3. Orthogonal Projection and Kalman Filter

The elementary approach to the derivation of the optimal Kalman filtering process discussed in Chapter 2 has the advantage that the optimal estimate $\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_{k|k}$ of the state vector $\mathbf{x}_k$ is easily understood to be a least-squares estimate of $\mathbf{x}_k$ with the properties that (i) the transformation that yields $\hat{\mathbf{x}}_k$ from the data $\overline{\mathbf{v}}_k = [\mathbf{v}_0^\top \cdots \mathbf{v}_k^\top]^\top$ is linear, (ii) $\hat{\mathbf{x}}_k$ is unbiased in the sense that $E(\hat{\mathbf{x}}_k) = E(\mathbf{x}_k)$, and (iii) it yields a minimum variance estimate with $(Var(\overline{\underline{\epsilon}}_{k,k}))^{-1}$ as the optimal weight. The disadvantage of this elementary approach is that certain matrices must be assumed to be nonsingular. In this chapter, we will drop the nonsingularity assumptions and give a rigorous derivation of the Kalman filtering algorithm.

3.1 Orthogonality Characterization of Optimal Estimates

Consider the linear stochastic system described by (2.1) such that Assumption 2.1 is satisfied. That is, consider the state-space description

$$\begin{cases} \mathbf{x}_{k+1} = A_k\mathbf{x}_k + \Gamma_k\underline{\xi}_k \\ \mathbf{v}_k = C_k\mathbf{x}_k + \underline{\eta}_k\,, \end{cases} \tag{3.1}$$

where A_k, Γ_k and C_k are known $n \times n$, $n \times p$ and $q \times n$ constant matrices, respectively, with $1 \le p, q \le n$, and

$$E(\underline{\xi}_k) = 0\,, \qquad E(\underline{\xi}_k\underline{\xi}_\ell^\top) = Q_k\delta_{k\ell}\,,$$

$$E(\underline{\eta}_k) = 0\,, \qquad E(\underline{\eta}_k\underline{\eta}_\ell^\top) = R_k\delta_{k\ell}\,,$$

$$E(\underline{\xi}_k\underline{\eta}_\ell^\top) = 0\,, \quad E(\mathbf{x}_0\underline{\xi}_k^\top) = 0\,, \quad E(\mathbf{x}_0\underline{\eta}_k^\top) = 0\,,$$

for all $k, \ell = 0, 1, \cdots$, with Q_k and R_k being positive definite and symmetric matrices.

Let $\mathbf{x}$ be a random n-vector and $\mathbf{w}$ a random q-vector. We define the "inner product" $\langle \mathbf{x}, \mathbf{w} \rangle$ to be the $n \times q$ matrix

$$\langle \mathbf{x}, \mathbf{w} \rangle = Cov(\mathbf{x}, \mathbf{w}) = E(\mathbf{x} - E(\mathbf{x}))(\mathbf{w} - E(\mathbf{w}))^\top .$$

Let $\|\mathbf{w}\|_q$ be the positive square root of $\langle \mathbf{w}, \mathbf{w} \rangle$. That is, $\|\mathbf{w}\|_q$ is a non-negative definite $q \times q$ matrix with

$$\|\mathbf{w}\|_q^2 = \|\mathbf{w}\|_q \|\mathbf{w}\|_q^\top = \langle \mathbf{w}, \mathbf{w} \rangle .$$

Similarly, let $\|\mathbf{x}\|_n$ be the positive square root of $\langle \mathbf{x}, \mathbf{x} \rangle$. Now, let $\mathbf{w}_0, \cdots, \mathbf{w}_r$ be random q-vectors and consider the "linear span":

$$\begin{aligned}
&Y(\mathbf{w}_0, \cdots, \mathbf{w}_r) \\
&=\{\mathbf{y} : \ \mathbf{y} = \sum_{i=0}^{r} P_i \mathbf{w}_i \, , \ P_0, \cdots, P_r, \ n \times q \ \text{ constant matrices } \} .
\end{aligned}$$

The first minimization problem we will study is to determine a $\hat{\mathbf{y}}$ in $Y(\mathbf{w}_0, \cdots, \mathbf{w}_r)$ such that $tr\|\mathbf{x}_k - \hat{\mathbf{y}}\|_n^2 = F_k$, where

$$F_k := min\{tr\|\mathbf{x}_k - \mathbf{y}\|_n^2 : \ \mathbf{y} \in Y(\mathbf{w}_0, \cdots, \mathbf{w}_r)\} . \tag{3.2}$$

The following result characterizes $\hat{\mathbf{y}}$.

Lemma 3.1. *$\hat{\mathbf{y}} \in Y(\mathbf{w}_0, \cdots, \mathbf{w}_r)$ satisfies $tr\|\mathbf{x}_k - \hat{\mathbf{y}}\|_n^2 = F_k$ if and only if*

$$\langle \mathbf{x}_k - \hat{\mathbf{y}}, \mathbf{w}_j \rangle = O_{n \times q}$$

for all $j = 0, 1, \cdots, r$. Furthermore, $\hat{\mathbf{y}}$ is unique in the sense that

$$tr\|\mathbf{x}_k - \hat{\mathbf{y}}\|_n^2 = tr\|\mathbf{x}_k - \tilde{\mathbf{y}}\|_n^2$$

only if $\hat{\mathbf{y}} = \tilde{\mathbf{y}}$.

To prove this lemma, we first suppose that $tr\|\mathbf{x}_k - \hat{\mathbf{y}}\|_n^2 = F_k$ but $\langle \mathbf{x}_k - \hat{\mathbf{y}}, \mathbf{w}_{j_0} \rangle = C \neq O_{n \times q}$ for some j_0 where $0 \leq j_0 \leq r$. Then $\mathbf{w}_{j_0} \neq 0$ so that $\|\mathbf{w}_{j_0}\|_q^2$ is a positive definite symmetric matrix and so is its inverse $\|\mathbf{w}_{j_0}\|_q^{-2}$. Hence, $C\|\mathbf{w}_{j_0}\|_q^{-2} C^\top \neq O_{n \times n}$ and is a non-negative definite and symmetric matrix. It can be shown that

$$tr\{C\|\mathbf{w}_{j_0}\|_q^{-2} C^\top\} > 0 \tag{3.3}$$

(cf. Exercise 3.1). Now, the vector $\hat{\mathbf{y}} + C\|\mathbf{w}_{j_0}\|_q^{-2}\mathbf{w}_{j_0}$ is in $Y(\mathbf{w}_0, \cdots, \mathbf{w}_r)$ and

$$\begin{aligned}
&tr\|\mathbf{x}_k - (\hat{\mathbf{y}} + C\|\mathbf{w}_{j_0}\|_q^{-2}\mathbf{w}_{j_0})\|_n^2 \\
&= tr\{\|\mathbf{x}_k - \hat{\mathbf{y}}\|_n^2 - \langle \mathbf{x}_k - \hat{\mathbf{y}}, \mathbf{w}_{j_0} \rangle (C\|\mathbf{w}_{j_0}\|_q^{-2})^\top - C\|\mathbf{w}_{j_0}\|_q^{-2} \langle \mathbf{w}_{j_0}, \mathbf{x}_k - \hat{\mathbf{y}} \rangle \\
&\quad + C\|\mathbf{w}_{j_0}\|_q^{-2} \|\mathbf{w}_{j_0}\|_q^2 (C\|\mathbf{w}_{j_0}\|_q^{-2})^\top\} \\
&= tr\{\|\mathbf{x}_k - \hat{\mathbf{y}}\|_n^2 - C\|\mathbf{w}_{j_0}\|_q^{-2} C^\top\} \\
&< tr\|\mathbf{x}_k - \hat{\mathbf{y}}\|_n^2 = F_k
\end{aligned}$$

by using (3.3). This contradicts the definition of F_k in (3.2).

Conversely, let $\langle \mathbf{x}_k - \hat{\mathbf{y}}, \mathbf{w}_j \rangle = O_{n \times q}$ for all $j = 0, 1, \cdots, r$. Let $\mathbf{y}$ be an arbitrary random n-vector in $Y(\mathbf{w}_0, \cdots, \mathbf{w}_r)$ and write $\mathbf{y}_0 = \mathbf{y} - \hat{\mathbf{y}} = \sum_{j=0}^r P_{0j} \mathbf{w}_j$ where P_{0j} are constant $n \times q$ matrices, $j = 0, 1, \cdots, r$. Then

$$
\begin{aligned}
& tr\|\mathbf{x}_k - \mathbf{y}\|_n^2 \\
&= tr\|(\mathbf{x}_k - \hat{\mathbf{y}}) - \mathbf{y}_0\|_n^2 \\
&= tr\{\|\mathbf{x}_k - \hat{\mathbf{y}}\|_n^2 - \langle \mathbf{x}_k - \hat{\mathbf{y}}, \mathbf{y}_0 \rangle - \langle \mathbf{y}_0, \mathbf{x}_k - \hat{\mathbf{y}} \rangle + \|\mathbf{y}_0\|_n^2\} \\
&= tr\left\{ \|\mathbf{x}_k - \hat{\mathbf{y}}\|_n^2 - \sum_{j=0}^r \langle \mathbf{x}_k - \hat{\mathbf{y}}, \mathbf{w}_j \rangle P_{0j}^\top - \sum_{j=0}^r P_{0j} \langle \mathbf{x}_k - \hat{\mathbf{y}}, \mathbf{w}_j \rangle^\top + \|\mathbf{y}_0\|_n^2 \right\} \\
&= tr\|\mathbf{x}_k - \hat{\mathbf{y}}\|_n^2 + tr\|\mathbf{y}_0\|_n^2 \\
&\geq tr\|\mathbf{x}_k - \hat{\mathbf{y}}\|_n^2 ,
\end{aligned}
$$

so that $tr\|\mathbf{x}_k - \hat{\mathbf{y}}\|_n^2 = F_k$. Furthermore, equality is attained if and only if $tr\|\mathbf{y}_0\|_n^2 = 0$ or $\mathbf{y}_0 = 0$ so that $\mathbf{y} = \hat{\mathbf{y}}$ (cf. Exercise 3.1). This completes the proof of the lemma.

3.2 Innovations Sequences

To use the data information, we require an "orthogonalization" process.

Definition 3.1. Given a random q-vector data sequence $\{\mathbf{v}_j\}$, $j = 0, \cdots, k$. The *innovations sequence* $\{\mathbf{z}_j\}$, $j = 0, \cdots, k$, of $\{\mathbf{v}_j\}$ (i.e., a sequence obtained by changing the original data sequence $\{\mathbf{v}_j\}$) is defined by

$$
\mathbf{z}_j = \mathbf{v}_j - C_j \hat{\mathbf{y}}_{j-1} , \quad j = 0, 1, \cdots, k , \tag{3.4}
$$

with $\hat{\mathbf{y}}_{-1} = 0$ and

$$
\hat{\mathbf{y}}_{j-1} = \sum_{i=0}^{j-1} \hat{P}_{j-1,i} \mathbf{v}_i \in Y(\mathbf{v}_0, \cdots, \mathbf{v}_{j-1}) , \quad j = 1, \cdots, k ,
$$

where the $q \times n$ matrices C_j are the observation matrices in (3.1) and the $n \times q$ matrices $\hat{P}_{j-1,i}$ are chosen so that $\hat{\mathbf{y}}_{j-1}$ solves the minimization problem (3.2) with $Y(\mathbf{w}_0, \cdots, \mathbf{w}_r)$ replaced by $Y(\mathbf{v}_0, \cdots, \mathbf{v}_{j-1})$.

We first give the correlation property of the innovations sequence.

Lemma 3.2. *The innovations sequence $\{\mathbf{z}_j\}$ of $\{\mathbf{v}_j\}$ satisfies the following property:*

$$\langle \mathbf{z}_j, \mathbf{z}_\ell \rangle = (R_\ell + C_\ell \|\mathbf{x}_\ell - \hat{\mathbf{y}}_{\ell-1}\|_n^2 C_\ell^\top)\delta_{j\ell}\,,$$

where $R_\ell = Var(\underline{\eta}_\ell) > 0$.

For convenience, we set

$$\hat{\mathbf{e}}_j = C_j(\mathbf{x}_j - \hat{\mathbf{y}}_{j-1})\,. \tag{3.5}$$

To prove the lemma, we first observe that

$$\mathbf{z}_j = \hat{\mathbf{e}}_j + \underline{\eta}_j\,, \tag{3.6}$$

where $\{\underline{\eta}_k\}$ is the observation noise sequence, and

$$\langle \underline{\eta}_\ell, \hat{\mathbf{e}}_j \rangle = O_{q\times q} \qquad \text{for all} \qquad \ell \geq j\,. \tag{3.7}$$

Clearly, (3.6) follows from (3.4), (3.5), and the observation equation in (3.1). The proof of (3.7) is left to the reader as an exercise (cf. Exercise 3.2). Now, for $j = \ell$, we have, by (3.6), (3.7), and (3.5) consecutively,

$$\begin{aligned}
\langle \mathbf{z}_\ell, \mathbf{z}_\ell \rangle &= \langle \hat{\mathbf{e}}_\ell + \underline{\eta}_\ell, \hat{\mathbf{e}}_\ell + \underline{\eta}_\ell \rangle \\
&= \langle \hat{\mathbf{e}}_\ell, \hat{\mathbf{e}}_\ell \rangle + \langle \underline{\eta}_\ell, \underline{\eta}_\ell \rangle \\
&= C_\ell \|\mathbf{x}_\ell - \hat{\mathbf{y}}_{\ell-1}\|_n^2 C_\ell^\top + R_\ell\,.
\end{aligned}$$

For $j \neq \ell$, since $\langle \hat{\mathbf{e}}_\ell, \hat{\mathbf{e}}_j \rangle^\top = \langle \hat{\mathbf{e}}_j, \hat{\mathbf{e}}_\ell \rangle$, we can assume without loss of generality that $j > \ell$. Hence, by (3.6), (3.7), and Lemma 3.1 we have

$$\begin{aligned}
\langle \mathbf{z}_j, \mathbf{z}_\ell \rangle &= \langle \hat{\mathbf{e}}_j, \hat{\mathbf{e}}_\ell \rangle + \langle \hat{\mathbf{e}}_j, \underline{\eta}_\ell \rangle + \langle \underline{\eta}_j, \hat{\mathbf{e}}_\ell \rangle + \langle \underline{\eta}_j, \underline{\eta}_\ell \rangle \\
&= \langle \hat{\mathbf{e}}_j, \hat{\mathbf{e}}_\ell + \underline{\eta}_\ell \rangle \\
&= \langle \hat{\mathbf{e}}_j, \mathbf{z}_\ell \rangle \\
&= \langle \hat{\mathbf{e}}_j, \mathbf{v}_\ell - C_\ell \hat{\mathbf{y}}_{\ell-1} \rangle \\
&= \left\langle C_j(\mathbf{x}_j - \hat{\mathbf{y}}_{j-1}),\, \mathbf{v}_\ell - C_\ell \sum_{i=0}^{\ell-1} \hat{P}_{\ell-1,i} \mathbf{v}_i \right\rangle \\
&= C_j \langle \mathbf{x}_j - \hat{\mathbf{y}}_{j-1}, \mathbf{v}_\ell \rangle - C_j \sum_{i=0}^{\ell-1} \langle \mathbf{x}_j - \hat{\mathbf{y}}_{j-1}, \mathbf{v}_i \rangle \hat{P}_{\ell-1,i}^\top C_\ell^\top \\
&= O_{q\times q}\,.
\end{aligned}$$

This completes the proof of the lemma.

Since $R_j > 0$, Lemma 3.2 says that $\{\mathbf{z}_j\}$ is an "orthogonal" sequence of nonzero vectors which we can normalize by setting

$$\mathbf{e}_j = \|\mathbf{z}_j\|_q^{-1}\mathbf{z}_j\,. \tag{3.8}$$

Then $\{\mathbf{e}_j\}$ is an "orthonormal" sequence in the sense that $\langle \mathbf{e}_i, \mathbf{e}_j\rangle = \delta_{ij}I_q$ for all i and j. Furthermore, it should be clear that

$$Y(\mathbf{e}_0, \cdots, \mathbf{e}_k) = Y(\mathbf{v}_0, \cdots, \mathbf{v}_k) \tag{3.9}$$

(cf. Exercise 3.3).

3.3 Minimum Variance Estimates

We are now ready to give the minimum variance estimate $\check{\mathbf{x}}_k$ of the state vector $\mathbf{x}_k$ by introducing the "Fourier expansion"

$$\check{\mathbf{x}}_k = \sum_{i=0}^{k} \langle \mathbf{x}_k, \mathbf{e}_i\rangle \mathbf{e}_i \tag{3.10}$$

of $\mathbf{x}_k$ with respect to the "orthonormal" sequence $\{\mathbf{e}_j\}$. Since

$$\langle \check{\mathbf{x}}_k, \mathbf{e}_j\rangle = \sum_{i=0}^{k} \langle \mathbf{x}_k, \mathbf{e}_i\rangle \langle \mathbf{e}_i, \mathbf{e}_j\rangle = \langle \mathbf{x}_k, \mathbf{e}_j\rangle\,,$$

we have

$$\langle \mathbf{x}_k - \check{\mathbf{x}}_k, \mathbf{e}_j\rangle = O_{n\times q}\,, \quad j = 0, 1, \cdots, k\,. \tag{3.11}$$

It follows from Exercise 3.3 that

$$\langle \mathbf{x}_k - \check{\mathbf{x}}_k, \mathbf{v}_j\rangle = O_{n\times q}\,, \quad j = 0, 1, \cdots, k\,, \tag{3.12}$$

so that by Lemma 3.1,

$$tr\|\mathbf{x}_k - \check{\mathbf{x}}_k\|_n^2 = min\{tr\|\mathbf{x}_k - \mathbf{y}\|_n^2 : \ \mathbf{y} \in Y(\mathbf{v}_0, \cdots, \mathbf{v}_k)\}\,.$$

That is, $\check{\mathbf{x}}_k$ is a minimum variance estimate of $\mathbf{x}_k$.

3.4 Kalman Filtering Equations

This section is devoted to the derivation of the Kalman filtering equations. From Assumption 2.1, we first observe that

$$\langle \underline{\xi}_{k-1}, \mathbf{e}_j \rangle = O_{n\times q}, \quad j = 0, 1, \cdots, k-1,$$

(cf. Exercise 3.4), so that

$$\begin{aligned}
\check{\mathbf{x}}_k &= \sum_{j=0}^{k} \langle \mathbf{x}_k, \mathbf{e}_j \rangle \mathbf{e}_j \\
&= \sum_{j=0}^{k-1} \langle \mathbf{x}_k, \mathbf{e}_j \rangle \mathbf{e}_j + \langle \mathbf{x}_k, \mathbf{e}_k \rangle \mathbf{e}_k \\
&= \sum_{j=0}^{k-1} \{ \langle A_{k-1}\mathbf{x}_{k-1}, \mathbf{e}_j \rangle \mathbf{e}_j + \langle \Gamma_{k-1}\underline{\xi}_{k-1}, \mathbf{e}_j \rangle \mathbf{e}_j \} + \langle \mathbf{x}_k, \mathbf{e}_k \rangle \mathbf{e}_k \\
&= A_{k-1} \sum_{j=0}^{k-1} \langle \mathbf{x}_{k-1}, \mathbf{e}_j \rangle \mathbf{e}_j + \langle \mathbf{x}_k, \mathbf{e}_k \rangle \mathbf{e}_k \\
&= A_{k-1}\check{\mathbf{x}}_{k-1} + \langle \mathbf{x}_k, \mathbf{e}_k \rangle \mathbf{e}_k .
\end{aligned}$$

Hence, by defining

$$\check{\mathbf{x}}_{k|k-1} = A_{k-1}\check{\mathbf{x}}_{k-1}, \tag{3.13}$$

where $\check{\mathbf{x}}_{k-1} := \check{\mathbf{x}}_{k-1|k-1}$, we have

$$\check{\mathbf{x}}_k = \check{\mathbf{x}}_{k|k} = \check{\mathbf{x}}_{k|k-1} + \langle \mathbf{x}_k, \mathbf{e}_k \rangle \mathbf{e}_k . \tag{3.14}$$

Obviously, if we can show that there exists a constant $n \times q$ matrix $\check{G}_k$ such that

$$\langle \mathbf{x}_k, \mathbf{e}_k \rangle \mathbf{e}_k = \check{G}_k(\mathbf{v}_k - C_k\check{\mathbf{x}}_{k|k-1}),$$

then the "prediction-correction" formulation of the Kalman filter is obtained. To accomplish this, we consider the random vector $(\mathbf{v}_k - C_k\check{\mathbf{x}}_{k|k-1})$ and obtain the following:

Lemma 3.3. *For* $j = 0, 1, \cdots, k$,

$$\langle \mathbf{v}_k - C_k\check{\mathbf{x}}_{k|k-1}, \mathbf{e}_j \rangle = \|\mathbf{z}_k\|_q \delta_{kj} .$$

To prove the lemma, we first observe that

$$\langle \hat{\mathbf{y}}_j, \mathbf{z}_k \rangle = O_{n\times q}, \quad j = 0, 1, \cdots, k-1, \tag{3.15}$$

(cf. Exercise 3.4). Hence, using (3.14), (3.11), and (3.15), we have

$$
\begin{aligned}
&\langle \mathbf{v}_k - C_k \check{\mathbf{x}}_{k|k-1}, \mathbf{e}_k \rangle \\
&= \langle \mathbf{v}_k - C_k(\check{\mathbf{x}}_{k|k} - \langle \mathbf{x}_k, \mathbf{e}_k \rangle \mathbf{e}_k), \mathbf{e}_k \rangle \\
&= \langle \mathbf{v}_k, \mathbf{e}_k \rangle - C_k \{ \langle \check{\mathbf{x}}_{k|k}, \mathbf{e}_k \rangle - \langle \mathbf{x}_k, \mathbf{e}_k \rangle \} \\
&= \langle \mathbf{v}_k, \mathbf{e}_k \rangle - C_k \langle \check{\mathbf{x}}_{k|k} - \mathbf{x}_k, \mathbf{e}_k \rangle \\
&= \langle \mathbf{v}_k, \mathbf{e}_k \rangle \\
&= \langle \mathbf{z}_k + C_k \hat{\mathbf{y}}_{k-1}, \|\mathbf{z}_k\|_q^{-1} \mathbf{z}_k \rangle \\
&= \langle \mathbf{z}_k, \mathbf{z}_k \rangle \|\mathbf{z}_k\|_q^{-1} + C_k \langle \hat{\mathbf{y}}_{k-1}, \mathbf{z}_k \rangle \|\mathbf{z}_k\|_q^{-1} \\
&= \|\mathbf{z}_k\|_q .
\end{aligned}
$$

On the other hand, using (3.14), (3.11), and (3.7), we have

$$
\begin{aligned}
&\langle \mathbf{v}_k - C_k \check{\mathbf{x}}_{k|k-1}, \mathbf{e}_j \rangle \\
&= \langle C_k \mathbf{x}_k + \underline{\eta}_k - C_k(\check{\mathbf{x}}_{k|k} - \langle \mathbf{x}_k, \mathbf{e}_k \rangle \mathbf{e}_k), \mathbf{e}_j \rangle \\
&= C_k \langle \mathbf{x}_k - \check{\mathbf{x}}_{k|k}, \mathbf{e}_j \rangle + \langle \underline{\eta}_k, \mathbf{e}_j \rangle + C_k \langle \mathbf{x}_k, \mathbf{e}_k \rangle \langle \mathbf{e}_k, \mathbf{e}_j \rangle \\
&= O_{q \times q}
\end{aligned}
$$

for $j = 0, 1, \cdots, k-1$. This completes the proof of the Lemma.

It is clear, by using Exercise 3.3 and the definition of $\check{\mathbf{x}}_{k-1} = \check{\mathbf{x}}_{k-1|k-1}$, that the random q-vector $(\mathbf{v}_k - C_k \check{\mathbf{x}}_{k|k-1})$ can be expressed as $\sum_{i=0}^{k} M_i \mathbf{e}_i$ for some constant $q \times q$ matrices M_i. It follows now from Lemma 3.3 that for $j = 0, 1, \cdots, k$,

$$
\left\langle \sum_{i=0}^{k} M_i \mathbf{e}_i, \mathbf{e}_j \right\rangle = \|\mathbf{z}_k\|_q \delta_{kj} ,
$$

so that $M_0 = M_1 = \cdots = M_{k-1} = 0$ and $M_k = \|\mathbf{z}_k\|_q$. Hence,

$$
\mathbf{v}_k - C_k \check{\mathbf{x}}_{k|k-1} = M_k \mathbf{e}_k = \|\mathbf{z}_k\|_q \mathbf{e}_k .
$$

Define

$$
\check{G}_k = \langle \mathbf{x}_k, \mathbf{e}_k \rangle \|\mathbf{z}_k\|_q^{-1} .
$$

Then we obtain

$$
\langle \mathbf{x}_k, \mathbf{e}_k \rangle \mathbf{e}_k = \check{G}_k (\mathbf{v}_k - C_k \check{\mathbf{x}}_{k|k-1}) .
$$

This, together with (3.14), gives the "prediction-correction" equation:

$$
\check{\mathbf{x}}_{k|k} = \check{\mathbf{x}}_{k|k-1} + \check{G}_k (\mathbf{v}_k - C_k \check{\mathbf{x}}_{k|k-1}) . \tag{3.16}
$$

We remark that $\check{\mathbf{x}}_{k|k}$ is an unbiased estimate of $\mathbf{x}_k$ by choosing an appropriate initial estimate. In fact,

$$\begin{aligned}&\mathbf{x}_k - \check{\mathbf{x}}_{k|k}\\ =&A_{k-1}\mathbf{x}_{k-1} + \Gamma_{k-1}\underline{\xi}_{k-1} - A_{k-1}\check{\mathbf{x}}_{k-1|k-1} - \check{G}_k(\mathbf{v}_k - C_kA_{k-1}\check{\mathbf{x}}_{k-1|k-1}).\end{aligned}$$

By using $\mathbf{v}_k = C_k\mathbf{x}_k + \underline{\eta}_k = C_kA_{k-1}\mathbf{x}_{k-1} + C_k\Gamma_{k-1}\underline{\xi}_{k-1} + \underline{\eta}_k$, we have

$$\begin{aligned}&\mathbf{x}_k - \check{\mathbf{x}}_{k|k}\\ =&(I - \check{G}_kC_k)A_{k-1}(\mathbf{x}_{k-1} - \check{\mathbf{x}}_{k-1|k-1})\\ &+ (I - \check{G}_kC_k)\Gamma_{k-1}\underline{\xi}_{k-1} - \check{G}_k\underline{\eta}_k.\end{aligned} \tag{3.17}$$

Since the noise sequences are of zero-mean, we have

$$E(\mathbf{x}_k - \check{\mathbf{x}}_{k|k}) = (I - \check{G}_kC_k)A_{k-1}E(\mathbf{x}_{k-1} - \check{\mathbf{x}}_{k-1|k-1}),$$

so that

$$E(\mathbf{x}_k - \check{\mathbf{x}}_{k|k}) = (I - \check{G}_kC_k)A_{k-1}\cdots(I - \check{G}_1C_1)A_0E(\mathbf{x}_0 - \check{\mathbf{x}}_{0|0}).$$

Hence, if we set

$$\check{\mathbf{x}}_{0|0} = E(\mathbf{x}_0), \tag{3.18}$$

then $E(\mathbf{x}_k - \check{\mathbf{x}}_{k|k}) = 0$ or $E(\check{\mathbf{x}}_{k|k}) = E(\mathbf{x}_k)$ for all k, i.e., $\check{\mathbf{x}}_{k|k}$ is indeed an unbiased estimate of $\mathbf{x}_k$.

Now what is left is to derive a recursive formula for $\check{G}_k$. Using (3.12) and (3.17), we first have

$$\begin{aligned}0 &= \langle \mathbf{x}_k - \check{\mathbf{x}}_{k|k}, \mathbf{v}_k\rangle\\ &= \langle (I - \check{G}_kC_k)A_{k-1}(\mathbf{x}_{k-1} - \check{\mathbf{x}}_{k-1|k-1}) + (I - \check{G}_kC_k)\Gamma_{k-1}\underline{\xi}_{k-1} - \check{G}_k\underline{\eta}_k,\\ &\quad C_kA_{k-1}((\mathbf{x}_{k-1} - \check{\mathbf{x}}_{k-1|k-1}) + \check{\mathbf{x}}_{k-1|k-1}) + C_k\Gamma_{k-1}\underline{\xi}_{k-1} + \underline{\eta}_k\rangle\\ &= (I - \check{G}_kC_k)A_{k-1}\|\mathbf{x}_{k-1} - \check{\mathbf{x}}_{k-1|k-1}\|_n^2A_{k-1}^\top C_k^\top\\ &\quad + (I - \check{G}_kC_k)\Gamma_{k-1}Q_{k-1}\Gamma_{k-1}^\top C_k^\top - \check{G}_kR_k,\end{aligned} \tag{3.19}$$

where we have used the facts that $\langle \mathbf{x}_{k-1} - \check{\mathbf{x}}_{k-1|k-1}, \check{\mathbf{x}}_{k-1|k-1}\rangle = O_{n\times n}$, a consequence of Lemma 3.1, and

$$\begin{aligned}&\langle \mathbf{x}_k, \underline{\xi}_k\rangle = O_{n\times n}, \quad \langle \check{\mathbf{x}}_{k|k}, \underline{\xi}_j\rangle = O_{n\times n},\\ &\langle \mathbf{x}_k, \underline{\eta}_j\rangle = O_{n\times q}, \quad \langle \check{\mathbf{x}}_{k-1|k-1}, \underline{\eta}_k\rangle = O_{n\times q},\end{aligned} \tag{3.20}$$

$j = 0, \cdots, k$ (cf. Exercise 3.5). Define

$$P_{k,k} = \|\mathbf{x}_k - \check{\mathbf{x}}_{k|k}\|_n^2$$

and

$$P_{k,k-1} = \|\mathbf{x}_k - \check{\mathbf{x}}_{k|k-1}\|_n^2 .$$

Then again by Exercise 3.5 we have

$$\begin{aligned} P_{k,k-1} &= \|A_{k-1}\mathbf{x}_{k-1} + \Gamma_{k-1}\underline{\xi}_{k-1} - A_{k-1}\check{\mathbf{x}}_{k-1|k-1}\|_n^2 \\ &= A_{k-1}\|\mathbf{x}_{k-1} - \check{\mathbf{x}}_{k-1|k-1}\|_n^2 A_{k-1}^\top + \Gamma_{k-1}Q_{k-1}\Gamma_{k-1}^\top \end{aligned}$$

or

$$P_{k,k-1} = A_{k-1}P_{k-1,k-1}A_{k-1}^\top + \Gamma_{k-1}Q_{k-1}\Gamma_{k-1}^\top . \tag{3.21}$$

On the other hand, from (3.19), we also obtain

$$\begin{aligned} &(I - \check{G}_kC_k)A_{k-1}P_{k-1,k-1}A_{k-1}^\top C_k^\top \\ &+ (I - \check{G}_kC_k)\Gamma_{k-1}Q_{k-1}\Gamma_{k-1}^\top C_k^\top - \check{G}_kR_k = 0 . \end{aligned}$$

In solving for $\check{G}_k$ from this expression, we write

$$\begin{aligned} &\check{G}_k[R_k + C_k(A_{k-1}P_{k-1,k-1}A_{k-1}^\top + \Gamma_{k-1}Q_{k-1}\Gamma_{k-1}^\top)C_k^\top] \\ &= [A_{k-1}P_{k-1,k-1}A_{k-1}^\top + \Gamma_{k-1}Q_{k-1}\Gamma_{k-1}^\top]C_k^\top \\ &= P_{k,k-1}C_k^\top . \end{aligned}$$

and obtain

$$\check{G}_k = P_{k,k-1}C_k^\top(R_k + C_kP_{k,k-1}C_k^\top)^{-1} , \tag{3.22}$$

where R_k is positive definite and $C_kP_{k,k-1}C_k^\top$ is non-negative definite so that their sum is positive definite (cf. Exercise 2.2).

Next, we wish to write $P_{k,k}$ in terms of $P_{k,k-1}$, so that together with (3.21), we will have a recursive scheme. This can be done as follows:

$$\begin{aligned} P_{k,k} &= \|\mathbf{x}_k - \check{\mathbf{x}}_{k|k}\|_n^2 \\ &= \|\mathbf{x}_k - (\check{\mathbf{x}}_{k|k-1} + \check{G}_k(\mathbf{v}_k - C_k\check{\mathbf{x}}_{k|k-1}))\|_n^2 \\ &= \|\mathbf{x}_k - \check{\mathbf{x}}_{k|k-1} - \check{G}_k(C_k\mathbf{x}_k + \underline{\eta}_k) + \check{G}_kC_k\check{\mathbf{x}}_{k|k-1}\|_n^2 \\ &= \|(I - \check{G}_kC_k)(\mathbf{x}_k - \check{\mathbf{x}}_{k|k-1}) - \check{G}_k\underline{\eta}_k\|_n^2 \\ &= (I - \check{G}_kC_k)\|\mathbf{x}_k - \check{\mathbf{x}}_{k|k-1}\|_n^2(I - \check{G}_kC_k)^\top + \check{G}_kR_k\check{G}_k^\top \\ &= (I - \check{G}_kC_k)P_{k,k-1}(I - \check{G}_kC_k)^\top + \check{G}_kR_k\check{G}_k^\top , \end{aligned}$$

where we have applied Exercise 3.5 to conclude that $\langle \mathbf{x}_k - \check{\mathbf{x}}_{k|k-1}, \underline{\eta}_k\rangle = O_{n\times q}$. This relation can be further simplified by using (3.22). Indeed, since

$$\begin{aligned} &(I - \check{G}_kC_k)P_{k,k-1}(\check{G}_kC_k)^\top \\ =&P_{k,k-1}C_k^\top\check{G}_k^\top - \check{G}_kC_kP_{k,k-1}C_k^\top\check{G}_k^\top \\ =&\check{G}_kC_kP_{k,k-1}C_k^\top\check{G}_k^\top + \check{G}_kR_k\check{G}_k^\top - \check{G}_kC_kP_{k,k-1}C_k^\top\check{G}_k^\top \\ =&\check{G}_kR_k\check{G}_k^\top , \end{aligned}$$

we have

$$\begin{aligned} P_{k,k} =& (I - \check{G}_k C_k) P_{k,k-1} (I - \check{G}_k C_k)^\top + (I - \check{G}_k C_k) P_{k,k-1} (\check{G}_k C_k)^\top \\ =& (I - \check{G}_k C_k) P_{k,k-1} . \end{aligned} \tag{3.23}$$

Therefore, combining (3.13), (3.16), (3.18), (3.21), (3.22) and (3.23), together with

$$P_{0,0} = \|\mathbf{x}_0 - \check{\mathbf{x}}_{0|0}\|_n^2 = Var(\mathbf{x}_0) , \tag{3.24}$$

we obtain the Kalman filtering equations which agree with the ones we derived in Chapter 2. That is, we have $\check{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k}$, $\check{\mathbf{x}}_{k|k-1} = \hat{\mathbf{x}}_{k|k-1}$ and $\check{G}_k = G_k$ as follows:

$$\begin{cases} P_{0,0} = Var(\mathbf{x}_0) \\ P_{k,k-1} = A_{k-1} P_{k-1,k-1} A_{k-1}^\top + \Gamma_{k-1} Q_{k-1} \Gamma_{k-1}^\top \\ G_k = P_{k,k-1} C_k^\top (C_k P_{k,k-1} C_k^\top + R_k)^{-1} \\ P_{k,k} = (I - G_k C_k) P_{k,k-1} \\ \hat{\mathbf{x}}_{0|0} = E(\mathbf{x}_0) \\ \hat{\mathbf{x}}_{k|k-1} = A_{k-1} \hat{\mathbf{x}}_{k-1|k-1} \\ \hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + G_k (\mathbf{v}_k - C_k \hat{\mathbf{x}}_{k|k-1}) \\ k = 1, 2, \cdots . \end{cases} \tag{3.25}$$

Of course, the Kalman filtering equations (2.18) derived in Section 2.4 for the general linear deterministic/stochastic system

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

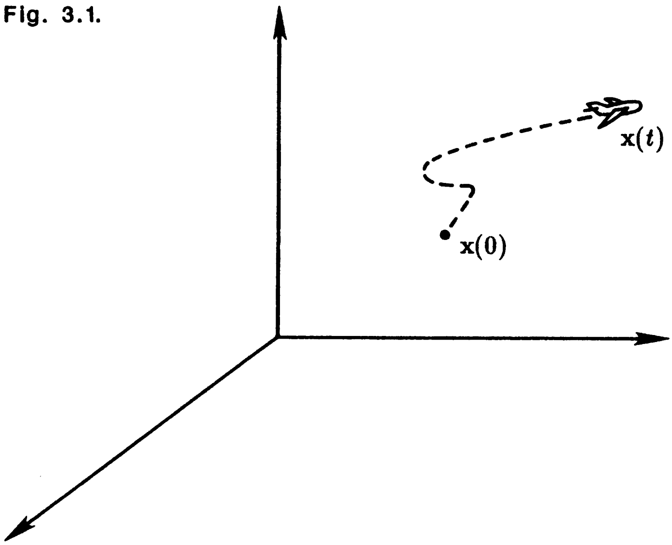
can also be obtained without the assumption on the invertibility of the matrices $A_k, Var(\overline{\underline{\xi}}_{k,j})$, etc. (cf. Exercise 3.6).

3.5 Real-Time Tracking

To illustrate the application of the Kalman filtering algorithm described by (3.25), let us consider an example of real-time tracking. Let $\mathbf{x}(t)$, $0 \leq t < \infty$, denote the trajectory in three-dimensional space of a flying object, where t denotes the time variable (cf. Fig.3.1). This vector-valued function is discretized by sampling and quantizing with sampling time $h > 0$ to yield

$$\mathbf{x}_k \doteq \mathbf{x}(kh), \qquad k = 0, 1, \cdots .$$

Fig. 3.1.



For practical purposes, $\mathbf{x}(t)$ can be assumed to have continuous first and second order derivatives, denoted by $\dot{\mathbf{x}}(t)$ and $\ddot{\mathbf{x}}(t)$, respectively, so that for small values of h, the position and velocity vectors $\mathbf{x}_k$ and $\dot{\mathbf{x}}_k \doteq \dot{\mathbf{x}}(kh)$ are governed by the equations

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{x}_k + h\dot{\mathbf{x}}_k + \frac{1}{2}h^2\ddot{\mathbf{x}}_k \\ \dot{\mathbf{x}}_{k+1} = \dot{\mathbf{x}}_k + h\ddot{\mathbf{x}}_k\,, \end{cases}$$

where $\ddot{\mathbf{x}}_k \doteq \ddot{\mathbf{x}}(kh)$ and $k = 0, 1, \cdots$. In addition, in many applications only the position (vector) of the flying object is observed at each time instant, so that $\mathbf{v}_k = C\mathbf{x}_k$ with $C = [\ I \quad 0 \quad 0\ ]$ is measured. In view of Exercise 3.8, to facilitate our discussion we only consider the tracking model

$$\begin{cases} \begin{bmatrix} x_{k+1}[1] \\ x_{k+1}[2] \\ x_{k+1}[3] \end{bmatrix} = \begin{bmatrix} 1 & h & h^2/2 \\ 0 & 1 & h \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_k[1] \\ x_k[2] \\ x_k[3] \end{bmatrix} + \begin{bmatrix} \xi_k[1] \\ \xi_k[2] \\ \xi_k[3] \end{bmatrix} \\ v_k = [\ 1 \quad 0 \quad 0\ ] \begin{bmatrix} x_k[1] \\ x_k[2] \\ x_k[3] \end{bmatrix} + \eta_k\,. \end{cases} \tag{3.26}$$

Here, $\{\underline{\xi}_k\}$, with $\underline{\xi}_k := [\ \xi_k[1] \quad \xi_k[2] \quad \xi_k[3]\ ]^\top$, and $\{\eta_k\}$ are assumed to be zero-mean Gaussian white noise sequences satisfying:

$$\begin{aligned}
&E(\underline{\xi}_k) = 0\,, \qquad E(\eta_k) = 0\,,\\
&E(\underline{\xi}_k \underline{\xi}_\ell^\top) = Q_k \delta_{k\ell}\,, \qquad E(\eta_k \eta_\ell) = r_k \delta_{k\ell}\,, \qquad E(\underline{\xi}_k \eta_\ell) = 0\,,\\
&E(\mathbf{x}_0 \underline{\xi}_k^\top) = 0\,, \qquad E(\mathbf{x}_0 \eta_k) = 0\,,
\end{aligned}$$

where Q_k is a non-negative definite symmetric matrix and $r_k > 0$ for all k. It is further assumed that initial conditions $E(\mathbf{x}_0)$ and $Var(\mathbf{x}_0)$ are given. For this tracking model, the Kalman filtering algorithm can be specified as follows: Let $P_k := P_{k,k}$ and let $P[i,j]$ denote the $(i,j)th$ entry of P. Then we have

$$\begin{aligned}
P_{k,k-1}[1,1] &= P_{k-1}[1,1] + 2hP_{k-1}[1,2] + h^2 P_{k-1}[1,3] + h^2 P_{k-1}[2,2]\\
&\qquad + h^3 P_{k-1}[2,3] + \frac{h^4}{4} P_{k-1}[3,3] + Q_{k-1}[1,1]\,,\\
P_{k,k-1}[1,2] &= P_{k,k-1}[2,1]\\
&= P_{k-1}[1,2] + hP_{k-1}[1,3] + hP_{k-1}[2,2] + \frac{3h^2}{2} P_{k-1}[2,3]\\
&\qquad + \frac{h^3}{2} P_{k-1}[3,3] + Q_{k-1}[1,2]\,,\\
P_{k,k-1}[2,2] &= P_{k-1}[2,2] + 2hP_{k-1}[2,3] + h^2 P_{k-1}[3,3] + Q_{k-1}[2,2]\,,\\
P_{k,k-1}[1,3] &= P_{k,k-1}[3,1]\\
&= P_{k-1}[1,3] + hP_{k-1}[2,3] + \frac{h^2}{2} P_{k-1}[3,3] + Q_{k-1}[1,3]\,,\\
P_{k,k-1}[2,3] &= P_{k,k-1}[3,2]\\
&= P_{k-1}[2,3] + hP_{k-1}[3,3] + Q_{k-1}[2,3]\,,\\
P_{k,k-1}[3,3] &= P_{k-1}[3,3] + Q_{k-1}[3,3]\,,
\end{aligned}$$

with $P_{0,0} = Var(\mathbf{x}_0)$,

$$G_k = \frac{1}{P_{k,k-1}[1,1] + r_k} \begin{bmatrix} P_{k,k-1}[1,1] \\ P_{k,k-1}[1,2] \\ P_{k,k-1}[1,3] \end{bmatrix},$$

$$P_k = P_{k,k-1} - \frac{1}{P_{k,k-1}[1,1] + r_k} \times$$

$$\begin{bmatrix} P_{k,k-1}^2[1,1] & P_{k,k-1}[1,1]P_{k,k-1}[1,2] & P_{k,k-1}[1,1]P_{k,k-1}[1,3] \\ P_{k,k-1}[1,1]P_{k,k-1}[1,2] & P_{k,k-1}^2[1,2] & P_{k,k-1}[1,2]P_{k,k-1}[1,3] \\ P_{k,k-1}[1,1]P_{k,k-1}[1,3] & P_{k,k-1}[1,2]P_{k,k-1}[1,3] & P_{k,k-1}^2[1,3] \end{bmatrix}$$

and the Kalman filtering algorithm is given by

$$\begin{bmatrix} \hat{x}_{k|k}[1] \\ \hat{x}_{k|k}[2] \\ \hat{x}_{k|k}[3] \end{bmatrix} = \begin{bmatrix} 1 - G_k[1] & (1 - G_k[1])h & (1 - G_k[1])h^2/2 \\ -G_k[2] & 1 - hG_k[2] & h - h^2 G_k[2]/2 \\ -G_k[3] & -hG_k[3] & 1 - h^2 G_k[3]/2 \end{bmatrix}$$

$$\cdot \begin{bmatrix} \hat{x}_{k-1|k-1}[1] \\ \hat{x}_{k-1|k-1}[2] \\ \hat{x}_{k-1|k-1}[3] \end{bmatrix} + \begin{bmatrix} G_k[1] \\ G_k[2] \\ G_k[3] \end{bmatrix} v_k \tag{3.27}$$

with $\hat{\mathbf{x}}_{0|0} = E(\mathbf{x}_0)$.

Exercises

3.1. Let $A \neq 0$ be a non-negative definite and symmetric constant matrix. Show that $\mathrm{tr}A > 0$. (Hint: Decompose A as $A = BB^\top$ with $B \neq 0$.)

3.2. Let

$$\hat{\mathbf{e}}_j = C_j(\mathbf{x}_j - \hat{\mathbf{y}}_{j-1}) = C_j\left(\mathbf{x}_j - \sum_{i=0}^{j-1} \hat{P}_{j-1,i}\mathbf{v}_i\right),$$

where $\hat{P}_{j-1,i}$ are some constant matrices. Use Assumption 2.1 to show that

$$\langle \underline{\eta}_\ell,\ \hat{\mathbf{e}}_j\rangle = O_{q\times q}$$

for all $\ell \geq j$.

3.3. For random vectors $\mathbf{w}_0, \cdots, \mathbf{w}_r$, define

$$Y(\mathbf{w}_0, \cdots, \mathbf{w}_r) \\ = \left\{\mathbf{y}:\ \mathbf{y} = \sum_{i=0}^{r} P_i\mathbf{w}_i\,,\quad P_0, \cdots, P_r,\ \textit{constant}\ \ \textit{matrices}\right\}.$$

Let

$$\mathbf{z}_j = \mathbf{v}_j - C_j\sum_{i=0}^{j-1}\hat{P}_{j-1,i}\mathbf{v}_i$$

be defined as in (3.4) and $\mathbf{e}_j = \|\mathbf{z}_j\|^{-1}\mathbf{z}_j$. Show that

$$Y(\mathbf{e}_0, \cdots, \mathbf{e}_k) = Y(\mathbf{v}_0, \cdots, \mathbf{v}_k)\,.$$

3.4. Let

$$\hat{\mathbf{y}}_{j-1} = \sum_{i=0}^{j-1}\hat{P}_{j-1,i}\mathbf{v}_i$$

and

$$\mathbf{z}_j = \mathbf{v}_j - C_j\sum_{i=0}^{j-1}\hat{P}_{j-1,i}\mathbf{v}_i\,.$$

Show that

$$\langle\hat{\mathbf{y}}_j,\ \mathbf{z}_k\rangle = O_{n\times q},\qquad j = 0, 1, \cdots, k-1\,.$$

3.5. Let $\mathbf{e}_j$ be defined as in Exercise 3.3. Also define

$$\check{\mathbf{x}}_k = \sum_{i=0}^{k} \langle \mathbf{x}_k,\ \mathbf{e}_i \rangle \mathbf{e}_i$$

as in (3.10). Show that

$$\langle \mathbf{x}_k,\ \underline{\xi}_k \rangle = O_{n\times n}\,, \qquad \langle \check{\mathbf{x}}_{k|k},\ \underline{\xi}_j \rangle = O_{n\times n}\,,$$

$$\langle \mathbf{x}_k,\ \underline{\eta}_j \rangle = O_{n\times q}\,, \qquad \langle \check{\mathbf{x}}_{k-1|k-1},\ \underline{\eta}_k \rangle = O_{n\times q}\,,$$

for $j = 0, 1, \cdots, k$.

3.6. 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. Derive the Kalman filtering algorithm for this model.

3.7. Consider a simplified radar tracking model where a large-amplitude and narrow-width impulse signal is transmitted by an antenna. The impulse signal propagates at the speed of light c, and is reflected by a flying object being tracked. The radar antenna receives the reflected signal so that a time-difference Δt is obtained. The range (or distance) d from the radar to the object is then given by $d = c\Delta t/2$. The impulse signal is transmitted periodically with period h. Assume that the object is traveling at a constant velocity w with random disturbance $\xi \sim N(0, q)$, so that the range d satisfies the difference equation

$$d_{k+1} = d_k + h(w_k + \xi_k)\,.$$

Suppose also that the measured range using the formula $d = c\Delta t/2$ has an inherent error Δd and is contaminated with noise η where $\eta \sim N(0, r)$, so that

$$v_k = d_k + \Delta d_k + \eta_k\,.$$

Assume that the initial target range is d_0 which is independent of ξ_k and η_k, and that $\{\xi_k\}$ and $\{\eta_k\}$ are also independent (cf. Fig.3.2). Derive a Kalman filtering algorithm as a range-estimator for this radar tracking system.

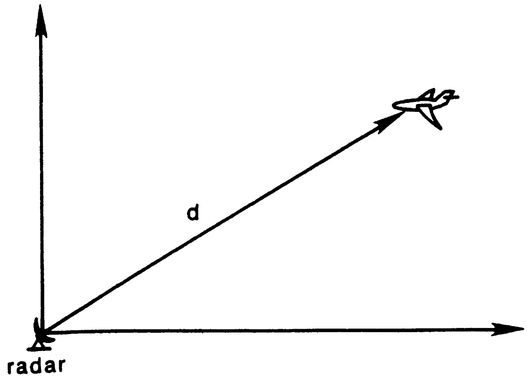


Fig. 3.2.

3.8. A linear stochastic system for radar tracking can be described as follows. Let Σ, ΔA, ΔE be the range, the azimuthal angular error, and the elevational angular error, respectively, of the target, with the radar being located at the origin (cf. Fig.3.3). Consider $\Sigma, \Delta A$, and ΔE as functions of time with first and second derivatives denoted by $\dot{\Sigma}$, $\Delta\dot{A}$, $\Delta\dot{E}$, $\ddot{\Sigma}$, $\Delta\ddot{A}$, $\Delta\ddot{E}$, respectively. Let $h > 0$ be the sampling time unit and set $\Sigma_k = \Sigma(kh)$, $\dot{\Sigma}_k = \dot{\Sigma}(kh)$, $\ddot{\Sigma}_k = \ddot{\Sigma}(kh)$, etc. Then, using the second degree Taylor polynomial approximation, the radar tracking model takes on the following linear stochastic state-space description:

$$\begin{cases} \mathbf{x}_{k+1} = \tilde{A}\mathbf{x}_k + \Gamma_k \underline{\xi}_k \\ \mathbf{v}_k = \tilde{C}\mathbf{x}_k + \underline{\eta}_k \,, \end{cases}$$

where

$$\mathbf{x}_k = \begin{bmatrix}\Sigma_k & \dot{\Sigma}_k & \ddot{\Sigma}_k & \Delta A_k & \Delta\dot{A}_k & \Delta\ddot{A}_k & \Delta E_k & \Delta\dot{E}_k & \Delta\ddot{E}_k\end{bmatrix}^\top,$$

$$\tilde{A} = \begin{bmatrix} 1 & h & h^2/2 & & & & & & \\ 0 & 1 & h & & & & & & \\ 0 & 0 & 1 & & & & & & \\ & & & 1 & h & h^2/2 & & & \\ & & & 0 & 1 & h & & & \\ & & & 0 & 0 & 1 & & & \\ & & & & & & 1 & h & h^2/2 \\ & & & & & & 0 & 1 & h \\ & & & & & & 0 & 0 & 1 \end{bmatrix},$$

$$\tilde{C} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix},$$

and $\{\underline{\xi}_k\}$ and $\{\underline{\eta}_k\}$ are independent zero-mean Gaussian white noise sequences with $Var(\underline{\xi}_k) = Q_k$ and $Var(\underline{\eta}_k) = R_k$. Assume that

$$\Gamma_k = \begin{bmatrix} \Gamma_k^1 & & \\ & \Gamma_k^2 & \\ & & \Gamma_k^3 \end{bmatrix},$$

$$Q_k = \begin{bmatrix} Q_k^1 & & \\ & Q_k^2 & \\ & & Q_k^3 \end{bmatrix}, \qquad R_k = \begin{bmatrix} R_k^1 & & \\ & R_k^2 & \\ & & R_k^3 \end{bmatrix},$$

where Γ_k^i are 3×3 submatrices, Q_k^i, 3×3 non-negative definite symmetric submatrices, and R_k^i, 3×3 positive definite symmetric submatrices, for $i = 1, 2, 3$. Show that this system can be decoupled into three subsystems with analogous state-space descriptions.

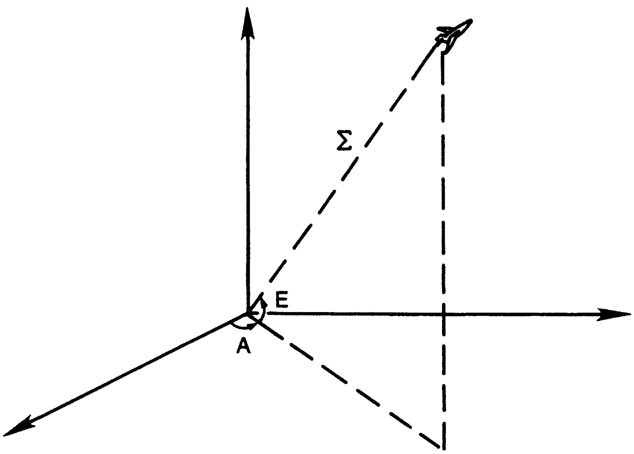


Fig. 3.3.