4

Discrete-Time Minimum-Variance Prediction and Filtering

4.1 Introduction

Kalman filters are employed wherever it is desired to recover data from the noise in an optimal way, such as satellite orbit estimation, aircraft guidance, radar, communication systems, navigation, medical diagnosis and finance. Continuous-time problems that possess differential equations may be easier to describe in a state-space framework, however, the filters have higher implementation costs because an additional integration step and higher sampling rates are required. Conversely, although discrete-time state-space models may be less intuitive, the ensuing filter difference equations can be realised immediately.

The discrete-time Kalman filter calculates predicted states via the linear recursion

$$\hat{x}_{k+1/k} = A_k \hat{x}_{k-1/k} + K_k (z_k - C_k \hat{x}_{k-1/k}) ,$$

where the predictor gain, K_k, is a function of the noise statistics and the model parameters. The above formula was reported by Rudolf E. Kalman in the 1960s [1], [2]. He has since received many awards and prizes, including the National Medal of Science, which was presented to him by President Barack Obama in 2009.

The Kalman filter calculations are simple and well-established. A possibly troublesome obstacle is expressing problems at hand within a state-space framework. This chapter derives the main discrete-time results to provide familiarity with state-space techniques and filter application. The continuous-time and discrete-time minimum-square-error Wiener filters were derived using a completing-the-square approach in Chapters 1 and 2, respectively. Similarly for time-varying continuous-time signal models, the derivation of the minimum-variance Kalman filter, presented in Chapter 3, relied on a least-mean-square (or conditional-mean) formula. This formula is used again in the solution of the discrete-time prediction and filtering problems. Predictions can be used when the measurements are irregularly spaced or missing at the cost of increased mean-square-error.

This chapter develops the prediction and filtering results for the case where the problem is nonstationary or time-varying. It is routinely assumed that the process and measurement noises are zero mean and uncorrelated. Nonzero mean cases can be accommodated by including deterministic inputs within the state prediction and filter output updates. Correlated noises can be handled by adding a term within the predictor gain and the underlying Riccati equation. The same approach is employed when the signal model

"Man will occasionally stumble over the truth, but most of the time he will pick himself up and continue on." *Winston Leonard Spencer-Churchill* possesses a direct-feedthrough term. A simplification of the generalised regulator problem from control theory is presented, from which the solutions of output estimation, input estimation (or equalisation), state estimation and mixed filtering problems follow immediately.

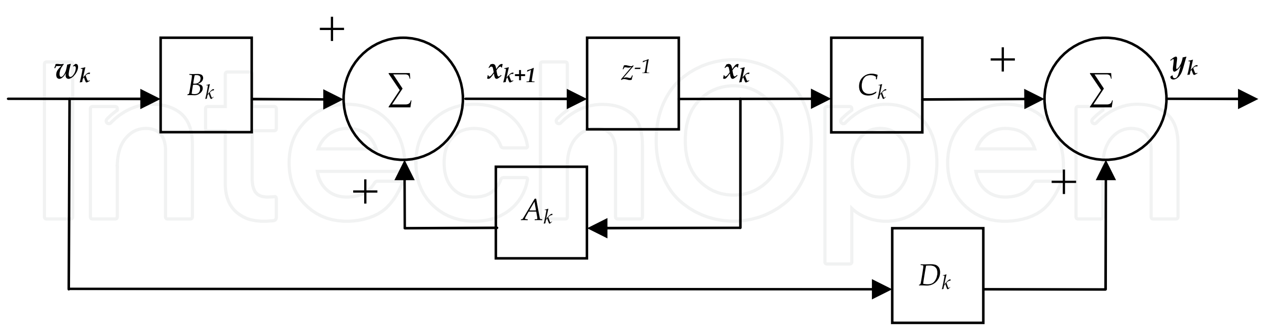


Figure 1. The discrete-time system $\mathcal{G}$ operates on the input signal $w_k \in \mathbb{R}^m$ and produces the output $y_k \in \mathbb{R}^p$.

4.2 The Time-varying Signal Model

A discrete-time time-varying system $\mathcal{G}: \mathbb{R}^m \rightarrow \mathbb{R}^p$ is assumed to have the state-space representation

$$x_{k+1} = A_k x_k + B_k w_k , \tag{1}$$

$$y_k = C_k x_k + D_k w_k , \tag{2}$$

where $A_k \in \mathbb{R}^{n \times n}$, $B_k \in \mathbb{R}^{n \times m}$, $C_k \in \mathbb{R}^{p \times n}$ and $D_k \in \mathbb{R}^{p \times p}$ over a finite interval $k \in [0, N]$. The w_k is a stochastic white process with

$$E\{w_k\} = 0, \; E\{w_j w_k^T\} = Q_k \delta_{jk} , \tag{3}$$

in which $\delta_{jk} = \begin{cases} 1 & \text{if } j = k \\ 0 & \text{if } j \neq k \end{cases}$ is the Kronecker delta function. This system is depicted in Fig. 1, in which z^{-1} is the unit delay operator. It is interesting to note that, at time k the current state

$$x_k = A_{k-1} x_{k-1} + B_{k-1} w_{k-1}, \tag{4}$$

does not involve w_k. That is, unlike continuous-time systems, here there is a one-step delay between the input and output sequences. The simpler case of $D_k = 0$, namely,

$$y_k = C_k x_k, \tag{5}$$

is again considered prior to the inclusion of a nonzero D_k.

"Rudy Kalman applied the state-space model to the filtering problem, basically the same problem discussed by Wiener. The results were astonishing. The solution was recursive, and the fact that the estimates could use only the past of the observations posed no difficulties." *Jan. C. Willems*

4.3 The State Prediction Problem

Suppose that observations of (5) are available, that is,

$$z_k = y_k + v_k, \tag{6}$$

where v_k is a white measurement noise process with

$$E\{v_k\} = 0,\ E\{v_j v_k^T\} = R_k \delta_{jk} \text{ and } E\{w_j v_k^T\} = 0. \tag{7}$$

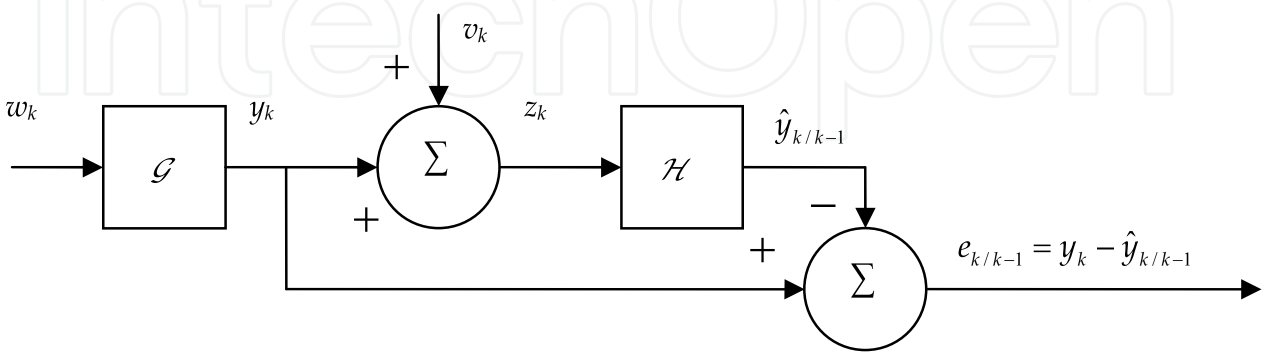


Figure 2. The state prediction problem. The objective is to design a predictor $\mathcal{H}$ which operates on the measurements and produces state estimates such that the variance of the error residual $e_{k/k-1}$ is minimised.

It is noted above for the state recursion (4), there is a one-step delay between the current state and the input process. Similarly, it is expected that there will be one-step delay between the current state estimate and the input measurement. Consequently, it is customary to denote $\hat{x}_{k/k-1}$ as the state estimate at time k, given measurements at time k – 1. The $\hat{x}_{k/k-1}$ is also known as the one-step-ahead state prediction. The objective here is to design a predictor $\mathcal{H}$ that operates on the measurements z_k and produces an estimate, $\hat{y}_{k/k-1} = C_k \hat{x}_{k/k-1}$, of $y_k = C_k y_k$, so that the covariance, $E\{e_{k/k-1} e_{k/k-1}^T\}$, of the error residual, $e_{k/k-1}$ = y_k – $\hat{y}_{k/k-1}$, is minimised. This problem is depicted in Fig. 2

4.4 The Discrete-time Conditional Mean Estimate

The predictor derivation that follows relies on the discrete-time version of the conditional-mean or least-mean-square estimate derived in Chapter 3, which is set out as follows. Consider a stochastic vector $[\alpha_k^T \quad \beta_k^T]^T$ having means and covariances

$$E\left\{\begin{bmatrix}\alpha_k \\ \beta_k\end{bmatrix}\right\} = \begin{bmatrix}\bar{\alpha} \\ \bar{\beta}\end{bmatrix} \tag{8}$$

"Prediction is very difficult, especially if it's about the future." *Niels Henrik David Bohr* and

$$E\left\{\begin{bmatrix}\alpha_k \\ \beta_k\end{bmatrix}\begin{bmatrix}\alpha_k^T & \beta_k^T\end{bmatrix}\right\} = \begin{bmatrix}\Sigma_{\alpha_k\alpha_k} & \Sigma_{\alpha_k\beta_k} \\ \Sigma_{\beta_k\alpha_k} & \Sigma_{\beta_k\beta_k}\end{bmatrix}. \tag{9}$$

respectively, where $\Sigma_{\beta_k\alpha_k} = \Sigma_{\alpha_k\beta_k}^T$. An estimate of α_k given β_k, denoted by $E\{\alpha_k \mid \beta_k\}$, which minimises $E\left\{(\alpha_k - E\{\alpha_k \mid \beta_k\})(\alpha_k - E\{\alpha_k \mid \beta_k\})^T\right\}$, is given by

$$E\{\alpha_k \mid \beta_k\} = \bar{\alpha} + \Sigma_{\alpha_k\beta_k}\Sigma_{\beta_k\beta_k}^{-1}(\beta_k - \bar{\beta}) . \tag{10}$$

The above formula is developed in [3] and established for Gaussian distributions in [4]. A derivation is requested in the problems. If a_k and β_k are scalars then (10) degenerates to the linear regression formula as is demonstrated below.

Example 1 (Linear regression [5]). The least-squares estimate $\hat{\alpha}_k = a\beta_k + b$ of α_k given data $a_k, \beta_k \in \mathbb{R}$ over [1, N], can be found by minimising the performance objective $J = \frac{1}{N}\sum_{k=1}^{N}(\alpha_k - \hat{\alpha})^2 = \frac{1}{N}\sum_{k=1}^{N}(\alpha_k - a\beta_k - b)^2$. Setting $\frac{dJ}{db} = 0$ yields $b = \bar{\alpha} - a\bar{\beta}$. Setting $\frac{dJ}{da} = 0$, substituting for b and using the definitions (8) – (9), results in $a = \Sigma_{\alpha_k\beta_k}\Sigma_{\beta_k\beta_k}^{-1}$.

4.5 Minimum-Variance Prediction

It follows from (1), (6), together with the assumptions $E\{w_k\} = 0$, $E\{v_k\} = 0$, that $E\{x_{k+1}\} = E\{A_k x_k\}$ and $E\{z_k\} = E\{C_k x_k\}$. It is assumed that similar results hold in the case of predicted state estimates, that is,

$$E\left\{\begin{bmatrix}\hat{x}_{k+1} \\ z_k\end{bmatrix}\right\} = \begin{bmatrix}A_k\hat{x}_{k/k-1} \\ C_k\hat{x}_{k/k-1}\end{bmatrix}. \tag{11}$$

Substituting (11) into (10) and denoting $\hat{x}_{k+1/k} = E\{\hat{x}_{k+1} \mid z_k\}$ yields the predicted state

$$\hat{x}_{k+1/k} = A_k\hat{x}_{k/k-1} + K_k(z_k - C_k\hat{x}_{k/k-1}) , \tag{12}$$

where $K_k = E\{\hat{x}_{k+1}z_k^T\}E\{z_k z_k^T\}^{-1}$ is known as the predictor gain, which is designed in the next section. Thus, the optimal one-step-ahead predictor follows immediately from the least-mean-square (or conditional mean) formula. A more detailed derivation appears in [4]. The structure of the optimal predictor is shown in Fig. 3. It can be seen from the figure that $\mathcal{H}$ produces estimates $\hat{y}_{k/k-1} = C_k\hat{x}_{k/k-1}$ from the measurements z_k.

"I admired Bohr very much. We had long talks together, long talks in which Bohr did practically all the talking." *Paul Adrien Maurice Dirac*

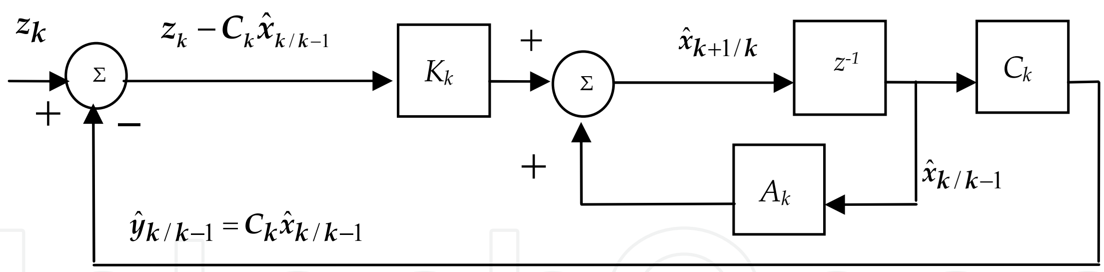


Figure 3. The optimal one-step-ahead predictor which produces estimates $\hat{x}_{k+1/k}$ of x_{k+1} given measurements z_k.

Let $\tilde{x}_{k/k-1} = x_k - \hat{x}_{k/k-1}$ denote the state prediction error. It is shown below that the expectation of the prediction error is zero, that is, the predicted state estimate is unbiased.

Lemma 1: *Suppose that* $\hat{x}_{0/0} = x_0$, *then*

$$E\{\tilde{x}_{k+1/k}\} = 0 \tag{13}$$

for all $k \in [0, N]$.

Proof: *The condition* $\hat{x}_{0/0} = x_0$ *is equivalent to* $\tilde{x}_{0/0} = 0$, *which is the initialisation step for an induction argument. Subtracting (12) from (1) gives*

$$\tilde{x}_{k+1/k} = (A_k - K_kC_k)\tilde{x}_{k/k-1} + B_kw_k - K_kv_k \tag{14}$$

and therefore

$$E\{\tilde{x}_{k+1/k}\} = (A_k - K_kC_k)E\{\tilde{x}_{k/k-1}\} + B_kE\{w_k\} - K_kE\{v_k\} . \tag{15}$$

From assumptions (3) and (7), the last two terms of the right-hand-side of (15) are zero. Thus, (13) follows by induction. □

4.6 Design of the Predictor Gain

It is shown below that the optimum predictor gain is that which minimises the prediction error covariance $E\{\tilde{x}_{k/k-1}\tilde{x}_{k/k-1}^T\}$.

Lemma 2: *In respect of the estimation problem defined by (1), (3), (5) - (7), suppose there exist solutions* $P_{k/k-1} = P_{k/k-1}^T \geq 0$ *to the Riccati difference equation*

$$P_{k+1/k} = A_kP_{k/k-1}A_k^T + B_kQ_kB_k^T - A_kP_{k/k-1}C_k^T(C_kP_{k/k-1}C_k^T + R_k)^{-1}C_kP_{k/k-1}A_k^T , \tag{16}$$

over [0, N], then the predictor gain

$$K_k = A_kP_{k/k-1}C_k^T(C_kP_{k/k-1}C_k^T + R_k)^{-1} , \tag{17}$$

within (12) minimises $P_{k/k-1} = E\{\tilde{x}_{k/k-1}\tilde{x}_{k/k-1}^T\}$.

"When it comes to the future, there are three kinds of people: those who let it happen, those who make it happen, and those who wondered what happened." *John M. Richardson Jr.*

Proof: *Constructing* $P_{k+1/k} = E\{\tilde{x}_{k+1/k}\tilde{x}_{k+1/k}^T\}$ *using (3), (7), (14),* $E\{\tilde{x}_{k/k-1}w_k^T\}$ *= 0 and* $E\{\tilde{x}_{k/k-1}v_k^T\}$ *= 0 yields*

$$P_{k+1/k} = (A_k - K_kC_k)P_{k/k-1}(A_k - K_kC_k)^T + B_kQ_kB_k^T + K_kR_kK_k^T , \tag{18}$$

which can be rearranged to give

$$\begin{aligned} P_{k+1/k} = {} & A_kP_{k/k-1}A_k^T - A_kP_{k/k-1}C_k^T(C_kP_{k/k-1}C_k^T + R_k)^{-1}C_kP_{k/k-1}A_k^T + B_kQ_kB_k^T \\ & + (K_k - A_kP_{k/k-1}C_k^T(C_kP_{k/k-1}C_k^T + R_k)^{-1})(C_kP_{k/k-1}C_k^T + R_k) \\ & \times (K_k - A_kP_{k/k-1}C_k^T(C_kP_{k/k-1}C_k^T + R_k)^{-1})^T , \end{aligned} \tag{19}$$

By inspection of (19), the predictor gain (17) minimises $P_{k+1/k}$. □

4.7 Minimum-Variance Filtering

It can be seen from (12) that the predicted state estimate $\hat{x}_{k/k-1}$ is calculated using the previous measurement z_{k-1} as opposed to the current data z_k. A state estimate, given the data at time k, which is known as the filtered state, can similarly be obtained using the linear least squares or conditional-mean formula. In Lemma 1 it was shown that the predicted state estimate is unbiased. Therefore, it is assumed that the expected value of the filtered state equals the expected value of the predicted state, namely,

$$E\left\{\begin{bmatrix} \hat{x}_{k/k} \\ z_k \end{bmatrix}\right\} = \begin{bmatrix} \hat{x}_{k/k-1} \\ C_k\hat{x}_{k/k-1} \end{bmatrix}. \tag{20}$$

Substituting (20) into (10) and denoting $\hat{x}_{k/k} = E\{\hat{x}_k \mid z_k\}$ yields the filtered estimate

$$\hat{x}_{k/k} = \hat{x}_{k/k-1} + L_k(z_k - C_k\hat{x}_{k/k-1}) , \tag{21}$$

where $L_k = E\{\hat{x}_kz_k^T\}E\{z_kz_k^T\}^{-1}$ is known as the filter gain, which is designed subsequently. Let $\tilde{x}_{k/k} = x_k - \hat{x}_{k/k}$ denote the filtered state error. It is shown below that the expectation of the filtered error is zero, that is, the filtered state estimate is unbiased.

Lemma 3: *Suppose that* $\hat{x}_{0/0} = x_0$*, then*

$$E\{\tilde{x}_{k/k}\} = 0 \tag{22}$$

for all $k \in [0, N]$.

"To be creative you have to contribute something different from what you've done before. Your results need not be original to the world; few results truly meet that criterion. In fact, most results are built on the work of others." *Lynne C. Levesque*

Proof: *Following the approach of [6], combining (4) - (6) results in* $z_k = C_kA_{k-1}x_{k-1} + C_kB_{k-1}w_{k-1} + v_k$, *which together with (21) yields*

$$\tilde{x}_{k/k} = (I - L_kC_k)A_{k-1}\tilde{x}_{k-1/k-1} + (I - L_kC_k)B_{k-1}w_{k-1} - L_kv_k \,. \quad (23)$$

From (23) and the assumptions (3), (7), it follows that

$$\begin{aligned} E\{\tilde{x}_{k/k}\} &= (I - L_kC_k)A_{k-1}E\{\tilde{x}_{k-1/k-1}\} \\ &= (I - L_kC_k)A_{k-1}\cdots(I - L_1C_1)A_0E\{\tilde{x}_{0/0}\} \,. \end{aligned} \quad (24)$$

Hence, with the initial condition $\hat{x}_{0/0} = x_0$, $E\{\tilde{x}_{k/k}\} = 0$. □

4.8 Design of the Filter Gain

It is shown below that the optimum filter gain is that which minimises the covariance $E\{\tilde{x}_{k/k}\tilde{x}_{k/k}^T\}$, where $\tilde{x}_{k/k} = x_k - \hat{x}_{k/k}$ is the filter error.

Lemma 4: *In respect of the estimation problem defined by (1), (3), (5) - (7), suppose there exists a solution* $P_{k/k} = P_{k/k}^T \geq 0$ *to the Riccati difference equation*

$$P_{k/k} = P_{k/k-1} - P_{k/k-1}C_k^T(C_kP_{k/k-1}C_k^T + R_k)^{-1}C_kP_{k/k-1} \,, \quad (25)$$

over [0, N], then the filter gain

$$L_k = P_{k/k-1}C_k^T(C_kP_{k/k-1}C_k^T + R_k)^{-1} \,, \quad (26)$$

within (21) minimises $P_{k/k} = E\{\tilde{x}_{k/k}\tilde{x}_{k/k}^T\}$.

Proof: *Subtracting* $\hat{x}_{k/k}$ *from* x_k *yields* $\tilde{x}_{k/k} = x_k - \hat{x}_{k/k} = x_k - \hat{x}_{k/k-1} - L_k(Cx_k + v_k - C\hat{x}_{k/k-1})$, *that is,*

$$\tilde{x}_{k/k} = (I - L_kC_k)\tilde{x}_{k/k-1} - L_kv_k \quad (27)$$

and

$$P_{k/k} = (I - L_kC_k)P_{k/k-1}(I - L_kC_k)^T + L_kR_kL_k^T \,, \quad (28)$$

which can be rearranged as

$$\begin{aligned} P_{k/k} &= P_{k/k-1} - P_{k/k-1}C_k^T(C_kP_{k/k-1}C_k^T + R_k)^{-1}C_kP_{k/k-1} \\ &\quad + (L_k - P_{k/k-1}C_k^T(C_kP_{k/k-1}C_k^T + R_k)^{-1})(C_kP_{k/k-1}C_k^T + R_k)(L_k - P_{k/k-1}C_k^T(C_kP_{k/k-1}C_k^T + R_k)^{-1})^T \end{aligned} \quad (29)$$

By inspection of (29), the filter gain (26) minimises $P_{k/k}$. □

Example 2 (Data Fusion). Consider a filtering problem in which there are two measurements of the same state variable (possibly from different sensors), namely A_k, B_k, $Q_k \in \mathbb{R}$, $C_k = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $R_k = \begin{bmatrix} R_{1,k} & 0 \\ 0 & R_{2,k} \end{bmatrix}$, with $R_{1,k}$, $R_{2,k} \in \mathbb{R}$. Let $P_{k/k-1}$ denote the solution of the Riccati difference equation (25). By applying Cramer's rule within (26) it can be found that the filter gain is given by

---

"A professor is one who can speak on any subject - for precisely fifty minutes." *Norbert Wiener*

$$L_k = \begin{bmatrix} \dfrac{R_{2,k}P_{k/k-1}}{R_{2,k}P_{k/k-1} + R_{1,k}P_{k/k-1} + R_{1,k}R_{2,k}} & \dfrac{R_{1,k}P_{k/k-1}}{R_{2,k}P_{k/k-1} + R_{1,k}P_{k/k-1} + R_{1,k}R_{2,k}} \end{bmatrix},$$

from which it follows that $\lim_{R_{1,k}\to 0 \quad R_{2,k}\neq 0} L_k = \begin{bmatrix} 1 & 0 \end{bmatrix}$ and $\lim_{R_{2,k}\to 0 \quad R_{1,k}\neq 0} L_k = \begin{bmatrix} 0 & 1 \end{bmatrix}$. That is, when the first measurement is noise free, the filter ignores the second measurement and *vice versa*. Thus, the Kalman filter weights the data according to the prevailing measurement qualities.

4.9 The Predictor-Corrector Form

The Kalman filter may be written in the following predictor-corrector form. The corrected (or filtered) error covariances and states are respectively given by

$$\begin{aligned} P_{k/k} &= P_{k/k-1} - P_{k/k-1}C_k^T(C_kP_{k/k-1}C_k^T + R_k)^{-1}C_kP_{k/k-1} \\ &= P_{k/k-1} - L_k(C_kP_{k/k-1}C_k^T + R_k)^{-1}L_k^T \\ &= (I - L_kC_k)P_{k/k-1}, \end{aligned} \tag{30}$$

$$\begin{aligned} \hat{x}_{k/k} &= \hat{x}_{k/k-1} + L_k(z_k - C_k\hat{x}_{k/k-1}) \\ &= (I - L_kC_k)\hat{x}_{k/k-1} + L_kz_k, \end{aligned} \tag{31}$$

where $L_k = P_{k/k-1}C_k^T(C_kP_{k/k-1}C_k^T + R_k)^{-1}$. Equation (31) is also known as the measurement update. The predicted state and error covariances are respectively given by

$$\begin{aligned} \hat{x}_{k+1/k} &= A_k\hat{x}_{k/k} \\ &= (A_k - K_kC_k)\hat{x}_{k/k-1} + K_kz_k, \end{aligned} \tag{32}$$

$$P_{k+1/k} = A_kP_{k/k}A_k^T + B_kQ_kB_k^T, \tag{33}$$

where $K_k = A_kP_{k/k-1}C_k^T(C_kP_{k/k-1}C_k^T + R_k)^{-1}$. It can be seen from (31) that the corrected estimate, $\hat{x}_{k/k}$, is obtained using measurements up to time k. This contrasts with the prediction at time k + 1 in (32), which is based on all previous measurements. The output estimate is given by

$$\begin{aligned} \hat{y}_{k/k} &= C_k\hat{x}_{k/k} \\ &= C_k\hat{x}_{k/k-1} + C_kL_k(z_k - C_k\hat{x}_{k/k-1}) \\ &= C_k(I - L_kC_k)\hat{x}_{k/k-1} + C_kL_kz_k. \end{aligned} \tag{34}$$

"Before the advent of the Kalman filter, most mathematical work was based on Norbert Wiener's ideas, but the 'Wiener filtering' had proved difficult to apply. Kalman's approach, based on the use of state space techniques and a recursive least-squares algorithm, opened up many new theoretical and practical possibilities. The impact of Kalman filtering on all areas of applied mathematics, engineering, and sciences has been tremendous." *Eduardo Daniel Sontag*

4.10 The *A* Posteriori Filter

The above predictor-corrector form is used in the construction of extended Kalman filters for nonlinear estimation problems (see Chapter 10). When state predictions are not explicitly required, the following one-line recursion for the filtered state can be employed. Substituting $\hat{x}_{k/k-1} = A_{k-1}\hat{x}_{k-1/k-1}$ into $\hat{x}_{k/k} = (I - L_k C_k)\hat{x}_{k/k-1} + L_k z_k$ yields $\hat{x}_{k/k} = (I - L_k C_k) A_{k-1}\hat{x}_{k-1/k-1} + L_k z_k$. Hence, the output estimator may be written as

$$\begin{bmatrix} \hat{x}_{k/k} \\ \hat{y}_{k/k} \end{bmatrix} = \begin{bmatrix} (I - L_k C_k)A_{k-1} & L_k \\ C_k & \end{bmatrix} \begin{bmatrix} \hat{x}_{k-1/k-1} \\ z_k \end{bmatrix}, \tag{35}$$

This form is called the *a posteriori* filter within [7], [8] and [9]. The absence of a direct feed-through matrix above reduces the complexity of the robust filter designs described in [7], [8] and [9].

4.11 The Information Form

Algebraically equivalent recursions of the Kalman filter can be obtained by propagating a so-called corrected information state

$$\underline{\hat{x}}_{k/k} = P_{k/k}^{-1}\hat{x}_{k/k}, \tag{36}$$

and a predicted information state

$$\underline{\hat{x}}_{k+1/k} = P_{k+1/k}^{-1}\hat{x}_{k+1/k}. \tag{37}$$

The expression

$$(A + BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}, \tag{38}$$

which is variously known as the Matrix Inversion Lemma, the Sherman-Morrison formula and Woodbury's identity, is used to derive the information filter, see [3], [4], [11], [14] and [15]. To confirm the above identity, premultiply both sides of (38) by $(A + BD^{-1}C)$ to obtain

$$\begin{aligned} I &= I + BCDA^{-1} - B(C^{-1} + DA^{-1}B)^{-1}DA^{-1} - BCDA^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1} \\ &= I + BCDA^{-1} - B(I + CDA^{-1}B)^{-1}(C^{-1} + DA^{-1}B)^{-1}DA^{-1} \\ &= I + BCDA^{-1} - BC(C^{-1} + DA^{-1}B)^{-1}(C^{-1} + DA^{-1}B)^{-1}DA^{-1}, \end{aligned}$$

"I have been aware from the outset that the deep analysis of something which is now called Kalman filtering was of major importance. But even with this immodesty I did not quite anticipate all the reactions to this work." *Rudolf Emil Kalman* from which the result follows. From the above Matrix Inversion Lemma and (30) it follows that

$$\begin{aligned} P_{k/k}^{-1} &= (P_{k/k-1} - P_{k/k-1}C_k^T(C_kP_{k/k-1}C_k^T + R_k)^{-1}C_kP_{k/k-1})^{-1} \\ &= P_{k/k-1}^{-1} + C_k^TR_k^{-1}C_k , \end{aligned} \tag{39}$$

assuming that $P_{k/k-1}^{-1}$ and R_k^{-1} exist. An expression for $P_{k+1/k}^{-1}$ can be obtained from the Matrix Inversion Lemma and (33), namely,

$$\begin{aligned} P_{k+1/k}^{-1} &= (A_kP_{k/k}A_k^T + B_kQ_kB_k^T)^{-1} \\ &= (F_k^{-1} + B_kQ_kB_k^T)^{-1} , \end{aligned} \tag{40}$$

where $F_k = (A_kP_{k/k}A_k^T)^{-1} = A_k^{-T}P_{k/k}^{-1}A_k^{-1}$, which gives

$$P_{k+1/k}^{-1} = (I - F_kB_k(B_k^TF_kB_k + Q_k^{-1})^{-1}B_k^T)F_k . \tag{41}$$

Another useful identity is

$$\begin{aligned} (A + BCD)^{-1}BC &= A^{-1}(I + BCDA^{-1})^{-1}BC \\ &= A^{-1}B(I + CDA^{-1}B)^{-1}C \\ &= A^{-1}B(C^{-1} + DA^{-1}B)^{-1} . \end{aligned} \tag{42}$$

From (42) and (39), the filter gain can be expressed as

$$\begin{aligned} L_k &= P_{k/k-1}C_k^T(C_kP_{k/k-1}C_k^T + R_k)^{-1} \\ &= (P_{k/k-1}^{-1} + C_k^TR_k^{-1}C_k)^{-1}C_k^TR_k^{-1} \\ &= P_{k/k}C_k^TR_k^{-1} . \end{aligned} \tag{43}$$

Premultiplying (39) by $P_{k/k}$ and rearranging gives

$$I - L_kC_k = P_{k/k}P_{k/k-1}^{-1} . \tag{44}$$

It follows from (31), (36) and (44) that the corrected information state is given by

$$\begin{aligned} \underline{\hat{x}}_{k/k} &= P_{k/k}^{-1}\hat{x}_{k/k} \\ &= P_{k/k}^{-1}(I - L_kC_k)\hat{x}_{k/k-1} + P_{k/k}^{-1}L_kz_k \\ &= \underline{\hat{x}}_{k/k-1} + C_k^TR_k^{-1}z_k . \end{aligned} \tag{45}$$

"Information is the oxygen of the modern age. It seeps through the walls topped by barbed wire, it wafts across the electrified borders." *Ronald Wilson Reagan*

The predicted information state follows from (37), (41) and the definition of F_k, namely,

$$\begin{aligned}\underline{\hat{x}}_{k+1/k} &= P_{k+1/k}^{-1}\hat{x}_{k+1/k} \\ &= P_{k+1/k}^{-1}A_k\hat{x}_{k/k} \\ &= (I - F_kB_k(B_k^TF_kB_k + Q_k^{-1})^{-1}B_k^T)F_kA_k\hat{x}_{k/k} \\ &= (I - F_kB_k(B_k^TF_kB_k + Q_k^{-1})^{-1}B_k^T)A_k^{-T}\underline{\hat{x}}_{k/k}\,.\end{aligned} \tag{46}$$

Recall from Lemma 1 and Lemma 3 that $E\{x_k - \hat{x}_{k+1/k}\}$ = 0 and $E\{x_k - \hat{x}_{k/k}\}$ = 0, provided $\hat{x}_{0/0}$ = x_0. Similarly, with $\hat{x}_{0/0}$ = $P_{0/0}^{-1}x_0$, it follows that $E\{x_k - P_{k+1/k}\underline{\hat{x}}_{k+1/k}\}$ = 0 and $E\{x_k - P_{k/k}\underline{\hat{x}}_{k/k}\}$ = 0. That is, the information states (scaled by the appropriate covariances) will be unbiased, provided that the filter is suitably initialised. The calculation cost and potential for numerical instability can influence decisions on whether to implement the predictor-corrector form (30) - (33) or the information form (39) - (46) of the Kalman filter. The filters have similar complexity, both require a $p \times p$ matrix inverse in the measurement updates (31) and (45). However, inverting the measurement covariance matrix for the information filter may be troublesome when the measurement noise is negligible.

4.12 Comparison with Recursive Least Squares

The recursive least squares (RLS) algorithm is equivalent to the Kalman filter designed with the simplifications $A_k = I$ and $B_k = 0$; see the derivations within [10], [11]. For convenience, consider a more general RLS algorithm that retains the correct A_k but relies on the simplifying assumption $B_k = 0$. Under these conditions, denote the RLS algorithm's predictor gain by

$$\underline{K}_k = A_k\underline{P}_{k/k-1}C_k^T(C_k\underline{P}_{k/k-1}C_k^T + R_k)^{-1}, \tag{47}$$

where $\underline{P}_{k/k-1}$ is obtained from the Riccati difference equation

$$\underline{P}_{k+1/k} = A_k\underline{P}_{k/k-1}A_k^T - A_k\underline{P}_{k/k-1}C_k^T(C_k\underline{P}_{k/k-1}C_k^T + R_k)^{-1}C_k\underline{P}_{k/k-1}A_k^T\,. \tag{48}$$

It is argued below that the cost of the above model simplification is an increase in mean-square-error.

Lemma 5: *Let $P_{k-1/k}$ denote the predicted error covariance within (33) for the optimal filter. Under the above conditions, the predicted error covariance, $\overline{P}_{k/k-1}$, exhibited by the RLS algorithm satisfies*

$$P_{k/k-1} \le \overline{P}_{k/k-1}\,. \tag{49}$$

"All of the books in the world contain no more information than is broadcast as video in a single large American city in a single year. Not all bits have equal value." *Carl Edward Sagan*

Proof: *From the approach of Lemma 2, the RLS algorithm's predicted error covariance is given by*

$$\begin{aligned}\overline{P}_{k+1/k} &= A_k\overline{P}_{k/k-1}A_k^T - A_k\overline{P}_{k/k-1}C_k^T(C_k\overline{P}_{k/k-1}C_k^T + R_k)^{-1}C_k\overline{P}_{k/k-1}A_k^T + B_kQ_kB_k^T \\ &+(\underline{K}_k - A_k\overline{P}_{k/k-1}C_k^T(C_k\overline{P}_{k/k-1}C_k^T + R_k)^{-1})(C_k\overline{P}_{k/k-1}C_k^T + R_k) \\ &\times(\underline{K}_k - A_k\overline{P}_{k/k-1}C_k^T(C_k\overline{P}_{k/k-1}C_k^T + R_k)^{-1})^T .\end{aligned} \tag{50}$$

The last term on the right-hand-side of (50) is nonzero since the above RLS algorithm relies on the erroneous assumption $B_kQ_kB_k^T = 0$*. Therefore (49) follows.* □

4.13 Repeated Predictions

When there are gaps in the data record, or the data is irregularly spaced, state predictions can be calculated an arbitrary number of steps ahead. The one-step-ahead prediction is given by (32). The two, three and *j*-step-ahead predictions, given data at time *k*, are calculated as

$$\hat{x}_{k+2/k} = A_{k+1}\hat{x}_{k+1/k} \tag{51}$$

$$\hat{x}_{k+3/k} = A_{k+2}\hat{x}_{k+2/k} \tag{52}$$

$$\vdots$$

$$\hat{x}_{k+j/k} = A_{k+j-1}\hat{x}_{k+j-1/k}, \tag{53}$$

see also [4], [12]. The corresponding predicted error covariances are given by

$$P_{k+2/k} = A_{k+1}P_{k+1/k}A_{k+1}^T + B_{k+1}Q_{k+1}B_{k+1}^T \tag{54}$$

$$P_{k+3/k} = A_{k+2}P_{k+2/k}A_{k+2}^T + B_{k+2}Q_{k+2}B_{k+2}^T \tag{55}$$

$$\vdots$$

$$P_{k+j/k} = A_{k+j-1}P_{k+j-1/k}A_{k+j-1}^T + B_{k+j-1}Q_{k+j-1}B_{k+j-1}^T . \tag{56}$$

Another way to handle missing measurements at time *i* is to set $C_i = 0$, which leads to the same predicted states and error covariances. However, the cost of relying on repeated predictions is an increased mean-square-error which is demonstrated below.

Lemma 6:

(i) $P_{k/k} \le P_{k/k-1}$.

(ii) *Suppose that*

$$A_kA_k^T + B_kQ_kB_k^T \ge I \tag{57}$$

for all k $\in$ *[0, N], then* $P_{k+j/k} \ge P_{k+j-1/k}$ *for all (j+k)* $\in$ *[0, N] .*

"Where a calculator on the ENIAC is equipped with 18,000 vacuum tubes and weighs 30 tones, computers in the future may have only 1,000 vacuum tubes and perhaps weigh 1.5 tons." *Popular Mechanics*, 1949

Proof:

(i) *The claim follows by inspection of (30) since* $L_{k-1}(C_{k-1}P_{k-1/k-2}C_{k-1}^T + R_{k-1})L_{k-1}^T \geq 0$. *Thus, the filter outperforms the one-step-ahead predictor.*

(ii) *For* $P_{k+j-1/k} \geq 0$*, condition (57) yields* $A_{k+j-1}P_{k+j-1/k}A_{k+j-1}^T + B_{k+j-1}Q_{k+j-1}B_{k+j-1}^T \geq P_{k+j-1/k}$ *which together with (56) results in* $P_{k+j/k} \geq P_{k+j-1/k}$. □

Example 3. Consider a filtering problem where $A = 0.9$ and $B = C = Q = R = 1$, for which $AA^T + BQB^T = 1.81 > 1$. The predicted error covariances, $P_{k+j/k}$, $j = 1 \ldots 10$, are plotted in Fig. 4. The monotonically increasing sequence of error variances shown in the figure demonstrates that degraded performance occurs during repeated predictions. Fig. 5 shows some sample trajectories of the model output (dotted line), filter output (crosses) and predictions (circles) assuming that $z_3 \ldots z_8$ are unavailable. It can be seen from the figure that the prediction error increases with time k, which illustrates Lemma 6.

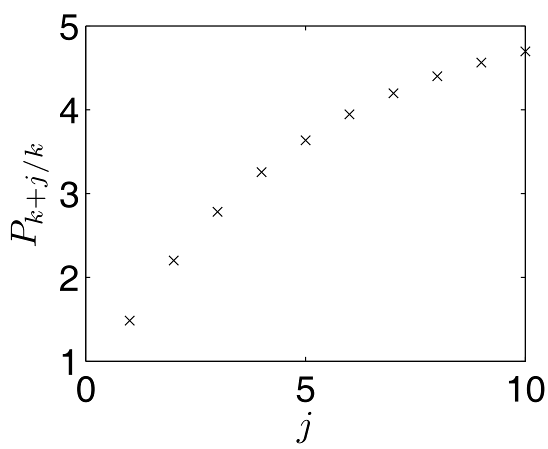


Figure 4. Predicted error variances for Example 3.

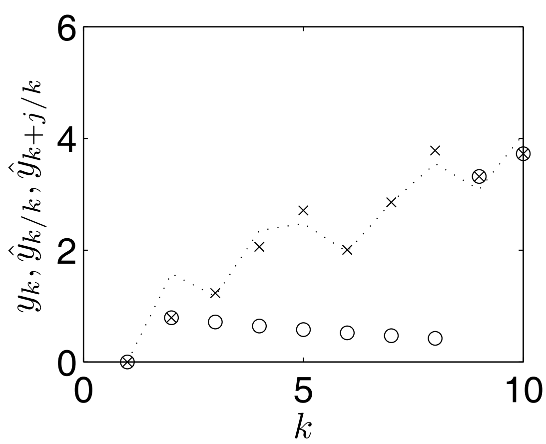


Figure 5. Sample trajectories for Example 3: y_k (dotted line), $\hat{y}_{k/k}$ (crosses) and $\hat{y}_{k+j/k}$ (circles).

4.14 Accommodating Deterministic Inputs

Suppose that the signal model is described by

$$x_{k+1} = A_k x_k + B_k w_k + \mu_k , \quad (58)$$

$$y_k = C_k x_k + \pi_k , \quad (59)$$

where μ_k and π_k are deterministic inputs (such as known non-zero means). The modifications to the Kalman recursions can be found by assuming $\hat{x}_{k+1/k} = A_k\hat{x}_{k/k} + \mu_k$ and $\hat{y}_{k/k-1} = C_k\hat{x}_{k/k-1} + \pi_k$. The filtered and predicted states are then given by

$$\hat{x}_{k/k} = \hat{x}_{k/k-1} + L_k(z_k - C_k\hat{x}_{k/k-1} - \pi_k) \quad (60)$$

"I think there is a world market for maybe five computers." *Thomas John Watson* and

$$\hat{x}_{k+1/k} = A_k \hat{x}_{k/k} + \mu_k \tag{61}$$

$$= A_k \hat{x}_{k/k-1} + K_k (z_k - C_k \hat{x}_{k/k-1} - \pi_k) + \mu_k , \tag{62}$$

respectively. Subtracting (62) from (58) gives

$$\begin{aligned} \tilde{x}_{k+1/k} &= A_k \tilde{x}_{k/k-1} - K_k (C_k \tilde{x}_{k/k-1} + \pi_k + v_k - \pi_k) + B_k w_k + \mu_k - \mu_k \\ &= (A_k - K_k C_k) \tilde{x}_{k/k-1} + B_k w_k - K_k v_k , \end{aligned} \tag{63}$$

where $\tilde{x}_{k/k-1} = x_k - \hat{x}_{k/k-1}$. Therefore, the predicted error covariance,

$$\begin{aligned} P_{k+1/k} &= (A_k - K_k C_k) P_{k/k-1} (A_k - K_k C_k)^T + B_k Q_k B_k^T + K_k R_k K_k^T \\ &= A_k P_{k/k-1} A_k^T - K_k (C_k P_{k/k-1} C_k^T + R_k) K_k^T + B_k Q_k B_k^T , \end{aligned} \tag{64}$$

is unchanged. The filtered output is given by

$$\hat{y}_{k/k} = C_k \hat{x}_{k/k} + \pi_k . \tag{65}$$

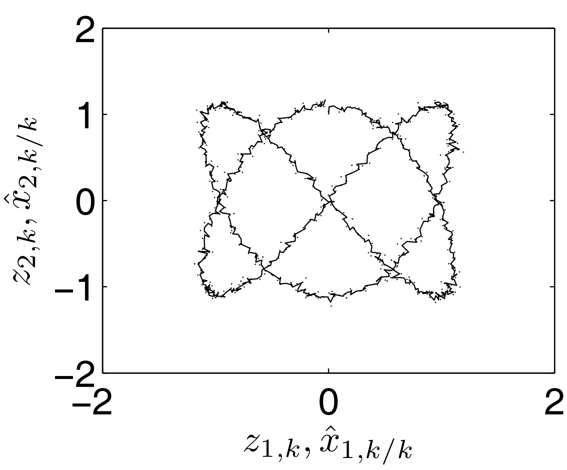


Figure 6. Measurements (dotted line) and filtered states (solid line) for Example 4.

Example 4. Consider a filtering problem where A = diag(0.1, 0.1), B = C = diag(1, 1), Q = R = diag(0.001, 0.001), with $\mu_k = \begin{bmatrix} \sin(2k) \\ \cos(3k) \end{bmatrix}$. The filtered states calculated from (60) are shown in Fig. 6. The resulting Lissajous figure illustrates that states having nonzero means can be modelled using deterministic inputs.

"There is no reason anyone would want a computer in their home." *Kenneth Harry Olson*

4.15 Correlated Process and Measurement Noises

Consider the case where the process and measurement noises are correlated

$$E\left\{\begin{bmatrix} w_j \\ v_j \end{bmatrix}\begin{bmatrix} w_k^T & v_k^T \end{bmatrix}\right\} = \begin{bmatrix} Q_k & S_k \\ S_k^T & R_k \end{bmatrix}\delta_{jk} . \tag{66}$$

The generalisation of the optimal filter that takes the above into account was published by Kalman in 1963 [2]. The expressions for the state prediction

$$\hat{x}_{k+1/k} = A_k\hat{x}_{k/k-1} + K_k(z_k - C_k\hat{x}_{k/k-1}) \tag{67}$$

and the state prediction error

$$\tilde{x}_{k+1/k} = (A_k - K_kC_k)\tilde{x}_{k/k-1} + B_kw_k - K_kv_k \tag{68}$$

remain the same. It follows from (68) that

$$E\{\tilde{x}_{k+1/k}\} = (A_k - K_kC_k)E\{\tilde{x}_{k/k-1}\} + \begin{bmatrix} B_k & -K_k \end{bmatrix}\begin{bmatrix} E\{w_k\} \\ E\{v_k\} \end{bmatrix}. \tag{69}$$

As before, the optimum predictor gain is that which minimises the prediction error covariance $E\{\tilde{x}_{k/k-1}\tilde{x}_{k/k-1}^T\}$.

Lemma 7: *In respect of the estimation problem defined by (1), (5), (6) with noise covariance (66), suppose there exist solutions* $P_{k/k-1} = P_{k/k-1}^T \ge 0$ *to the Riccati difference equation*

$$P_{k+1/k} = A_kP_{k/k-1}A_k^T + B_kQ_kB_k^T - (A_kP_{k/k-1}C_k^T + B_kS_k)(C_kP_{k/k-1}C_k^T + R_k)^{-1}(A_kP_{k/k-1}C_k^T + B_kS_k)^T \tag{70}$$

over [0, N], then the state prediction (67) with the gain

$$K_k = (A_kP_{k/k-1}C_k^T + B_kS_k)(C_kP_{k/k-1}C_k^T + R_k)^{-1} , \tag{71}$$

minimises $P_{k/k-1} = E\{\tilde{x}_{k/k-1}\tilde{x}_{k/k-1}^T\}$.

Proof: *It follows from (69) that*

$$\begin{aligned} E\{\tilde{x}_{k+1/k}\tilde{x}_{k+1/k}^T\} &= (A_k - K_kC_k)E\{\tilde{x}_{k/k-1}\tilde{x}_{k/k-1}^T\}(A_k - K_kC_k)^T \\ &\quad + \begin{bmatrix} B_k & -K_k \end{bmatrix}\begin{bmatrix} Q_k & S_k \\ S_k^T & R_k \end{bmatrix}\begin{bmatrix} B_k^T \\ -K_k^T \end{bmatrix} \\ &= (A_k - K_kC_k)E\{\tilde{x}_{k/k-1}\tilde{x}_{k/k-1}^T\}(A_k - K_kC_k)^T \\ &\quad + B_kQ_kB_k^T + K_kR_kK_k^T - B_kS_kK_k^T - K_kS_kB_k^T . \end{aligned} \tag{72}$$

"640K ought to be enough for anybody." *William Henry (Bill) Gates III*

Expanding (72) and denoting $P_{k/k-1} = E\{\tilde{x}_{k/k-1}\tilde{x}_{k/k-1}^T\}$ *gives*

$$P_{k+1/k} = A_k P_{k/k-1} A_k^T + B_k Q_k B_k^T - (A_k P_{k/k-1} C_k^T + B_k S_k)(C_k P_{k/k-1} C_k^T + R_k)^{-1}(A_k P_{k/k-1} C_k^T + B_k S_k)^T$$
$$+ \left(K_k - (A_k P_{k/k-1} C_k^T + B_k S_k)(C_k P_{k/k-1} C_k^T + R_k)^{-1}\right)\left(C_k P_{k/k-1} C_k^T + R_k\right)$$
$$\times \left(K_k - (A_k P_{k/k-1} C_k^T + B_k S_k)(C_k P_{k/k-1} C_k^T + R_k)^{-1}\right)^T. \quad (73)$$

By inspection of (73), the predictor gain (71) minimises $P_{k+1/k}$. □

Thus, the predictor gain is calculated differently when w_k and v_k are correlated. The calculation of the filtered state and filtered error covariance are unchanged, *viz.*

$$\hat{x}_{k/k} = (I - L_k C_k)\hat{x}_{k/k-1} + L_k z_k, \quad (74)$$

$$P_{k/k} = (I - L_k C_k) P_{k/k-1} (I - L_k C_k)^T + L_k R_k L_k^T, \quad (75)$$

where

$$L_k = P_{k/k-1} C_k^T (C_k P_{k/k-1} C_k^T + R_k)^{-1}. \quad (76)$$

However, $P_{k/k-1}$ is now obtained from the Riccati difference equation (70).

4.16 Including a Direct-Feedthrough Matrix

Suppose now that the signal model possesses a direct-feedthrough matrix, D_k, namely

$$x_{k+1} = A_k x_k + B_k w_k, \quad (77)$$

$$y_k = C_k x_k + D_k w_k. \quad (78)$$

Let the observations be denoted by

$$z_k = C_k x_k + \underline{v}_k, \quad (79)$$

where $\underline{v}_k = D_k w_k + v_k$, under the assumptions (3) and (7). It follows that

$$E\left\{\begin{bmatrix} w_j \\ \underline{v}_j \end{bmatrix}\begin{bmatrix} w_k^T & \underline{v}_k^T \end{bmatrix}\right\} = \begin{bmatrix} Q_k & Q_k D_k^T \\ D_k Q_k & D_k Q_k D_k^T + R_k \end{bmatrix}\delta_{jk}. \quad (80)$$

The approach of the previous section may be used to obtain the minimum-variance predictor for the above system. Using (80) within Lemma 7 yields the predictor gain

$$K_k = (A_k P_{k/k-1} C_k^T + B_k Q_k D_k^T)\Omega_k^{-1}, \quad (81)$$

"Everything that can be invented has been invented." *Charles Holland Duell* where

$$\Omega_k = C_k P_{k/k-1} C_k^T + D_k Q_k D_k^T + R_k \quad (82)$$

and $P_{k/k-1}$ is the solution of the Riccati difference equation

$$P_{k+1/k} = A_k P_{k/k-1} A_k^T - K_k \Omega_k K_k^T + B_k Q_k B_k^T \,. \quad (83)$$

The filtered states can be calculated from (74) , (82), (83) and $L_k = P_{k/k-1} C_k^T \Omega_k^{-1}$.

4.17 Solution of the General Filtering Problem

The general filtering problem is shown in Fig. 7, in which it is desired to develop a filter $\mathcal{H}$ that operates on noisy measurements of $\mathcal{G}_2$ and estimates the output of $\mathcal{G}_1$. Frequency domain solutions for time-invariant systems were developed in Chapters 1 and 2. Here, for the time-varying case, it is assumed that the system $\mathcal{G}_2$ has the state-space realisation

$$x_{k+1} = A_k x_k + B_k w_k \,, \quad (84)$$

$$y_{2,k} = C_{2,k} x_k + D_{2,k} w_k \,. \quad (85)$$

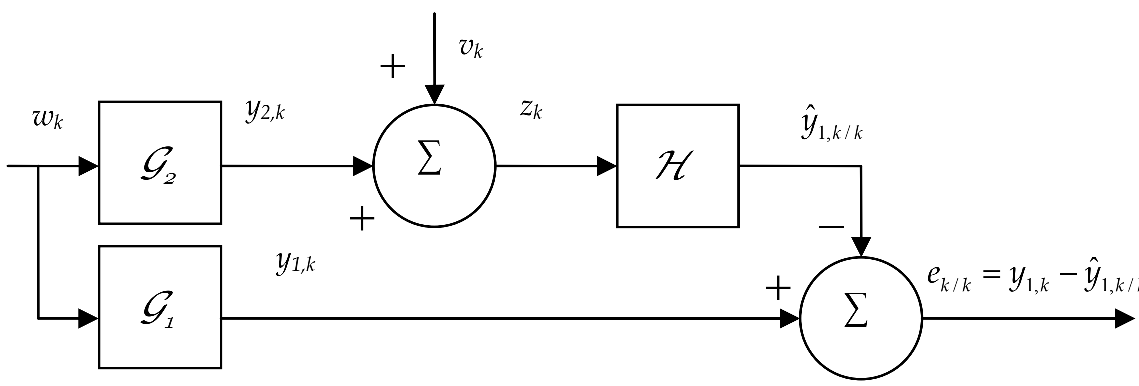


Figure 7. The general filtering problem. The objective is to estimate the output of $\mathcal{G}_1$ from noisy measurements of $\mathcal{G}_2$.

Suppose that the system $\mathcal{G}_1$ has the realisation (84) and

$$y_{1,k} = C_{1,k} x_k + D_{1,k} w_k \,. \quad (86)$$

The objective is to produce estimates $\hat{y}_{1,k/k}$ of $y_{1,k}$ from the measurements

$$z_k = C_{2,k} x_k + \underline{v}_k \,, \quad (87)$$

"He was a multimillionaire. Wanna know how he made all of his money? He designed the little diagrams that tell which way to put batteries on." *Stephen Wright* where $\underline{v}_k = D_{2,k} w_k + v_k$, so that the variance of the estimation error,

$$e_{k/k} = y_{1,k} - \hat{y}_{1,k/k} , \tag{88}$$

is minimised. The predicted state follows immediately from the results of the previous sections, namely,

$$\begin{aligned} \hat{x}_{k+1/k} &= A_k \hat{x}_{k/k-1} + K_k (z_k - C_{2,k} \hat{x}_{k/k-1}) \\ &= (A_k - K_k C_{2,k}) \hat{x}_{k/k-1} + K_k z_k \end{aligned} \tag{89}$$

where

$$K_k = (A_k P_{k/k-1} C_{2,k}^T + B_k Q_k D_{2,k}^T) \Omega_k^{-1} \tag{90}$$

and

$$\Omega_k = C_{2,k} P_{k/k-1} C_{2,k}^T + D_{2,k} Q_k D_{2,k}^T + R_k , \tag{91}$$

in which $P_{k/k-1}$ evolves from

$$P_{k+1/k} = A_k P_{k/k-1} A_k^T - K_k \Omega_k K_k^T + B_k Q_k B_k^T . \tag{92}$$

In view of the structure (89), an output estimate of the form

$$\begin{aligned} \hat{y}_{1,k/k} &= C_{1,k} \hat{x}_{k/k-1} + L_k (z_k - C_{2,k} \hat{x}_{k/k-1}) \\ &= (C_{1,k} - L_k C_{2,k}) \hat{x}_{k/k-1} + L_k z_k , \end{aligned} \tag{93}$$

is sought, where L_k is a filter gain to be designed. Subtracting (93) from (86) gives

$$\begin{aligned} e_{k/k} &= y_{1,k} - \hat{y}_{1,k/k} \\ &= (C_{1,k} - L_k C_{2,k}) \tilde{x}_{k/k-1} + \begin{bmatrix} D_{1,k} & -L_k \end{bmatrix} \begin{bmatrix} w_k \\ \underline{v}_k \end{bmatrix} . \end{aligned} \tag{94}$$

It is shown below that an optimum filter gain can be found by minimising the output error covariance $E\{e_{k/k} e_{k/k}^T\}$.

Lemma 8: *In respect of the estimation problem defined by (84) - (88), the output estimate* $\hat{y}_{1,k/k}$ *with the filter gain*

$$L_k = (C_{1,k} P_{k/k-1} C_{2,k}^T + D_{1,k} Q_k D_{2,k}^T) \Omega_k^{-1} \tag{95}$$

minimises $E\{e_{k/k} e_{k/k}^T\}$.

"This 'telephone' has too many shortcomings to be seriously considered as a means of communication. The device is inherently of no value to us." *Western Union* memo, 1876

Proof: *It follows from (94) that*

$$E\{e_{k/k}e_{k/k}^T\} = (C_{1,k} - L_kC_{2,k})P_{k/k-1}(C_{1,k}^T - C_{2,k}^TL_k^T)$$
$$+\begin{bmatrix} D_{1,k} & -L_k \end{bmatrix}\begin{bmatrix} Q_k & Q_kD_{2,k}^T \\ D_{2,k}Q_k^T & D_{2,k}Q_kD_{2,k}^T + R_k \end{bmatrix}\begin{bmatrix} D_{1,k}^T \\ -L_k^T \end{bmatrix} \tag{96}$$
$$= C_{1,k}P_{k/k-1}C_{1k}^T - C_{2,k}\bar{L}_k\Omega_k\bar{L}_k^TC_{2,k}^T,$$

which can be expanded to give

$$E\{e_{k/k}e_{k/k}^T\} = C_{1,k}P_{k/k-1}C_{1,k}^T + D_{2,k}Q_kD_{2,k}^T - (C_{1,k}P_{k/k-1}C2,_k^T + D_{1,k}Q_kD_{2,k}^T)\Omega_k^{-1}(C_{1,k}P_{k/k-1}C_{2,k}^T + D_{1,k}Q_kD_{2,k}^T)^T$$
$$-(C_{1,k}P_{k/k-1}C2,_k^T + D_{1,k}Q_kD_{2,k}^T)\Omega_k^{-1}(C_{1,k}P_{k/k-1}C_{2,k}^T + D_{1,k}Q_kD_{2,k}^T)^T$$
$$+(L_k - (C_{1,k}P_{k/k-1}C_{2,k}^T + D_{1,k}Q_kD_{2,k}^T)\Omega_k^{-1})\Omega_k$$
$$+(L_k - (C_{1,k}P_{k/k-1}C_{2,k}^T + D_{1,k}Q_kD_{2,k}^T)\Omega_k^{-1})^T . \tag{97}$$

By inspection of (97), the filter gain (95) minimises $E\{e_{k/k}e_{k/k}^T\}$ *.* □

The filter gain (95) has been generalised to include arbitrary $C_{1,k}$, $D_{1,k}$, and $D_{2,k}$. For state estimation, $C_2 = I$ and $D_2 = 0$, in which case (95) reverts to the simpler form (26). The problem (84) – (88) can be written compactly in the following generalised regulator framework from control theory [13].

$$\begin{bmatrix} x_{k+1} \\ e_{k/k} \\ z_k \end{bmatrix} = \begin{bmatrix} A_k & B_{1,1,k} & 0 \\ C_{1,1,k} & D_{1,1,k} & D_{1,2,k} \\ C_{2,1,k} & D_{2,1,k} & 0 \end{bmatrix}\begin{bmatrix} x_k \\ \begin{bmatrix} v_k \\ w_k \end{bmatrix} \\ -\hat{y}_{1,k/k} \end{bmatrix}, \tag{98}$$

where $B_{1,1,k} = \begin{bmatrix} 0 & B_k \end{bmatrix}$, $C_{1,1,k} = C_{1,k}$, $C_{2,1,k} = C_{2,k}$ $D_{1,1,k} = \begin{bmatrix} 0 & D_{1,k} \end{bmatrix}$, $D_{1,2,k} = I$ and $D_{2,1,k} = \begin{bmatrix} I & D_{2,k} \end{bmatrix}$. With the above definitions, the minimum-variance solution can be written as

$$\hat{x}_{k+1/k} = A_k\hat{x}_{k/k-1} + K_k(z_k - C_{2,1,k}\hat{x}_{k/k-1}), \tag{99}$$

$$\hat{y}_{1,k/k} = C_{1,1,k}\hat{x}_{k/k-1} + L_k(z_k - C_{2,1,k}\hat{x}_{k/k-1}), \tag{100}$$

"The wireless music box has no imaginable commercial value. Who would pay for a message sent to nobody in particular?" *David Sarnoff* where

$$K_k = \left(A_k P_{k/k-1} C_{2,1,k}^T + B_{1,1,k} \begin{bmatrix} R_k & 0 \\ 0 & Q_k \end{bmatrix} D_{2,1,k}^T \right) \left(C_{2,1,k} P_{k/k-1} C_{2,1,k}^T + D_{2,1,k} \begin{bmatrix} R_k & 0 \\ 0 & Q_k \end{bmatrix} D_{2,1,k}^T \right)^{-1}, \quad (101)$$

$$L_k = \left(C_{1,1,k} P_{k/k-1} C_{2,1,k}^T + D_{1,1,k} \begin{bmatrix} R_k & 0 \\ 0 & Q_k \end{bmatrix} D_{2,1,k}^T \right) \left(C_{2,1,k} P_{k/k-1} C_{2,1,k}^T + D_{2,1,k} \begin{bmatrix} R_k & 0 \\ 0 & Q_k \end{bmatrix} D_{2,1,k}^T \right)^{-1}, \quad (102)$$

in which $P_{k/k-1}$ is the solution of the Riccati difference equation

$$P_{k+1/k} = A_k P_{k/k-1} A_k^T - K_k (C_{2,1,k} P_{k/k-1} C_{2,1,k}^T + D_{2,1,k} \begin{bmatrix} R_k & 0 \\ 0 & Q_k \end{bmatrix} D_{2,1,k}^T) K_k^T + B_{1,1,k} \begin{bmatrix} R_k & 0 \\ 0 & Q_k \end{bmatrix} B_{1,1,k}^T. \quad (103)$$

The application of the solution (99) – (100) to output estimation, input estimation (or equalisation), state estimation and mixed filtering problems is demonstrated in the example below.

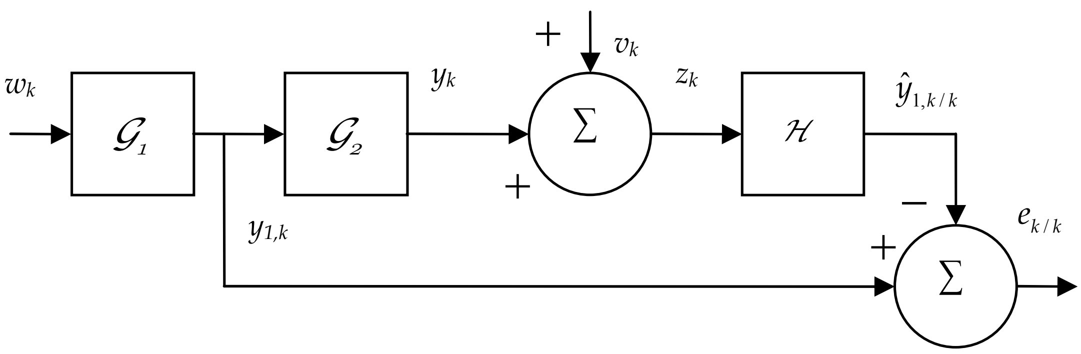


Figure 8. The mixed filtering and equalisation problem considered in Example 5. The objective is to estimate the output of the plant $\mathcal{G}_1$ which has been corrupted by the channel $\mathcal{G}_2$ and the measurement noise v_k.

Example 5.

(i) For output estimation problems, where $C_{1,k} = C_{2,k}$ and $D_{1,k} = D_{2,k}$, the predictor gain (101) and filter gain (102) are identical to the previously derived (90) and (95), respectively.

(ii) For state estimation problems, set $C_{1,k} = I$ and $D_{1,k} = 0$.

(iii) For equalisation problems, set $C_{1,k} = 0$ and $D_{1,k} = I$.

(iv) Consider a mixed filtering and equalisation problem depicted in Fig. 8, where the output of the plant $\mathcal{G}_1$ has been corrupted by the channel $\mathcal{G}_2$. Assume that $\mathcal{G}_1$ has the realisation $\begin{bmatrix} x_{1,k+1} \\ y_{1,k} \end{bmatrix} = \begin{bmatrix} A_{1,k} & B_{1,k} \\ C_{1,k} & D_{1,k} \end{bmatrix} \begin{bmatrix} x_{1,k} \\ w_k \end{bmatrix}$. Noting the realisation of

"Video won't be able to hold on to any market it captures after the first six months. People will soon get tired of staring at a plywood box every night." *Daryl Francis Zanuck* the cascaded system $\mathcal{G}_1\mathcal{G}_2$ (see Problem 7), the minimum-variance solution can be found by setting $A_k = \begin{bmatrix} A_{2,k} & B_{2,k}C_{1,k} \\ 0 & A_{1,k} \end{bmatrix}$, $B_{1,1k} = \begin{bmatrix} 0 & B_{2,k}D_{1,k} \\ 0 & B_{1,k} \end{bmatrix}$, $B_{1,2,k} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $C_{1,1,k} = \begin{bmatrix} C_{1,k} & 0 \end{bmatrix}$, $C_{2,1,k} = \begin{bmatrix} C_{2,k} & D_{2,k}C_{1,k} \end{bmatrix}$, $D_{1,1,k} = \begin{bmatrix} 0 & D_{1,k} \end{bmatrix}$ and $D_{2,1,k} = \begin{bmatrix} I & D_{2,k}D_{1,k} \end{bmatrix}$.

4.18 Hybrid Continuous-Discrete Filtering

Often a system's dynamics evolve continuously but measurements can only be observed in discrete time increments. This problem is modelled in [20] as

$$\dot{x}(t) = A(t)x(t) + B(t)w(t)\,, \tag{104}$$

$$z_k = C_k x_k + v_k\,, \tag{105}$$

where $E\{w(t)\} = 0$, $E\{w(t)w^T(\tau)\} = Q(t)\delta(t - \tau)$, $E\{v_k\} = 0$, $E\{v_j v_k^T\} = R_k\delta_{jk}$ and $x_k = x(kT_s)$, in which T_s is the sampling interval. Following the approach of [20], state estimates can be obtained from a hybrid of continuous-time and discrete-time filtering equations. The predicted states and error covariances are obtained from

$$\dot{\hat{x}}(t) = A(t)\hat{x}(t)\,, \tag{106}$$

$$\dot{P}(t) = A(t)P(t) + P(t)A^T(t) + B(t)Q(t)B^T(t)\,. \tag{107}$$

Define $\hat{x}_{k/k-1} = \hat{x}(t)$ and $P_{k/k-1} = P(t)$ at $t = kT_s$. The corrected states and error covariances are given by

$$\hat{x}_{k/k} = \hat{x}_{k/k-1} + L_k(z_k - C_k\hat{x}_{k/k-1})\,, \tag{108}$$

$$P_{k/k} = (I - L_k C_k)P_{k/k-1}\,, \tag{109}$$

where $L_k = P_{k/k-1}C_k^T(C_k P_{k/k-1}C_k^T + R_k)^{-1}$. The above filter is a linear system having jumps at the discrete observation times. The states evolve according to the continuous-time dynamics (106) in-between the sampling instants. This filter is applied in [20] for recovery of cardiac dynamics from medical image sequences.

4.19 Conclusion

A linear, time-varying system $\mathcal{G}_2$ is assumed to have the realisation $x_{k+1} = A_k x_k + B_k w_k$ and $y_{2,k} = C_{2,k}x_k + D_{2,k}w_k$. In the general filtering problem, it is desired to estimate the output of a second reference system $\mathcal{G}_1$ which is modelled as $y_{1,k} = C_{1,k}x_k + D_{1,k}w_k$. The Kalman filter which estimates $y_{1,k}$ from the measurements $z_k = y_{2,k} + v_k$ at time k is listed in Table 1.

"Louis Pasteur's theory of germs is ridiculous fiction." *Pierre Pachet*, Professor of Physiology at Toulouse, 1872

If the state-space parameters are known exactly then this filter minimises the predicted and corrected error covariances $E\{(x_k - \hat{x}_{k/k-1})(x_k - \hat{x}_{k/k-1})^T\}$ and $E\{(x_k - \hat{x}_{k/k})(x_k - \hat{x}_{k/k})^T\}$, respectively. When there are gaps in the data record, or the data is irregularly spaced, state predictions can be calculated an arbitrary number of steps ahead, at the cost of increased mean-square-error.

	ASSUMPTIONS	MAIN RESULTS
Signals and system	$E\{w_k\} = E\{v_k\} = 0$. $E\{w_k w_k^T\} = Q_k$ and $E\{v_k v_k^T\} = R_k$ are known. A_k, B_k, $C_{1,k}$, $C_{2,k}$, $D_{1,k}$, $D_{2,k}$ are known.	$x_{k+1} = A_k x_k + B_k w_k$ $y_{2,k} = C_{2,k} x_k + D_{2,k} w_k$ $z_k = y_{2,k} + v_k$ $y_{1,k} = C_{1,k} x_k + D_{1,k} w_k$
Filtered state and output factorisation		$\hat{x}_{k+1/k} = (A_k - K_k C_{2,k})\hat{x}_{k/k-1} + K_k z_k$ $\hat{y}_{1,k/k} = (C_{1,k} - L_k C_{2,k})\hat{x}_{k/k-1} + L_k z_k$
Predictor gain, filter gain and Riccati difference equation	$Q_k > 0$, $R_k > 0$. $C_{2,k} P_{k/k-1} C_{2,k}^T + D_{2,k} Q_k D_{2,k}^T + R_k > 0$.	$K_k = (A_k P_{k/k-1} C_{2,k}^T + B_k Q_k D_{2,k}^T)$ $\times (C_{2,k} P_{k/k-1} C_{2,k}^T + D_{2,k} Q_k D_{2,k}^T + R_k)^{-1}$ $L_k = (C_{1,k} P_{k/k-1} C_{2,k}^T + D_{1,k} Q_k D_{2,k}^T)$ $\times (C_{2,k} P_{k/k-1} C_{2,k}^T + D_{2,k} Q_k D_{2,k}^T + R_k)^{-1}$ $P_{k+1/k} = A_k P_{k/k-1} A_k^T + B_k Q_k B_k^T$ $-K_k (C_{2,k} P_{k/k-1} C_{2,k}^T + D_{2,k} Q_k D_{2,k}^T + R_k) K_k^T$

Table 1.1. Main results for the general filtering problem.

The filtering solution is specialised to output estimation with $C_{1,k} = C_{2,k}$ and $D_{1,k} = D_{2,k}$.

In the case of input estimation (or equalisation), $C_{1,k} = 0$ and $D_{1,k} = I$, which results in $\hat{w}_{k/k} = -L_k C_{2,k}\hat{x}_{k/k-1} + L_k z_k$, where the filter gain is instead calculated as $L_k = Q_k D_{2,k}^T (C_{2,k} P_{k/k-1} C_{2,k}^T + D_{2,k} Q_k D_{2,k}^T + R_k)^{-1}$.

For problems where $C_{1,k} = I$ (state estimation) and $D_{1,k} = D_{2,k} = 0$, the filtered state calculation simplifies to $\hat{x}_{k/k} = (I - L_k C_{2,k})\hat{x}_{k/k-1} + L_k z_k$, where $\hat{x}_{k/k-1} = A_k \hat{x}_{k-1/k-1}$ and $L_k =$

"Heavier-than-air flying machines are impossible." *Baron William Thomson Kelvin*

$P_{k/k-1}C_{2,k}^T(C_{2,k}P_{k/k-1}C_{2,k}^T + R_k)^{-1}$. This predictor-corrector form is used to obtain robust, hybrid and extended Kalman filters. When the predicted states are not explicitly required, the state corrections can be calculated from the one-line recursion $\hat{x}_{k/k} = (I - L_kC_{2,k})A_{k-1}\hat{x}_{k-1/k-1} + L_kz_k$.

If the simplifications $B_k = D_{2,k} = 0$ are assumed and the pair $(A_k, C_{2,k})$ is retained, the Kalman filter degenerates to the RLS algorithm. However, the cost of this model simplification is an increase in mean-square-error.

4.20 Problems

Problem 1. Suppose that $E\left\{\begin{bmatrix}\alpha_k \\ \beta_k\end{bmatrix}\right\} = \begin{bmatrix}\bar{\alpha} \\ \bar{\beta}\end{bmatrix}$ and $E\left\{\begin{bmatrix}\alpha_k \\ \beta_k\end{bmatrix}\begin{bmatrix}\alpha_k^T & \beta_k^T\end{bmatrix}\right\} = \begin{bmatrix}\Sigma_{\alpha_k\alpha_k} & \Sigma_{\alpha_k\beta_k} \\ \Sigma_{\beta_k\alpha_k} & \Sigma_{\beta_k\beta_k}\end{bmatrix}$. Show that an estimate of α_k given β_k, which minimises $E\{(\alpha_k - E\{\alpha_k \mid \beta_k\})(\alpha_k - E\{\alpha_k \mid \beta_k\})^T\}$, is given by $E\{\alpha_k \mid \beta_k\} = \bar{\alpha} + \Sigma_{\alpha_k\beta_k}\Sigma_{\beta_k\beta_k}^{-1}(\beta_k - \bar{\beta})$.

Problem 2. Derive the predicted error covariance $P_{k+1/k} = A_kP_{k/k-1}A_k^T - A_kP_{k/k-1}C_k^T(C_kP_{k/k-1}C_k^T + R_k)^{-1}C_kP_{k/k-1}A_k^T + B_kQ_kB_k^T$ from the state prediction $\hat{x}_{k+1/k} = A_k\hat{x}_{k/k-1} + K_k(z_k - C_k\hat{x}_{k/k-1})$, the model $x_{k+1} = A_kx_k + B_kw_k$, $y_k = C_kx_k$ and the measurements $z_k = y_k + v_k$.

Problem 3. Assuming the state correction $\hat{x}_{k/k} = \hat{x}_{k/k-1} + L_k(z_k - C_k\hat{x}_{k/k-1})$, show that the corrected error covariance is given by $P_{k/k} = P_{k/k-1} - L_k(C_kP_{k/k-1}C_k^T + R_k)L_k^T$.

Problem 4 [11], [14], [17], [18], [19]. Consider the standard discrete-time filter equations

$$\hat{x}_{k/k-1} = A_k\hat{x}_{k-1/k-1},$$

$$\hat{x}_{k/k} = \hat{x}_{k/k-1} + L_k(z_k - C_k\hat{x}_{k/k-1}),$$

$$P_{k/k-1} = A_kP_{k-1/k-1}A_k^T + B_kQ_kB_k^T,$$

$$P_{k/k} = P_{k/k-1} - L_k(C_kP_{k/k-1}C_k^T + R_k)L_k^T,$$

where $L_k = P_{k/k-1}C_k^T(C_kP_{k/k-1}C_k^T + R_k)^{-1}$. Derive the continuous-time filter equations, namely

$$\dot{\hat{x}}(t_k) = A(t_k)\hat{x}(t_k) + K(t_k)\big(z(t_k) - C(t_k)\hat{x}(t_k)\big),$$

$$\dot{P}(t_k) = A(t_k)P(t_k) + P(t_k)A^T(t_k) - P(t_k)C^T(t_k)R^{-1}(t_k)C(t_k)P(t_k) + B(t_k)Q(t_k)B^T(t_k),$$

"But what is it good for?" Engineer at the Advanced Computing Systems Division of IBM, commenting on the micro chip, 1968 where $K(t_k) = P(t_k)C(t_k)R^{-1}(t_k)$. (Hint: Introduce the quantities $A_k = (I + A(t_k))\Delta t$, $B(t_k) = B_k$, $C(t_k) = C_k$, $P_{k/k}$, $Q(t_k) = Q_k/\Delta t$, $R(t_k) = R_k\Delta t$, $\hat{x}(t_k) = \hat{x}_{k/k}$, $P(t_k) = P_{k/k}$, $\dot{\hat{x}}(t_k) = \lim_{\Delta t \to 0} \dfrac{\hat{x}_{k/k} - \hat{x}_{k-1/k-1}}{\Delta t}$, $\dot{P}(t_k) = \lim_{\Delta t \to 0} \dfrac{P_{k+1/k} - P_{k/k-1}}{\Delta t}$ and $\Delta t = t_k - t_{k-1}$.)

Problem 5. Derive the two-step-ahead predicted error covariance $P_{k+2/k} = A_{k+1}P_{k+1/k}A_{k+1}^T + B_{k+1}Q_{k+1}B_{k+1}^T$.

Problem 6. Verify that the Riccati difference equation $P_{k+1/k} = A_kP_{k/k-1}A_k^T - K_k(C_kP_{k/-1}C_k^T + R_k)K_k^T + B_kQ_kB_k^T$, where $K_k = (A_kP_{k/k-1}C_k + B_kS_k)(C_kP_{k/k-1}C_k^T + R_k)^{-1}$, is equivalent to $P_{k+1/k} = (A_k - K_kC_k)P_{k/k-1}(A_k - K_kC_k)^T + K_kR_kK_k^T + B_kQ_kB_k^T - B_kS_kK_k^T - K_kS_kB_k^T$.

Problem 7 [16]. Suppose that the systems $y_{1,k} = \mathcal{G}_1 w_k$ and $y_{2,k} = \mathcal{G}_2 w_k$ have the state-space realisations

$$\begin{bmatrix} x_{1,k+1} \\ y_{1,k} \end{bmatrix} = \begin{bmatrix} A_{1,k} & B_{1,k} \\ C_{1,k} & D_{1,k} \end{bmatrix} \begin{bmatrix} x_{1,k} \\ w_k \end{bmatrix} \text{ and } \begin{bmatrix} x_{2,k+1} \\ y_{2,k} \end{bmatrix} = \begin{bmatrix} A_{2,k} & B_{2,k} \\ C_{2,k} & D_{2,k} \end{bmatrix} \begin{bmatrix} x_{2,k} \\ w_k \end{bmatrix}.$$

Show that the system $y_{3,k} = \mathcal{G}_2\mathcal{G}_1 w_k$ is given by

$$\begin{bmatrix} x_{1,k+1} \\ y_{3,k} \end{bmatrix} = \begin{bmatrix} A_{1,k} & 0 & B_{1,k} \\ B_{2,k}C_{1,k} & A_{2,k} & B_{2,k}D_{1,k} \\ D_{2,k}C_{1,k} & C_{2,k} & D_{2,k}D_{1,k} \end{bmatrix} \begin{bmatrix} x_{1,k} \\ x_{2,k} \\ w_k \end{bmatrix}.$$

4.21 Glossary

In addition to the notation listed in Section 2.6, the following nomenclature has been used herein.

$\mathcal{G}$	A system that is assumed to have the realisation $x_{k+1} = A_kx_k + B_kw_k$ and $y_k = C_kx_k + D_kw_k$ where A_k, B_k, C_k and D_k are time-varying matrices of appropriate dimension.
Q_k, R_k	Time-varying covariance matrices of stochastic signals w_k and v_k, respectively.
$\mathcal{G}^H$	Adjoint of $\mathcal{G}$. The adjoint of a system having the state-space parameters $\{A_k, B_k, C_k, D_k\}$ is a system parameterised by $\{A_k^T, -C_k^T, -B_k^T, D_k^T\}$.
$\hat{x}_{k/k}$	Filtered estimate of the state x_k given measurements at time k.
$\tilde{x}_{k/k}$	Filtered state estimation error which is defined by $\tilde{x}_{k/k} = x_k - \hat{x}_{k/k}$.
$P_{k/k}$	Corrected error covariance matrix at time k given measurements at time k.

"What sir, would you make a ship sail against the wind and currents by lighting a bonfire under her deck? I pray you excuse me. I have no time to listen to such nonsense." *Napoléon Bonaparte*

L_k	Time-varying filter gain matrix.
$\hat{x}_{k+1/k}$	Predicted estimate of the state x_{k+1} given measurements at time k.
$\tilde{x}_{k+1/k}$	Predicted state estimation error which is defined by $\tilde{x}_{k+1/k}$ = x_{k+1} – $\hat{x}_{k+1/k}$.
$P_{k+1/k}$	Predicted error covariance matrix at time k + 1 given measurements at time k.
K_k	Time-varying predictor gain matrix.
RLS	Recursive Least Squares.

4.22 References

[1] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems", *Transactions of the ASME, Series D, Journal of Basic Engineering*, vol 82, pp. 35 – 45, 1960.

[2] R. E. Kalman, "New Methods in Wiener Filtering Theory", *Proc. First Symposium on Engineering Applications of Random Function Theory and Probability*, Wiley, New York, pp. 270 – 388, 1963.

[3] T. Söderström, *Discrete-time Stochastic Systems: Estimation and Control*, Springer-Verlag London Ltd., 2002.

[4] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*, Prentice-Hall Inc,Englewood Cliffs, New Jersey, 1979.

[5] G. S. Maddala, *Introduction to Econometrics*, Second Edition, Macmillan Publishing Co., New York, 1992.

[6] C. K. Chui and G. Chen, *Kalman Filtering with Real-Time Applications*, 3rd Ed., Springer-Verlag, Berlin, 1999.

[7] I. Yaesh and U. Shaked, "H_∞-Optimal Estimation – The Discrete Time Case", *Proceedings of the MTNS*, pp. 261 – 267, Jun. 1991.

[8] U. Shaked and Y. Theodor, "H_∞ Optimal Estimation: A Tutorial", *Proceedings 31st IEEE Conference on Decision and Control*, pp. 2278 – 2286, Tucson, Arizona, Dec. 1992.

[9] F. L. Lewis, L. Xie and D. Popa, *Optimal and Robust Estimation: With an Introduction to Stochastic Control Theory*, Second Edition, Series in Automation and Control Engineering, Taylor & Francis Group, LLC, 2008.

[10] T. Kailath, A. H. Sayed and B. Hassibi, *Linear Estimation*, Prentice-Hall, Inc., Upper Saddle River, New Jersey, 2000.

[11] D. Simon, *Optimal State Estimation, Kalman H_∞ and Nonlinear Approaches*, John Wiley & Sons, Inc., Hoboken, New Jersey, 2006.

[12] P. J. Brockwell and R. A. Davis, *Time Series: Theory and Methods,* Second Edition, Springer-Verlag New York, Inc., 1991.

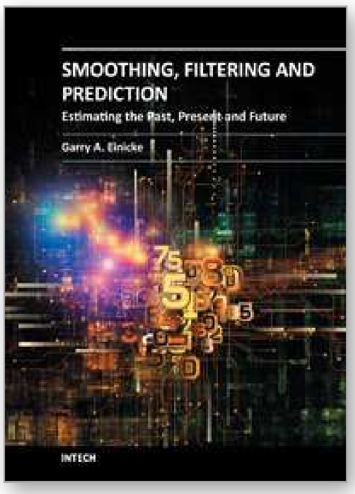
[13] D. J. N. Limebeer, M. Green and D. Walker, "Discrete-time H_∞ Control", *Proceedings 28th IEEE Conference on Decision and Control*, Tampa, pp. 392 – 396, Dec., 1989.

[14] R. G. Brown and P. Y. C. Hwang, *Introduction to Random Signals and Applied Kalman Filtering*, Second Edition, John Wiley & Sons, Inc., New York, 1992.

"The horse is here today, but the automobile is only a novelty - a fad." President of *Michigan Savings Bank*

[15] K. Ogata, *Discrete-time Control Systems*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1987.
[16] M. Green and D. J. N. Limebeer, *Linear Robust Control*, Prentice-Hall Inc. Englewood Cliffs, New Jersey, 1995.
[17] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*, Academic Press, Inc., New York, 1970.
[18] A. P. Sage and J. L. Melsa, *Estimation Theory with Applications to Communications and Control*, McGraw-Hill Book Company, New York, 1971.
[19] A. Gelb, *Applied Optimal Estimation*, The Analytic Sciences Corporation, USA, 197
[20] S. Tong and P. Shi, "Sampled-Data Filtering Framework for Cardiac Motion Recovery: Optimal Estimation of Continuous Dynamics From Discrete Measurements", *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 10, pp. 1750 – 1761, Oct. 2007.

"Airplanes are interesting toys but of no military value." *Marechal Ferdinand Foch*



Smoothing, Filtering and Prediction - Estimating The Past, Present and Future
Edited by

ISBN 978-953-307-752-9
Hard cover, 276 pages
Publisher InTech
Published online 24, February, 2012
Published in print edition February, 2012

This book describes the classical smoothing, filtering and prediction techniques together with some more recently developed embellishments for improving performance within applications. It aims to present the subject in an accessible way, so that it can serve as a practical guide for undergraduates and newcomers to the field. The material is organised as a ten-lecture course. The foundations are laid in Chapters 1 and 2, which explain minimum-mean-square-error solution construction and asymptotic behaviour. Chapters 3 and 4 introduce continuous-time and discrete-time minimum-variance filtering. Generalisations for missing data, deterministic inputs, correlated noises, direct feedthrough terms, output estimation and equalisation are described. Chapter 5 simplifies the minimum-variance filtering results for steady-state problems. Observability, Riccati equation solution convergence, asymptotic stability and Wiener filter equivalence are discussed. Chapters 6 and 7 cover the subject of continuous-time and discrete-time smoothing. The main fixed-lag, fixed-point and fixed-interval smoother results are derived. It is shown that the minimum-variance fixed-interval smoother attains the best performance. Chapter 8 attends to parameter estimation. As the above-mentioned approaches all rely on knowledge of the underlying model parameters, maximum-likelihood techniques within expectation-maximisation algorithms for joint state and parameter estimation are described. Chapter 9 is concerned with robust techniques that accommodate uncertainties within problem specifications. An extra term within Riccati equations enables designers to trade-off average error and peak error performance. Chapter 10 rounds off the course by applying the afore-mentioned linear techniques to nonlinear estimation problems. It is demonstrated that step-wise linearisations can be used within predictors, filters and smoothers, albeit by forsaking optimal performance guarantees.

InTech Europe
University Campus STeP Ri

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai

Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821