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

\[
\begin{align*}
    \text{\begin{cases} 
        y_{k+1} &= A_k y_k + B_k u_k + \Gamma_k \xi_k \\
        w_k &= C_k y_k + D_k u_k + \eta_k,
    \end{cases}} \end{align*}
\]

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 \(1 \leq m, p, q \leq n\), \(\{u_k\}\) a (known) sequence of \(m\)-vectors (called a deterministic input sequence), and \(\{\xi_k\}\) and \(\{\eta_k\}\) are, respectively, (unknown) system and observation noise sequences, with known statistical information such as mean, variance, and covariance. Since both the deterministic input \(\{u_k\}\) and noise sequences \(\{\xi_k\}\) and \(\{\eta_k\}\) 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:

\[
\begin{align*}
    \text{\begin{cases} 
        z_{k+1} &= A_k z_k + B_k u_k \\
        s_k &= C_k z_k + D_k u_k,
    \end{cases}} \end{align*}
\]
and a linear (purely) stochastic system:

\[
\begin{align*}
x_{k+1} &= A_k x_k + \Gamma_k \xi_k \\
v_k &= C_k x_k + \eta_k,
\end{align*}
\]

(2.1)

with \(w_k = s_k + v_k\) and \(y_k = z_k + x_k\). The advantage of the decomposition is that the solution of \(z_k\) in the linear deterministic system is well known and is given by the so-called transition equation

\[
z_k = (A_{k-1} \cdots A_0) z_0 + \sum_{i=1}^{k} (A_{k-1} \cdots A_{i-1}) B_{i-1} u_{i-1}.
\]

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

\[
\hat{y}_k = z_k + \hat{x}_k
\]

becomes the optimal estimate of the state vector \(y_k\) 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 \(\{\xi_k\}\) and \(\{\eta_k\}\) be sequences of zero-mean Gaussian white noise such that \(\text{Var}(\xi_k) = Q_k\) and \(\text{Var}(\eta_k) = R_k\) are positive definite matrices and \(E(\xi_k \eta_{k\ell}^\top) = 0\) for all \(k\) and \(\ell\). The initial state \(x_0\) is also assumed to be independent of \(\xi_k\) and \(\eta_k\) in the sense that \(E(x_0 \xi_{k\ell}^\top) = 0\) and \(E(x_0 \eta_{k\ell}^\top) = 0\) for all \(k\).

## 2.2 Optimality Criterion

In determining the optimal estimate \(\hat{x}_k\) of \(x_k\), 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 \(v_j, j = 0, 1, \cdots, k\), in determining the estimate \(\hat{x}_k\) of \(x_k\) (instead of just using \(v_k\) as discussed in Sect. 1.3). To accomplish this, we introduce the vectors

\[
\bar{v}_j = \begin{bmatrix} v_0 \\ \vdots \\ v_j \end{bmatrix}, \quad j = 0, 1, \cdots,
\]
2.2 Optimality Criterion

and obtain $\hat{x}_k$ from the data vector $\bar{v}_k$. 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

$$v_j = H_{k,j} x_k + \bar{\epsilon}_{k,j},$$

(2.2)

where

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

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

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

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

and this yields

$$\begin{bmatrix} C_0 x_0 + \eta_0 \\ \vdots \\ C_j x_j + \eta_j \end{bmatrix} = \begin{bmatrix} v_0 \\ \vdots \\ v_j \end{bmatrix} = \bar{v}_j.$$

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

$$x_k = \Phi_{k\ell} x_\ell + \sum_{i=\ell+1}^{k} \Phi_{ki} \Gamma_{i-1} \bar{\epsilon}_{i-1},$$

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

$$x_\ell = \Phi_{k\ell} x_k - \sum_{i=\ell+1}^{k} \Phi_{ki} \Gamma_{i-1} \bar{\epsilon}_{i-1};$$
which is (2.2).

Now, using the least-squares estimate discussed in Chap. 1, Sect. 1.3, with weight
\[ W_{kj} = (\text{Var}(\xi_{k,j}))^{-1} \], 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 \( \hat{x}_{k|j} \) of \( x_k \) using the data \( v_0, \ldots, v_j \).

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

We will only discuss digital filtering. However, since \( \hat{x}_k = \hat{x}_{k|k} \) is determined by using all data \( v_0, \ldots, v_k \), 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 \( \hat{x}_k = \hat{x}_{k|k} \) from the “prediction” \( \hat{x}_{k|k-1} \) and \( \hat{x}_{k|k-1} \) from the estimate \( \hat{x}_{k-1} = \hat{x}_{k-1|k-1} \). 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 \( \hat{x}_k \) in real-time, we will derive the recursive formula

\[
\begin{align*}
\hat{x}_{k|k} &= \hat{x}_{k|k-1} + G_k (v_k - C_k \hat{x}_{k|k-1}) \\
\hat{x}_{k|k-1} &= A_{k-1} \hat{x}_{k-1|k-1},
\end{align*}
\]

where \( G_k \) will be called the Kalman gain matrices. The starting point is the initial estimate \( \hat{x}_0 = \hat{x}_{0|0} \). Since \( \hat{x}_0 \) is an unbiased estimate of the initial state \( x_0 \), we could use \( \hat{x}_0 = E(x_0) \), 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 \( \hat{x}_{k|j} \) be the (optimal) least-squares estimate of \( x_k \) with minimum variance by choosing the weight matrix to be

\[ W_{k,j} = (\text{Var}(\xi_{k,j}))^{-1} \]

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

\[
W_{k,k-1}^{-1} = \begin{bmatrix} R_0 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & R_{k-1} & \end{bmatrix} + \text{Var} \begin{bmatrix} C_0 \sum_{i=1}^{k} \Phi_{0i} \Gamma_{i-1} \xi_{j-1} \\ \vdots \\ C_{k-1} \Phi_{k-1,k} \Gamma_{k-1} \xi_{k-1} \end{bmatrix}
\] (2.3)
and

\[ W_{-1}^{-1} W_{k,k} = \begin{bmatrix} W_{-1}^{-1} & 0 \\ 0 & R_k \end{bmatrix} \]  (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

\[(H^T_{k,j} W_{k,j} H_{k,j}), \quad j = k - 1 \quad \text{and} \quad k,\]

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

\[ \hat{x}_{k|j} = (H^T_{k,j} W_{k,j} H_{k,j})^{-1} H^T_{k,j} W_{k,j} \nu_j. \]  (2.5)

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

\[ H^T_{k,k-1} W_{k,k-1} H_{k,k-1} + C_k^T R_k^{-1} C_k \]

and

\[ H^T_{k,k} W_{k,k} \nu_k = H^T_{k,k-1} W_{k,k-1} \nu_{k-1} + C_k^T R_k^{-1} \nu_k. \]

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

\[ (H^T_{k,k-1} W_{k,k-1} H_{k,k-1} + C_k^T R_k^{-1} C_k) \hat{x}_{k|k-1} \]

\[ = H^T_{k,k-1} W_{k,k-1} \nu_{k-1} + C_k^T R_k^{-1} \nu_k. \]

A simple subtraction gives

\[ (H^T_{k,k-1} W_{k,k-1} H_{k,k-1} + C_k^T R_k^{-1} C_k) (\hat{x}_{k|k} - \hat{x}_{k|k-1}) \]

\[ = C_k^T R_k^{-1} (\nu_k - C_k \hat{x}_{k|k-1}). \]

Now define

\[ G_k = (H^T_{k,k-1} W_{k,k-1} H_{k,k-1} + C_k^T R_k^{-1} C_k)^{-1} C_k^T R_k^{-1} \]

\[ = (H^T_{k,k} W_{k,k} H_{k,k})^{-1} C_k^T R_k^{-1}. \]
Then we have
\[ \hat{x}_{k|k} = \hat{x}_{k|k-1} + G_k (v_k - C_k \hat{x}_{k|k-1}). \tag{2.6} \]

Since \( \hat{x}_{k|k-1} \) is a one-step prediction and \((v_k - C_k \hat{x}_{k|k-1})\) 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 \( \hat{x}_{k|k-1} \) from \( \hat{x}_{k-1|k-1} \). This is simply the equation
\[ \hat{x}_{k|k-1} = A_{k-1} \hat{x}_{k-1|k-1}. \tag{2.7} \]

To prove this, we first note that
\[ \bar{e}_{k,k-1} = \bar{e}_{k-1,k-1} - H_{k,k-1} \Gamma_{k-1} \bar{\xi}_{k-1} \]
so that
\[
W_{k,k-1}^{-1} = W_{k-1,k-1}^{-1} + H_{k-1,k-1} \Phi_{k-1,k} \Gamma_{k-1} Q_{k-1}^{-1} \Gamma_{k-1,k}^{\top} \Phi_{k-1,k}^{\top} H_{k-1,k-1} \tag{2.8}
\]
(cf. Exercise 2.3). Hence, by Lemma 1.2, we have
\[
W_{k,k-1} = W_{k-1,k-1} - W_{k-1,k-1} H_{k-1,k-1} \Phi_{k-1,k} \Gamma_{k-1} (Q_{k-1}^{-1} + \Gamma_{k-1}^{\top} \Phi_{k-1,k}^{\top} H_{k-1,k-1} W_{k-1,k-1} H_{k-1,k-1} \Gamma_{k-1}^{-1})^{-1} \Gamma_{k-1}^{\top} \Phi_{k-1,k} W_{k-1,k-1} \tag{2.9}
\]
(cf. Exercise 2.4). Then by the transition relation
\[ H_{k,k-1} = H_{k-1,k-1} \Phi_{k-1,k} \]
we have
\[
H_{k,k-1}^{-1} W_{k,k-1} = \Phi_{k-1,k}^{\top} W_{k,k-1} = \Phi_{k-1,k}^{\top} \{ I - H_{k-1,k-1} W_{k-1,k-1} H_{k-1,k-1} \} G_{k-1} (Q_{k-1}^{-1} + \Gamma_{k-1}^{\top} \Phi_{k-1,k}^{\top} H_{k-1,k-1} W_{k-1,k-1} H_{k-1,k-1} \Gamma_{k-1}^{-1})^{-1} \Gamma_{k-1}^{\top} \Phi_{k-1,k} W_{k-1,k-1} \tag{2.10}
\]
(cf. Exercise 2.5). It follows that
\[
(H_{k,k-1}^{-1} W_{k,k-1}) \Phi_{k,k-1} (H_{k-1,k-1}^{-1} W_{k-1,k-1} H_{k-1,k-1})^{-1} = H_{k,k-1}^{-1} W_{k,k-1} \tag{2.11}
\]
(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

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

(cf. Exercise 2.7), so that

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

(2.13)

Furthermore, we can show that

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

(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

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

(2.15)

(cf. Exercise 2.9) and that

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

(2.16)
In particular, we have

\[ P_{0,0} = E(x_0 - E(x_0))(x_0 - E(x_0))^\top = Var(x_0). \]

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):

\[
\begin{align*}
P_{0,0} &= Var(x_0) \\
\gamma_{k,k-1} &= A_{k-1}P_{k-1,k-1}A_{k-1}^\top + \Gamma_{k-1}Q_{k-1}\Gamma_{k-1}^\top \\
G_k &= P_{k,k-1}C_k^\top(C_kP_{k-1,k-1}C_k^\top + R_k)^{-1} \\
P_k,k &= (I - G_kC_k)P_{k,k-1} \\
\hat{x}_{0|0} &= E(x_0) \\
\hat{x}_{k|k-1} &= A_{k-1}\hat{x}_{k-1|k-1} \\
\hat{x}_{k|k} &= \hat{x}_{k|k-1} + G_k(v_k - C_k\hat{x}_{k|k-1}) \\
k &= 1, 2, \ldots.
\end{align*}
\]

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

\section*{2.4 Kalman Filtering Process}

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

\[
\begin{align*}
x_{k+1} &= A_kx_k + B_ku_k + \Gamma_k\xi_k \\
v_k &= C_kx_k + D_ku_k + \eta_k,
\end{align*}
\]

where \( \{u_k\} \) is a sequence of \( m \)-vectors with \( 1 \leq m \leq n \). 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

\[
\begin{align*}
P_{0,0} &= Var(x_0) \\
P_{k,k-1} &= A_{k-1}P_{k-1,k-1}A_{k-1}^T + \Gamma_{k-1}Q_{k-1}\Gamma_{k-1}^T \\
G_k &= P_{k,k-1}C_k^T(C_kP_{k,k-1}C_k^T + R_k)^{-1} \\
P_{k,k} &= (I - G_kC_k)P_{k,k-1} \\
\hat{x}_{k|0} &= E(x_0) \\
\hat{x}_{k|k-1} &= A_{k-1}\hat{x}_{k-1|k-1} + B_{k-1}u_{k-1} \\
\hat{x}_{k|k} &= \hat{x}_{k|k-1} + G_k(v_k - D_ku_k - C_k\hat{x}_{k|k-1}) \\
k &= 1, 2, \ldots
\end{align*}
\]  

(2.18)

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

**Exercises**

2.1 Let

\[
\begin{bmatrix}
\xi_{k,0} \\
\vdots \\
\xi_{k,j}
\end{bmatrix}
\]

and

\[
\begin{bmatrix}
\eta_{k,\ell} - C_\ell \sum_{i=\ell+1}^{k} \Phi_{\ell i}\Gamma_{i-1}\xi_{j-1}
\end{bmatrix}
\]

where \(\{\xi_{k}\}\) and \(\{\eta_{k}\}\) are both zero-mean Gaussian white noise sequences with \(Var(\xi_{k}) = Q_k\) and \(Var(\eta_{k}) = R_k\). Define \(W_{k,j} = (Var(\xi_{k,j}))^{-1}\). Show that

\[
W_{k,k-1}^{-1} = \begin{bmatrix}
R_0 & 0 \\
\vdots & \ddots \\
0 & R_{k-1}
\end{bmatrix} + Var \begin{bmatrix}
C_0 \sum_{i=1}^{k} \Phi_{0i}\Gamma_{i-1}\xi_{j-1} \\
\vdots \\
C_{k-1}\Phi_{k-1,k}\Gamma_{k-1}\xi_{k-1}
\end{bmatrix}
\]
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 $\bar{e}_{k,j}$ and $W_{k,j}$ be defined as in Exercise 2.1. Verify the relation

$$\bar{e}_{k,k-1} = \bar{e}_{k-1,k-1} - H_{k,k-1} \Gamma_{k-1} \xi_{k-1}$$

where

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

and then show that

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

2.4 Use Exercise 2.3 and Lemma 1.2 to show that

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

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

$$H_{k,k-1}^T W_{k,k-1} = \Phi_{k-1,k}^T \{ I - H_{k,k-1}^T W_{k-1,k-1} H_{k-1,k-1} \Phi_{k-1,k} \Gamma_{k-1} (Q_{k-1}^{-1} + \Gamma_{k-1}^T \Phi_{k-1,k}^T H_{k-1,k-1} W_{k-1,k-1} H_{k-1,k-1} \Phi_{k-1,k} \Gamma_{k-1})^{-1} \Gamma_{k-1}^T \Phi_{k-1,k}^T H_{k-1,k-1} W_{k-1,k-1} \}.$$ 

2.6 Use Exercise 2.5 to derive the identity:

$$(H_{k,k-1}^T W_{k,k-1} H_{k,k-1}) \Phi_{k,k-1} (H_{k,k-1}^T W_{k,k-1} H_{k,k-1})^{-1} \cdot H_{k,k-1}^T W_{k,k-1} = H_{k,k-1}^T W_{k,k-1}.$$ 

2.7 Use Lemma 1.2 to show that

$$P_{k,k-1} C_k^T (C_k P_{k,k-1} C_k^T + R_k)^{-1} = P_{k,k} C_k^T R_k^{-1} = G_k.$$
2.8 Start with
\[ P_{k-1} = (H_{k-1}^T W_{k-1} H_{k-1})^{-1} \]
Use Lemma 1.2, (2.8), and the
definition of
\[ P_{k} = (H_{k}^T W_{k} H_{k})^{-1} \]
to show that
\[ P_{k-1} = A_{k-1} P_{k-1, k-1} A_{k-1}^T + \Gamma_{k-1} Q_{k-1} \Gamma_{k-1}^T. \]

2.9 Use (2.5) and (2.2) to prove that
\[ E(x_k - \hat{x}_{k|k-1})(x_k - \hat{x}_{k|k-1})^T = P_{k,k-1} \]
and
\[ E(x_k - \hat{x}_{k|k})(x_k - \hat{x}_{k|k})^T = P_{k,k}. \]

2.10 Consider the one-dimensional linear stochastic dynamic system
\[ x_{k+1} = ax_k + \xi_k, \quad x_0 = 0, \]
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
\[ \begin{cases} x_{k+1} = x_k & \vspace{0.3cm} \\
v_k = x_k + \eta_k & \end{cases} \]
with \( E(\eta_k) = 0, Var(\eta_k) = \sigma^2, E(\varepsilon_0) = 0 \) and \( Var(x_0) = \mu^2. \) Show that
\[ \hat{x}_{k|k} = \hat{x}_{k-1|k-1} + \frac{\mu^2}{\sigma^2 + k \mu^2} (v_k - \hat{x}_{k-1|k-1}) \]
and that \( \hat{x}_{k|k} \to c \) for some constant \( c \) as \( k \to \infty. \)

2.12 Let \( \{v_k\} \) be a sequence of data obtained from the observation of a zero-
mean random vector \( y \) with unknown variance \( Q. \) The variance of \( y \) can be
estimated by
\[ \hat{Q}_N = \frac{1}{N} \sum_{k=1}^{N} (v_k v_k^T). \]
Derive a prediction-correction recursive formula for this estimation.

2.13 Consider the linear deterministic/stochastic system
\[ \begin{cases} x_{k+1} = A_k x_k + B_k u_k + \Gamma_k \xi_k & \\
v_k = C_k x_k + D_k u_k + \xi_k, & \end{cases} \]
where \( \{u_k\} \) is a given sequence of deterministic control input \( m \)-vectors, \( 1 \leq m \leq n. \) Suppose that Assumption 2.1 is satisfied and the matrix \( Var(\xi_k) \) is
nonsingular (cf. (2.2) for the definition of \( \xi_k, j \)). Derive the Kalman filtering
equations for this model.
In digital signal processing, a widely used mathematical model is the following so-called ARMA (autoregressive moving-average) process:

\[ v_k = \sum_{i=1}^{N} B_i v_{k-i} + \sum_{i=0}^{M} A_i u_{k-i}, \]

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 \{\( u_k \)\} and \{\( v_k \)\} are input and output digital signal sequences, respectively (cf. Fig. 2.3). Assuming that \( M \leq N \), show that the input-output relationship can be described as a state-space model

\[
\begin{align*}
\mathbf{x}_{k+1} &= A\mathbf{x}_k + B\mathbf{u}_k \\
\mathbf{v}_k &= C\mathbf{x}_k + D\mathbf{u}_k
\end{align*}
\]

Fig. 2.3 Block diagram of the ARMA model
with $x_0 = 0$, where

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

$$C = [I \ 0 \ \cdots \ 0] \quad and \quad D = [A_0].$$
Kalman Filtering
with Real-Time Applications
Chui, C.K.; Chen, G.
2017, XVIII, 247 p. 34 illus., Hardcover
ISBN: 978-3-319-47610-0