Borrower: TYC

Lending String: *UCW,CTD,FAU,CTN

Patron: Cruz-Uribe, David

Journal Title: International journal of mathematical education in science and technology.

Volume: 17 **Issue:** 3
Month/Year: 1986**Pages:** 347 - 366

Article Author:

Article Title: Lawrence R. Weill; Paul N. De Land; The Kalman filter; an introduction to the mathematics of linear least mean square recursive estimation

Imprint: London ; New York ; Wiley & Sons, 1970-

ILL Number: 57409711

Call #:

Location:

ARIEL
Charge
Maxcost: $50IFM

Shipping Address:
Trinity College
Library - ILL Dept
300 Summit St
Hartford, CT 06106-3100

Fax: 860 297-2251
Ariel: ariel.trincoll.edu

INT. J. MATH. EDUC. SCI. TECHNOL., 1986, VOL. 17, NO. 3, 347–366

The Kalman filter: an introduction to the mathematics of linear least mean square recursive estimation

by LAWRENCE R. WEILL and PAUL N. DE LAND

Department of Mathematics, California State University, Fullerton, California 92634, U.S.A.

(*Received 17 April 1984*)

This paper introduces the fundamental ideas of Kalman filtering, a recursive estimation technique widely used for continuous estimation of the state of a dynamic system. The estimation problem is posed within the well known Hilbert space framework of classical linear analysis. This permits an easily grasped geometric interpretation which is stripped of cumbersome details that tend to obscure the essential notions.

Considerable emphasis is placed on the development of the mathematical models for the state and measurement equations. A practical example of a real-world dynamic system (the motion of a ship) is used to motivate the form of the state equation required in Kalman filtering, as well as the measurement equation.

The recursive estimator and error covariance equations are derived for a one-dimensional dynamic system using a sequence of geometric visualizations as the derivation proceeds. The more tedious algebraic manipulations, which are not needed for an essential understanding of the derivation, are relegated to Appendices. This approach removes distracting details which often accompany derivations found in the engineering literature, and makes it apparent that Kalman filtering is based on elegant classical mathematics.

1. Introduction

In almost all branches of science, mathematical techniques are needed for estimating quantities on the basis of 'noisy' or imperfect measurement data. A very important estimation technique, called *Kalman filtering*, evolved in the early 1960s and since then, has found an ever-widening range of applications, notably in the aerospace and electronics industries. Typical applications include the continuous estimation of the position and velocity of a missile, a ship or an orbiting satellite, based on a continuing sequence of measurements, such as range or bearing data from a radar system.

The purpose of this paper is to introduce the fundamental ideas of Kalman filtering by posing the problem within the well known Hilbert space framework of classical linear analysis. This permits a largely geometric interpretation which in turn allows a rapid grasp of the estimation problem and the solution strategy, which is stripped of cumbersome details that tend to obscure the essential notions. The approach presented here also makes it apparent that Kalman filtering is based on elegant classical mathematics, a fact often not emphasized in the detailed Kalman filter derivations commonly found in current literature.

2. A simple recursive estimation problem

An introduction to the type of estimation problem solved by Kalman filtering is provided by the following example. Consider a motionless ship which is located at an unknown distance X offshore. By choosing an appropriate origin, it may be assumed that X is a zero-mean random variable having finite variance σ^2. Suppose that a shore-based radar makes a sequence of measurements of X, denoted by $Z(1)$, $Z(2), \ldots, Z(k)$. Since these measurements are noisy, it is assumed that they have the form

$$Z(i) = X + V(i), \quad i = 1,2,3,\ldots \tag{1}$$

where i is regarded as a time index, and $\{V(i)\}$ is a sequence of zero-mean random variables each of which is uncorrelated with X and satisfies $E[V(i)\,V(j)] = \delta_{ij} R$ for known $R > 0$.

The objective in this situation is to use the measurements $Z(1), Z(2), \ldots, Z(k)$ to obtain the best linear estimate $\hat{X}(k) = \sum_{i=1}^{k} a_i Z(i)$ of X at time k. By best we mean that $\hat{X}(k)$ is to minimize the mean square error $E([X - \hat{X}(k)]^2)$ for each k.

The solution technique in this simple setting is well known. The full power of the Kalman filter is not necessary, and in fact, a straightforward application of the projection theorem, sometimes called the orthogonality principle (see § 5 and appendix A), yields

$$\hat{X}(k) = \frac{\sigma^2}{k\sigma^2 + R} \sum_{i=1}^{k} Z(i) \tag{2}$$

As new measurements are received (that is, as k increases in value), (2) can be re-evaluated for each k. However, observe that the current estimate can be updated recursively by the relation

$$\hat{X}(k+1) = \frac{\sigma^2}{(k+1)\sigma^2 + R} Z(k+1) + \frac{k\sigma^2 + R}{(k+1)\sigma^2 + R} \hat{X}(k) \tag{3}$$

Thus, the recursive estimate is of the form

$$\hat{X}(k+1) = A(k+1)Z(k+1) + B(k+1)\hat{X}(k). \tag{4}$$

The advantage of the recursive estimate (3) is that it is computationally more efficient than the nonrecursive form. In fact, the nonrecursive estimate (2) involves a sum in which the number of terms grows without bound as the current time index k increases. The difference in computational burden becomes more pronounced in more complex situations. As we shall see in § 6, (3) represents the Kalman filter solution for our relatively simple estimation problem. In general, the Kalman filter provides numerically efficient recursive solutions to a large class of more complex estimation problems, in which the method of solution is not obvious. For example, suppose the ship in our simple example were moving with a constant, but unknown speed. How would we optimally estimate its position at each time instant k in a computationally efficient manner? A Kalman solution to this problem will be given in §6.

3. Generalizing the stationary ship problem

The stationary ship problem of the previous section can be generalized in several interesting ways. In these generalizations it is far less obvious how one might obtain a recursive best linear estimate of the ship's position.

Assume that the ship's position is subject to random perturbations caused by wave motion. The position might then be modelled by the recursive equation

$$X(k+1)=X(k)+U(k) \tag{5}$$

where $X(k)$ is the ship's position at time k, and $\{U(k)\}$ is a random sequence having known second-order statistics. The noisy measurements of the ship's position would be modelled by

$$Z(k)=X(k)+V(k) \tag{6}$$

Equation (5) will be called the *state equation*, which mod s the ship's dynamic behaviour, and equation (6) is the *measurement equation*. All Kalman filtering problems are modelled with a state and a measurement equation.

A generalization of (5) and (6) is possible if we introduce scalar constants F and H to obtain

$$X(k+1)=F\,X(k)+U(k) \tag{7}$$

and

$$Z(k)=H\,X(k)+V(k) \tag{8}$$

When $F=H=1$, we obtain (5) and (6). Equations (7) and (8) are the scalar version of the Kalman filter problem. A detailed solution in this case will be presented in §5.

The scalar model (7) and (8) is readily extended to the situation where the states $X(k)$ and measurements $Z(k)$ are random vectors, and F and H are known matrices. As an example of such a generalization, assume that a ship is moving at a constant but unknown speed $\dot{X}$ in a direction at right angles to the shoreline. The state equation (5) would then become

$$X(k+1)=X(k)+\dot{X}+U_1(k) \tag{9}$$

since there is a unit time increment. If $\dot{X}$ were also subject to random perturbations, it would no longer be time independent, and we would have the additional equation

$$\dot{X}(k+1)=\dot{X}(k)+U_2(k) \tag{10}$$

where $\{U_2(k)\}$ is a random sequence having known statistics. Equation (9) would become

$$X(k+1)=X(k)+\dot{X}(k)+U_1(k) \tag{11}$$

Equations (10) and (11) can be written as a single vector state equation:

$$\begin{bmatrix} X(k+1) \\ \dot{X}(k+1) \end{bmatrix}=\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} X(k) \\ \dot{X}(k) \end{bmatrix}+\begin{bmatrix} U_1(k) \\ U_2(k) \end{bmatrix} \tag{12}$$

or

$$\bar{X}(k+1)=F\,\bar{X}(k)+\bar{U}(k) \tag{13}$$

where F is now a 2×2 matrix.

Suppose we make noisy measurements of the ship's velocity, as well as position. We could rewrite (6) as

$$Z_1(k)=X(k)+V_1(k) \tag{14}$$

and introduce another measurement equation

$$Z_2(k)=\dot{X}(k)+V_2(k) \tag{15}$$

Equations (14) and (15) can be combined into the single vector measurement equation

$$\begin{bmatrix} Z_1(k) \\ Z_2(k) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X(k) \\ \dot{X}(k) \end{bmatrix} + \begin{bmatrix} V_1(k) \\ V_2(k) \end{bmatrix} \tag{16}$$

or

$$\bar{Z}(k) = H\bar{X}(k) + \bar{V}(k), \tag{17}$$

where H is now a 2×2 matrix. Note that in (16), the matrix H would become

$$H = [1 \quad 0] \tag{18}$$

if only the ship's position $X(k)$ were measured.

Further generalizations are possible. For example, the ship's position and velocity in two dimensions can be modelled by the state equation

$$\begin{bmatrix} X(k+1) \\ \dot{X}(k+1) \\ Y(k+1) \\ \dot{Y}(k+1) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X(k) \\ \dot{X}(k) \\ Y(k) \\ \dot{Y}(k) \end{bmatrix} + \begin{bmatrix} U_1(k) \\ U_2(k) \\ U_3(k) \\ U_4(k) \end{bmatrix} \tag{19}$$

If only the horizontal and vertical position coordinates $X(k)$ and $Y(k)$ are measured, the measurement equation would be

$$\begin{bmatrix} Z_1(k) \\ Z_2(k) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X(k) \\ \dot{X}(k) \\ Y(k) \\ \dot{Y}(k) \end{bmatrix} + \begin{bmatrix} V_1(k) \\ V_2(k) \end{bmatrix} \tag{20}$$

As will be seen, the Kalman estimator provides a recursive least mean-square estimate of *each component* of the state vector $\bar{X}(k)$ based on measurements $\bar{Z}(1)$, $\bar{Z}(2), \ldots, \bar{Z}(k)$. Thus, in (19), not only can the ship's position coordinates $X(k)$ and $Y(k)$ be estimated, but also $\dot{X}(k)$ and $\dot{Y}(k)$ can be estimated. Furthermore, the matrices F and H are permitted to be functions of the time index k.

4. General statement of the Kalman filtering problem

The preceding discussion leads to the following general statement of the Kalman filtering problem:

(*a*) We are given the vector state equation

$$\bar{X}(k+1) = F(k)\bar{X}(k) + \bar{U}(k), \quad k = 1, 2, 3, \ldots \tag{21}$$

where $\bar{X}(k)$ is an n-dimensional state vector, $F(k)$ is a known $n \times n$ matrix for each k, and $\bar{U}(k)$ is an n-dimensional random vector of mean zero satisfying

$$E\{\bar{U}(k)\bar{U}^T(j)\} = Q(k)\delta_{kj}, \tag{22}$$

where $Q(k)$ is a known $n \times n$ covariance matrix for each k.

(*b*) We are also given the vector measurement equation

$$\bar{Z}(k) = H(k)\bar{X}(k) + \bar{V}(k) \tag{23}$$

where $H(k)$ is an $m \times n$ matrix, and $\bar{V}(k)$ is an m-dimensional random error vector of mean zero satisfying

$$E\{\bar{V}(k)\bar{V}^T(j)\} = R(k)\delta_{kj} \tag{24}$$

where $R(k)$ is a known positive definite covariance matrix for each k.

(*c*) The initial state vector $\bar{X}(1)$ is assumed to have zero mean with known covariance matrix P_1.

(*d*) It is assumed that the random vectors $\bar{X}(1)$, $\bar{U}(j)$, and $\bar{V}(k)$ are all uncorrelated for $j \geqslant 1$, $k \geqslant 1$.

The objective is to determine the best linear estimate $\hat{\bar{X}}(k)$ of $\bar{X}(k)$ in the recursive form

$$\hat{\bar{X}}(k+1) = A(k+1)\bar{Z}(k+1) + B(k+1)\hat{\bar{X}}(k) \tag{25}$$

where $A(k+1)$ and $B(k+1)$ are matrices which are relatively easy to calculate. $\hat{\bar{X}}(k)$ is said to be a best estimate of $\bar{X}(k)$ if

$$E\{[\bar{X}(k) - \hat{\bar{X}}(k)]^T[\bar{X}(k) - \hat{\bar{X}}(k)]\}$$

is minimized.

5. The Kalman solution in the scalar case

In §§2 and 3, examples were discussed in which the state equation (21) and measurement equation (23) are reduced to their scalar versions (7) and (8). It is in this scalar setting that the solution for the Kalman filter model will be derived. The solution procedure in this scalar case is readily extended to the general multidimensional case. The key ideas in each case are parallel, and the presentation of the scalar case solution allows a progression of these ideas without the cumbersome multidimensional notation and calculations. We also impose one further simplification of the Kalman model as presented in section 4. We shall assume that the known quantities are constant in time, so $F(k) = F$, $H(k) = H$, $Q(k) = Q$ and $R(k) = R$ are taken as known time-invariant constants. Again, the fundamentals of the solution procedure remain intact with this simplification.

5.1. *The mathematical setting*

The development of the Kalman solution takes place in a Hilbert space context which is imposed on the random variables $X(1), X(2), \ldots, X(k), Z(1), Z(2), \ldots$, and $Z(k)$. This setting is available to any finite set of finite variance random variables, say $\{Y_1, Y_2, \ldots, Y_n\}$. Let the finite dimensional vector space $V = [Y_1, \ldots, Y_n]$ be the linear span of the Y_i's over the reals. An inner product on V is given by $\langle Y, W\rangle = E(YW)$. V is now a finite dimensional Hilbert space with its norm given by

$$\| Y \| = (E(Y^2))^{1/2} = \| Y \|_2$$

the L_2 norm.

The Kalman filter solution relies in large part on the application of some fundamental results from Hilbert space theory. We shall now state these results (with proofs provided in Appendix B).

Theorem 1 (The Projection Theorem)

Let V be a finite dimensional Hilbert space and S be a subspace of V. For each $Y \in V$ there exists a unique $\hat{Y} \in S$ such that

$$\| Y - \hat{Y} \| \leqslant \| Y - Z \| \quad \text{for all } Z \in S$$

Furthermore, $\hat{Y}$ is the unique minimizing vector if and only if $Y - \hat{Y}$ is orthogonal to Z for all $Z \in S$, that is, $\langle Y - \hat{Y}, Z \rangle = 0$ for all $Z \in S$.

$\hat{Y}$ is termed the linear minimum mean squared error (LMMSE) estimate of Y in S.

Theorem 2 (Linearity of estimators)

With V and S as above, suppose that X, Y, $W \in V$ with $W = aX + bY$, a, b real numbers. Let $\hat{X}$, $\hat{Y}$ and $\hat{W}$ be the LMMSE estimators of X, Y and W, respectively, in S. Then $\hat{W} = a\hat{X} + b\hat{Y}$.

Theorem 3 (Orthogonal decomposition)

Let S_1 and S_2 be orthogonal subspaces of the finite-dimensional Hilbert space V. For $Y \in V$, let $\hat{Y}_1$ and $\hat{Y}_2$ be the LMMSE estimates of Y in S_1 and S_2, respectively. If $\hat{Y}$ is the LMMSE estimate of Y in $S_1 \oplus S_2$, then $\hat{Y} = \hat{Y}_1 + \hat{Y}_2$.

5.2. *A geometric view of the solution*

Before we become embroiled in the calculation of the Kalman solution, we want to provide a geometric setting for the entire solution process. The geometry is informally depicted in figures 1–4, and along with the accompanying narrative, these figures provide the framework for understanding the essence of the solution to the Kalman filter problem.

In the narrative which follows, $\hat{X}(k|j)$ denotes the LMMSE estimate of $X(k)$ based on the measurements $Z(1), \ldots, Z(j)$. Thus the estimate $\hat{X}(k)$ in previous sections will be denoted by $\hat{X}(k|k)$.

5.2.1. *Figure 1*:

Suppose that measurements $Z(1)$, $Z(2), \ldots, Z(k)$ of the respective states $X(1)$,

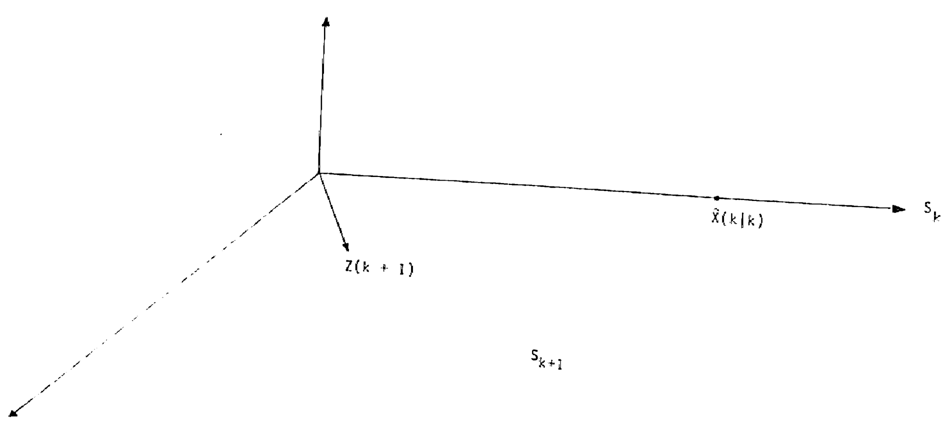

Figure 1.

$X(2), \ldots, X(k)$ have been made, and that the best linear estimate $\hat{X}(k|k)$ of $X(k)$, based on these measurements, has been computed. The estimate $\hat{X}(k|k)$ is a linear combination of the measurements $Z(1), Z(2), \ldots, Z(k)$; hence, it lies in the subspace S_k generated by these measurements. This subspace is represented by the horizontal axis in the figure, and $\hat{X}(k|k)$ is shown as a point on the axis.

Now suppose that the next measurement $Z(k+1)$ becomes available. In general, this measurement will lie outside of S_k as shown in the figure. At this time we wish to compute $\hat{X}(k+1|k+1)$, the best linear estimate of $X(k+1)$, given the previous measurements $Z(1), Z(2), \ldots, Z(k)$ and the latest measurement $Z(k+1)$. This estimate will be a linear combination of the measurements $Z(1), Z(2), \ldots, Z(k+1)$, and therefore will lie in the subspace S_{k+1} obtained by annexing the measurement $Z(k+1)$ to the subspace S_k of 'old' measurements $Z(1), Z(2), \ldots, Z(k)$. Thus in the figure, S_{k+1} is shown as the plane determined by the horizontal axis S_k and the measurement vector $Z(k+1)$.

5.2.2. *Figure 2*:

By the projection theorem, the estimate $\hat{X}(k+1|k+1)$ is the orthogonal projection of the state $X(k+1)$ onto the subspace S_{k+1} generated by $Z(1), Z(2), \ldots, Z(k+1)$, but it is certainly not obvious how one might compute this projection. The basic strategy is to decompose S_{k+1} into the direct sum $S_k \oplus \tilde{S}$ of the subspace S_k and another subspace $\tilde{S}$ which is orthogonal to S_k. By the orthogonal decomposition theorem, $\hat{X}(k+1|k+1)$ will then be the sum of the orthogonal projections of $X(k+1)$ onto the subspaces S_k and $\tilde{S}$. As will be seen shortly, these latter two projections can be determined if we first characterize the subspace $\tilde{S}$. In the figure, $\tilde{S}$ is shown as an axis perpendicular to S_k, and lying in the horizontal plane. $\tilde{S}$ is easily found by again using the projection theorem, which guarantees that the orthogonal projection $\hat{Z}(k+1|k)$ of the measurement $Z(k+1)$ onto S_k has the property that the difference vector $\tilde{Z}(k+1) \triangleq Z(k+1) - \hat{Z}(k+1|k)$ is orthogonal to S_k. If we let $\tilde{S}$ be the subspace generated by $\tilde{Z}(k+1)$, it is evident that $\tilde{S}$ is orthogonal to S_k and that $S_{k+1} = S_k \oplus \tilde{S}$, as desired.

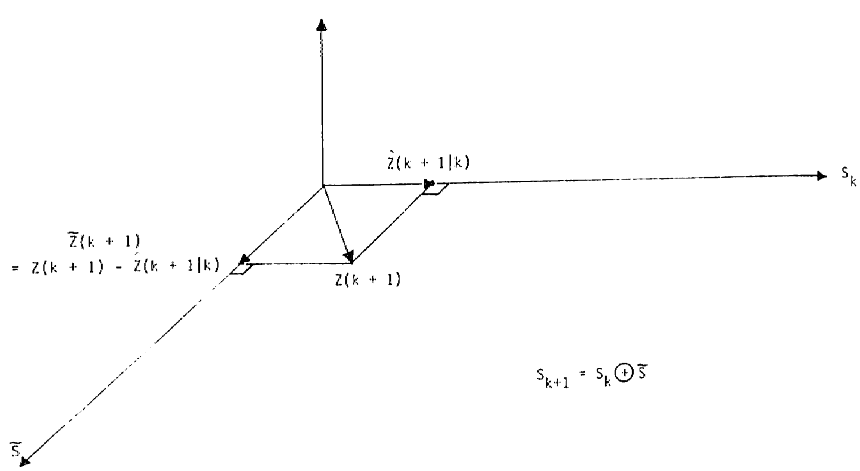

Figure 2.

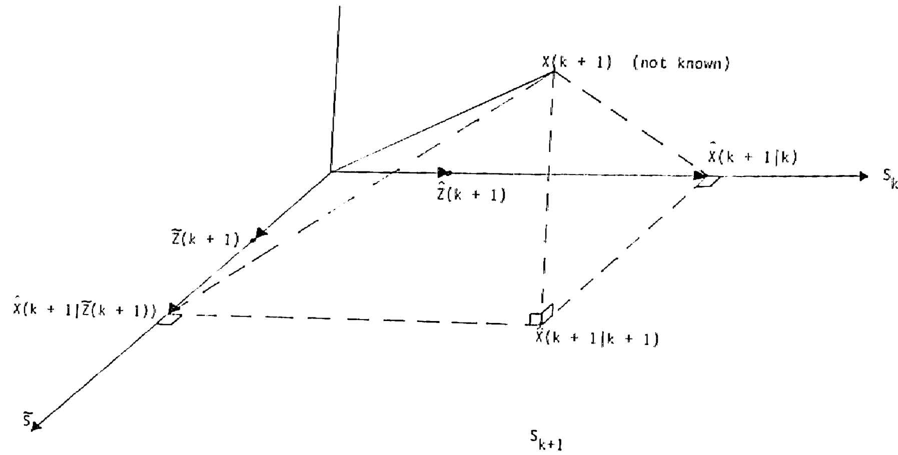

Figure 3.

5.2.3. *Figure 3*:

Since the desired estimate $\hat{X}(k+1|k+1)$ is the orthogonal projection of the state $X(k+1)$ onto the subspace $S_{k+1} = S_k \oplus \tilde{S}$, the orthogonal decomposition theorem guarantees that $\hat{X}(k+1|k+1)$ is the sum of the orthogonal projection $\hat{X}(k+1|k)$ of $X(k+1)$ onto S_k and the orthogonal projection $\hat{X}(k+1|\tilde{Z}(k+1))$ on to $\tilde{S}$, that is

$$\hat{X}(k+1|k+1) = \hat{X}(k+1|k) + \hat{X}(k+1|\tilde{Z}(k+1)) \tag{26}$$

This relationship is shown geometrically in the figure. Thus the problem is reduced to that of computing $\hat{X}(k+1|k)$ and $\hat{X}(k+1|\tilde{Z}(k+1))$. The figure seems to suggest that the state $X(k+1)$ must be known in order to compute these two projections. However, a fact of fundamental importance in Kalman filtering is that the state equation (7) and measurement equation (8) permit calculation of $\hat{X}(k+1|k)$ and $\hat{X}(k+1|\tilde{Z}(k+1))$ *without* explicit knowledge of $X(k+1)$.

5.2.4. *Figure 4*:

The computation of the projections $\hat{X}(k+1|k)$ and $\hat{X}(k+1|\tilde{Z}(k+1))$ require the use of the state equation (7) and measurement equation (8). It will be proved in

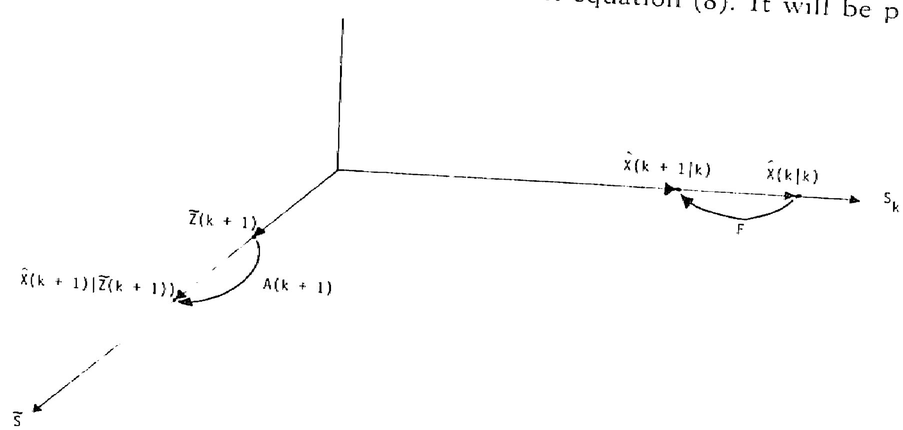

Figure 4.

proposition 1 that $\hat{X}(k+1|k)=F\hat{X}(k|k)$. This relationship is shown by the arrow labelled 'F' in the figure. A more extensive derivation is required for $\hat{X}(k+1|\tilde{Z}(k+1))$. The starting point for the derivation is as follows. Since we are considering each measurement $Z(k)$ to be a scalar random variable, it is evident that $\tilde{S}$ is a one-dimensional subspace spanned by the single vector $\tilde{Z}(k+1)$. Therefore, $\hat{X}(k+1|\tilde{Z}(k+1))=A(k+1)\tilde{Z}(k+1)$ where $A(k+1)$ is a scalar which needs to be determined. This scalar is often called the *Kalman gain*.

5.3. *The Kalman solution*

The geometric view which was just discussed has established that the LMMSE $\hat{X}(k+1|k+1)$ of $X(k+1)$ is given by

$$\hat{X}(k+1|k+1)=\hat{X}\ (k+1|k)+A(k+1)\tilde{Z}(k+1). \tag{27}$$

At this point, the goal is to show that (27) leads to the scalar version of (25), namely

$$\hat{X}(k+1|k+1)=B(k+1)\hat{X}(k|k)+A(k+1)Z(k+1) \tag{28}$$

where $B(k+1)$ and $A(k+1)$ are scalars which are determined recursively.

We shall focus on the summands in (27) separately beginning with $\hat{X}(k+1|k)$.

Proposition 1: $\hat{X}(k+1|k)=F\hat{X}(k|k)$.

Proof

Since $X(k+1)=FX(k)+U(k)$, applying theorem 2 yields

$$\hat{X}(k+1|k)=F\hat{X}(k|k)+\hat{U}(k|k)$$

where $\hat{U}(k|k)$ is the LMMSE estimator of $U(k)$ based on $Z(1),\ldots, Z(k)$. From Appendix C, $\hat{U}(k|k)=0$, and the result is established.

The preliminary form of the solution (27) is now replaced by

$$\hat{X}(k+1|k+1)=F\hat{X}(k|k)+A(k+1)\tilde{Z}(k+1) \tag{29}$$

The development now focuses on the second summand in (29). We have

Proposition 2: $\tilde{Z}(k+1)=Z(k+1)-HF\hat{X}(k|k)$

Proof

By definition,

$$\tilde{Z}(k+1)=Z(k+1)-\hat{Z}(k+1|k)$$

Since $Z(k+1)=HX(k+1)+V(k+1)$, theorem 2 gives

$$\hat{Z}(k+1|k)=H\hat{X}(k+1|k)+\hat{V}(k+1|k)$$

Since $\hat{V}(k+1|k)=0$ (see Appendix C), using proposition 1 gives $\hat{Z}(k+1|k) = H\hat{X}(k+1|k)=\mathrm{HF}\hat{\mathrm{X}}(k|k)$. Hence

$$\tilde{Z}(k+1)=Z(k+1)-HF\hat{X}(k|k)$$

which concludes the proof.

The form of the solution (29) is now revised to give

$$\begin{aligned}\hat{X}(k+1|k+1)&=F\hat{X}(k|k)+A(k+1)(Z(k+1)-HF\hat{X}(k|k))\\ &=F(1-HA(k+1))\hat{X}(k|k)+A(k+1)Z(k+1)\end{aligned} \tag{30}$$

Equation (30) is now in the desired form (28). The goal now is to give the explicit form of the Kalman gain $A(k+1)$, which is done below in proposition 3. The proof, which is somewhat detailed, appears in Appendix C. The recursive solution for $A(k+1)$ is given in terms of the error covariance of $\hat{X}(k|j)$, that is

$$P(k|j)=E[(X(k)-\hat{X}(k|j))^2]$$

We can now state:

Proposition 3: The Kalman gain is given by

$$A(k+1)=\frac{HP(k+1|k)}{H^2P(k+1|k)+R} \tag{31}$$

where

$$P(k+1|k)=F^2P(k|k)+Q \tag{32}$$

and

$$P(k|k)=(1-HA(k))P(k|k-1) \tag{33}$$

Equations (30)–(33) comprise the solution to the Kalman filter along with the initial conditions $\hat{X}(1|0)=0$ and $P(1|0)=E(X(1)^2)=P_1$, where $P_1 \geqslant 0$ is known.

5.4. *The Kalman filter equations for the vector case*

When $\hat{X}(k)$ is a vector and F, H, Q and R are matrices (examples are given in §3), the Kalman solution is obtained in essentially the same way as in the scalar case, and the resulting equations which are virtually identical to equations (30)–(33) above, are as follows:

$$\hat{X}(k+1|k+1)=(\mathrm{I}-A(k+1)H)F\hat{X}(k|k)+A(k+1)\bar{Z}(k+1) \tag{34}$$

$$A(k+1)=P(k+1|k)H^{\mathrm{T}}[HP(k+1|k)H^{\mathrm{T}}+R]^{-1} \tag{35}$$

$$P(k+1|k)=FP(k|k)F^{\mathrm{T}}+Q \tag{36}$$

$$P(k|k)=[\mathrm{I}-A(k)H]P(k|k-1) \tag{37}$$

6. The Kalman filter solution of the stationary and moving ship problems

An appreciation of the utility of the Kalman solution can be gained by applying it to the stationary ship problem of §2 and to the generalization given by equations (12) and (16), in which the ship is assumed to have constant speed. In the first case, we shall show that the estimator equation reduces to the previously obtained recursive estimator given by (3). However, we shall get as a bonus the covariance of the estimation error, also expressed in recursive form. In the case of the moving ship, a closed-form solution is difficult to obtain, so results of a computer simulation will be discussed.

6.1. *The stationary ship case*

The state and measurement equations for the stationary ship in §2 are

$$X(k+1)=X(k) \tag{38}$$

and

$$Z(k)=X(k)+V(k) \tag{39}$$

Thus, in the Kalman filter solution (30)–(33), all matrices are scalar constants with values $F=H=1$, $Q=E\{U^2(k)\}=0$ (no noise in the state equation). The scalar $R=E\{V^2(k)\}$ has a known positive value. The initial variance $P(1|0)$ in the ship's position is σ^2. Substituting these quantities into (30)–(33) gives

$$\hat{X}(k+1|k+1)=(1-A(k+1))\hat{X}(k|k)+A(k+1)Z(k+1) \tag{40}$$

$$A(k+1)=\frac{P(k+1|k)}{P(k+1|k)+\mathrm{R}} \tag{41}$$

$$P(k+1|k)=P(k|k) \tag{42}$$

and

$$P(k|k)=(1-A(k))P(k|k-1) \tag{43}$$

Since $P(1|0)=\sigma^2$, we obtain from (41),

$$A(1)=\frac{\sigma^2}{\sigma^2+R}$$

Starting with $P(1|0)$ and $A(1)$, repeated cycling through equations (43), (42) and (41) (in that order) yields $A(k+1)$ and $P(k+1|k)$ on the kth cycle. Both of these quantities are used in the $(k+1)$st cycle, and $A(k+1)$ is used in (40), along with measurement $(k+1)$ and the previous estimate $\hat{X}(k|k)$, to produce the updated estimate $\hat{X}(k+1|k+1)$ of the ship's position. After several iterations of (43), (42) and (41), it is evident that

$$A(k)=\frac{R\sigma^2}{k\sigma^2+R} \tag{44}$$

and

$$P(k|k)=P(k+1|k)=\frac{R\sigma^2}{k\sigma^2+R} \tag{45}$$

both of which are easily verified by induction. Substituting (44) into (40) yields

$$\hat{X}(k+1|k+1)=\frac{k\sigma^2+R}{(k+1)\sigma^2+R}X(k|k)+\frac{\sigma^2}{(k+1)\sigma^2+R}Z(k+1)$$

which is precisely the estimator equation (3) obtained earlier. In addition, however, equation (45) gives the variance $P(k|k)$ of the error of the kth estimate. In this case $P(k|k)$ tends to zero with increasing k which makes sense intuitively, since the ship is stationary and the number of independent measurements is increasing without bound.

6.2. *The moving ship case*

If we assume that the ship moves at a constant but unknown speed $\dot{X}$ at right angles to the shoreline, the appropriate state equation is

$$\begin{bmatrix} X(k+1) \\ \dot{X}(k+1) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X(k) \\ \dot{X}(k) \end{bmatrix} \tag{46}$$

Equation (46) is the same as equation (12), except that there is no perturbation of the ship's position by noise. If we measure the ship's position $X(k)$ at each time instant, the measurement equation is

$$Z(k)=[1 \quad 0]\begin{bmatrix} X(k) \\ \dot{X}(k) \end{bmatrix} + V(k) \tag{47}$$

Thus, in the Kalman filter equations (34)–(37), all matrices are constant, and

$$F=\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \qquad H=[1 \quad 0]$$

and

$$Q=E\{[U_1(k),\ U_2(k)]^T[U_1(k),\ U_2(k)]\}=\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Q is the zero matrix because there is no perturbing noise in the state equation. The scalar $R=E\{V(k)\}$ has a known positive value. The initial convariance matrix $P(1|0)$ of the ship's state vector is

$$\begin{aligned} P(1|0) &= E\{[X(1),\ \dot{X}(1)]^T[X(1),\ \dot{X}(1)]\} \\ &= \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix} \end{aligned}$$

where σ_1^2 and σ_2^2 are assumed known. We have assumed a diagonal covariance matrix in the interest of simplicity, although it is not necessary to do so. Substituting the above quantities into equations (34)–(37) and iterating in exactly the same manner as was done for the stationary ship case gives recursive solutions for the state vector estimate

$$\hat{\bar{X}}(k|k)=\begin{bmatrix} \hat{X}(k|k) \\ \hat{\dot{X}}(k|k) \end{bmatrix}$$

and the error covariance matrix

$$\begin{aligned} P(k|k) &= \{[\bar{X}(k)-\hat{\bar{X}}(k|k)][\bar{X}(k)-\hat{\bar{X}}(k|k)]\} \\ &= \begin{bmatrix} P_{11}(k) & P_{12}(k) \\ P_{21}(k) & P_{22}(k) \end{bmatrix} \end{aligned}$$

The derivation of a closed-form solution appears to be intractable, but the interative form readily lends itself to a numerical solution. Figure 5(a)–(d) shows the behaviour of the Kalman solution for four cases.

Case 1: (figure 5(a))

The upper diagonal element

$$P_{11}(k) \triangleq E\{[X(k)-\hat{X}(k)]^2\}$$

of the estimation error covariance matrix $P(k|k)$ is the variance of the position estimation error, which is shown as a function of k by a solid line in the figure. Here the diagonal terms of $P(1|0)$, representing the initial uncertainty in the state vector, were chosen to be 1 (position) and 0 (velocity). A value of $R=1$ was used for the

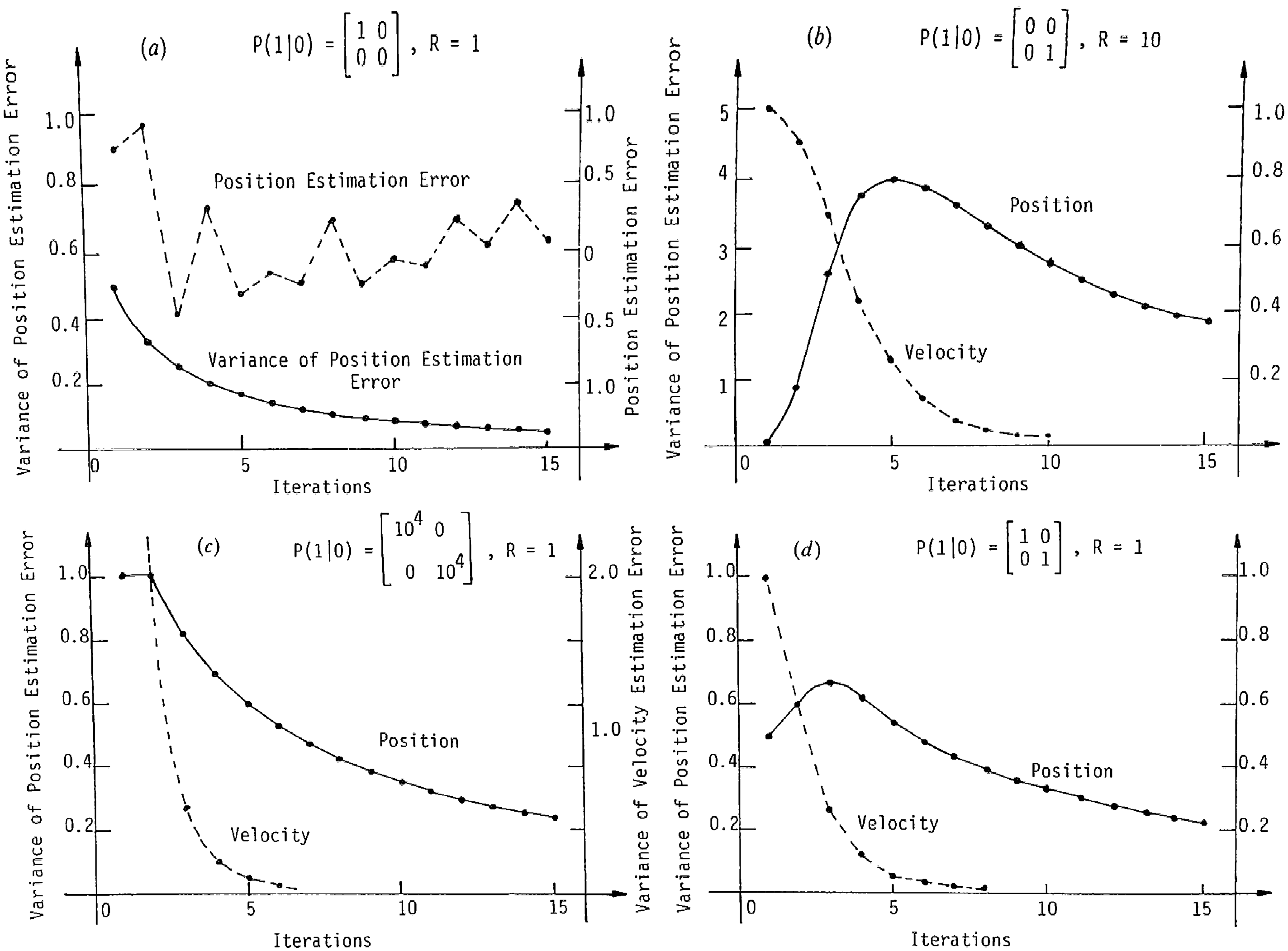

Figure 5. The behaviour of the Kalman solution is shown for four different combinations of R, the variance of the measurement noise, and $P(1|0)$, the initial error covariance matrix.

variance of the measurement noise. The dashed line is the actual position estimation error that occurred in a computer simulation. It is interesting to note that, since the velocity is known precisely (and is therefore zero since all random variables are assumed to have zero mean), the problem reduces to the stationary ship case. In figure 5(a) it can be seen that

$$P_{11}(k)=\frac{1}{k+1}, \quad k=1,2,\dots$$

which is precisely of the variance of the estimation error given by (45) in the stationary ship case where $\sigma^2=R=1$.

Case 2: (figure 5(b))

In this case the diagonal terms of $P(1|0)$ are 0 (position) and 1 (velocity), with $R=10$. The variance $P_{11}(k)$ of the position estimation error is again plotted as a solid line. The dashed line is the variance of the velocity estimation error, which is the element

$$P_{22}(k)\triangleq E\{[\dot{X}(k)-\dot{\hat{X}}(k)]^2\}$$

of $P(k|k)$. Note that although the initial position $X(1)$ of the ship is known exactly in this case, the uncertainty in position is actually *increasing* for a short time. This occurs because the initial velocity uncertainty is mapped into a position uncertainty at subsequent times. In the absence of measurements (or equivalently, with measurements having infinite variance in the measurement error), the resulting position uncertainty would increase linearly with time. However, the variance of the velocity estimation error is eventually reduced to a value which permits the successive position measurements to maintain a continuing reduction in the variance of the position estimation error.

Case 3: (figure 5(c))

This represents a commonly encountered situation where the initial uncertainty in the position and velocity is not known. In such cases it is common practice to assume that the initial uncertainty is very large compared with the accuracy of the measurements; thus the diagonal elements of $P(1|0)$ were chosen to be 10^4 for both position and velocity, with $R=1$. An unexpected result is that for $k=1$ and 2, the variance of the position estimation error is essentially the same. It appears that the second measurement at $k=2$ gives us no information with which to improve our position estimate. This strange behaviour can be explained as follows: The first measurement results in a position estimate with an error variance nearly equal to one, which would be expected, since $R=1$ and the initial position variance is 10^4. However, since all measurements are of position only, there remains the large velocity uncertainty $P_{22}(1)=10^4$ after the first measurement is made. This causes a large position uncertainty just prior to the next measurement at $k=2$, and when the second measurement is made, the variance in position estimation error is still nearly equal to one. At this point, however, the large velocity uncertainty no longer exists since two position measurements are sufficient to determine velocity.

Case 4: (figure 5(d))

In this case the diagonal elements of $P(1|0)$ are 1 for both position and velocity, and $R=1$. It is interesting to compare the variance of position estimation error with the corresponding curve in figure 5(a) (case 1), where the initial velocity uncertainty was zero. In that case, the position estimates are better. This illustrates the general principle that uncertainties in a given component of the state vector can cause uncertainties in another component, that is, the states can be *coupled*.

7. Concluding remarks

In general, the Kalman filter provides numerically efficient recursive solutions to a large class of estimation problems. These solutions are characterized by the following properties:

(*a*) The estimate is a current linear least mean square estimate of a sequence of random variables (or random vectors) which can be updated recursively as time evolves, based on a continuing sequence of noisy measurements.
(*b*) The covariance matrix of the estimation error is also computed recursively and does not depend on the values of the measurements. This permits the performance of the estimator to be analysed *prior* to its use with actual data.
(*c*) Only the first and second moments of all random variables need be known. Thus, the estimates are optimum even when the underlying random processes are not normally distributed.
(*d*) The statistics of the underlying state and measurement random processes are permitted to change with time.

However, standard Kalman filter theory requires that the random sequence to be estimated be modelled as a certain type of linear difference equation. Examples of such processes were given in section 3. Furthermore, the measurements must be linear, as was also exemplified in section 3.

Prior to the introduction of Kalman filtering in the early 1960s, a nonrecursive linear least-mean square estimation technique called *Wiener filtering* was commonly used for estimating random processes. However, this approach, based on pioneering work by Norbert Wiener in the 1940s, has restricted applicability, primarily because the amount of computation required for each estimate grows without bound in time. Furthermore, the solution for the estimator does not suggest how to implement the solution on a computer. The beauty of the Kalman approach is that the computational complexity of the recursive solution does not grow with time, and the estimator equations explicitly show the sequence of operations to be performed by a computer to update each estimate.

Although the essence of Kalman filtering has been explored in this paper, a great amount of additional knowledge is required to bridge the gap between theory and practical application. For example, we have not even touched on the problem of modelling errors. In many practical situations the underlying processes are generated by nonlinear phenomena, so the linear model is only an approximation. Serious problems can arise if the linear approximation is not good enough. The *extended Kalman filter*, which allows some nonlinearity in the model, is an attempt to solve this problem. There are also potential problems caused by lack of accurate knowledge of the covariance matrices in the state and measurement models. In such cases, a thorough analysis of the *robustness* (that is, the insensitivity to inaccurate assumptions) of the Kalman filter may be required. Sometimes there may also be computational difficulties, such as those of inverting an ill-conditioned matrix in the covariance equation, that must be anticipated and dealt with. Despite these difficulties, the Kalman filter has proved to be a powerful estimation technique that has wide application.

Appendix A
The direct solution for the stationary ship problem.

The objective in this problem is to use the measurement data $Z(1), Z(2), \ldots, Z(k)$, to give the best linear estimate $\hat{X}(k) = \Sigma_{i=1}^{k} a_i Z(i)$ of X for each k. The projection theorem (theorem 1) implies that the estimation error is orthogonal to the data, that is,

$$0 = E[(X - \hat{X}(k))Z(j)], \quad j = 1, 2, \ldots, k$$

Using the calculations

$$E[XZ(j)] = \sigma^2$$

and

$$E[Z(i)Z(j)] = \sigma^2 + \delta_{ij} R$$

we obtain the system of equations

$$\sum_{i=1}^{k} a_i(\sigma^2 + \delta_{ij} R) = \sigma^2, \quad j = 1, \ldots, k$$

The solution for this system is

$$a_i = \frac{\sigma^2}{k\sigma^2 + R}, \quad i = 1, 2, \ldots, k$$

and so

$$X(k) = \frac{\sigma^2}{k\sigma^2 + R} \sum_{i=1}^{k} Z(i)$$

which is (2).

Appendix B
Proofs of theorems 1–3

(a) Proof of theorem 1, the projection theorem.

In this finite dimensional setting, each $Y \in V$ can be written as $Y = \hat{Y} + \tilde{Y}$, where $\hat{Y} \in S$ and $\langle Z, \tilde{Y} \rangle = 0$ for any $Z \in S$. Then for $Z \in S$,

$$\begin{aligned} \| Y - Z \|^2 &= \langle \hat{Y} + \tilde{Y} - Z, \hat{Y} + \tilde{Y} - Z \rangle \\ &= \| \tilde{Y} \|^2 + \| \hat{Y} - Z \|^2 \\ &\geqslant \| \tilde{Y} \|^2 \\ &= \| Y - \hat{Y} \|^2 \end{aligned}$$

So $\| Y - Z \| \geqslant \| Y - \hat{Y} \|$, and $\hat{Y}$ is unique. The equivalence statement in the theorem is now clear.

(b) Proof of theorem 2, linearity of estimators.

This follows from the projection theorem since, for any $Z \in S$

$$\langle W - \hat{W}, Z \rangle = a \langle X - \hat{X}, Z \rangle + b \langle Y - \hat{Y}, Z \rangle = 0$$

(c) Proof of theorem 3, orthogonal decomposition.

Let $\hat{Y} = T_1 + T_2$ where $T_1 \in S_1$ and $T_2 \in S_2$. Now $Y - T_1 = Y - \hat{Y} + T_2$, but $Y - \hat{Y}$

and T_2 are both orthogonal to S_1. So $T_1 = \hat{Y}_1$ by the projection theorem (theorem 1). Similarly, $T_2 = \hat{Y}_2$.

Appendix C
Supporting results for the Kalman solution

We begin with a collection of additional second-order results.

Proposition C1: $E[X(k)U(j)] = \chi\{k \geqslant j+1\}F^{k-j-1}Q$, where $\chi(A)$ is the indicator function for the set A.

Proof

Since $X(k) = FX(k-1) + U(k-1)$,

$$E(X(k)U(j)) = FE[X(k-1)U(j)] + E[U(k-1)U(j)]$$

Since $E[X(1)U(j)] = 0$ and $E[U(k-1)U(j)] = \delta_{k-1,j}Q$, we obtain inductively

$$E[X(k)U(j)] = 0 \quad \text{for } k = 1, 2, \ldots, j$$

Also

$$E[X(j+1)U(j)] = Q$$

Proceeding inductively again, for $r > 1$,

$$\begin{aligned} E[X(j+r)U(j)] &= FE[X(j+r-1)U(j)] \\ &= F^{r-1}E[X(j+1)U)(j)] \\ &= F^{r-1}Q \end{aligned}$$

Proposition C2: $E[X(j)V(k)] = 0$ for all j and k.

Proof

Since $X(j) = FX(j-1) + U(j-1)$,

$$E[X(j)V(k)] = FE[X(j-1)V(k)] + E[U(j-1)V(k)].$$

Recall that $E[U(j-1)V(k)] = 0$ for all j and k, and $E[X(1)V(k)] = 0$ for all k, so the result follows immediately by induction.

The following well known result is now given.

Theorem C1:

Let $X, Z_1 \ldots, Z_m$ be random variables with finite variance and let $\bar{a} \in R^m$. Let $\bar{Z} = (Z_1, \ldots, Z_m)^T$, and assume that $[E(\bar{Z}\bar{Z}^T)]^{-1}$ exists. Then the linear estimate $\hat{X} = \bar{a}^T\bar{Z}$ such that $E[(X - \hat{X})^2]$ is minimum is

$$\hat{X} = E(X\bar{Z}^T)[E(\bar{Z}\bar{Z}^T)]^{-1}\bar{Z}$$

Proof

By the projection theorem,

$$E[(X - \bar{a}^T\bar{Z})\bar{Z}^T] = 0$$

which gives

$$\bar{a} = E(X\bar{Z}^T)[E(\bar{Z}\bar{Z}^T)]^{-1}$$

The proofs of propositions 1 and 2 in § 5.3 can now be completed.

Proposition C3: $\hat{U}(k|k)=0$ for all $k\geqslant 1$.

Proof

Let $\tilde{Z}=(Z(1),\ldots,Z(k))^{\mathrm{T}}$. Then from theorem C1

$$\hat{U}(k|k)=E(U(k)\tilde{Z}^{\mathrm{T}})[E(\tilde{Z}\tilde{Z}^{\mathrm{T}})]^{-1}\tilde{Z}=0$$

since for $1\leqslant r\leqslant k$,

$$\begin{aligned} E[U(k)Z(r)] &= E[U(k)(HX(r)+V(r))] \\ &= HE[U(k)X(r)]+E[U(k)V(r)]=0 \end{aligned}$$

Proposition C4: $\hat{V}(k+1|k)=0$ for all $k\geqslant 1$.

Proof

Analogous to the proof of proposition C3,

$$\hat{V}(k+1|k)=E(V(k+1)\tilde{Z}^{\mathrm{T}})[E(\tilde{Z}\tilde{Z}^{\mathrm{T}})]^{-1}\tilde{Z}=0,$$

since for $1\leqslant r\leqslant k$,

$$\begin{aligned} E[V(k+1)Z(r)] &= HE[V(k+1)X(r)]+E[V(r)V(k+1)] \\ &= 0 \end{aligned}$$

We now begin the justification of proposition 3 in which the recursive form of the Kalman gain $A(k)$ is given.

Proposition C5: $A(k+1)=\dfrac{E[X(k+1)\tilde{Z}(k+1)]}{E[\tilde{Z}(k+1)^2]}$.

Proof

Since $X(k+1)-A(k+1)\ \tilde{Z}(k+1)$ is orthogonal to $Z(k+1)$,

$$\begin{aligned} 0 &= E[(X(k+1)-A(k+1)\ \tilde{Z}(k+1))\ \tilde{Z}(k+1)] \\ &= E(X(k+1)\ \tilde{Z}(k+1))-A)(k+1)\ E[\tilde{Z}(k+1)^2]. \end{aligned}$$

Solving for $A(k+1)$ gives the result.

Proposition C6: $E[\tilde{Z}(k+1)^2]=H^2P(k+1|k)+R$.

Proof

In the proof of proposition 2, we saw that $\hat{Z}(k+1|k)=H\hat{X}(k+1|k)$. Hence,

$$\begin{aligned} \tilde{Z}(k+1) &= Z(k+1)-\hat{Z}(k+1|k) \\ &= Z(k+1)-H\hat{X}(k+1|k) \\ &= HX(k+1)+V(k+1)-H\hat{X}(k+1|k) \\ &= H[X(k+1)-\hat{X}(k+1|k)]+V(k+1) \end{aligned}$$

Then

$$\begin{aligned} E[\tilde{Z}(k+1)^2] &= E[(H(X(k+1)-\hat{X}(k+1|k))+V(k+1))^2] \\ &= H^2P(k+1|k)+R+2HE[(X(k+1)-\hat{X}(k+1|k))\,V(k+1)] \end{aligned}$$

But $E(X(k+1)V(k+1))=0$, and since $\hat{X}(k+1|k)=\Sigma_{i=1}^{k}a_i Z(i)$,

$$E[\hat{X}(k+1|k)V(k+1)=\sum_{i=1}^{k} a_i E[Z(i)V(k+1)]$$
$$=\sum_{i=1}^{k} a_i\{HE[X(i)V(k+1)]+E[V(i)V(k+1)]\}$$
$$=0$$

by proposition C2 and (24). Consequently, the desired result follows.

Proposition C7: $E[X(k+1)\tilde{Z}(k+1)]=HP(k+1|k)$.

Proof

Again using $\tilde{Z}(k+1)=H[X(k+1)-\hat{X}(k+1|k)]+V(k+1)$, we have

$$E[X(k+1)\ \tilde{Z}(k+1)]=E[X(k+1)(H[X(k+1)-\hat{X}(k+1|k)]+V(k+1)]$$
$$=E\{[(X(k+1)-\hat{X}(k+1|k))+\hat{X}(k+1|k)][H(X(k+1)-\hat{X}(k+1|k))+V(k+1)]\}$$
$$=HP(k+1|k)+E[X(k+1)V(k+1)]+HE[\hat{X}(k+1|k)\{\hat{X}(k+1)-X(k+1|k)\}]$$
$$=HP(k+1|k)$$

since $E[X(k+1)V(k+1)]=0$ from proposition C2, and

$$E[X(k+1|k)\{X(k+1)-X(k+1|k)\}]=0$$

because $X(k+1|k)$ and $\hat{X}(k+1)-\hat{X}(k+1|k)$ are orthogonal.

In order to justify (32) and (33) in proposition 3 we have these two results:

Proposition C8: $P(k+1|k)=F^2P(k|k)+Q$.

Proof

$$P(k+1-k)=E[\{X(k+1)-\hat{X}(k+1|k)\}^2]$$
$$=E(\{FX(k)+U(k)-F\hat{X}(k|k)\}^2)$$
$$=E[\{F(X(k)-\hat{X}(k|k))+U(k)\}^2]$$
$$=F^2P(k|k)+Q+2FE[(X(k)-\hat{X}(k|k))U(k)]$$
$$=F^2P(k|k)+Q$$

since $E[(X(k)-\hat{X}(k|k))U(k)]=0$ by proposition C1.

Proposition C9: $P(k+1|k+1)=(1-HA(k+1))P(k+1|k)$.

Proof

$$P(k+1|k+1)=E[(\hat{X}(k+1)-X(k+1|k+1))^2]$$
$$=E[(X(k+1)-\hat{X}(k+1|k)-A(k+1)\tilde{Z}(k+1))^2] \quad \text{(see (27))}$$
$$=E[(X(k+1)-\hat{X}(k+1|k)-A(k+1)(H[X(k+1)-\hat{X}(k+1|k)]+V(k+1)^2]$$
$$=E[(1-HA(k+1)(X(k+1)-\hat{X}(k+1|k))-A(k+1)V(k+1))^2]$$
$$=(1-HA(k+1))^2P(k+1|k)+RA(k+1)^2$$
$$=(1-HA(k+1))P(k+1|k)-HA(k+1)P(k+1|k)$$
$$+H^2A(k+1)^2P(k+1|k)+RA(k+1)^2$$

Now observe that since

$$A(k+1)=\frac{HP(k+1|k)}{H^2P(k+1|k)+R}$$

then

$$H^2A(k+1)^2P(k+1|k)+RA(k+1)^2=A(k+1)\frac{HP(k+1|k)}{H^2P(k+1|k)+R}[H^2P(k+1|k)+R]$$

$$=HA(k+1)P(k+1|k)$$

Hence the result follows.

References

[1] ANDERSON, B. D. O. and MOORE, J. B., 1979, *Optimal Filtering*, (Englewood Cliffs, NJ: Prentice-Hall).
[2] GELB, A. (ed.), 1974, *Applied Optimal Estimation* (Cambridge, MA: MIT Press).
[3] KAILATH, T., 1974, *IEEE Trans. on Information Theory*, **IT–20**, (2), March.
[4] LUENBERGER, D. G., 1969, *Optimization by Vector Space Methods* (New York: Wiley).
[5] MEDITCH, J. S., 1969, *Stochastic Optimal Linear Estimation and Control* (New York: McGraw-Hill).
[6] SORENSON, H. W., 1970, *IEEE Spectrum*, **7** (7), July.
[7] VAN TREES, H. L., 1968, *Detection, Estimation, and Modulation Theory*, Part I, (New York: Wiley).