

Chapter 17

Kalman Filters

17.1 Introduction

This section provides a brief description of Kalman filter that can be considered an extension of the Wiener filtering concept [4]. The Kalman filter has as objective the minimization of the estimation square error of a nonstationary signal buried in noise. The estimated signal itself is modeled utilizing the state–space formulation [1] describing its dynamical behavior. In summary, Kalman filtering deals with random processes described using state–space modeling which generate signals that can be measured and processed utilizing time recursive estimation formulas. The presentation here is brief and addresses the case of signals and noises represented in vector form; for more details in this subject the reader can consult many books available presenting Kalman filtering, including [3, 5]. There are many different ways to describe the Kalman filtering problem, and to derive its corresponding relations, here we follow the presentations of [2, 6].

17.2 State–Space Model

A convenient form of representing some dynamic systems is through what is called the state–space representation [1]. In such description, the outputs of the memory elements are considered as the system states. The state signals are collected in a vector denoted as $\mathbf{x}(k)$ which are in turn generated from its previous state $\mathbf{x}(k - 1)$ and from an external signal vector denoted as $\mathbf{n}(k)$. The observed or measured signals are collected in another vector denoted as $\mathbf{y}(k)$ whose elements originate from linear combinations of the previous state variables and of external signals represented in $\mathbf{n}_1(k)$. If we know the values of the external signals $\mathbf{n}(k)$ and $\mathbf{n}_1(k)$, we can determine the current values of the system states, which will be the delay inputs, and the system observation vector as follows:

$$\begin{cases} \mathbf{x}(k) = \mathbf{A}(k - 1)\mathbf{x}(k - 1) + \mathbf{B}(k)\mathbf{n}(k) \\ \mathbf{y}(k) = \mathbf{C}^T(k)\mathbf{x}(k - 1) + \mathbf{D}(k)\mathbf{n}_1(k) \end{cases} \quad (17.1)$$

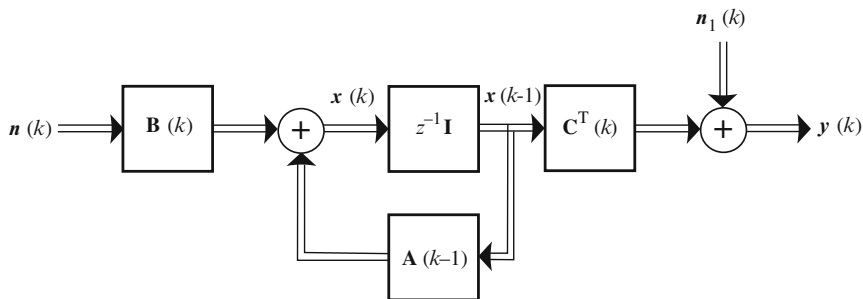


Fig. 17.1 State-space model for Kalman filtering formulation

where $\mathbf{x}(k)$ is the $(N + 1) \times 1$ vector of the state variables. If M is the number of system inputs and L is the number of system outputs, we then have that $\mathbf{A}(k - 1)$ is $(N + 1) \times (N + 1)$, $\mathbf{B}(k)$ is $(N + 1) \times M$, $\mathbf{C}(k)$ is $(N + 1) \times L$, and $\mathbf{D}(k)$ is $L \times L$.¹

Figure 17.1 shows the state–space system which generates the observation vector $\mathbf{y}(k)$ having as inputs the noise vectors $\mathbf{n}(k)$ and $\mathbf{n}_1(k)$, where the state variables $\mathbf{x}(k)$ are processes generated with excitation noise $\mathbf{n}(k)$.

The recursive solution of (17.1) can be described as

$$\mathbf{x}(k) = \prod_{l=0}^{k-1} \mathbf{A}(l)\mathbf{x}(0) + \sum_{i=1}^k \left[\prod_{l=i}^{k-1} \mathbf{A}(l) \right] \mathbf{B}(i)\mathbf{n}(i) \tag{17.2}$$

where $\prod_{l=k}^{k-1} \mathbf{A}(l) = 1$.

17.2.1 Simple Example

Let’s describe a particular example where we assume the signal $x(k)$ is a sample of an autoregressive process generated from the output of a system described by a linear difference equation given by

$$x(k) = \sum_{i=1}^{N+1} -a_i(k - 1)x(k - i) + n(k) \tag{17.3}$$

where $n(k)$ is a white noise. The coefficients $a_i(k - 1)$, for $i = 1, 2, \dots, N + 1$, are the time-varying parameters of the AR process. As part of the Kalman filtering procedure is the estimation of $x(k)$ from noisy measurements denoted as $y_l(k)$ for $l = 1, 2, \dots, L$.

¹In standard state–space formulation the matrix $\mathbf{D}(k)$ represents a feedforward connection between the input and the output of the dynamic system, in this discussion this matrix is not a feedforward matrix and is considered to be identity.

We can collect a sequence of signals to be estimated and noise measurements in vector forms as

$$\mathbf{x}(k) = \begin{bmatrix} x(k) \\ x(k-1) \\ \vdots \\ x(k-N) \end{bmatrix}$$

$$\mathbf{y}(k) = \begin{bmatrix} y_1(k) \\ y_2(k) \\ \vdots \\ y_L(k) \end{bmatrix} \quad (17.4)$$

where L represents the number of observations collected in $\mathbf{y}(k)$.

Each entry of the observation vector is considered to be generated through the following model:

$$y_l(k) = \mathbf{c}_l^T(k)\mathbf{x}(k-1) + n_{1,l}(k) \quad (17.5)$$

where $n_{1,l}(k)$ for $l = 1, 2, \dots, L$ are also white noises uncorrelated with each other and with $n(k)$.

Applying the state-space formulation to the particular set of (17.3) and (17.5) leads to a block of state variables originating from an autoregressive process described by

$$\mathbf{x}(k) = \begin{bmatrix} x(k) \\ x(k-1) \\ \vdots \\ x(k-N) \end{bmatrix}$$

$$= \begin{bmatrix} -a_1(k-1) & -a_2(k-1) & \cdots & -a_N(k-1) & -a_{N+1}(k-1) \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} x(k-1) \\ x(k-2) \\ \vdots \\ x(k-N-1) \end{bmatrix}$$

$$+ \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} n(k)$$

$$\mathbf{y}(k) = \begin{bmatrix} \mathbf{c}_1^T(k) \\ \mathbf{c}_2^T(k) \\ \vdots \\ \mathbf{c}_L^T(k) \end{bmatrix} \begin{bmatrix} x(k-1) \\ x(k-2) \\ \vdots \\ x(k-N-1) \end{bmatrix} + \mathbf{n}_1(k) \quad (17.6)$$

where for this case of single-input and multiple-output system $\mathbf{B}(k)$ is $(N + 1) \times M$ whose only nonzero element is the entry $(1, 1)$ that equals one, $\mathbf{C}(k)$ is $(N + 1) \times L$, and $\mathbf{D}(k)$ is just an identity matrix since the measurement noise contributes to the elements of the observation vector in an uncoupled form.

17.3 Kalman Filtering

In the following discussion we derive the Kalman filter for the general state–space description of (17.1). For that it is assumed we know

$$\mathbf{R}_{n_1}(k) = E[\mathbf{n}_1(k)\mathbf{n}_1^T(k)] \quad (17.7)$$

$$\mathbf{R}_n(k) = E[\mathbf{n}(k)\mathbf{n}^T(k)] \quad (17.8)$$

$\mathbf{A}(k - 1)$ and $\mathbf{C}(k)$, and that $\mathbf{n}(k)$ and $\mathbf{n}_1(k)$ are zero-mean white processes and uncorrelated with each other.

By assuming that we have the measurements $\mathbf{y}(k)$ available and that we employ all the data available up to a given iteration, we seek the optimal estimate of the state vector $\mathbf{x}(k)$, denoted by $\hat{\mathbf{x}}(k|k)$. As justified along the Kalman filtering derivation, the optimal solution has the following general form:

$$\hat{\mathbf{x}}(k|k) = \mathbf{A}(k-1)\hat{\mathbf{x}}(k-1|k-1) + \mathbf{K}(k) [\mathbf{y}(k) - \mathbf{C}^T(k)\mathbf{A}(k-1)\hat{\mathbf{x}}(k-1|k-1)] \quad (17.9)$$

where $\mathbf{K}(k)$ is the $(N + 1) \times L$ matrix called Kalman gain. The reader can notice that:

- The term $\mathbf{A}(k - 1)\hat{\mathbf{x}}(k - 1|k - 1)$ tries to bring the contribution of the previous estimation of the state variable to the current one, as suggests the state–space equation (17.1).
- The term $[\mathbf{y}(k) - \mathbf{C}^T(k)\mathbf{A}(k - 1)\hat{\mathbf{x}}(k - 1|k - 1)]$ is a correction term consisting of the difference between the observation vector and its estimate given by $\mathbf{C}^T(k)\mathbf{A}(k - 1)\hat{\mathbf{x}}(k - 1|k - 1)$, which in turn is a function the previous state-variable estimate.
- The Kalman gain aims at filtering out estimation errors and noise so that the state variable gets the best possible correction term, which minimizes the MSE.

In order to derive the optimal solution for the Kalman gain, let's first consider two cases where the estimate of $\mathbf{x}(k)$ is computed using observation data available until iteration k and another until iteration $k - 1$, denoted by $\hat{\mathbf{x}}(k|k)$ and $\hat{\mathbf{x}}(k|k - 1)$, respectively. The estimation error vectors in these cases are defined by

$$\mathbf{e}(k|k) = \mathbf{x}(k) - \hat{\mathbf{x}}(k|k) \quad (17.10)$$

$$\mathbf{e}(k|k - 1) = \mathbf{x}(k) - \hat{\mathbf{x}}(k|k - 1) \quad (17.11)$$

These errors have covariance matrices defined as

$$\mathbf{R}_e(k|k) = E[\mathbf{e}(k|k)\mathbf{e}^T(k|k)] \quad (17.12)$$

$$\mathbf{R}_e(k|k-1) = E[\mathbf{e}(k|k-1)\mathbf{e}^T(k|k-1)] \quad (17.13)$$

Given an instant $k-1$ when the information $\hat{\mathbf{x}}(k-1|k-1)$ and $\mathbf{R}_e(k-1|k-1)$ are available, we first try to estimate $\hat{\mathbf{x}}(k|k-1)$ which does not require the current observation. Whenever a new observation $\mathbf{y}(k)$ is available, $\hat{\mathbf{x}}(k|k)$ is estimated.

According to (17.1), at a given iteration the actual state-space vector evolves as

$$\mathbf{x}(k) = \mathbf{A}(k-1)\mathbf{x}(k-1) + \mathbf{B}(k)\mathbf{n}(k) \quad (17.14)$$

Since the elements of $\mathbf{n}(k)$ are zero mean, a possible unbiased MSE estimate for $\mathbf{x}(k)$ is provided by

$$\hat{\mathbf{x}}(k|k-1) = \mathbf{A}(k-1)\hat{\mathbf{x}}(k-1|k-1) \quad (17.15)$$

since the previous estimate $\hat{\mathbf{x}}(k-1|k-1)$ is available and $\mathbf{A}(k-1)$ is assumed known.

As a result, the state-variable estimation error when the last available observation is related to iteration $k-1$ is given by

$$\begin{aligned} \mathbf{e}(k|k-1) &= \mathbf{x}(k) - \hat{\mathbf{x}}(k|k-1) \\ &= \mathbf{A}(k-1)\mathbf{x}(k-1) + \mathbf{B}(k)\mathbf{n}(k) - \mathbf{A}(k-1)\hat{\mathbf{x}}(k-1|k-1) \\ &= \mathbf{A}(k-1)\mathbf{e}(k-1|k-1) + \mathbf{B}(k)\mathbf{n}(k) \end{aligned} \quad (17.16)$$

Assuming that $E[\mathbf{e}(k-1|k-1)] = \mathbf{0}$, meaning that $\hat{\mathbf{x}}(k-1|k-1)$ is an unbiased estimate of $\mathbf{x}(k-1)$, and recalling that the elements of $\mathbf{n}(k)$ are white noise with zero mean, then it is possible to conclude that

$$E[\mathbf{e}(k|k-1)] = \mathbf{0} \quad (17.17)$$

so that $\hat{\mathbf{x}}(k|k-1)$ is also an unbiased estimate of $\mathbf{x}(k)$.

The covariance matrix of $\mathbf{e}(k|k-1)$ can be expressed as follows:

$$\begin{aligned} \mathbf{R}_e(k|k-1) &= E[\mathbf{e}(k|k-1)\mathbf{e}^T(k|k-1)] \\ &= \mathbf{A}(k-1)E[\mathbf{e}(k-1|k-1)\mathbf{e}^T(k-1|k-1)]\mathbf{A}^T(k-1) \\ &\quad + \mathbf{B}(k)E[\mathbf{n}(k)\mathbf{n}^T(k)]\mathbf{B}^T(k) \\ &= \mathbf{A}(k-1)\mathbf{R}_e(k-1|k-1)\mathbf{A}^T(k-1) + \mathbf{B}(k)\mathbf{R}_n(k)\mathbf{B}^T(k) \end{aligned} \quad (17.18)$$

The next step is to estimate $\hat{\mathbf{x}}(k|k)$ from $\hat{\mathbf{x}}(k|k-1)$. In this case we use a linear filtering of the most recent estimate of the state variable $\hat{\mathbf{x}}(k|k-1)$ properly combined with another linear filtered contribution of the most recent measurement vector $\mathbf{y}(k)$. The resulting estimation expression for $\hat{\mathbf{x}}(k|k)$ has the following form

$$\hat{\mathbf{x}}(k|k) = \tilde{\mathbf{K}}(k)\hat{\mathbf{x}}(k|k-1) + \mathbf{K}(k)\mathbf{y}(k) \quad (17.19)$$

The challenge now is to compute the optimal expressions for the linear filtering matrices $\tilde{\mathbf{K}}(k)$ and $\mathbf{K}(k)$.

The state-variable estimation error $\mathbf{e}(k|k)$ that includes the last available observation can then be described as

$$\mathbf{e}(k|k) = \mathbf{x}(k) - \tilde{\mathbf{K}}(k)\hat{\mathbf{x}}(k|k-1) - \mathbf{K}(k)\mathbf{y}(k) \quad (17.20)$$

This expression can be rewritten in a more convenient form by replacing $\hat{\mathbf{x}}(k|k-1)$ using the first relation of (17.16) and replacing $\mathbf{y}(k)$ by its state-space formulation of (17.6). The resulting relation is

$$\begin{aligned} \mathbf{e}(k|k) &= \mathbf{x}(k) + \tilde{\mathbf{K}}(k) [\mathbf{e}(k|k-1) - \mathbf{x}(k)] - \mathbf{K}(k) [\mathbf{C}^T(k)\mathbf{x}(k) + \mathbf{n}_1(k)] \\ &= [\mathbf{I} - \tilde{\mathbf{K}}(k) - \mathbf{K}(k)\mathbf{C}^T(k)] \mathbf{x}(k) + \tilde{\mathbf{K}}(k)\mathbf{e}(k|k-1) - \mathbf{K}(k)\mathbf{n}_1(k) \end{aligned} \quad (17.21)$$

We know that $E[\mathbf{n}_1(k)] = \mathbf{0}$ and that $E[\mathbf{e}(k|k-1)] = \mathbf{0}$ since $\hat{\mathbf{x}}(k|k-1)$ is an unbiased estimate of $\mathbf{x}(k)$. However, $\hat{\mathbf{x}}(k|k)$ should also be an unbiased estimate of $\mathbf{x}(k)$, that is, $E[\mathbf{e}(k|k)] = \mathbf{0}$. The latter relation is true if we choose

$$\tilde{\mathbf{K}}(k) = \mathbf{I} - \mathbf{K}(k)\mathbf{C}^T(k) \quad (17.22)$$

so that the first term in the last expression of (17.21) becomes zero.

By replacing (17.22) in (17.19), the estimate of the state variable using the current measurements becomes

$$\begin{aligned} \hat{\mathbf{x}}(k|k) &= [\mathbf{I} - \mathbf{K}(k)\mathbf{C}^T(k)] \hat{\mathbf{x}}(k|k-1) + \mathbf{K}(k)\mathbf{y}(k) \\ &= \hat{\mathbf{x}}(k|k-1) + \mathbf{K}(k) [\mathbf{y}(k) - \mathbf{C}^T(k)\hat{\mathbf{x}}(k|k-1)] \end{aligned} \quad (17.23)$$

where according to (17.21) and (17.22) the corresponding estimation error vector is described by

$$\begin{aligned} \mathbf{e}(k|k) &= [\mathbf{I} - \mathbf{K}(k)\mathbf{C}^T(k)] \mathbf{e}(k|k-1) - \mathbf{K}(k)\mathbf{n}_1(k) \\ &= \tilde{\mathbf{K}}(k)\mathbf{e}(k|k-1) - \mathbf{K}(k)\mathbf{n}_1(k) \end{aligned} \quad (17.24)$$

where the last equality highlights the connection with (17.19).

The covariance matrix of $\mathbf{e}(k|k)$ can then be expressed as

$$\begin{aligned}\mathbf{R}_e(k|k) &= E[\mathbf{e}(k|k)\mathbf{e}^T(k|k)] \\ &= [\mathbf{I} - \mathbf{K}(k)\mathbf{C}^T(k)]\mathbf{R}_e(k|k-1)[\mathbf{I} - \mathbf{K}(k)\mathbf{C}^T(k)]^T + \mathbf{K}(k)\mathbf{R}_{n_1}(k)\mathbf{K}^T(k) \\ &= [\mathbf{I} - \mathbf{K}(k)\mathbf{C}^T(k)]\mathbf{R}_e(k|k-1) \\ &\quad - \{[\mathbf{I} - \mathbf{K}(k)\mathbf{C}^T(k)]\mathbf{R}_e(k|k-1)\mathbf{C}(k) - \mathbf{K}(k)\mathbf{R}_{n_1}(k)\}\mathbf{K}^T(k) \quad (17.25)\end{aligned}$$

The trace of this covariance matrix determines how good is the estimate of the state variables at a given iteration. As a result, the Kalman gain should be designed in order to minimize the trace of $\mathbf{R}_e(k|k)$ shown as follows, since it corresponds to the estimation error variance. Defining

$$\xi_K = \text{tr}[\mathbf{R}_e(k|k)] \quad (17.26)$$

it then follows that²

$$\frac{\partial \xi_K}{\partial \mathbf{K}(k)} = -2[\mathbf{I} - \mathbf{K}(k)\mathbf{C}^T(k)]\mathbf{R}_e(k|k-1)\mathbf{C}(k) + 2\mathbf{K}(k)\mathbf{R}_{n_1}(k) \quad (17.27)$$

By equating this derivative with zero it is possible to simplify (17.25) since its last term becomes zero, allowing the update to the covariance matrix to have a rather simple form given by

$$\mathbf{R}_e(k|k) = [\mathbf{I} - \mathbf{K}(k)\mathbf{C}^T(k)]\mathbf{R}_e(k|k-1) \quad (17.28)$$

The main purpose of (17.27) is of course to calculate the Kalman gain whose expression is given by

$$\mathbf{K}(k) = \mathbf{R}_e(k|k-1)\mathbf{C}(k)[\mathbf{C}^T(k)\mathbf{R}_e(k|k-1)\mathbf{C}(k) + \mathbf{R}_{n_1}(k)]^{-1} \quad (17.29)$$

Now we have all the expressions required to describe the Kalman filtering algorithm. First we should initialize $\hat{\mathbf{x}}(0|0)$ with $\mathbf{x}(0)$ if available, otherwise generate a zero-mean white Gaussian noise vector. Then initialize the error covariance matrix as $\mathbf{R}_e(0|0) = \mathbf{x}(0)\mathbf{x}^T(0)$. After initialization the algorithm computes $\hat{\mathbf{x}}(k|k-1)$ as per (17.15) then the error covariance $\mathbf{R}_e(k|k-1)$ using (17.18). Next we calculate the Kalman gain as in (17.29) and update the estimate $\hat{\mathbf{x}}(k|k)$ using (17.23) which now takes the form

$$\begin{aligned}\hat{\mathbf{x}}(k|k) &= \hat{\mathbf{x}}(k|k-1) + \mathbf{K}(k)[\mathbf{y}(k) - \mathbf{C}^T(k)\hat{\mathbf{x}}(k|k-1)] \\ &= \hat{\mathbf{x}}(k|k-1) + \mathbf{K}(k)[\mathbf{y}(k) - \hat{\mathbf{y}}(k|k-1)] \quad (17.30)\end{aligned}$$

²It was used the facts that $\frac{\partial \text{tr}[\mathbf{AB}]}{\partial \mathbf{A}} = \mathbf{B}^T$ and $\frac{\partial \text{tr}[\mathbf{ABA}^T]}{\partial \mathbf{A}} = 2\mathbf{AB}$, and that $\mathbf{R}_e(k|k-1)$ and $\mathbf{R}_{n_1}(k)$ are symmetric matrices.

Algorithm 17.1 Kalman filter

Initialization

$$\hat{\mathbf{x}}(0|0) = \mathbf{x}(0) \quad \mathbf{R}_e(0|0) = \mathbf{x}(0)\mathbf{x}^T(0)$$

Do for $k \geq 1$

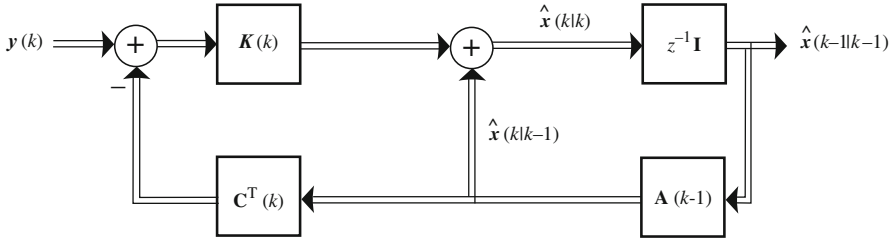
$$\hat{\mathbf{x}}(k|k-1) = \mathbf{A}(k-1)\hat{\mathbf{x}}(k-1|k-1)$$

$$\mathbf{R}_e(k|k-1) = \mathbf{A}(k-1)\mathbf{R}_e(k-1|k-1)\mathbf{A}^T(k-1) + \mathbf{B}(k)\mathbf{R}_{n_1}(k)\mathbf{B}^T(k)$$

$$\mathbf{K}(k) = \mathbf{R}_e(k|k-1)\mathbf{C}^T(k) [\mathbf{C}^T(k)\mathbf{R}_e(k|k-1)\mathbf{C}^T(k) + \mathbf{R}_{n_1}(k)]^{-1}$$

$$\hat{\mathbf{x}}(k|k) = \hat{\mathbf{x}}(k|k-1) + \mathbf{K}(k)(\mathbf{y}(k) - \mathbf{C}^T(k)\hat{\mathbf{x}}(k|k-1))$$

$$\mathbf{R}_e(k|k) = [\mathbf{I} - \mathbf{K}(k)\mathbf{C}^T(k)]\mathbf{R}_e(k|k-1)$$

**Fig. 17.2** Kalman filtering structure

where in the first expression we used (17.15), and in the second expression we observe that the term $\mathbf{C}^T(k)\hat{\mathbf{x}}(k|k-1)$ represents an unbiased estimate of $\mathbf{y}(k)$ denoted as $\hat{\mathbf{y}}(k|k-1)$. Finally (17.28) updates the error covariance $\mathbf{R}_e(k|k)$ to include the current measurement contribution. Algorithm 17.1 describes the Kalman filtering procedure. Figure 17.2 illustrates how the building blocks of the Kalman filtering algorithm interact among themselves. As can be observed, from the measurement signal $\mathbf{y}(k)$ we perform the best possible estimate of the state variable $\hat{\mathbf{x}}(k|k)$. The Kalman filter solution corresponds to the optimal minimum MSE estimator whenever the noise and the state signal are jointly Gaussian, otherwise it is the optimal linear minimum MSE solution, see [5] for details.

The complex version of the Kalman filter algorithm is almost identical to Algorithm 17.1 and can be derived by replacing $\mathbf{x}^T(0)$ by $\mathbf{x}^H(0)$, $\mathbf{C}^T(k)$ by $\mathbf{C}^H(k)$, and $\mathbf{A}^T(k-1)$ by $\mathbf{A}^H(k-1)$.

Example 17.1. In a nonstationary environment the optimal coefficient vector is described by

$$w_o(k) = 0.9w_o(k-1) - 0.81w_o(k-2) + n_w(k)$$

for $k \geq 1$, where $n_w(k)$ is a zero-mean Gaussian white processes with variance 0.64. Assume $w_o(0) = w_o(-1) = 0$.

Assume this time-varying coefficient is observed through a noisy measurement described by

$$y(k) = 0.9w_o(k) + n_1(k)$$

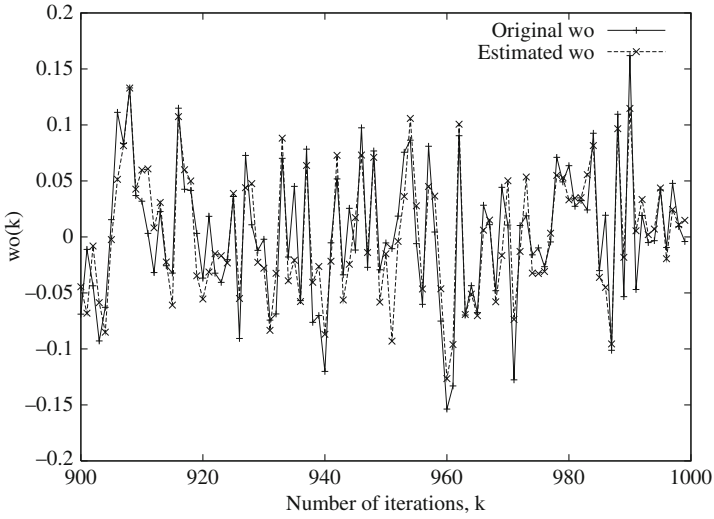


Fig. 17.3 Tracking performance of the Kalman filter

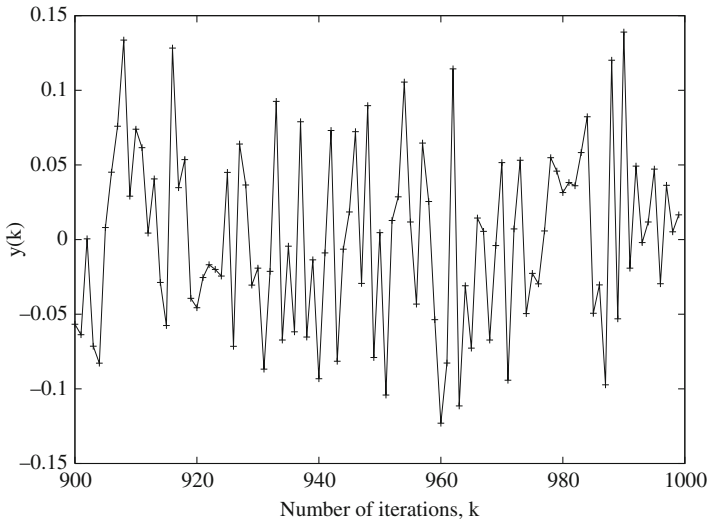


Fig. 17.4 Noisy measurement signal

where $n_1(k)$ is another zero-mean Gaussian white processes with variance 0.16.

Run the Kalman filter algorithm to estimate $w_o(k)$ from $y(k)$. Plot $w_o(k)$, its estimate $\hat{w}_o(k)$ and $y(k)$.

Solution. The results presented correspond to the average of 200 independent runs of the Kalman filter algorithm. Figure 17.3 shows the signal $w_o(k)$ being tracked

by its estimate $\hat{w}_o(k)$ from iteration 900 to 1,000, whereas Fig. 17.4 illustrates the measurement signal $y(k)$ from where $\hat{w}_o(k)$ was computed. As can be observed, the Kalman filter algorithm is able to track quite closely the signal $w_o(k)$ from noisy measurements given by $y(k)$. \square

17.4 Kalman Filter and RLS

As observed in the previous section, the Kalman filtering formulation requires the knowledge of the state–space model generating the observation vector. Such information is not available in a number of adaptive-filtering setups but is quite common in problems related to tracking targets, positioning of dynamic systems, and prediction and estimation of time-varying phenomena, just to mention a few. However, a proper analysis of the Kalman filtering setup allows us to disclose some links with the RLS algorithms. These links are the subject of this section.

Let's start by observing that in the RLS context one tries to estimate the unknown system parameters denoted as $\mathbf{w}_o(k)$ through the adaptive-filtering coefficients $\mathbf{w}(k)$. The equivalent operation in Kalman filtering is the estimation of $\mathbf{x}(k)$ given by $\hat{\mathbf{x}}(k|k)$. The reference signal in the RLS case is $d(k)$ corresponding to the scalar version of $\mathbf{y}(k)$ denoted as $y(k)$ in the Kalman case. The estimate of $y(k)$ is given by $\hat{y}(k|k-1) = \mathbf{c}^T(k)\hat{\mathbf{x}}(k|k-1)$ since matrix $\mathbf{C}^T(k)$ is a row vector in the single output case. As such, it is easy to infer that $\hat{y}(k|k-1)$ corresponds to the adaptive-filter output denoted as $y(k)$ in the RLS case.

Equation (5.9) repeated here for convenience

$$\mathbf{w}(k) = \mathbf{w}(k-1) + e(k)\mathbf{S}_D(k)\mathbf{x}(k) \quad (17.31)$$

is meant for coefficient update in the RLS algorithms. This equation is equivalent to

$$\begin{aligned} \hat{\mathbf{x}}(k|k) &= \hat{\mathbf{x}}(k|k-1) + \mathbf{k}(k) (y(k) - \mathbf{c}^T(k)\hat{\mathbf{x}}(k|k-1)) \\ &= \hat{\mathbf{x}}(k|k-1) + \mathbf{k}(k) (y(k) - \hat{y}(k|k-1)) \\ &= \hat{\mathbf{x}}(k|k-1) + \mathbf{k}(k)e_y(k) \end{aligned} \quad (17.32)$$

where $e_y(k)$ is an a priori error in the estimate of $y(k)$. It can be observed that the Kalman gain matrix $\mathbf{K}(k)$ becomes a vector denoted as $\mathbf{k}(k)$. By comparing (17.32) with (17.31), we can infer that $\mathbf{k}(k)$ is equivalent to $\mathbf{S}_D(k)\mathbf{x}(k)$.

The updating of the Kalman gain in the scalar output case is given by

$$\mathbf{k}(k) = \mathbf{R}_e(k|k-1)\mathbf{c}(k) [\mathbf{c}^T(k)\mathbf{R}_e(k|k-1)\mathbf{c}(k) + r_{n_1}(k)]^{-1} \quad (17.33)$$

where $r_{n_1}(k)$ is the additional noise variance. Again by comparing (17.32) with (5.5), we can infer that $\mathbf{k}(k)$ is equivalent to

$$\begin{aligned} \mathbf{S}_D(k)\mathbf{x}(k) &= \frac{1}{\lambda} \left[\mathbf{S}_D(k-1) - \frac{\mathbf{S}_D(k-1)\mathbf{x}(k)\mathbf{x}^T(k)\mathbf{S}_D(k-1)}{\lambda + \mathbf{x}^T(k)\mathbf{S}_D(k-1)\mathbf{x}(k)} \right] \mathbf{x}(k) \\ &= \frac{\mathbf{S}_D(k-1)\mathbf{x}(k)}{\lambda + \mathbf{x}^T(k)\mathbf{S}_D(k-1)\mathbf{x}(k)} \\ &= \frac{\frac{1}{\lambda}\mathbf{S}_D(k-1)\mathbf{x}(k)}{1 + \frac{1}{\lambda}\mathbf{x}^T(k)\mathbf{S}_D(k-1)\mathbf{x}(k)} \end{aligned} \quad (17.34)$$

Now if we assume that the measurement noise in (17.33) has unit variance, it is straightforward to observe by comparing (17.33) and (17.34) that $\mathbf{R}_e(k|k-1)$ plays the role of $\frac{1}{\lambda}\mathbf{S}_D(k-1)$ in the RLS algorithm.

The related quantities in the specialized Kalman filter and the RLS algorithm disclosed so far are

$$\begin{aligned} \mathbf{x}(k) &\iff \mathbf{w}_o(k) \\ y(k) &\iff d(k) \\ \hat{y}(k|k-1) &\iff y(k) \\ \hat{\mathbf{x}}(k|k) &\iff \mathbf{w}(k) \\ e_y(k) &\iff e(k) \\ \mathbf{k}(k) &\iff \mathbf{S}_D(k)\mathbf{x}(k) \\ \mathbf{R}_e(k|k-1) &\iff \frac{1}{\lambda}\mathbf{S}_D(k-1) \end{aligned} \quad (17.35)$$

These relations show that given that $\mathbf{x}(k)$ in the Kalman filter algorithm follows the pattern of $\mathbf{w}_o(k)$ and $r_{n_1}(k)$ has unit variance (compare (17.33) and (17.34)), the Kalman filter and the RLS algorithms should lead to similar solutions.

As happens with the conventional RLS algorithm, the Kalman filter algorithm faces stability problems when implemented in finite precision mainly related to the ill-conditioning of the estimation error covariance matrix $\mathbf{R}_e(k|k)$. In practical implementations this matrix could be updated in a factorized form such as $\mathbf{U}_e(k|k)\mathbf{D}_e(k|k)\mathbf{U}_e^T(k|k)$, where $\mathbf{U}_e(k|k)$ is upper triangular with ones on the diagonal and $\mathbf{D}_e(k|k)$ is a diagonal matrix.

References

1. P.S.R. Diniz, E.A.B. da Silva, S.L. Netto, *Digital Signal Processing: System Analysis and Design*, 2nd edn. (Cambridge University Press, Cambridge, 2010)
2. M.H. Hayes, *Statistical Digital Signal Processing and Modeling* (Wiley, New York, 1996)

3. T. Kailath, A.H. Sayed, B. Hassibi, *Linear Estimation* (Prentice Hall, Englewood Cliffs, 2000)
4. R.E. Kalman, A new approach to linear filtering and prediction problem. *Trans. ASME J. Basic Eng.* **82**, 34–45 (1960)
5. S.M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory* (Prentice Hall, Englewood Cliffs, 1993)
6. D.G. Manolakis, V.K. Ingle, S.M. Kogon, *Statistical and Adaptive Signal Processing* (McGraw Hill, New York, 2000)