

$$\begin{aligned} S(k) &= A^{\mathrm{T}}(k)S(k+1)A(k) + D^{\mathrm{T}}(k)D(k) \\ &\quad - [A^{\mathrm{T}}(k)S(k+1)B(k) + D^{\mathrm{T}}(k)F(k)] \\ &\quad [B^{\mathrm{T}}(k)S(k+1)B(k) + F^{\mathrm{T}}(k)F(k)]^{-1} \\ &\quad [B^{\mathrm{T}}(k)S(k+1)A(k) + F^{\mathrm{T}}(k)D(k)] \\ S(N) &= Q. \end{aligned} \tag{6.3.22}$$

This determines the feedback matrices

$$M(k) = [B^{\mathrm{T}}(k)S(k+1)B(k) + F^{\mathrm{T}}(k)F(k)]^{-1} [B^{\mathrm{T}}(k)S(k+1)A(k) + F^{\mathrm{T}}(k)D(k)].$$

(b) Solve the matrix Riccati equation of Kalman filtering forwards from the initial time to give matrices $P(0), \ldots, P(N)$:

$$\begin{aligned} P(k+1) = {} & A(k)P(k)A^{\mathrm{T}}(k) + C(k)C^{\mathrm{T}}(k) - [A(k)P(k)H^{\mathrm{T}}(k) + C(k)G^{\mathrm{T}}(k)] \\ & H(k)P(k)H^{\mathrm{T}}(k) + G(k)G^{\mathrm{T}}(k)]^{-1} \\ & [H(k)P(k)A^{\mathrm{T}}(k) + G(k)C^{\mathrm{T}}(k)] \\ P(0) = {} & P_0. \end{aligned} \tag{6.3.23}$$

This determines the Kalman gain matrices

$$K(k) = [A(k)P(k)H^{\mathrm{T}}(k) + C(k)G^{\mathrm{T}}(k)][H(k)P(k)H^{\mathrm{T}}(k) + G(k)G^{\mathrm{T}}(k)]^{-1}.$$

It is important to notice that these computations refer *independently* to the control and filtering problems respectively, in that (a) involves the 'cost' parameters Q, $D(k)$, $F(k)$ but not the 'noise' parameters P_0, $C(k)$, $G(k)$, whereas the converse is true in the case of (b).

The property that the optimal control takes the form $\hat{u}^1(k) = -M(k)\hat{x}_{k|k-1}$ where $M(k)$ is the same as in the deterministic or complete observation cases, expresses the so-called 'certainty-equivalence principle' which, put in another way, states that, optimally, the controller *acts as if* the state estimate $\hat{x}_{k|k-1}$ were equal to the true state x_k *with certainty*. Of course, the controller knows that this is not the case, but no other admissible strategy will give better performance.

That $M(k)$ is unchanged in the presence of observation noise is entirely due to the quadratic cost criterion which ensures that the cost function for the problem in innovations form is, apart from a fixed constant, the same as that in the original form. On the other hand, the fact that the intermediate statistic to be computed is $\hat{x}_{k|k-1}$, regardless of cost parameters, is a property which extends to more general forms of cost function. To see this, recall that whatever admissible control is applied, the conditional distribution of x_k given y^{k-1} is $N(\hat{x}_{k|k-1}, P(k))$. Now suppose that the cost to be minimized takes a general form similar to (6.1.14), i.e.

$$C_N(u) = E\left(\sum_{k=0}^{N-1} l(k, x_k, u_k) + g(x_N)\right)$$

where l and g are, say, bounded functions. Introducing intermediate conditional expectations, we can express $C_N(u)$ as

$$C_N(u) = E\left(\sum_{k=0}^{N-1} E[l(k, x_k, u_k)|y^{k-1}] + E[g(x_N)|y^{N-1}]\right).$$

The conditional expectation can now be evaluated by integrating with respect to the conditional distribution. This gives

$$E[l(k, x_k, u_k)|y^{k-1}] = \hat{l}(k, \hat{x}_{k|k-1}, u_k)$$

and

$$E[g(x_N)|y^{N-1}] = \hat{g}(\hat{x}_{N|N-1})$$

where

$$\hat{l}(k, \hat{x}, u) = \int_{\mathbb{R}^n} l(k, z, u) \frac{1}{(2\pi)^{n/2}(\det(P(k)))^{1/2}} \cdot \exp((z - \hat{x})^{\mathrm{T}} P^{-1}(k)(z - \hat{x}))\,\mathrm{d}z$$

$$\hat{g}(\hat{x}) = \int_{\mathbb{R}^n} g(z) \frac{1}{(2\pi)^{n/2}(\det(P(N)))^{1/2}} \cdot \exp((z - \hat{x})^{\mathrm{T}} P^{-1}(N)(z - \hat{x}))\,\mathrm{d}z.$$

Thus

$$C_N(u) = E\left(\sum_{k=0}^{N-1} \hat{l}(k, \hat{x}_{k|k-1}, u_k) + \hat{g}(\hat{x}_{N|N-1}) \right). \qquad (6.3.24)$$

The problem (6.3.19), (6.3.24) is now in innovations form and can be solved by dynamic programming. Define functions $W_0, \ldots, W_N$ by

$$W_N(\hat{x}) = \hat{g}(\hat{x})$$
$$W_k(\hat{x}) = \min_v \{\hat{l}(k, \hat{x}, v) + E^{(v)} W_k(A\hat{x} + Bv + K(k)v_k)$$
$$k = N - 1, \ldots, 0 \qquad (6.3.25)$$

where $E^{(v)}$ denotes expectation taken over the distribution of v_k, which is $N(0, HP(k)H^{\mathrm{T}} + GG^{\mathrm{T}})$. Let $\hat{u}^1(k, \hat{x})$ be a value of v which achieves the minimum in (6.3.25). Then the optimal control is

$$\hat{u}_k^1 = \hat{u}^1(k, \hat{x}_{k|k-1})$$

with minimal cost

$$C_N(\hat{u}^1) = W_0(m_0).$$

This can be checked by the same sort of 'verification theorem' proved earlier. Thus is this general problem the 'data processing' still consists of calculating $\hat{x}_{k|k-1}$ via the Kalman filter, but the control function $\hat{u}^1(k, \hat{x})$ is not related in any simple way to the control function $u^1(k, x)$ which is optimal in the case of complete observations.

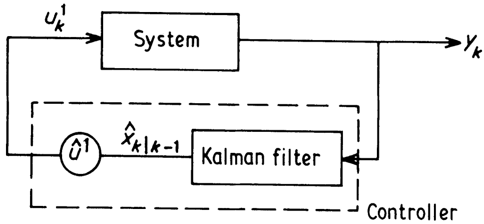

Fig. 6.2

In summary, we see that the optimal controller *separates* into two parts, a *filtering stage* and a *control stage* as shown in Fig. 6.2. The filtering stage is *always the same* regardless of the control objective. This is the *separation principle.* The certainty-equivalence principle applies when $\hat{u}^1(k, x_k)$ is the optimal completely observable control, but this is a much more special property which holds only in the quadratic cost case.

These results point to a general cybernetic principle, namely that when systems are to be controlled on the basis of noisy measurements the true 'state' of the system which is relevant for control is the *conditional distribution of the original state given the observations.* Note that in the LQG problem this is completely determined by $\hat{x}_{k|k-1}$ since the conditional distribution is $N(\hat{x}_{k|k-1}, P(k))$ and $P(k)$ does not depend on the observations. Thus the Kalman filter in effect updates the conditional distribution of x_k given y^{k-1}. The problem can be solved in an effective way because of the simple parametrization of the conditional density and the fact that there is an efficient algorithm – the Kalman filter – for updating the parameter $\hat{x}_{k|k-1}$. More general problems typically involve extensive computation due to the lack of any low-dimensional statistic characterizing the conditional distributions.

6.3.3 Discounted costs and the infinite-time problem

In this section we will assume that the system matrices A, B, H, C, G are time-invariant, that $D(k) = \rho^{k/2}D$, $F(k) = \rho^{k/2}F$, and that Q is replaced by $\rho^N Q$ for some $\rho < 1$, so that the cost function becomes

$$C_N^\rho(u) = E\left[\sum_{k=0}^{N-1} \rho^k \| Dx_k + Fu_k \|^2 + \rho^N x_N^{\mathrm{T}} Q x_N\right].$$

In view of the 'separation property', the Kalman filter matrices $P(k)$, $K(k)$ are unaffected by the discount factor ρ. By specializing the preceding results, or by using an argument involving x_k^ρ, u_k^ρ as in Section 6.2, one can verify that the control which minimizes $C_N^\rho(u)$ is

$$\hat{u}_k^\rho = -M^\rho(k)\hat{x}_{k|k-1}$$

with $M^\rho(k)$ as before. The cost corresponding to $\hat{u}^\rho$ is

$$\begin{aligned} C_N^\rho(\hat{u}^\rho) = {} & m_0^{\mathrm{T}} S^\rho(0) m_0 + \rho^N \operatorname{tr}[P(N)Q] \\ & + \sum_{k=0}^{N-1} \rho^k \operatorname{tr}[DP(k)D^{\mathrm{T}} \\ & + \rho K(k)[HP(k)H^{\mathrm{T}} + GG^{\mathrm{T}}]K^{\mathrm{T}}(k)S^\rho(k+1)]. \end{aligned}$$

Thus if a discount factor is introduced, the filtering computation (b) is unchanged while, in the control computation (a), A and B are replaced by $\rho^{1/2}A$, $\rho^{1/2}B$ respectively.

Turning now to the minimization of the infinite-time cost,

$$C_\infty^\rho(u) = E\left[\sum_{k=0}^{\infty} \rho^k \| Dx_k + Fu_k \|^2\right],$$

we have to consider the asymptotic properties of *both* Riccati equations (6.3.32) and (6.3.23). The conditions required are as follows

$$\left.\begin{array}{l}(A, B) \\ (\check{A}, \check{C})\end{array}\right\} \text{ stabilizable} \tag{6.3.26}$$

$$\left.\begin{array}{l}(\hat{D}, \hat{A}) \\ (H, A)\end{array}\right\} \text{ detectable}$$

where

$$\begin{aligned} \check{A} &= A - CG^{\mathrm{T}}(GG^{\mathrm{T}})^{-1}H & \qquad \hat{A} &= A - B(F^{\mathrm{T}}F)^{-1}F^{\mathrm{T}}D \\ \check{C} &= C[I - G^{\mathrm{T}}(GG^{\mathrm{T}})^{-1}G] & \qquad \hat{D} &= [I - F(F^{\mathrm{T}}F)^{-1}F^{\mathrm{T}}]D. \end{aligned}$$

These conditions simplify under the additional conditions, assumed at the outset in most treatments of LQG control, that $CG^{\mathrm{T}} = 0$ (no correlation between state and observation noise) and $F^{\mathrm{T}}D = 0$ (no 'cross-term' in the cost criterion). Under these conditions, $\check{A} = \hat{A} = A$, $\check{C} = C$ and $\hat{D} = D$; thus conditions (6.3.26) stipulate that the system be stabilizable from either the control or the noise input, and that it be detectable either via the output Hx_k or via the 'output' Dx_k appearing in the cost function.

According to the results in Appendix B, conditions (6.3.26) guarantee that the algebraic Riccati equations corresponding to (6.3.22), (6.3.23) have unique non-negative definite solutions S, P respectively and that the solutions of (6.3.22), (6.3.23) converge to S, P for arbitrary non-negative definite terminal condition Q and initial condition P_0 respectively. The optimal control for the infinite-time problem can now be obtained by applying the results of Section 6.2 concerning the completely observable case. Indeed, the innovations representation is, as above,

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k-1} + Bu_k + \tilde{C}(k)\tilde{v}_k \tag{6.3.27}$$

where $\tilde{v}_k$ is the normalized innovations process and

$$\tilde{C}(k) = K(k)[HP(k)H^{\mathrm{T}} + GG^{\mathrm{T}}]^{1/2}.$$

Note that, as $k \to \infty$,

$$\tilde{C}(k) \to \tilde{C} = K[HPH^{\mathrm{T}} + GG^{\mathrm{T}}]^{1/2}.$$

where P is the solution of the algebraic Riccati equation and K the corresponding Kalman gain. As in (6.3.17) the cost expressed in terms of $\hat{x}_{k|k-1}$ is

$$C_{\infty}^{\rho}(u) = E\left[\sum_{k=0}^{\infty} \rho^k \| D\hat{x}_{k|k-1} + Fu_k \|^2\right] + \sum_{k=0}^{\infty} \rho^k \operatorname{tr}[DP(k)D^{\mathrm{T}}] \tag{6.3.28}$$

and the final sum is finite since $\operatorname{tr}[DP(k)D^{\mathrm{T}}] \to \operatorname{tr}[DPD^{\mathrm{T}}]$ as $k \to \infty$. We now apply the results of Section 6.2 to the infinite-time completely observable problem constituted by (6.3.27), (6.3.28), and conclude that the optimal control is

$$\hat{u}_k^{\rho} = -M^{\rho}\hat{x}_{k|k-1} \tag{6.3.29}$$

with cost, as in (6.2.16),

$$m_0^{\mathrm{T}} S^{\rho} m_0 + \sum_{k=0}^{\infty} \rho^{k+1} \operatorname{tr}[\tilde{C}^{\mathrm{T}}(k) S^{\rho} \tilde{C}(k)] + \sum_{k=0}^{\infty} \rho^k \operatorname{tr}[DP(k)D^{\mathrm{T}}].$$

Substituting for $\tilde{C}(k)$ gives the final cost expression

$$C_{\infty}^{\rho}(\hat{u}^{\rho}) = m_0^{\mathrm{T}} S^{\rho} m_0 + \sum_{k=0}^{\infty} \rho^k [DP(k)D^{\mathrm{T}} + \rho K(k)[HP(k)H^{\mathrm{T}} + GG^{\mathrm{T}}]K^{\mathrm{T}}(k)S^{\rho}]].$$

Appearances to the contrary, $\hat{u}_k^\rho$ given by (6.3.29) is not a constant-coefficient controller since the gain $K(k)$ in the Kalman filter depends on $P(k)$ which is not constant unless P_0 happens to be equal to the stationary value P. A simpler control algorithm is obtained if $K(k)$ is replaced by its stationary value $K = [APH^T + CG^T]\ [HPH^T + GG^T]^{-1}$, that is we apply the control value

$$\hat{v}_k^\rho := -M^\rho z_k \tag{6.3.30}$$

where z_k is generated by

$$\begin{aligned} z_{k+1} &= Az_k - BM^\rho z_k + K(y_k - Hz_k) \\ z_0 &= m_0 \end{aligned} \tag{6.3.31}$$

(this is the Kalman filter algorithm with $P(k)$ replaced by P). Of course, z_k is in general not equal to $\hat{x}_{k|k-1}$. Control $\hat{v}^\rho$ is not optimal for the discounted cost problem, but $\hat{v}^1$ *is* optimal in the sense of minimizing the average cost per unit time,

$$C_{av}(u) = \lim_{N\to\infty} \frac{1}{N} E\left[\sum_{k=0}^{N} \|Dx_k + Fu_k\|^2\right]. \tag{6.3.32}$$

As remarked earlier, this criterion is insensitive to the behaviour of the process for small k; and, for large K, z_k and $\hat{x}_{k|k-1}$ are practically indistinguishable.

Theorem 6.3.3

Suppose conditions (6.3.26) hold. Then the control $\hat{v}^1$ given by (6.3.30), (6.3.31) with $\rho = 1$ minimizes $C_{av}(u)$ in the class of all output feedback controls such that $C_{av}(u)$ exists and $E\|x_k\|^2$ is bounded. The minimal cost is

$$C_{av}(\hat{v}^1) = \mathrm{tr}[DPD^T + K(HPH^T + GG^T)K^TS]. \tag{6.3.33}$$

PROOF It follows from the arguments above and Theorem 6.3.2 that the control $\hat{u}_k^1$ of (6.3.20) is optimal for C_{av} and that its cost is given by the expression in (6.3.33). Thus it remains to show that $\hat{v}^1$ is admissible and that its cost coincides with that of $\hat{u}^1$.

Define $\xi_k := x_k - z_k$. Recalling that $y_k = Hx_k + Gw_k$ and hence that $y_k - Hz_k = H(x_k - z_k) + Gw_k$, we see that the joint process (z_k, ξ_k) satisfies:

$$\begin{bmatrix} z_{k+1} \\ \xi_{k+1} \end{bmatrix} = \begin{bmatrix} A - BM & KH \\ 0 & A - KH \end{bmatrix} \begin{bmatrix} z_k \\ \xi_k \end{bmatrix} + \begin{bmatrix} KG \\ C + KG \end{bmatrix} w_k$$

$$=: \bar{A} \begin{bmatrix} z_k \\ \xi_k \end{bmatrix} + \bar{C} w_k. \tag{6.3.34}$$

Under conditions (6.3.26) both $A - BM$ and $A - KH$ are stable. This implies that $\bar{A}$ is stable since the eigenvalues of $\bar{A}$ are those of $(A - BM)$ together with those of $(A - KH)$. Thus the covariance matrix Ξ_k of (x_k, ξ_k) is convergent to Ξ satisfying $\Xi = \bar{A}\Xi\bar{A}^{\mathrm{T}} + \bar{C}\bar{C}^{\mathrm{T}}$. Since $x_k = \xi_k + z_k$, this shows that $E\|x_k\|^2$ is bounded and $C_{\mathrm{av}}(\hat{v}^1)$ exists. Note that

$$Dx_k + F\hat{v}_k^1 = \bar{D} \begin{bmatrix} z_k \\ \xi_k \end{bmatrix}$$

where $\bar{D} = (D - FM, D)$, so that

$$C_{\mathrm{av}}(\hat{v}^1) = \mathrm{tr}[\bar{D}\Xi\bar{D}^{\mathrm{T}}].$$

The process $\eta_k := \mathrm{col}\{\hat{x}_{k|k-1}, \tilde{x}_{k|k-1}\}$ satisfies (6.3.34) with $\bar{A}$ and $\bar{C}$ replaced by $\bar{A}(k)$ and $\bar{C}(k)$ obtained by substituting $K(k)$ for K in $\bar{A}$ and $\bar{C}$. Denote $\Gamma(k) := \mathrm{cov}(\eta_k)$. Then $\Gamma(k)$ satisfies

$$\Gamma(k+1) = \bar{A}(k)\Gamma(k)\bar{A}^{\mathrm{T}}(k) + \bar{C}(k)\bar{C}^{\mathrm{T}}(k) \tag{6.3.35}$$

We know that $\Gamma := \lim_{k\to\infty} \Gamma(k)$ exists and that

$$C_{\mathrm{av}}(\hat{u}^1) = \mathrm{tr}[\bar{D}\Gamma\bar{D}^{\mathrm{T}}].$$

Taking the limit as $k \to \infty$ in (6.3.35) we see that Γ satisfies $\Gamma = \bar{A}\Gamma\bar{A}^{\mathrm{T}} + \bar{C}\bar{C}^{\mathrm{T}}$, i.e. $\Gamma = \Xi$. This completes the proof. □

Finally, a remark on the stabilizability and detectability conditions (6.3.26). The conditions on (A, B), $(\hat{D}, A)$ ensure that S^ρ, the solution to the 'discounted' algebraic Riccati equation, exists for any $\rho < 1$, but if these conditions are not met then S^ρ may only exist for $\rho < \rho_0$ for some $\rho_0 < 1$. According to the separation principle, however, discounting has no effect on the Riccati equation (6.3.23) generating $P(k)$ so that no weakening of the conditions on $(\check{A}, \check{C})$ and (H, A) is possible. The reason for this minor asymmetry in the problem is of course that, while we are free to select the cost function coefficients D, F in any manner we choose, their counterparts C and G in the filtering problem are part of the system specification.

As in the complete observations case, little can be said about the average cost problem if conditions (6.3.26) are not met.

Notes

Dynamic programming was introduced in its modern form by Bellman (1957). Recent texts describing various aspects of it include Bertsekas (1976) and Whittle (1981). The linear regulator problem was solved by Kalman (1960) who also noted the filtering/control duality. For references on properties of the Riccati equation and the algebraic Riccati equation, see Chapter 3. The use of linear/quadratic control as a design methodology for multivariable systems has been pioneered by Harvey and Stein (1978); see also Kwakernaak (1976).

The 'certainty-equivalence principle' was first enunciated in the economics literature, by Simon (1956). The 'separation principle' is clearly presented (for continuous-time systems) in Wonham (1968) and is also discussed in Fleming and Rishel (1975). The stochastic linear regulator is discussed in one form or another in most texts on stochastic control, including Bertsekas (1976) and Whittle (1981).

References

Bellman, R. (1957) *Dynamic Programming*, Princeton University Press, New Jersey.

Bertsekas, D. P. (1976) *Dynamic Programming and Stochastic Control*, Academic Press, New York.

Fleming, W. H. and Rishel, R. W. (1975) *Deterministic and Stochastic Optimal Control*, Springer-Verlag, Berlin.

Harvey, C. A. and Stein, G. (1978) Quadratic weights for asymptotic regulator properties. *IEEE Trans. Automatic Control*, **AC-23**, 378–387.

Kalman, R. E. (1960) Contributions to the theory of optimal control. *Bol. Soc. Math. Mexicana*, **5**, 102–119.

Kwakernaak, H. (1976) Asymptotic root loci for multivariable linear optimal regulators. *IEEE Trans. Automatic Control*, **AC-21**, 378–381.

Simon, H. A. (1956) Dynamic programming under uncertainty with a quadratic criterion function. *Econometrica*, **24**, 74–81.

Whittle, P. (1981) *Optimization over Time*, vols 1 and 2, John Wiley, Chichester.

Wonham, W. M. (1968) On the separation theorem of stochastic control. *SIAM J. Control*, **6**, 312–326.