2 Kalman Filter: An Elementary Approach

This chapter is devoted to a most elementary introduction to the Kalman filtering algorithm. By assuming invertibility of certain matrices, the Kalman filtering "prediction-correction" algorithm will be derived based on the optimality criterion of least-squares unbiased estimation of the state vector with the optimal weight, using all available data information. The filtering algorithm is first obtained for a system with no deterministic (control) input. By superimposing the deterministic solution, we then arrive at the general Kalman filtering algorithm.

2.1 The Model

Consider a linear system with state-space description

$$\begin{cases} \mathbf{y}_{k+1} = & A_k\mathbf{y}_k + B_k\mathbf{u}_k + \Gamma_k\underline{\xi}_k \\ \mathbf{w}_k = & C_k\mathbf{y}_k + D_k\mathbf{u}_k + \underline{\eta}_k, \end{cases}$$

where $A_k, B_k, \Gamma_k, C_k, D_k$ are $n \times n, n \times m, n \times p, q \times n, q \times m$ (known) constant matrices, respectively, with $1 \leq m, p, q \leq n$, $\{\mathbf{u}_k\}$ a (known) sequence of m-vectors (called a *deterministic input sequence*), and $\{\underline{\xi}_k\}$ and $\{\underline{\eta}_k\}$ are, respectively, (unknown) system and observation noise sequences, with known statistical information such as mean, variance, and covariance. Since both the deterministic input $\{\mathbf{u}_k\}$ and noise sequences $\{\underline{\xi}_k\}$ and $\{\underline{\eta}_k\}$ are present, the system is usually called a *linear deterministic/stochastic system*. This system can be decomposed into the sum of a linear deterministic system:

$$\begin{cases} \mathbf{z}_{k+1} = & A_k\mathbf{z}_k + B_k\mathbf{u}_k \\ \mathbf{s}_k = & C_k\mathbf{z}_k + D_k\mathbf{u}_k, \end{cases}$$

C.K. Chui and G. Chen, *Kalman Filtering*, DOI 10.1007/978-3-319-47612-4_2

and a linear (purely) stochastic system:

$$\begin{cases} \mathbf{x}_{k+1} = A_k \mathbf{x}_k + \Gamma_k \underline{\xi}_k \\ \mathbf{v}_k = C_k \mathbf{x}_k + \underline{\eta}_k, \end{cases} \tag{2.1}$$

with $\mathbf{w}_k = \mathbf{s}_k + \mathbf{v}_k$ and $\mathbf{y}_k = \mathbf{z}_k + \mathbf{x}_k$. The advantage of the decomposition is that the solution of $\mathbf{z}_k$ in the linear deterministic system is well known and is given by the so-called *transition equation*

$$\mathbf{z}_k = (A_{k-1} \cdots A_0)\mathbf{z}_0 + \sum_{i=1}^{k} (A_{k-1} \cdots A_{i-1}) B_{i-1} \mathbf{u}_{i-1}.$$

Hence, it is sufficient to derive the optimal estimate $\hat{\mathbf{x}}_k$ of $\mathbf{x}_k$ in the stochastic state-space description (2.1), so that

$$\hat{\mathbf{y}}_k = \mathbf{z}_k + \hat{\mathbf{x}}_k$$

becomes the optimal estimate of the state vector $\mathbf{y}_k$ in the original linear system. Of course, the estimate has to depend on the statistical information of the noise sequences. In this chapter, we will only consider zero-mean Gaussian white noise processes.

Assumption 2.1 Let $\{\underline{\xi}_k\}$ and $\{\underline{\eta}_k\}$ be sequences of zero-mean Gaussian white noise such that $Var(\underline{\xi}_k) = Q_k$ and $Var(\underline{\eta}_k) = R_k$ are positive definite matrices and $E(\underline{\xi}_k \underline{\eta}_\ell^\top) = 0$ for all k and ℓ. The initial state $\mathbf{x}_0$ is also assumed to be independent of $\underline{\xi}_k$ and $\underline{\eta}_k$ in the sense that $E(\mathbf{x}_0 \underline{\xi}_k^\top) = 0$ and $E(\mathbf{x}_0 \underline{\eta}_k^\top) = 0$ for all k.

2.2 Optimality Criterion

In determining the optimal estimate $\hat{\mathbf{x}}_k$ of $\mathbf{x}_k$, it will be seen that the optimality is in the sense of least-squares followed by choosing the optimal weight matrix that gives a minimum variance estimate as discussed in Sect. 1.3. However, we will incorporate the information of all data $\mathbf{v}_j$, $j = 0, 1, \cdots, k$, in determining the estimate $\hat{\mathbf{x}}_k$ of $\mathbf{x}_k$ (instead of just using $\mathbf{v}_k$ as discussed in Sect. 1.3). To accomplish this, we introduce the vectors

$$\overline{\mathbf{v}}_j = \begin{bmatrix} \mathbf{v}_0 \\ \vdots \\ \mathbf{v}_j \end{bmatrix}, \qquad j = 0, 1, \cdots,$$

and obtain $\hat{\mathbf{x}}_k$ from the data vector $\overline{\mathbf{v}}_k$. For this approach, we assume for the time being that all the system matrices A_j are nonsingular. Then it can be shown that the state-space description of the linear stochastic system can be written as

$$\overline{\mathbf{v}}_j = H_{k,j}\mathbf{x}_k + \overline{\underline{\epsilon}}_{k,j}, \tag{2.2}$$

where

$$H_{k,j} = \begin{bmatrix} C_0\Phi_{0k} \\ \vdots \\ C_j\Phi_{jk} \end{bmatrix} \quad and \quad \overline{\underline{\epsilon}}_{k,j} = \begin{bmatrix} \underline{\epsilon}_{k,0} \\ \vdots \\ \underline{\epsilon}_{k,j} \end{bmatrix},$$

with $\Phi_{\ell k}$ being the transition matrices defined by

$$\Phi_{\ell k} = \begin{cases} A_{\ell-1}\cdots A_k & \text{if } \ell > k, \\ I & \text{if } \ell = k, \end{cases}$$

$\Phi_{\ell k} = \Phi_{k\ell}^{-1}$ if $\ell < k$, and

$$\underline{\epsilon}_{k,\ell} = \underline{\eta}_\ell - C_\ell \sum_{i=\ell+1}^{k} \Phi_{\ell i}\Gamma_{i-1}\underline{\xi}_{i-1}.$$

Indeed, by applying the inverse transition property of Φ_{ki} described above and the transition equation

$$\mathbf{x}_k = \Phi_{k\ell}\mathbf{x}_\ell + \sum_{i=\ell+1}^{k} \Phi_{ki}\Gamma_{i-1}\underline{\xi}_{i-1},$$

which can be easily obtained from the first recursive equation in (2.1), we have

$$\mathbf{x}_\ell = \Phi_{\ell k}\mathbf{x}_k - \sum_{i=\ell+1}^{k} \Phi_{\ell i}\Gamma_{i-1}\underline{\xi}_{i-1};$$

and this yields

$$\begin{aligned} & H_{k,j}\mathbf{x}_k + \overline{\underline{\epsilon}}_{k,j} \\ &= \begin{bmatrix} C_0\Phi_{0k} \\ \vdots \\ C_j\Phi_{jk} \end{bmatrix}\mathbf{x}_k + \begin{bmatrix} \underline{\eta}_0 - C_0\sum_{i=1}^{k}\Phi_{0i}\Gamma_{i-1}\underline{\xi}_{i-1} \\ \vdots \\ \underline{\eta}_j - C_j\sum_{i=j+1}^{k}\Phi_{ji}\Gamma_{i-1}\underline{\xi}_{i-1} \end{bmatrix} \\ &= \begin{bmatrix} C_0\mathbf{x}_0 + \underline{\eta}_0 \\ \vdots \\ C_j\mathbf{x}_j + \underline{\eta}_j \end{bmatrix} = \begin{bmatrix} \mathbf{v}_0 \\ \vdots \\ \mathbf{v}_j \end{bmatrix} = \overline{\mathbf{v}}_j \end{aligned}$$

which is (2.2).

Now, using the least-squares estimate discussed in Chap. 1, Sect. 1.3, with weight $W_{kj} = (Var(\underline{\epsilon}_{k,j}))^{-1}$, where the inverse is assumed only for the purpose of illustrating the optimality criterion, we arrive at the linear, unbiased, minimum variance least-squares estimate $\hat{\mathbf{x}}_{k|j}$ of $\mathbf{x}_k$ using the data $\mathbf{v}_0, \cdots, \mathbf{v}_j$.

Definition 2.1 (1) For $j = k$, we denote $\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_{k|k}$ and call the estimation process a *digital filtering process*. (2) For $j < k$, we call $\hat{\mathbf{x}}_{k|j}$ an *optimal prediction* of $\mathbf{x}_k$ and the process a *digital prediction process*. (3) For $j > k$, we call $\hat{\mathbf{x}}_{k|j}$ a *smoothing estimate* of $\mathbf{x}_k$ and the process a *digital smoothing process*.

We will only discuss digital filtering. However, since $\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_{k|k}$ is determined by using *all* data $\mathbf{v}_0, \cdots, \mathbf{v}_k$, the process is not applicable to real-time problems for very large values of k, since the need for storage of the data and the computational requirement grow with time. Hence, we will derive a recursive formula that gives $\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_{k|k}$ from the "prediction" $\hat{\mathbf{x}}_{k|k-1}$ and $\hat{\mathbf{x}}_{k|k-1}$ from the estimate $\hat{\mathbf{x}}_{k-1} = \hat{\mathbf{x}}_{k-1|k-1}$. At each step, we only use the incoming bit of the data information so that very little storage of the data is necessary. This is what is usually called the *Kalman filtering algorithm*.

2.3 Prediction-Correction Formulation

To compute $\hat{\mathbf{x}}_k$ in real-time, we will derive the recursive formula

$$\begin{cases} \hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + G_k(\mathbf{v}_k - C_k\hat{x}_{k|k-1}) \\ \hat{\mathbf{x}}_{k|k-1} = A_{k-1}\hat{\mathbf{x}}_{k-1|k-1}, \end{cases}$$

where G_k will be called the *Kalman gain* matrices. The starting point is the initial estimate $\hat{\mathbf{x}}_0 = \hat{\mathbf{x}}_{0|0}$. Since $\hat{\mathbf{x}}_0$ is an unbiased estimate of the initial state $\mathbf{x}_0$, we could use $\hat{\mathbf{x}}_0 = E(\mathbf{x}_0)$, which is a constant vector. In the actual Kalman filtering, G_k must also be computed recursively. The two recursive processes together will be called the *Kalman filtering process*.

Let $\hat{\mathbf{x}}_{k|j}$ be the (optimal) least-squares estimate of $\mathbf{x}_k$ with minimum variance by choosing the weight matrix to be

$$W_{k,j} = (Var(\underline{\epsilon}_{k,j}))^{-1}$$

using $\overline{\mathbf{v}}_j$ in (2.2) (see Sect. 1.3 for details). It is easy to verify that

$$W_{k,k-1}^{-1} = \begin{bmatrix} R_0 & & 0 \\ & \ddots & \\ 0 & & R_{k-1} \end{bmatrix} + Var \begin{bmatrix} C_0 \sum_{i=1}^{k} \Phi_{0i}\Gamma_{i-1}\underline{\xi}_{i-1} \\ \vdots \\ C_{k-1}\Phi_{k-1,k}\Gamma_{k-1}\underline{\xi}_{k-1} \end{bmatrix} \tag{2.3}$$

and

$$W_{k,k}^{-1} = \begin{bmatrix} W_{k,k-1}^{-1} & 0 \\ 0 & R_k \end{bmatrix} \tag{2.4}$$

(cf. Exercise 2.1). Hence, $W_{k,k-1}$ and $W_{k,k}$ are positive definite (cf. Exercise 2.2).

In this chapter, we also assume that the matrices

$$(H_{k,j}^\top W_{k,j} H_{k,j}), \quad j = k-1 \quad and \quad k,$$

are nonsingular. Then it follows from Chap. 1, Sect. 1.3, that

$$\hat{\mathbf{x}}_{k|j} = (H_{k,j}^\top W_{k,j} H_{k,j})^{-1} H_{k,j}^\top W_{k,j} \overline{\mathbf{v}}_j. \tag{2.5}$$

Our first goal is to relate $\hat{\mathbf{x}}_{k|k-1}$ with $\hat{\mathbf{x}}_{k|k}$. To do so, we observe that

$$\begin{aligned} H_{k,k}^\top W_{k,k} H_{k,k} &= [H_{k,k-1}^\top C_k^\top] \begin{bmatrix} W_{k,k-1} & 0 \\ 0 & R_k^{-1} \end{bmatrix} \begin{bmatrix} H_{k,k-1} \\ C_k \end{bmatrix} \\ &= H_{k,k-1}^\top W_{k,k-1} H_{k,k-1} + C_k^\top R_k^{-1} C_k \end{aligned}$$

and

$$H_{k,k}^\top W_{k,k} \overline{\mathbf{v}}_k = H_{k,k-1}^\top W_{k,k-1} \overline{\mathbf{v}}_{k-1} + C_k^\top R_k^{-1} \mathbf{v}_k.$$

Using (2.5) and the above two equalities, we have

$$\begin{aligned} &(H_{k,k-1}^\top W_{k,k-1} H_{k,k-1} + C_k^\top R_k^{-1} C_k) \hat{\mathbf{x}}_{k|k-1} \\ =& H_{k,k-1}^\top W_{k,k-1} \overline{\mathbf{v}}_{k-1} + C_k^\top R_k^{-1} C_k \hat{\mathbf{x}}_{k|k-1} \end{aligned}$$

and

$$\begin{aligned} &(H_{k,k-1}^\top W_{k,k-1} H_{k,k-1} + C_k^\top R_k^{-1} C_k) \hat{\mathbf{x}}_{k|k} \\ &= (H_{k,k}^\top W_{k,k} H_{k,k}) \hat{\mathbf{x}}_{k|k} \\ &= H_{k,k-1}^\top W_{k,k-1} \overline{\mathbf{v}}_{k-1} + C_k^\top R_k^{-1} \mathbf{v}_k. \end{aligned}$$

A simple subtraction gives

$$\begin{aligned} &(H_{k,k-1}^\top W_{k,k-1} H_{k,k-1} + C_k^\top R_k^{-1} C_k)(\hat{\mathbf{x}}_{k|k} - \hat{\mathbf{x}}_{k|k-1}) \\ =& C_k^\top R_k^{-1} (\mathbf{v}_k - C_k \hat{\mathbf{x}}_{k|k-1}). \end{aligned}$$

Now define

$$\begin{aligned} G_k &= (H_{k,k-1}^\top W_{k,k-1} H_{k,k-1} + C_k^\top R_k^{-1} C_k)^{-1} C_k^\top R_k^{-1} \\ &= (H_{k,k}^\top W_{k,k} H_{k,k})^{-1} C_k^\top R_k^{-1}. \end{aligned}$$

Then we have

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + G_k(\mathbf{v}_k - C_k\hat{\mathbf{x}}_{k|k-1}). \tag{2.6}$$

Since $\hat{\mathbf{x}}_{k|k-1}$ is a one-step prediction and $(\mathbf{v}_k - C_k\hat{\mathbf{x}}_{k|k-1})$ is the error between the real data and the prediction, (2.6) is in fact a "prediction-correction" formula with the Kalman gain matrix G_k as a weight matrix. To complete the recursive process, we need an equation that gives $\hat{\mathbf{x}}_{k|k-1}$ from $\hat{\mathbf{x}}_{k-1|k-1}$. This is simply the equation

$$\hat{\mathbf{x}}_{k|k-1} = A_{k-1}\hat{\mathbf{x}}_{k-1|k-1}. \tag{2.7}$$

To prove this, we first note that

$$\underline{\epsilon}_{k,k-1} = \underline{\epsilon}_{k-1,k-1} - H_{k,k-1}\Gamma_{k-1}\underline{\xi}_{k-1}$$

so that

$$W_{k,k-1}^{-1} = W_{k-1,k-1}^{-1} + H_{k-1,k-1}\Phi_{k-1,k}\Gamma_{k-1}Q_{k-1}\Gamma_{k-1}^{\top}\Phi_{k-1,k}^{\top}H_{k-1,k-1}^{\top} \tag{2.8}$$

(cf. Exercise 2.3). Hence, by Lemma 1.2, we have

$$\begin{aligned} W_{k,k-1} = W_{k-1,k-1} - &W_{k-1,k-1}H_{k-1,k-1}\Phi_{k-1,k}\Gamma_{k-1}(Q_{k-1}^{-1} \\ &+ \Gamma_{k-1}^{\top}\Phi_{k-1,k}^{\top}H_{k-1,k-1}^{\top}W_{k-1,k-1}H_{k-1,k-1}\Phi_{k-1,k}\Gamma_{k-1})^{-1} \\ &\cdot \Gamma_{k-1}^{\top}\Phi_{k-1,k}^{\top}H_{k-1,k-1}^{\top}W_{k-1,k-1} \end{aligned} \tag{2.9}$$

(cf. Exercise 2.4). Then by the transition relation

$$H_{k,k-1} = H_{k-1,k-1}\Phi_{k-1,k}$$

we have

$$\begin{aligned} &H_{k,k-1}^{\top}W_{k,k-1} \\ &= \Phi_{k-1,k}^{\top}\{I - H_{k-1,k-1}^{\top}W_{k-1,k-1}H_{k-1,k-1}\Phi_{k-1,k}\Gamma_{k-1}(Q_{k-1}^{-1} \\ &\quad + \Gamma_{k-1}^{\top}\Phi_{k-1,k}^{\top}H_{k-1,k-1}^{\top}W_{k-1,k-1}H_{k-1,k-1}\Phi_{k-1,k}\Gamma_{k-1})^{-1} \\ &\quad \cdot \Gamma_{k-1}^{\top}\Phi_{k-1,k}^{\top}\}H_{k-1,k-1}^{\top}W_{k-1,k-1} \end{aligned} \tag{2.10}$$

(cf. Exercise 2.5). It follows that

$$\begin{aligned} &(H_{k,k-1}^{\top}W_{k,k-1}H_{k,k-1})\Phi_{k,k-1}(H_{k-1,k-1}^{\top}W_{k-1,k-1}H_{k-1,k-1})^{-1} \\ &\cdot H_{k-1,k-1}^{\top}W_{k-1,k-1} = H_{k,k-1}^{\top}W_{k,k-1} \end{aligned} \tag{2.11}$$

(cf. Exercise 2.6). This, together with (2.5) with $j = k-1$ and k, gives (2.7).

Our next goal is to derive a recursive scheme for calculating the Kalman gain matrices G_k. Write

$$G_k = P_{k,k} C_k^\top R_k^{-1}$$

where

$$P_{k,k} = (H_{k,k}^\top W_{k,k} H_{k,k})^{-1}$$

and set

$$P_{k,k-1} = (H_{k,k-1}^\top W_{k,k-1} H_{k,k-1})^{-1}.$$

Then, since

$$P_{k,k}^{-1} = P_{k,k-1}^{-1} + C_k^\top R_k^{-1} C_k,$$

we obtain, using Lemma 1.2,

$$P_{k,k} = P_{k,k-1} - P_{k,k-1} C_k^\top (C_k P_{k,k-1} C_k^\top + R_k)^{-1} C_k P_{k,k-1}.$$

It can be proved that

$$G_k = P_{k,k-1} C_k^\top (C_k P_{k,k-1} C_k^\top + R_k)^{-1} \tag{2.12}$$

(cf. Exercise 2.7), so that

$$P_{k,k} = (I - G_k C_k) P_{k,k-1}. \tag{2.13}$$

Furthermore, we can show that

$$P_{k,k-1} = A_{k-1} P_{k-1,k-1} A_{k-1}^\top + \Gamma_{k-1} Q_{k-1} \Gamma_{k-1}^\top \tag{2.14}$$

(cf. Exercise 2.8). Hence, using (2.13) and (2.14) with the initial matrix $P_{0,0}$, we obtain a recursive scheme to compute $P_{k-1,k-1}$, $P_{k,k-1}$, G_k and $P_{k,k}$ for $k = 1, 2, \cdots$. Moreover, it can be shown that

$$\begin{aligned} P_{k,k-1} &= E(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})^\top \\ &= Var(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}) \end{aligned} \tag{2.15}$$

(cf. Exercise 2.9) and that

$$P_{k,k} = E(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})^\top = Var(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}). \tag{2.16}$$

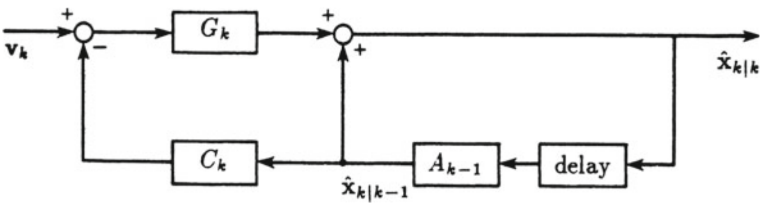

Fig. 2.1 Block diagram of the Kalman filtering algorithm

In particular, we have

$$P_{0,0} = E(\mathbf{x}_0 - E\mathbf{x}_0)(\mathbf{x}_0 - E\mathbf{x}_0)^\top = Var(\mathbf{x}_0).$$

Finally, combining all the results obtained above, we arrive at the following Kalman filtering process for the linear stochastic system with state-space description (2.1):

$$\begin{cases} P_{0,0} = Var(\mathbf{x}_0) \\ P_{k,k-1} = A_{k-1}P_{k-1,k-1}A_{k-1}^\top + \Gamma_{k-1}Q_{k-1}\Gamma_{k-1}^\top \\ G_k = P_{k,k-1}C_k^\top(C_kP_{k,k-1}C_k^\top + R_k)^{-1} \\ P_{k,k} = (I - G_kC_k)P_{k,k-1} \\ \hat{\mathbf{x}}_{0|0} = E(\mathbf{x}_0) \\ \hat{\mathbf{x}}_{k|k-1} = A_{k-1}\hat{\mathbf{x}}_{k-1|k-1} \\ \hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + G_k(\mathbf{v}_k - C_k\hat{\mathbf{x}}_{k|k-1}) \\ k = 1, 2, \cdots . \end{cases} \tag{2.17}$$

This algorithm may be realized as shown in Fig. 2.1.

2.4 Kalman Filtering Process

Let us now consider the general linear deterministic/stochastic system where the deterministic control input $\{\mathbf{u}_k\}$ is present. More precisely, let us consider the state-space description

$$\begin{cases} \mathbf{x}_{k+1} = A_k\mathbf{x}_k + B_k\mathbf{u}_k + \Gamma_k\underline{\xi}_k \\ \mathbf{v}_k = C_k\mathbf{x}_k + D_k\mathbf{u}_k + \underline{\eta}_k, \end{cases}$$

where $\{\mathbf{u}_k\}$ is a sequence of m-vectors with $1 \le m \le n$. Then by superimposing the deterministic solution with (2.17), the Kalman filtering process for this system

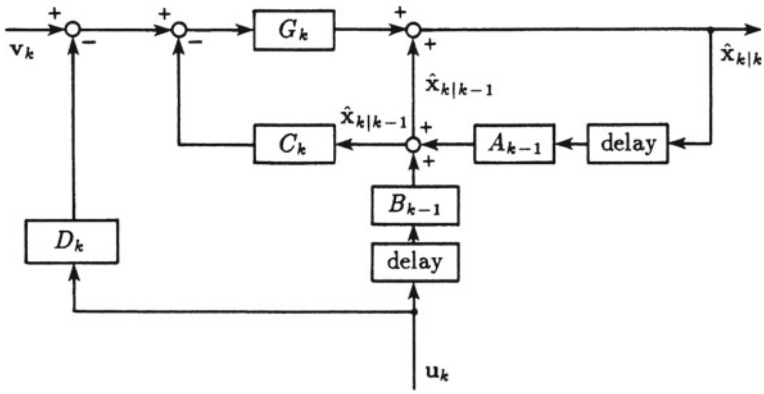

Fig. 2.2 Block diagram of the Kalman filtering algorithm with control input

is given by

$$\begin{cases} P_{0,0} = Var(\mathbf{x}_0) \\ P_{k,k-1} = A_{k-1}P_{k-1,k-1}A_{k-1}^\top + \Gamma_{k-1}Q_{k-1}\Gamma_{k-1}^\top \\ G_k = P_{k,k-1}C_k^\top(C_kP_{k,k-1}C_k^\top + R_k)^{-1} \\ P_{k,k} = (I - G_kC_k)P_{k,k-1} \\ \hat{\mathbf{x}}_{0|0} = E(\mathbf{x}_0) \\ \hat{\mathbf{x}}_{k|k-1} = A_{k-1}\hat{\mathbf{x}}_{k-1|k-1} + B_{k-1}\mathbf{u}_{k-1} \\ \hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + G_k(\mathbf{v}_k - D_k\mathbf{u}_k - C_k\hat{\mathbf{x}}_{k|k-1}) \\ k = 1, 2, \cdots, \end{cases} \tag{2.18}$$

(cf. Exercise 2.13). This algorithm may be implemented as shown in Fig. 2.2.

Exercises

2.1 Let

$$\overline{\underline{\epsilon}}_{k,j} = \begin{bmatrix} \underline{\epsilon}_{k,0} \\ \vdots \\ \underline{\epsilon}_{k,j} \end{bmatrix} \qquad and \qquad \underline{\epsilon}_{k,\ell} = \underline{\eta}_\ell - C_\ell \sum_{i=\ell+1}^{k} \Phi_{\ell i}\Gamma_{i-1}\underline{\xi}_{i-1},$$

where $\{\underline{\xi}_k\}$ and $\{\underline{\eta}_k\}$ are both zero-mean Gaussian white noise sequences with $Var(\underline{\xi}_k) = Q_k$ and $Var(\underline{\eta}_k) = R_k$. Define $W_{k,j} = (Var(\overline{\underline{\epsilon}}_{k,j}))^{-1}$. Show that

$$W_{k,k-1}^{-1} = \begin{bmatrix} R_0 & & 0 \\ & \ddots & \\ 0 & & R_{k-1} \end{bmatrix} + Var \begin{bmatrix} C_0 \sum_{i=1}^{k} \Phi_{0i}\Gamma_{i-1}\underline{\xi}_{i-1} \\ \vdots \\ C_{k-1}\Phi_{k-1,k}\Gamma_{k-1}\underline{\xi}_{k-1} \end{bmatrix}$$

and

$$W_{k,k}^{-1} = \begin{bmatrix} W_{k,k-1}^{-1} & 0 \\ 0 & R_k \end{bmatrix}.$$

2.2 Show that the sum of a positive definite matrix A and a non-negative definite matrix B is positive definite.

2.3 Let $\underline{\epsilon}_{k,j}$ and $W_{k,j}$ be defined as in Exercise 2.1. Verify the relation

$$\underline{\epsilon}_{k,k-1} = \underline{\epsilon}_{k-1,k-1} - H_{k,k-1}\Gamma_{k-1}\underline{\xi}_{k-1}$$

where

$$H_{k,j} = \begin{bmatrix} C_0\Phi_{0k} \\ \vdots \\ C_j\Phi_{jk} \end{bmatrix},$$

and then show that

$$W_{k,k-1}^{-1} = W_{k-1,k-1}^{-1} + H_{k-1,k-1}\Phi_{k-1,k}\Gamma_{k-1}Q_{k-1}\Gamma_{k-1}^\top\Phi_{k-1,k}^\top H_{k-1,k-1}^\top.$$

2.4 Use Exercise 2.3 and Lemma 1.2 to show that

$$\begin{aligned} W_{k,k-1} = W_{k-1,k-1} - W_{k-1,k-1}H_{k-1,k-1}\Phi_{k-1,k}\Gamma_{k-1}(Q_{k-1}^{-1} \\ + \Gamma_{k-1}^\top\Phi_{k-1,k}^\top H_{k-1,k-1}^\top W_{k-1,k-1}H_{k-1,k-1}\Phi_{k-1,k}\Gamma_{k-1})^{-1} \\ \cdot \Gamma_{k-1}^\top\Phi_{k-1,k}^\top H_{k-1,k-1}^\top W_{k-1,k-1}. \end{aligned}$$

2.5 Use Exercise 2.4 and the relation $H_{k,k-1} = H_{k-1,k-1}\Phi_{k-1,k}$ to show that

$$\begin{aligned} & H_{k,k-1}^\top W_{k,k-1} \\ =& \Phi_{k-1,k}^\top\{I - H_{k-1,k-1}^\top W_{k-1,k-1}H_{k-1,k-1}\Phi_{k-1,k}\Gamma_{k-1}(Q_{k-1}^{-1} \\ & + \Gamma_{k-1}^\top\Phi_{k-1,k}^\top H_{k-1,k-1}^\top W_{k-1,k-1}H_{k-1,k-1}\Phi_{k-1,k}\Gamma_{k-1})^{-1} \\ & \cdot \Gamma_{k-1}^\top\Phi_{k-1,k}^\top\}H_{k-1,k-1}^\top W_{k-1,k-1}. \end{aligned}$$

2.6 Use Exercise 2.5 to derive the identity:

$$\begin{aligned} & (H_{k,k-1}^\top W_{k,k-1}H_{k,k-1})\Phi_{k,k-1}(H_{k-1,k-1}^\top W_{k-1,k-1}H_{k-1,k-1})^{-1} \\ & \cdot H_{k-1,k-1}^\top W_{k-1,k-1} = H_{k,k-1}^\top W_{k,k-1}. \end{aligned}$$

2.7 Use Lemma 1.2 to show that

$$P_{k,k-1}C_k^\top(C_kP_{k,k-1}C_k^\top + R_k)^{-1} = P_{k,k}C_k^\top R_k^{-1} = G_k.$$

2.8 Start with $P_{k,k-1} = (H_{k,k-1}^{\top} W_{k,k-1} H_{k,k-1})^{-1}$. Use Lemma 1.2, (2.8), and the definition of $P_{k,k} = (H_{k,k}^{\top} W_{k,k} H_{k,k})^{-1}$ to show that

$$P_{k,k-1} = A_{k-1} P_{k-1,k-1} A_{k-1}^{\top} + \Gamma_{k-1} Q_{k-1} \Gamma_{k-1}^{\top}.$$

2.9 Use (2.5) and (2.2) to prove that

$$E(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})^{\top} = P_{k,k-1}$$

and

$$E(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})^{\top} = P_{k,k}.$$

2.10 Consider the one-dimensional linear stochastic dynamic system

$$x_{k+1} = a x_k + \xi_k, \qquad x_0 = 0,$$

where $E(x_k) = 0$, $Var(x_k) = \sigma^2$, $E(x_k \xi_j) = 0$, $E(\xi_k) = 0$, and $E(\xi_k \xi_j) = \mu^2 \delta_{kj}$. Prove that $\sigma^2 = \mu^2/(1-a^2)$ and $E(x_k x_{k+j}) = a^{|j|}\sigma^2$ for all integers j.

2.11 Consider the one-dimensional stochastic linear system

$$\begin{cases} x_{k+1} = x_k \\ v_k = x_k + \eta_k \end{cases}$$

with $E(\eta_k) = 0$, $Var(\eta_k) = \sigma^2$, $E(x_0) = 0$ and $Var(x_0) = \mu^2$. Show that

$$\begin{cases} \hat{x}_{k|k} = \hat{x}_{k-1|k-1} + \frac{\mu^2}{\sigma^2 + k\mu^2}(v_k - \hat{x}_{k-1|k-1}) \\ \hat{x}_{0|0} = 0 \end{cases}$$

and that $\hat{x}_{k|k} \to c$ for some constant c as $k \to \infty$.

2.12 Let $\{\mathbf{v}_k\}$ be a sequence of data obtained from the observation of a zero-mean random vector $\mathbf{y}$ with unknown variance Q. The variance of $\mathbf{y}$ can be estimated by

$$\hat{Q}_N = \frac{1}{N} \sum_{k=1}^{N} (\mathbf{v}_k \mathbf{v}_k^{\top}).$$

Derive a prediction-correction recursive formula for this estimation.

2.13 Consider the linear deterministic/stochastic system

$$\begin{cases} \mathbf{x}_{k+1} = A_k \mathbf{x}_k + B_k \mathbf{u}_k + \Gamma_k \underline{\xi}_k \\ \mathbf{v}_k = C_k \mathbf{x}_k + D_k \mathbf{u}_k + \underline{\eta}_k, \end{cases}$$

where $\{\mathbf{u}_k\}$ is a given sequence of deterministic control input m-vectors, $1 \le m \le n$. Suppose that Assumption 2.1 is satisfied and the matrix $Var(\overline{\underline{\epsilon}}_{k,j})$ is nonsingular (cf. (2.2) for the definition of $\overline{\underline{\epsilon}}_{k,j}$). Derive the Kalman filtering equations for this model.

2.14 In digital signal processing, a widely used mathematical model is the following so-called *ARMA* (*autoregressive moving-average*) process:

$$\mathbf{v}_k = \sum_{i=1}^{N} B_i \mathbf{v}_{k-i} + \sum_{i=0}^{M} A_i \mathbf{u}_{k-i},$$

where the $n \times n$ matrices $B_1, \cdots, B_N$ and the $n \times q$ matrices $A_0, A_1, \cdots, A_M$ are independent of the time variable k, and $\{\mathbf{u}_k\}$ and $\{\mathbf{v}_k\}$ are input and output digital signal sequences, respectively (cf. Fig. 2.3). Assuming that $M \le N$, show that the input-output relationship can be described as a state-space model

$$\begin{cases} \mathbf{x}_{k+1} = A\mathbf{x}_k + B\mathbf{u}_k \\ \mathbf{v}_k = C\mathbf{x}_k + D\mathbf{u}_k \end{cases}$$

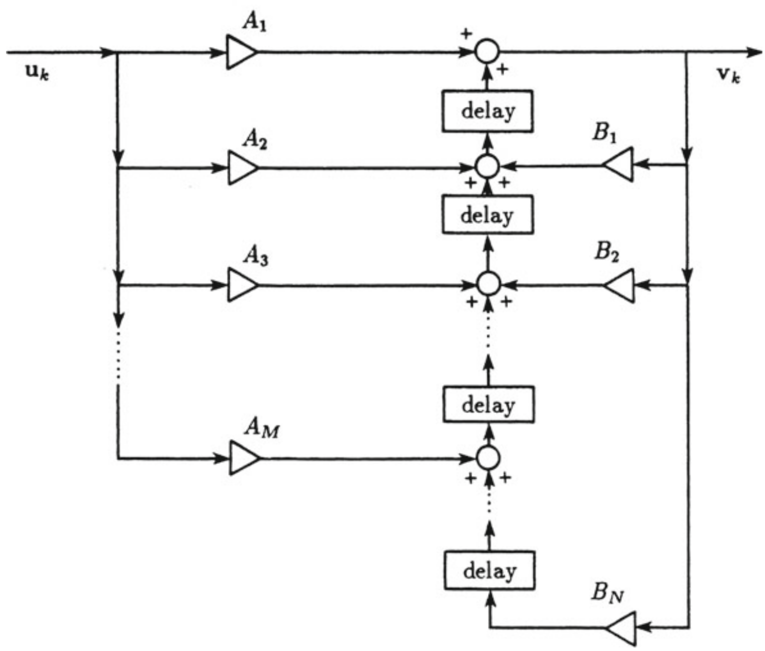

Fig. 2.3 Block diagram of the ARMA model

with $\mathbf{x}_0 = 0$, where

$$A = \begin{bmatrix} B_1 & I & 0 & \cdots & 0 \\ B_2 & 0 & I & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ B_{N-1} & 0 & \cdots & \cdots & I \\ B_N & 0 & \cdots & \cdots & 0 \end{bmatrix}, \quad B = \begin{bmatrix} A_1 + B_1 A_0 \\ A_2 + B_2 A_0 \\ \vdots \\ A_M + B_M A_0 \\ B_{M+1} A_0 \\ \vdots \\ B_N A_0 \end{bmatrix},$$

$$C = [I \; 0 \; \cdots \; 0] \qquad and \qquad D = [A_0].$$

Springer
http://www.springer.com/978-3-319-47610-0

Kalman Filtering
with Real-Time Applications
Chui, C.K.; Chen, G.
2017, XVIII, 247 p. 34 illus., Hardcover
ISBN: 978-3-319-47610-0