

CHAPTER 9

Optimal smoothing

In a *post mortem* (after the fact) analysis, it is possible to wait for more observations to accumulate. In that case, the estimate can be improved by smoothing.

—Andrew Jazwinski [Jaz70, p. 143]

In previous chapters, we discussed how to obtain the optimal *a priori* and *a posteriori* state estimates. The *a priori* state estimate at time k , \hat{x}_k^- , is the state estimate at time k based on all the measurements up to (but not including) time k . The *a posteriori* state estimate at time k , \hat{x}_k^+ , is the state estimate at time k based on all the measurements up to and including time k :

$$\begin{aligned}\hat{x}_k^- &= E(x_k | y_1, \dots, y_{k-1}) \\ \hat{x}_k^+ &= E(x_k | y_1, \dots, y_k)\end{aligned}\tag{9.1}$$

There are often situations in which we want to obtain other types of state estimates. We will define $\hat{x}_{k,j}$ as the estimate of x_k given all measurements up to and including time j . With this notation, we see that

$$\begin{aligned}\hat{x}_{k,k-1} &= \hat{x}_k^- \\ \hat{x}_{k,k} &= \hat{x}_k^+\end{aligned}\tag{9.2}$$

Now suppose, for example, that we have recorded measurements up to time index 54 and we want to obtain an estimate of the state at time index 33. Our theory in

the previous chapters tells us how to obtain \hat{x}_{33}^- or \hat{x}_{33}^+ , but those estimates only use the measurements up to and including times 32 and 33, respectively. If we have more measurements (e.g., measurements up to time 54) it stands to reason that we should be able to get an even better estimate of x_{33} . This chapter discusses some ways of obtaining better estimates.

In another scenario, it may be that we are interested in obtaining an estimate of the state at a fixed time j . As measurements keep rolling in, we want to keep updating our estimate \hat{x}_j . In other words, we want to obtain $\hat{x}_{j,j+1}$, $\hat{x}_{j,j+2}$, \dots . This could be the case, for example, if a satellite takes a picture at time j . In order to more accurately process the photograph at time j we need an estimate of the satellite state (position and velocity) at time j . As the satellite continues to orbit, we may obtain additional range measurements of the satellite, so we can continue to update the estimate of x_j and thus improve the quality of the processed photograph. This situation is called fixed-point smoothing because the time point for which we want to obtain a state estimate (time j in this example) is fixed, but the number of measurements that are available to improve that estimate continually changes. Fixed-point smoothing is depicted in Figure 9.1 and is discussed in Section 9.2.

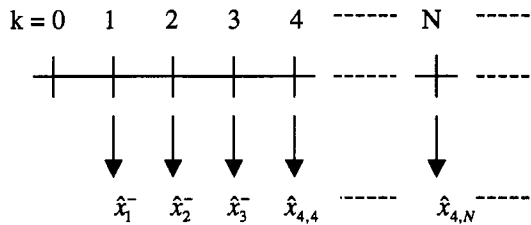


Figure 9.1 Fixed-point smoothing. We desire an estimate of x_4 . Up until $k = 4$, the standard Kalman filter operates. At $k = 4$, we have $\hat{x}_4^- = \hat{x}_{4,4}$, which is the estimate of x_4 based on measurements up to and including y_3 . As time progresses, we continue to refine our estimate of x_4 based on an increasing number of measurements. At time $k = N$, we have $\hat{x}_{4,N}$, which is the estimate of x_4 based on measurements up to and including time $N - 1$.

Another type of smoothing is fixed-lag smoothing. In this situation, we want to obtain an estimate of the state at time $(k - N)$ given measurements up to and including time k , where the time index k continually changes as we obtain new measurements, but the lag N is a constant. In other words, at each time point we have N future measurements available for our state estimate. We therefore want to obtain $\hat{x}_{k-N,k}$ for $k = N, N + 1, \dots$, where N is a fixed positive integer. This could be the case, for example, if a satellite is continually taking photographs that are to be displayed or transmitted N time steps after the photograph is taken. In this case, since the photograph is processed N time steps after it is taken, we have N additional measurements after each photograph that are available to update the estimate of the satellite state and hence improve the quality of the photograph. Fixed-lag smoothing is depicted in Figure 9.2 and is discussed in Section 9.3.

The final type of smoothing is fixed-interval smoothing. In this situation, we have a fixed interval of measurements (y_1, y_2, \dots, y_M) that are available, and we want to obtain the optimal state estimates at all the times in that interval. For each state estimate we want to use all of the measurements in the time interval. That is, we want to obtain $\hat{x}_{0,M}$, $\hat{x}_{1,M}$, \dots , $\hat{x}_{M,M}$. This is the case when we have recorded

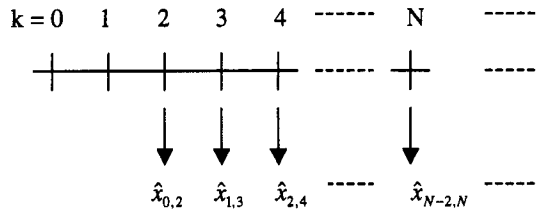


Figure 9.2 Fixed-lag smoothing. We desire an estimate of the state at each time step based on measurements two time steps ahead. After processing y_2 , we form the estimate $\hat{x}_{0,2}$, which is the estimate of x_0 based on measurements up to and including y_2 . Similarly, $\hat{x}_{1,3}$ is the estimate of x_1 based on measurements up to and including y_3 .

some data that are available for post-processing. For example, if a manufacturing process has run over the weekend and we have recorded all of the data, and now we want to plot a time history of the best estimate of the process state, we can use all of the recorded data to estimate the states at each of the time points. Fixed-interval smoothing is depicted in Figure 9.3 and is discussed in Section 9.4.

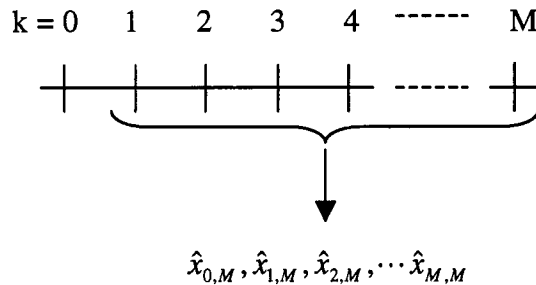


Figure 9.3 Fixed-interval smoothing. We desire an estimate of the state at each time step based on all of the measurements in some interval. After processing all of the measurements from y_1 to y_M , we form the estimate $\hat{x}_{0,M}$, which is the estimate of x_0 based on all the measurements. Similarly, $\hat{x}_{1,M}$ is the estimate of x_1 based on all the measurements.

Our derivation of these optimal smoothers will be based on a form for the Kalman filter different than we have seen in previous chapters. Therefore, before we can discuss the optimal smoothers, we will first present an alternate Kalman filter form in Section 9.1.

9.1 AN ALTERNATE FORM FOR THE KALMAN FILTER

In order to put ourselves in position to derive optimal smoothers, we first need to derive yet another form for the Kalman filter. This is the form presented in [And79]. The equations describing the system and the Kalman filter were derived in Section 5.1 as follows:

$$\begin{aligned}
x_k &= F_{k-1}x_{k-1} + w_{k-1} \\
y_k &= H_k x_k + v_k \\
P_{k+1}^- &= F_k P_k^+ F_k^T + Q_k \\
K_k &= P_k^- H_k^T (H_k P_k^- H_k^T + R_k)^{-1} \\
P_k^+ &= (I - K_k H_k) P_k^- (I - K_k H_k)^T + K_k R_k K_k^T \\
\hat{x}_k^- &= F_{k-1} \hat{x}_{k-1}^+ \\
\hat{x}_k^+ &= \hat{x}_k^- + K_k (y_k - H_k \hat{x}_k^-)
\end{aligned} \tag{9.3}$$

Now if we define L_k as

$$\begin{aligned}
L_k &= F_k K_k \\
&= F_k P_k^- H_k^T (H_k P_k^- H_k^T + R_k)^{-1}
\end{aligned} \tag{9.4}$$

and substitute the expression for \hat{x}_k^+ into the expression for \hat{x}_{k+1}^- , then we obtain

$$\begin{aligned}
\hat{x}_{k+1}^- &= F_k \hat{x}_k^- + F_k K_k (y_k - H_k \hat{x}_k^-) \\
&= F_k \hat{x}_k^- + L_k (y_k - H_k \hat{x}_k^-)
\end{aligned} \tag{9.5}$$

Expanding the expression for P_k^+ gives

$$P_k^+ = P_k^- - K_k H_k P_k^- - P_k^- H_k^T K_k^T + K_k H_k P_k^- H_k^T K_k^T + K_k R_k K_k^T \tag{9.6}$$

Substituting for K_k gives

$$\begin{aligned}
P_k^+ &= P_k^- - P_k^- H_k^T (H_k P_k^- H_k^T + R_k)^{-1} H_k P_k^- - \\
&\quad P_k^- H_k^T (H_k P_k^- H_k^T + R_k)^{-1} H_k P_k^- + \\
&\quad P_k^- H_k^T (H_k P_k^- H_k^T + R_k)^{-1} H_k P_k^- H_k^T (H_k P_k^- H_k^T + R_k)^{-1} H_k P_k^- + \\
&\quad P_k^- H_k^T (H_k P_k^- H_k^T + R_k)^{-1} R_k (H_k P_k^- H_k^T + R_k)^{-1} H_k P_k^-
\end{aligned} \tag{9.7}$$

Performing some factoring and collection of like terms on this equation gives

$$\begin{aligned}
P_k^+ &= P_k^- + P_k^- H_k^T [-(H_k P_k^- H_k^T + R_k)^{-1} - (H_k P_k^- H_k^T + R_k)^{-1} + \\
&\quad (H_k P_k^- H_k^T + R_k)^{-1} H_k P_k^- H_k^T (H_k P_k^- H_k^T + R_k)^{-1} + \\
&\quad (H_k P_k^- H_k^T + R_k)^{-1} R_k (H_k P_k^- H_k^T + R_k)^{-1}] H_k P_k^- \\
&= P_k^- + P_k^- H_k^T [-(H_k P_k^- H_k^T + R_k)^{-1} - (H_k P_k^- H_k^T + R_k)^{-1} + \\
&\quad (H_k P_k^- H_k^T + R_k)^{-1} (H_k P_k^- H_k^T + R_k) (H_k P_k^- H_k^T + R_k)^{-1}] H_k P_k^- \\
&= P_k^- - P_k^- H_k^T (H_k P_k^- H_k^T + R_k)^{-1} H_k P_k^-
\end{aligned} \tag{9.8}$$

Substituting this expression for P_k^+ into the expression for P_{k+1}^- gives

$$\begin{aligned}
P_{k+1}^- &= F_k P_k^+ F_k^T + Q_k \\
&= F_k [P_k^- - P_k^- H_k^T (H_k P_k^- H_k^T + R_k)^{-1} H_k P_k^-] F_k^T + Q_k \\
&= F_k P_k^- (F_k - L_k H_k)^T + Q_k
\end{aligned} \tag{9.9}$$

Combining Equations (9.4), (9.5), and (9.9) gives the alternate form for the one-step *a priori* Kalman filter, which can be summarized as follows:

$$\begin{aligned} L_k &= F_k P_k^- H_k^T (H_k P_k^- H_k^T + R_k)^{-1} \\ P_{k+1}^- &= F_k P_k^- (F_k - L_k H_k)^T + Q_k \\ \hat{x}_{k+1}^- &= F_k \hat{x}_k^- + L_k (y_k - H_k \hat{x}_k^-) \end{aligned} \quad (9.10)$$

where L_k is the redefined Kalman gain. This form of the filter obtains only *a priori* state estimates and covariances. Note that the Kalman gain, L_k , for this form of the filter is not the same as the Kalman gain, K_k , for the form of the filter that we derived in Section 5.1. However, the two forms result in identical state estimates and estimation-error covariances.

9.2 FIXED-POINT SMOOTHING

The objective in fixed-point smoothing is to obtain *a priori* state estimates of x_j at times $j+1, j+2, \dots, k, k+1, \dots$. We will use the notation $\hat{x}_{j,k}$ to refer to the estimate of x_j that is obtained by using all of the measurements up to and including time $(k-1)$. That is, $\hat{x}_{j,k}$ can be thought of as the *a priori* estimate of x_j at time k :

$$\hat{x}_{j,k} = E(x_j | y_1, \dots, y_{k-1}) \quad k \geq j \quad (9.11)$$

With this definition we see that

$$\begin{aligned} \hat{x}_{j,j} &= E(x_j | y_1, \dots, y_{j-1}) \\ &= \hat{x}_j^- \end{aligned} \quad (9.12)$$

In other words, $\hat{x}_{j,j}$ is just the normal *a priori* state estimate at time j that we derived in Section 5.1. We also see that

$$\begin{aligned} \hat{x}_{j,j+1} &= E(x_j | y_1, \dots, y_j) \\ &= \hat{x}_j^+ \end{aligned} \quad (9.13)$$

In other words, $\hat{x}_{j,j+1}$ is just the normal *a posteriori* state estimate at time j that we derived in Section 5.1. The question addressed by fixed-point smoothing is as follows: When we get the next measurement at time $(j+1)$, how can we incorporate that information to obtain an improved estimate (along with its covariance) for the state at time j ? Furthermore, when we get additional measurements at times $(j+2), (j+3)$, etc., how can we incorporate that information to obtain an improved estimate (along with its covariance) for the state at time j ?

In order to derive the fixed-point smoother, we will define a new state variable x' . This new state variable will be initialized as $x'_j = x_j$, and will have the dynamics $x'_{k+1} = x'_k$ ($k = j, j+1, \dots$). With this definition, we see that $x'_k = x_j$ for all $k \geq j$. So if we can use the standard Kalman filter to find the *a priori* estimate of x'_k then we will, by definition, have a smoothed estimate of x_j given measurements up to and including time $(k-1)$. In other words, the *a priori* estimate of x'_k will be equal to $\hat{x}_{j,k}$. This idea is depicted in Figure 9.4.

Our original system is given as

$$\begin{aligned} x_k &= F_{k-1} x_{k-1} + w_{k-1} \\ y_k &= H_k x_k + v_k \end{aligned} \quad (9.14)$$

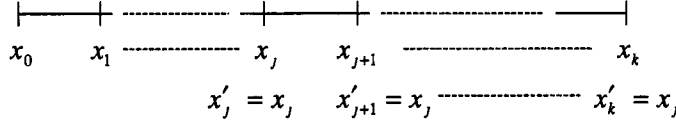


Figure 9.4 This illustrates the idea that is used to obtain the fixed-point smoother. A fictitious state variable x' is initialized as $x'_j = x_j$ and from that point on has an identity state transition matrix. The *a priori* estimate of x'_k is then equal to $\hat{x}_{j,k}$.

Augmenting the dynamics of our newly defined state x' to the original system results in the following:

$$\begin{aligned} \begin{bmatrix} x_k \\ x'_k \end{bmatrix} &= \begin{bmatrix} F_{k-1} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} x_{k-1} \\ x'_{k-1} \end{bmatrix} + \begin{bmatrix} I \\ 0 \end{bmatrix} w_{k-1} \\ y_k &= \begin{bmatrix} H_k & 0 \end{bmatrix} \begin{bmatrix} x_k \\ x'_k \end{bmatrix} + v_k \end{aligned} \quad (9.15)$$

If we use a standard Kalman filter to obtain an *a priori* estimate of the augmented state, the covariance of the estimation error can be written as

$$E \left[\begin{pmatrix} x_k - \hat{x}_k^- \\ x_j - \hat{x}_{j,k} \end{pmatrix} \begin{pmatrix} (x_k - \hat{x}_k^-)^T & (x_j - \hat{x}_{j,k})^T \end{pmatrix} \right] = \begin{bmatrix} P_k & \Sigma_k^T \\ \Sigma_k & \Pi_k \end{bmatrix} \quad (9.16)$$

The covariance P_k above is the normal *a priori* covariance of the estimate of x_k . We have dropped the minus superscript for ease of notation, and we will also feel free to drop the minus superscript on all other quantities in this section with the understanding that all estimates and covariances are *a priori*. The Σ_k and Π_k matrices are defined by the above equation. Note that at time $k = j$, Σ_k and Π_k are given as

$$\begin{aligned} \Sigma_j &= E[(x_j - \hat{x}_{j,j})(x_j - \hat{x}_j^-)^T] \\ &= E[(x_j - \hat{x}_j^-)(x_j - \hat{x}_j^-)^T] \\ &= P_j \\ \Pi_j &= E[(x_j - \hat{x}_{j,j})(x_j - \hat{x}_{j,j})^T] \\ &= E[(x_j - \hat{x}_j^-)(x_j - \hat{x}_j^-)^T] \\ &= P_j \end{aligned} \quad (9.17)$$

The Kalman filter summarized in Equation (9.10) can be written for the augmented system as follows:

$$\begin{bmatrix} \hat{x}_{k+1}^- \\ \hat{x}_{j,k+1}^- \end{bmatrix} = \begin{bmatrix} F_{k-1} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} \hat{x}_k^- \\ \hat{x}_{j,k}^- \end{bmatrix} + \begin{bmatrix} L_k \\ \lambda_k \end{bmatrix} \left(y_k - \begin{bmatrix} H_k & 0 \end{bmatrix} \begin{bmatrix} \hat{x}_k^- \\ \hat{x}_{j,k}^- \end{bmatrix} \right) \quad (9.18)$$

where L_k is the normal Kalman filter gain given in Equation (9.10), and λ_k is the additional part of the Kalman gain, which will be determined later in this section. Writing Equation (9.18) as two separate equations gives

$$\begin{aligned} \hat{x}_{k+1}^- &= F_{k-1} \hat{x}_k^- + L_k (y_k - H_k \hat{x}_k^-) \\ \hat{x}_{j,k+1}^- &= \hat{x}_{j,k}^- + \lambda_k (y_k - H_k \hat{x}_k^-) \end{aligned} \quad (9.19)$$

The Kalman gain can be written from Equation (9.10) as follows:

$$\begin{aligned} \begin{bmatrix} L_k \\ \lambda_k \end{bmatrix} &= \begin{bmatrix} F_{k-1} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} P_k & \Sigma_k^T \\ \Sigma_k & \Pi_k \end{bmatrix} \begin{bmatrix} H_k^T \\ 0 \end{bmatrix} \times \\ &\quad \left(\begin{bmatrix} H_k & 0 \end{bmatrix} \begin{bmatrix} P_k & \Sigma_k^T \\ \Sigma_k & \Pi_k \end{bmatrix} \begin{bmatrix} H_k^T \\ 0 \end{bmatrix} + R_k \right)^{-1} \\ &= \begin{bmatrix} F_{k-1} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} P_k & \Sigma_k^T \\ \Sigma_k & \Pi_k \end{bmatrix} \begin{bmatrix} H_k^T \\ 0 \end{bmatrix} (H_k P_k H_k^T + R_k)^{-1} \end{aligned} \quad (9.20)$$

Writing this equation as two separate equations gives

$$\begin{aligned} L_k &= F_k P_k H_k^T (H_k P_k H_k^T + R_k)^{-1} \\ \lambda_k &= \Sigma_k H_k^T (H_k P_k H_k^T + R_k)^{-1} \end{aligned} \quad (9.21)$$

The Kalman filter estimation-error covariance-update equation can be written from Equation (9.10) as follows:

$$\begin{aligned} \begin{bmatrix} P_{k+1} & \Sigma_{k+1}^T \\ \Sigma_{k+1} & \Pi_{k+1} \end{bmatrix} &= \begin{bmatrix} F_k & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} P_k & \Sigma_k^T \\ \Sigma_k & \Pi_k \end{bmatrix} \times \\ &\quad \left(\begin{bmatrix} F_k^T & 0 \\ 0 & I \end{bmatrix} - \begin{bmatrix} H_k^T \\ 0 \end{bmatrix} \begin{bmatrix} L_k^T & \lambda_k^T \end{bmatrix} \right) + \begin{bmatrix} Q_k & 0 \\ 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} F_k P_k F_k^T - F_k P_k H_k^T L_k^T & -F_k P_k H_k^T \lambda_k^T + F_k \Sigma_k^T \\ \Sigma_k F_k^T - \Sigma_k H_k^T L_k^T & -\Sigma_k H_k^T \lambda_k^T + \Pi_k \end{bmatrix} + \\ &\quad \begin{bmatrix} Q_k & 0 \\ 0 & 0 \end{bmatrix} \end{aligned} \quad (9.22)$$

Writing this equation as three separate equations gives

$$\begin{aligned} P_{k+1} &= F_k P_k (F_k - L_k H_k)^T + Q_k \\ \Pi_{k+1} &= \Pi_k - \Sigma_k H_k^T \lambda_k^T \\ \Sigma_{k+1}^T &= -F_k P_k H_k^T \lambda_k^T + F_k \Sigma_k^T \\ \Sigma_{k+1} &= \Sigma_k (F_k - L_k H_k)^T \end{aligned} \quad (9.23)$$

It is not immediately apparent from the above expressions that Σ_{k+1}^T is really the transpose of Σ_{k+1} , but the equality can be established by substituting for P_k and L_k .

Equations (9.19) – (9.23) completely define the fixed-point smoother. The fixed-point smoother, which is used for obtaining $\hat{x}_{j,k} = E(x_j | y_1, \dots, y_{k-1})$ for $k \geq j$, can be summarized as follows.

The fixed-point smoother

1. Run the standard Kalman filter up until time j , at which point we have \hat{x}_j^- and P_j^- . In the algorithm below, we omit the minus superscript on P_j^- for ease of notation.
2. Initialize the filter as follows:

$$\begin{aligned} \Sigma_j &= P_j \\ \Pi_j &= P_j \\ \hat{x}_{j,j} &= \hat{x}_j^- \end{aligned} \quad (9.24)$$

3. For $k = j, j + 1, \dots$, perform the following:

$$\begin{aligned}
 L_k &= F_k P_k H_k^T (H_k P_k H_k^T + R_k)^{-1} \\
 \lambda_k &= \Sigma_k H_k^T (H_k P_k H_k^T + R_k)^{-1} \\
 \hat{x}_{j,k+1} &= \hat{x}_{j,k} + \lambda_k (y_k - H_k \hat{x}_k^-) \\
 \hat{x}_{k+1}^- &= F_k \hat{x}_k^- + L_k (y_k - H_k \hat{x}_k^-) \\
 P_{k+1} &= F_k P_k (F_k - L_k H_k)^T + Q_k \\
 \Pi_{k+1} &= \Pi_k - \Sigma_k H_k^T \lambda_k^T \\
 \Sigma_{k+1} &= \Sigma_k (F_k - L_k H_k)^T
 \end{aligned} \tag{9.25}$$

As we recall from Equation (9.16), P_k is the *a priori* covariance of the standard Kalman filter estimate, Π_k is the covariance of the smoothed estimate of x_j at time k , and Σ_k is the cross covariance between the two.

9.2.1 Estimation improvement due to smoothing

Now we will look at the improvement in the estimate of x_j due to smoothing. The estimate \hat{x}_j^- is the standard *a priori* Kalman filter estimate of x_j , and the estimate $\hat{x}_{j,k+1}$ is the smoothed estimate after measurements up to and including time k have been processed. In other words, $\hat{x}_{j,k}$ uses $(k + 1 - j)$ more measurements to obtain the estimate of x_j than \hat{x}_j^- uses. How much more accurate can we expect our estimate to be with the use of these additional $(k + 1 - j)$ measurements? The estimation accuracy can be measured by the covariance. The improvement in estimation accuracy due to smoothing is equal to the standard estimation covariance P_j minus the smoothed estimation covariance Π_{k+1} . We can use Equations (9.24) and (9.25) to write this improvement as

$$\begin{aligned}
 P_j - \Pi_{k+1} &= \Pi_j - \left(\Pi_j - \sum_{i=j}^k \Sigma_i H_i^T \lambda_i^T \right) \\
 &= \sum_{i=j}^k \Sigma_i H_i^T \lambda_i^T
 \end{aligned} \tag{9.26}$$

Now assume for purposes of additional analysis that the system is time-invariant and the covariance of the standard filter has reached steady state at time j . Then we have

$$\lim_{k \rightarrow \infty} P_k^- = P \tag{9.27}$$

From Equation (9.25) we see that

$$\Sigma_{k+1} = \Sigma_k (F - LH)^T \tag{9.28}$$

where Σ is initialized as $\Sigma_j = P$. Combining this expression for Σ_{k+1} with its initial value, we see that

$$\begin{aligned}
 \Sigma_{k+1} &= P [(F - LH)^T]^{k+1-j} \\
 &= P (\tilde{F}^T)^{k+1-j}
 \end{aligned} \tag{9.29}$$

where \tilde{F} is defined by the above equation. Now substitute this expression, and the expression for λ from Equation (9.25), into Equation (9.26) to obtain

$$\begin{aligned} P_j - \Pi_{k+1} &= \sum_{i=j}^k \Sigma_i H^T \lambda^T \\ &= P \left[\sum_{i=j}^k \left(\tilde{F}^T \right)^{i-j} H^T (H P H^T + R)^{-1} H \tilde{F}^{i-j} \right] P \quad (9.30) \end{aligned}$$

The quantity on the right side of this equation is positive definite, which shows that the smoothed estimate of x_j is always better than the standard Kalman filter estimate. In other words, $(P_j - \Pi_{k+1}) > 0$, which implies that $\Pi_{k+1} < P_j$. Furthermore, the quantity on the right side is a sum of positive definite matrices, which shows that the larger the value of k (i.e., the more measurements that we use to obtain our smoothed estimate), the greater the improvement in the estimation accuracy. Also note from the above that the quantity $(H P H^T + R)$ inside the summation is inverted. This shows that as R increases, the quantity on the right side decreases. In the limit we see from Equation (9.30) that

$$\lim_{R \rightarrow \infty} (P_j - \Pi_{k+1}) = 0 \quad (9.31)$$

This illustrates the general principle that the larger the measurement noise, the smaller the improvement in estimation accuracy that we can obtain by smoothing. This is intuitive because large measurement noise means that additional measurements will not provide much improvement to our estimation accuracy.

■ EXAMPLE 9.1

In this example, we will see the improvement due to smoothing that can be obtained for a vehicle navigation problem. This is a second-order Newtonian system where $x(1)$ is position and $x(2)$ is velocity. The input is comprised of a commanded acceleration u plus acceleration noise \tilde{u} . The measurement y is a noisy measurement of position. After discretizing with a step size of T , the system equations can be written as

$$\begin{aligned} x_{k+1} &= \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} x_k + \begin{bmatrix} T^2/2 \\ T \end{bmatrix} (u_k + \tilde{u}_k) \\ &= \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} x_k + \begin{bmatrix} T^2/2 \\ T \end{bmatrix} u_k + w_k \\ y_k &= \begin{bmatrix} 1 & 0 \end{bmatrix} x_k + v_k \end{aligned} \quad (9.32)$$

Note that the process noise w_k is given as

$$w_k = \begin{bmatrix} T^2/2 \\ T \end{bmatrix} \tilde{u}_k \quad (9.33)$$

Now suppose the acceleration noise \tilde{u}_k has a standard deviation of a . We obtain the process noise covariance as follows:

$$\begin{aligned}
Q_k &= E(w_k w_k^T) \\
&= \begin{bmatrix} T^4/4 & T^3/2 \\ T^3/2 & T^2 \end{bmatrix} E(\tilde{u}_k^2) \\
&= a^2 \begin{bmatrix} T^4/4 & T^3/2 \\ T^3/2 & T^2 \end{bmatrix}
\end{aligned} \tag{9.34}$$

The percent improvement due to smoothing can be defined as

$$\text{Percent Improvement} = \frac{100 \operatorname{Tr}(P_j - \Pi_{k+1})}{\operatorname{Tr}(P_j)} \tag{9.35}$$

where j is the point which is being smoothed, and k is the number of measurements that are processed by the smoother. We can run the fixed-point smoother given by Equation (9.25) in order to smooth the position and velocity estimate at any desired time. Suppose we use the smoother equations to smooth the estimate at the second time step ($k = 1$). If we use measurements at times up to and including 10 seconds to estimate x_1 , then our estimate is denoted as $\hat{x}_{1,101}$. In this case, Table 9.1 shows the percent improvement due to smoothing after 10 seconds when the time step $T = 0.1$ and the acceleration noise standard deviation $a = 0.2$. As expected from the results of the previous subsection, we see that the improvement due to smoothing is more dramatic for small measurement noise.

Table 9.1 Improvement due to smoothing the state at the first time step after 10 seconds for Example 9.1. The improvement due to smoothing is more noticeable when the measurement noise is small.

| Measurement noise standard deviation | Percent Improvement |
|---|------------------------|
| 0.1 | 99.7 |
| 1 | 96.6 |
| 10 | 59.3 |
| 100 | 13.7 |
| 1000 | 0.2 |

Figure 9.5 shows the trace of Π_k , which is the covariance of the estimation error of the state at the first time step. As time progresses, our estimate of the state at the first time step improves. After 10 seconds of additional measurements, the estimate of the state at the first time step has improved by 96.6% relative to the standard Kalman filter estimate. Figure 9.6 shows the smoothed estimation error of the position and velocity of the first time step. We see that processing more measurements decreases the estimation-error covariance.

In general, the smoothed estimation errors shown in Figure 9.6 will converge to nonzero values. The estimation errors are zero-mean, but not for

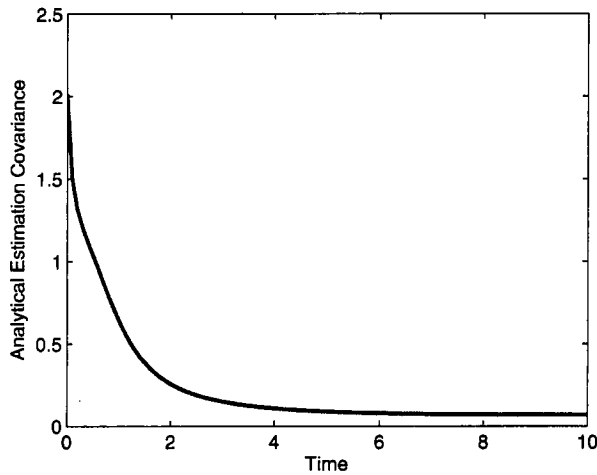


Figure 9.5 This shows the trace of the estimation-error covariance of the smoothed estimate of the state at the first time step for Example 9.1. As time progresses and we process more measurements, the covariance decreases, eventually reaching steady state.

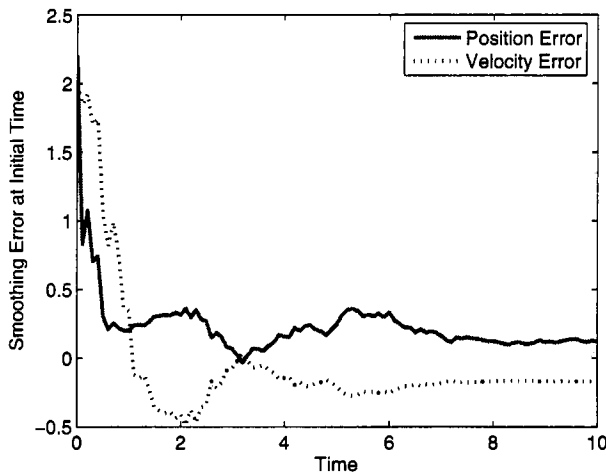


Figure 9.6 This shows typical estimation errors of the smoothed estimate of the state at the first time step for Example 9.1. As time progresses and we process more measurements, the estimation error decreases, and its standard deviation eventually reaches steady state.

any particular simulation. The estimation errors are zero-mean when averaged over many simulations. The system discussed here was simulated 1000 times and the variance of the estimation errors $(x_1 - \hat{x}_{1,101})$ were computed numerically to be equal to 0.054 and 0.012 for the two states. The diagonal elements of Π_{101} were equal to 0.057 and 0.012.

▽▽▽

9.2.2 Smoothing constant states

Now we will think about the improvement (due to smoothing) in the estimation accuracy of constant states. If the system states are constant then $F_k = I$ and $Q = 0$. Equation (9.25) shows that

$$\begin{aligned} P_{k+1} &= F_k P_k (F_k - L_k H_k)^T + Q_k \\ &= P_k (I - L_k H_k)^T \\ \Sigma_{k+1} &= \Sigma_k (F_k - L_k H_k)^T \\ &= \Sigma_k (I - L_k H_k)^T \end{aligned} \quad (9.36)$$

Comparing these expressions for P_{k+1} and Σ_{k+1} , and realizing from Equation (9.24) that the initial value of $\Sigma_j = P_j$, we see that $\Sigma_k = P_k$ for $k \geq j$. This means that the expression for L_k from Equation (9.25) can be written as

$$\begin{aligned} L_k &= F_k P_k H_k^T (H_k P_k H_k^T + R_k)^{-1} \\ &= \Sigma_k H_k^T (H_k P_k H_k^T + R_k)^{-1} \\ &= \lambda_k \end{aligned} \quad (9.37)$$

Substituting these results into the expression for Π_{k+1} from Equation (9.25) we see that

$$\begin{aligned} \Pi_{k+1} &= \Pi_k - \Sigma_k H_k^T \lambda_k^T \\ &= \Pi_k - P_k H_k^T L_k^T \end{aligned} \quad (9.38)$$

Realizing that the initial value of $\Pi_j = P_j$, and comparing this expression for Π_{k+1} with Equation (9.36) for P_{k+1} , we see that $\Pi_k = P_k$ for $k \geq j$. Recall that P_k is the covariance of the estimate of x_k from the standard Kalman filter, and Π_k is the covariance of the estimate of x_j given measurements up to and including time $(k - 1)$.

This result shows that constant states are not smoothable. Additional measurements are still helpful for refining an estimate of a constant state. However, there is no point to using smoothing for estimation of a constant state. If we want to estimate a constant state at time j using measurements up to time $k > j$, then we may as well simply run the standard Kalman filter up to time k . Implementing the smoothing equations will not gain any improvement in estimation accuracy.

9.3 FIXED-LAG SMOOTHING

In fixed-lag smoothing we want to obtain an estimate of the state at time $(k - N)$ given measurements up to and including time k , where the time index k continually changes as we obtain new measurements, but the lag N is a constant. In other words, at each time point we have N future measurements available for our state estimate. We therefore want to obtain $\hat{x}_{k-N,k}$ for $k = N, N + 1, \dots$, where N is a fixed positive integer. This could be the case, for example, if a satellite is continually taking photographs that are to be displayed or transmitted N time steps after the photograph is taken. In this case, since the photograph is processed N time steps after it is taken, we have N additional measurements after each photograph that

are available to update the estimate of the satellite state and hence improve the quality of the photograph. In this section we use the notation

$$\begin{aligned}\hat{x}_{k-N,k} &= E(x_{k-N}|y_1, \dots, y_k) \\ \Pi_{k-N} &= E[(x_{k-N} - \hat{x}_{k-N,k})(x_{k-N} - \hat{x}_{k-N,k})^T]\end{aligned}\quad (9.39)$$

Note that the notation has changed slightly from the previous section. In the previous section we used the notation $\hat{x}_{k,m}$ to refer to the estimate of x_k given measurements up to and including time $(m-1)$. In this section (and in the remainder of this chapter) we use $\hat{x}_{k,m}$ to refer to the estimate of x_k given measurements up to and including time m .

Let us define $x_{k,m}$ as the state x_{k-m} propagated with an identity transition matrix and zero process noise to time k . With this definition we see that

$$\begin{aligned}x_{k+1,1} &= x_k \\ x_{k+1,2} &= x_{k-1} \\ &= x_{k,1} \\ x_{k+1,3} &= x_{k-2} \\ &= x_{k,2} \\ &\text{etc.}\end{aligned}\quad (9.40)$$

We can therefore define the augmented system

$$\begin{aligned}\begin{bmatrix} x_{k+1} \\ x_{k+1,1} \\ \vdots \\ x_{k+1,N+1} \end{bmatrix} &= \begin{bmatrix} F_k & 0 & \cdots & 0 \\ I & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & I & 0 \end{bmatrix} \begin{bmatrix} x_k \\ x_{k,1} \\ \vdots \\ x_{k,N+1} \end{bmatrix} + \begin{bmatrix} I \\ 0 \\ \vdots \\ 0 \end{bmatrix} w_k \\ y_k &= [H_k \ 0 \ \cdots \ 0] \begin{bmatrix} x_k \\ x_{k,1} \\ \vdots \\ x_{k,N+1} \end{bmatrix} + v_k\end{aligned}\quad (9.41)$$

The Kalman filter estimates of the components of this augmented state vector are given as

$$\begin{aligned}E(x_{k+1}|y_1 \cdots y_k) &= \hat{x}_{k+1}^- \\ &= \hat{x}_{k+1,k} \\ E(x_{k+1,1}|y_1 \cdots y_k) &= E(x_k|y_1 \cdots y_k) \\ &= \hat{x}_k^+ \\ &= \hat{x}_{k,k} \\ E(x_{k+1,2}|y_1 \cdots y_k) &= E(x_{k-1}|y_1 \cdots y_k) \\ &= \hat{x}_{k-1,k} \\ &\vdots \\ E(x_{k+1,N+1}|y_1 \cdots y_k) &= \hat{x}_{k-N,k}\end{aligned}\quad (9.42)$$

We see that if we can use a Kalman filter to estimate the states of the augmented system (using measurements up to and including time k), then the estimate of the

last element of the augmented state vector, $x_{k+1,N+1}$, will be equal to the estimate of x_{k-N} given measurements up to and including time k . This is the estimate that we are looking for in fixed-lag smoothing. This idea is illustrated in Figure 9.7.

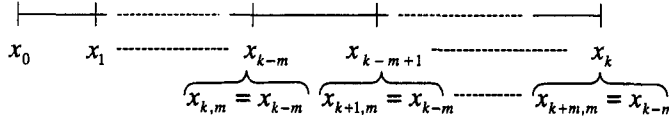


Figure 9.7 This illustrates the idea that is used to obtain the fixed-lag smoother. A fictitious state variable $x_{k,m}$ is initialized as $x_{k,m} = x_{k-m}$ and from that point on has an identity state transition matrix. The *a posteriori* estimate of $x_{k+m,m}$ is then equal to $\hat{x}_{k-m,k}$.

From Equation (9.10) we can write the Kalman filter for the augmented system of Equation (9.41) as follows:

$$\begin{bmatrix} \hat{x}_{k+1}^- \\ \hat{x}_{k,k} \\ \vdots \\ \hat{x}_{k-N,k} \end{bmatrix} = \begin{bmatrix} F_k & 0 & \cdots & 0 \\ I & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & I & 0 \end{bmatrix} \begin{bmatrix} \hat{x}_k^- \\ \hat{x}_{k-1,k-1} \\ \vdots \\ \hat{x}_{k-(N+1),k-1} \end{bmatrix} + \begin{bmatrix} L_{k,0} \\ L_{k,1} \\ \vdots \\ L_{k,N+1} \end{bmatrix} \left(y_k - \begin{bmatrix} H_k & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} \hat{x}_k^- \\ \hat{x}_{k-1,k-1} \\ \vdots \\ x_{k-(N+1),k-1} \end{bmatrix} \right) \quad (9.43)$$

where the $L_{k,i}$ matrices are components of the smoother gain that will be determined in this section. Note that $L_{k,0}$ is the standard Kalman gain. The smoother gain L_k is defined as

$$L_k = \begin{bmatrix} L_{k,0} \\ L_{k,1} \\ \vdots \\ L_{k,N+1} \end{bmatrix} \quad (9.44)$$

From Equation (9.10) we see that the L_k gain matrix is given by

$$L_k = \begin{bmatrix} F_k & 0 & \cdots & 0 \\ I & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & I & 0 \end{bmatrix} \begin{bmatrix} P_k^{0,0} & \cdots & (P_k^{0,N+1})^T \\ \vdots & \ddots & \vdots \\ P_k^{0,N+1} & \cdots & P_k^{N+1,N+1} \end{bmatrix} \begin{bmatrix} H_k^T \\ 0 \\ \vdots \\ 0 \end{bmatrix} \times \left(\begin{bmatrix} H_k & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} P_k^{0,0} & \cdots & (P_k^{0,N+1})^T \\ \vdots & \ddots & \vdots \\ P_k^{0,N+1} & \cdots & P_k^{N+1,N+1} \end{bmatrix} \begin{bmatrix} H_k^T \\ 0 \\ \vdots \\ 0 \end{bmatrix} + R_k \right)^{-1} \quad (9.45)$$

where the $P_k^{i,j}$ covariance matrices are defined as

$$P_k^{i,j} = E[(x_{k-j} - \hat{x}_{k-j,k-1})(x_{k-i} - \hat{x}_{k-i,k-1})^T] \quad (9.46)$$

The L_k expression above can be simplified to

$$L_k = \begin{bmatrix} F_k P_k^{0,0} H_k^T \\ P_k^{0,0} H_k^T \\ \vdots \\ P_k^{0,N} H_k^T \end{bmatrix} (H_k P_k^{0,0} H_k^T + R_k)^{-1} \quad (9.47)$$

From Equation (9.10) we see that the covariance-update equation for the Kalman filter for our augmented system can be written as

$$\begin{bmatrix} P_{k+1}^{0,0} & \cdots & (P_{k+1}^{0,N+1})^T \\ \vdots & \ddots & \vdots \\ P_{k+1}^{0,N+1} & \cdots & P_{k+1}^{N+1,N+1} \end{bmatrix} = \begin{bmatrix} F_k & 0 & \cdots & 0 \\ I & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & I & 0 \end{bmatrix} \begin{bmatrix} P_k^{0,0} & \cdots & (P_k^{0,N+1})^T \\ \vdots & \ddots & \vdots \\ P_k^{0,N+1} & \cdots & P_k^{N+1,N+1} \end{bmatrix} \times \\ \left(\begin{bmatrix} F_k^T & I & \cdots & 0 \\ 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & I \\ 0 & \cdots & \cdots & 0 \end{bmatrix} - \begin{bmatrix} H_k^T \\ 0 \\ \vdots \\ 0 \end{bmatrix} L_k^T \right) + \begin{bmatrix} Q_k & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 \end{bmatrix} \quad (9.48)$$

Substituting for L_k from Equation (9.47) and multiplying out gives

$$\begin{bmatrix} P_{k+1}^{0,0} & \cdots & (P_{k+1}^{0,N+1})^T \\ \vdots & \ddots & \vdots \\ P_{k+1}^{0,N+1} & \cdots & P_{k+1}^{N+1,N+1} \end{bmatrix} = \begin{bmatrix} F_k P_k^{0,0} & F_k (P_k^{0,1})^T & \cdots & F_k (P_k^{0,N+1})^T \\ P_k^{0,0} & (P_k^{0,1})^T & \cdots & (P_k^{0,N+1})^T \\ \vdots & \ddots & \ddots & \vdots \\ P_k^{0,N} & P_k^{1,N} & \cdots & (P_k^{N,N+1})^T \end{bmatrix} \times \\ \left(\begin{bmatrix} F_k^T & I & \cdots & 0 \\ 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & I \\ 0 & \cdots & \cdots & 0 \end{bmatrix} - H_k^T (H_k P_k^{0,0} H_k^T + R_k)^{-1} H_k \times \right. \\ \left. \begin{bmatrix} P_k^{0,0} F_k^T & P_k^{0,0} & \cdots & P_k^{0,N} \\ 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \right) + \begin{bmatrix} Q_k & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 \end{bmatrix} \quad (9.49)$$

This gives us the update equations for the P matrices. The equations for the first column of the P matrix are as follows:

$$\begin{aligned} P_{k+1}^{0,0} &= F_k P_k^{0,0} \left[F_k^T - H_k^T (H_k P_k^{0,0} H_k^T + R_k)^{-1} H_k P_k^{0,0} F_k^T \right] + Q_k \\ &= F_k P_k^{0,0} (F_k - L_{k,0} H_k)^T + Q_k \\ P_{k+1}^{0,1} &= P_k^{0,0} (F_k - L_{k,0} H_k)^T \\ &\vdots \\ P_{k+1}^{0,N+1} &= P_k^{0,N} (F_k - L_{k,0} H_k)^T \end{aligned} \quad (9.50)$$

The equations for the diagonal elements of the P matrix are as follows:

$$\begin{aligned}
 P_{k+1}^{1,1} &= P_k^{0,0} \left[I - H_k^T (H_k P_k^{0,0} H_k^T + R_k)^{-1} H_k P_k^{0,0} F_k^T \right] \\
 &= P_k^{0,0} - P_k^{0,0} H_k^T L_{k,1}^T F_k^T \\
 P_{k+1}^{2,2} &= P_k^{0,1} [-H_k^T L_{k,2}^T F_k^T] + P_k^{1,1} \\
 &= P_k^{1,1} - P_k^{0,1} H_k^T L_{k,2}^T F_k^T \\
 &\vdots \\
 P_{k+1}^{i,i} &= P_k^{i-1,i-1} - P_k^{0,i-1} H_k^T L_{k,i}^T F_k^T
 \end{aligned} \tag{9.51}$$

These equations give us the formulas that we can use for fixed-lag smoothing. This gives us the estimate $E(x_{k-N}|y_1, \dots, y_k)$ for a fixed N as k continually increments. The fixed-lag smoother is summarized as follows.

The fixed-lag smoother

1. Run the standard Kalman filter of Equation (9.10) to obtain \hat{x}_{k+1}^- , L_k , and P_k^- .
2. Initialize the fixed-lag smoother as follows:

$$\begin{aligned}
 \hat{x}_{k+1,k} &= \hat{x}_{k+1}^- \\
 L_{k,0} &= L_k \\
 P_k^{0,0} &= P_k^-
 \end{aligned} \tag{9.52}$$

3. For $i = 1, \dots, N+1$, perform the following:

$$\begin{aligned}
 L_{k,i} &= P_k^{0,i-1} H_k^T (H_k P_k^{0,0} H_k^T + R_k)^{-1} \\
 P_{k+1}^{i,i} &= P_k^{i-1,i-1} - P_k^{0,i-1} H_k^T L_{k,i}^T F_k^T \\
 P_{k+1}^{0,i} &= P_k^{0,i-1} (F_k - L_{k,0} H_k)^T \\
 \hat{x}_{k+1-i,k} &= \hat{x}_{k+2-i,k} + L_{k,i} (y_k - H_k \hat{x}_k^-)
 \end{aligned} \tag{9.53}$$

Note that the first time through this loop is the measurement update of the standard Kalman filter. At the end of this loop we have the smoothed estimates of *each* state with delays between 0 and N , given measurements up to and including time k . These estimates are denoted $\hat{x}_{k,k}, \dots, \hat{x}_{k-N,k}$. We also have the estimation-error covariances, denoted $P_{k+1}^{1,1}, \dots, P_{k+1}^{N+1,N+1}$.

The percent improvement due to smoothing can be computed as

$$\text{Percent Improvement} = \frac{100 \operatorname{Tr}(P_k^{0,0} - P_k^{N+1,N+1})}{\operatorname{Tr}(P_k^{0,0})} \tag{9.54}$$

■ EXAMPLE 9.2

Consider the same two state system as described in Example 9.1. Suppose we are trying to estimate the state of the system with a fixed time lag. The discretization time step $T = 0.1$ and the standard deviation of the acceleration noise is 10. Figure 9.8 shows the percent improvement in state estimation that is available with fixed-lag smoothing. The figure shows percent improvement as a function of lag size, and for two different values of measurement noise. The values on the plot are based on the theoretical estimation-error covariance. As expected, the improvement in estimation accuracy is more dramatic as the measurement noise decreases. This was discussed at the end of Section 9.2.

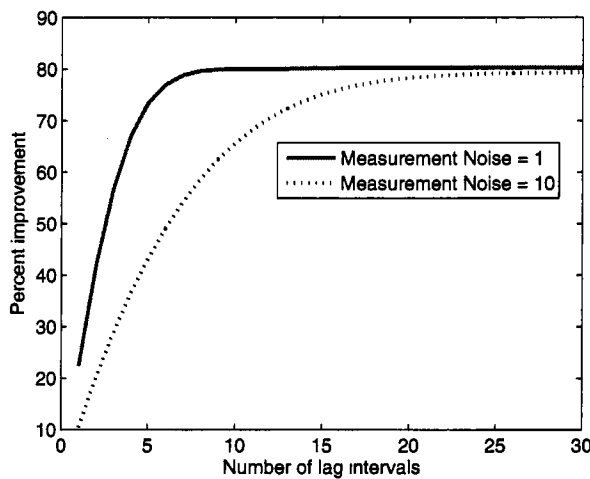


Figure 9.8 This shows the percent improvement of the trace of the estimation-error covariance of the smoothed estimate of the state (relative to the standard Kalman filter) for Example 9.2. As the number of lag intervals increases, the estimation error of the smoother decreases and the percent improvement increases. Also, as the measurement noise decreases, the improvement due to smoothing is more dramatic.

▽▽▽

9.4 FIXED-INTERVAL SMOOTHING

Suppose we have measurements for a fixed time interval. In fixed-interval smoothing we seek an estimate of the state at some of the interior points of the time interval. During the smoothing process we do not obtain any new measurements. Section 9.4.1 discusses the forward-backward approach to smoothing, which is perhaps the most straightforward smoothing algorithm. Section 9.4.2 discusses the RTS smoother, which is conceptually more difficult but is computationally cheaper than forward-backward smoothing.

9.4.1 Forward-backward smoothing

Suppose we want to estimate the state x_m based on measurements from $k = 1$ to $k = N$, where $N > m$. The forward-backward approach to smoothing obtains two estimates of x_m . The first estimate, \hat{x}_f , is based on the standard Kalman filter that operates from $k = 1$ to $k = m$. The second estimate, \hat{x}_b , is based on a Kalman filter that runs backward in time from $k = N$ back to $k = m$. The forward-backward approach to smoothing combines the two estimates to form an optimal smoothed estimate. This approach was first suggested in [Fra69].

Suppose that we combine a forward estimate \hat{x}_f of the state and a backward estimate \hat{x}_b of the state to get a smoothed estimate of x as follows:

$$\hat{x} = K_f \hat{x}_f + K_b \hat{x}_b \quad (9.55)$$

where K_f and K_b are constant matrix coefficients to be determined. Note that \hat{x}_f and \hat{x}_b are both unbiased since they are both outputs from Kalman filters. Therefore, if \hat{x} is to be unbiased, we require $K_f + K_b = I$ (see Problem 9.9). This gives

$$\hat{x} = K_f \hat{x}_f + (I - K_f) \hat{x}_b \quad (9.56)$$

The covariance of the estimate can then be found as

$$\begin{aligned} P &= E[(x - \hat{x})(x - \hat{x})^T] \\ &= E\{[x - K_f \hat{x}_f - (I - K_f) \hat{x}_b][\cdot \cdot]^T\} \\ &= E\{[K_f(e_f - e_b) + e_b][\cdot \cdot]^T\} \\ &= E\{K_f(e_f e_f^T + e_b e_b^T) K_f^T + e_b e_b^T - K_f e_b e_b^T - e_b e_b^T K_f^T\} \end{aligned} \quad (9.57)$$

where $e_f = x - x_f$, $e_b = x - x_b$, and we have used the fact that $E(e_f e_b^T) = 0$. The estimates \hat{x}_f and \hat{x}_b are both unbiased, and e_f and e_b are independent (since they depend on separate sets of measurements). We can minimize the trace of P with respect to K_f using results from Equation (1.66) and Problem 1.4:

$$\begin{aligned} \frac{\partial \text{Tr}(P)}{\partial K_f} &= 2E\{K_f(e_f e_f^T + e_b e_b^T) - e_b e_b^T\} \\ &= 2[K_f(P_f + P_b) - P_b] \end{aligned} \quad (9.58)$$

where $P_f = E(e_f e_f^T)$ is the covariance of the forward estimate, and $P_b = E(e_b e_b^T)$ is the covariance of the backward estimate. Setting this equal to zero to find the optimal value of K_f gives

$$\begin{aligned} K_f &= P_b(P_f + P_b)^{-1} \\ K_b &= P_f(P_f + P_b)^{-1} \end{aligned} \quad (9.59)$$

The inverse of $(P_f + P_b)$ always exists since both covariance matrices are positive definite. We can substitute this result into Equation (9.57) to find the covariance of the fixed-interval smoother as follows:

$$\begin{aligned} P &= P_b(P_f + P_b)^{-1}(P_f + P_b)(P_f + P_b)^{-1}P_b + \\ &\quad P_b - P_b(P_f + P_b)^{-1}P_b - P_b(P_f + P_b)^{-1}P_b \end{aligned} \quad (9.60)$$

Using the identity $(A+B)^{-1} = B^{-1}(AB^{-1} + I)^{-1}$ (see Problem 9.2), we can write the above equation as

$$\begin{aligned} P &= (P_f P_b^{-1} + I)^{-1} (P_f + P_b) (P_b^{-1} P_f + I)^{-1} + \\ &\quad P_b - (P_f P_b^{-1} + I)^{-1} P_b - (P_f P_b^{-1} + I)^{-1} P_b \end{aligned} \quad (9.61)$$

Multiplying out the first term, and again using the identity $(A+B)^{-1} = B^{-1}(AB^{-1} + I)^{-1}$ on the last two terms, results in

$$\begin{aligned} P &= \left[(P_b^{-1} + P_f^{-1})^{-1} + (P_b^{-1} P_f P_b^{-1} + P_b^{-1})^{-1} \right] (P_b^{-1} P_f + I)^{-1} + \\ &\quad P_b - 2(P_b^{-1} P_f P_b^{-1} + P_b^{-1})^{-1} \end{aligned} \quad (9.62)$$

From the matrix inversion lemma of Equation (1.39) we see that

$$(P_b^{-1} P_f P_b^{-1} + P_b^{-1})^{-1} = P_b - (P_f^{-1} + P_b^{-1})^{-1} \quad (9.63)$$

Substituting this into Equation (9.62) gives

$$\begin{aligned} P &= P_b (P_b^{-1} P_f + I)^{-1} + P_b - 2P_b + 2(P_f^{-1} + P_b^{-1})^{-1} \\ &= (P_b^{-1} P_f P_b^{-1} + P_b^{-1})^{-1} - P_b + 2(P_f^{-1} + P_b^{-1})^{-1} \\ &= P_b - (P_f^{-1} + P_b^{-1})^{-1} - P_b + 2(P_f^{-1} + P_b^{-1})^{-1} \\ &= (P_f^{-1} + P_b^{-1})^{-1} \end{aligned} \quad (9.64)$$

These results form the basis for the fixed-interval smoothing problem. The system model is given as

$$\begin{aligned} x_k &= F_{k-1} x_{k-1} + w_{k-1} \\ y_k &= H_k x_k + v_k \\ w_k &\sim (0, Q_k) \\ v_k &\sim (0, R_k) \end{aligned} \quad (9.65)$$

Suppose we want a smoothed estimate at time index m . First we run the forward Kalman filter normally, using measurements up to and including time m .

1. Initialize the forward filter as follows:

$$\begin{aligned} \hat{x}_{f0}^+ &= E(x_0) \\ P_{f0}^+ &= E \left[(x_0 - \hat{x}_{f0}^+) (x_0 - \hat{x}_{f0}^+)^T \right] \end{aligned} \quad (9.66)$$

2. For $k = 1, \dots, m$, perform the following:

$$\begin{aligned} P_{fk}^- &= F_{k-1} P_{f,k-1}^+ F_{k-1}^T + Q_{k-1} \\ K_{fk} &= P_{fk}^- H_k^T (H_k P_{fk}^- H_k^T + R_k)^{-1} \\ &= P_{fk}^+ H_k^T R_k^{-1} \\ \hat{x}_{fk}^- &= F_{k-1} \hat{x}_{f,k-1}^+ \\ \hat{x}_{fk}^+ &= \hat{x}_{fk}^- + K_{fk} (y_k - H_k \hat{x}_{fk}^-) \\ P_{fk}^+ &= (I - K_{fk} H_k) P_{fk}^- \end{aligned} \quad (9.67)$$

At this point we have a forward estimate for x_m , along with its covariance. These quantities are obtained using measurements up to and including time m .

The backward filter needs to run backward in time, starting at the final time index N . Since the forward and backward estimates must be independent, none of the information that was used in the forward filter is allowed to be used in the backward filter. Therefore, P_{bN}^- must be infinite:

$$P_{bN}^- = \infty \quad (9.68)$$

We are using the minus superscript on P_{bN}^- to indicate the backward covariance at time N before the measurement at time N is processed. (Recall that the filtering is performed backward in time.) So P_{bN}^- will be updated to obtain P_{bN}^+ after the measurement at time N is processed. Then it will be extrapolated backward in time to obtain $P_{b,N-1}^-$, and so on.

Now the question arises how to initialize the backward state estimate \hat{x}_{bk}^- at the final time $k = N$. We can solve this problem by introducing the new variable

$$s_k = P_{bk}^{-1} \hat{x}_{bk} \quad (9.69)$$

A minus or plus superscript can be added on all the quantities in the above equation to indicate values before or after the measurement at time k is taken into account. Since $P_{bN}^- = \infty$ it follows that

$$s_N^- = 0 \quad (9.70)$$

The infinite boundary condition on P_{bk}^- means that we cannot run the standard Kalman filter backward in time because we have to begin with an infinite covariance. Instead we run the information filter from Section 6.2 backward in time. This can be done by writing the system of Equation (9.65) as

$$\begin{aligned} x_{k-1} &= F_{k-1}^{-1} x_k + F_{k-1}^{-1} w_{k-1} \\ &= F_{k-1}^{-1} x_k + w_{b,k-1} \\ y_k &= H_k x_k + v_k \\ w_{bk} &\sim (0, F_k^{-1} Q_k F_k^{-T}) \\ v_k &\sim (0, R_k) \end{aligned} \quad (9.71)$$

Note that F_k^{-1} should always exist if it comes from a real system, because F_k comes from a matrix exponential that is always invertible (see Sections 1.2 and 1.4). The backward information filter can be written as follows.

1. Initialize the filter with $\mathcal{I}_{bN}^- = 0$.
2. For $k = N, N-1, \dots$, perform the following:

$$\begin{aligned} \mathcal{I}_{bk}^+ &= \mathcal{I}_{bk}^- + H_k^T R_k^{-1} H_k \\ K_{bk} &= (\mathcal{I}_{bk}^+)^{-1} H_k^T R_k^{-1} \\ \hat{x}_{bk}^+ &= \hat{x}_{bk}^- + K_{bk} (y_k - H_k \hat{x}_{bk}^-) \\ \mathcal{I}_{b,k-1}^- &= [F_{k-1}^{-1} (\mathcal{I}_{bk}^+)^{-1} F_{k-1}^{-T} + F_{k-1}^{-1} Q_{k-1} F_{k-1}^{-T}]^{-1} \\ &= F_{k-1}^T [(\mathcal{I}_{bk}^+)^{-1} + Q_{k-1}]^{-1} F_{k-1} \\ &= F_{k-1}^T [Q_{k-1}^{-1} - Q_{k-1}^{-1} (\mathcal{I}_{bk}^+ + Q_{k-1}^{-1})^{-1} Q_{k-1}^{-1}] F_{k-1} \\ \hat{x}_{b,k-1}^- &= F_{k-1}^{-1} \hat{x}_{bk}^+ \end{aligned} \quad (9.72)$$

The first form for $\mathcal{I}_{b,k-1}^-$ above requires the inversion of \mathcal{I}_{bk}^+ . Consider the first time step for the backward filter (i.e., at $k = N$). The information matrix \mathcal{I}_{bN}^- is initialized to zero, and then the first time through the above loop we set $\mathcal{I}_{bk}^+ = \mathcal{I}_{bk}^- + H_k^T R_k^{-1} H_k$. If there are fewer measurements than states, $H_k^T R_k^{-1} H_k$ will always be singular and, therefore, \mathcal{I}_{bk}^+ will be singular at $k = N$. Therefore, the first form given above for $\mathcal{I}_{b,k-1}^-$ will not be computable. In practice we can get around this by initializing \mathcal{I}_{bN}^- to a small nonzero matrix instead of zero.

The third form for $\mathcal{I}_{b,k-1}^-$ above has its own problems. It does not require the inversion of \mathcal{I}_{bk}^+ , but it does require the inversion of Q_{k-1} . So the third form of $\mathcal{I}_{b,k-1}^-$ is not computable unless Q_{k-1} is nonsingular. Again, in practice we can get around this by making a small modification to Q_{k-1} so that it is numerically nonsingular.

Since we need to update $s_k = \mathcal{I}_{bk} \hat{x}_{bk}$ instead of \hat{x}_{bk} (because of initialization issues) as defined in Equation (9.69), we rewrite the update equations for the state estimate as follows:

$$\begin{aligned}\hat{x}_{bk}^+ &= \hat{x}_{bk}^- + K_{bk}(y_k - H_k \hat{x}_{bk}^-) \\ s_k^+ &= \mathcal{I}_{bk}^+ \hat{x}_{bk}^+ \\ &= \mathcal{I}_{bk}^+ \hat{x}_{bk}^- + \mathcal{I}_{bk}^+ K_{bk}(y_k - H_k \hat{x}_{bk}^-)\end{aligned}\quad (9.73)$$

Now note from Equation (6.33) that we can write $\mathcal{I}_{bk}^+ = \mathcal{I}_{bk}^- + H_k^T R_k^{-1} H_k$, and $K_{bk} = P_{bk}^+ H_k^T R_k^{-1}$. Substituting these expressions into the above equation for s_k^+ gives

$$\begin{aligned}s_k^+ &= \mathcal{I}_{bk}^- \hat{x}_{bk}^- + H_k^T R_k^{-1} H_k \hat{x}_{bk}^- + H_k^T R_k^{-1} (y_k - H_k \hat{x}_{bk}^-) \\ &= s_k^- + H_k^T R_k^{-1} y_k\end{aligned}\quad (9.74)$$

We combine this with Equation (9.72) to write the backward information filter as follows.

1. Initialize the filter as follows:

$$\begin{aligned}s_N^- &= 0 \\ \mathcal{I}_{bN}^- &= 0\end{aligned}\quad (9.75)$$

2. For $k = N, N-1, \dots, m+1$, perform the following:

$$\begin{aligned}\mathcal{I}_{bk}^+ &= \mathcal{I}_{bk}^- + H_k^T R_k^{-1} H_k \\ s_k^+ &= s_k^- + H_k^T R_k^{-1} y_k \\ \mathcal{I}_{b,k-1}^- &= [F_{k-1}^{-1} (\mathcal{I}_{bk}^+)^{-1} F_{k-1}^{-T} + F_{k-1}^{-1} Q_{k-1} F_{k-1}^{-T}]^{-1} \\ &= F_{k-1}^T [(\mathcal{I}_{bk}^+)^{-1} + Q_{k-1}]^{-1} F_{k-1} \\ &= F_{k-1}^T [Q_{k-1}^{-1} - Q_{k-1}^{-1} (\mathcal{I}_{bk}^+ + Q_{k-1}^{-1})^{-1} Q_{k-1}^{-1}] F_{k-1} \\ s_{k-1}^- &= \mathcal{I}_{b,k-1}^- F_{k-1}^{-1} (\mathcal{I}_{bk}^+)^{-1} s_k^+\end{aligned}\quad (9.76)$$

3. Perform one final time update to obtain the backward estimate of x_m :

$$\begin{aligned}
\mathcal{I}_{bm}^- &= Q_m^{-1} - Q_m^{-1} F_m^{-1} (\mathcal{I}_{b,m+1}^+ + F_m^{-T} Q_m^{-1} F_m^{-1})^{-1} F_m^{-T} Q_m^{-1} \\
P_{bm}^- &= (\mathcal{I}_{bm}^-)^{-1} \\
s_m^- &= \mathcal{I}_{bm}^- F_m^{-1} (\mathcal{I}_{b,m+1}^+)^{-1} s_{m+1}^+ \\
\hat{x}_{bm}^- &= (\mathcal{I}_{bm}^-)^{-1} s_m^-
\end{aligned} \tag{9.77}$$

Now we have the backward estimate \hat{x}_{bm}^- and its covariance P_{bm}^- . These quantities are obtained from measurements $m+1, m+2, \dots, N$.

After we obtain the backward quantities as outlined above, we combine them with the forward quantities from Equation (9.67) to obtain the final state estimate and covariance:

$$\begin{aligned}
K_f &= P_{bm}^- (P_{fm}^+ + P_{bm}^-)^{-1} \\
\hat{x}_m &= K_f \hat{x}_{fm}^+ + (I - K_f) \hat{x}_{bm}^- \\
P_m &= \left[(P_{fm}^+)^{-1} + (P_{bm}^-)^{-1} \right]^{-1}
\end{aligned} \tag{9.78}$$

We can obtain an alternative equation for \hat{x}_m by manipulating the above equations. If we substitute for K_f in the above expression for \hat{x}_m then we obtain

$$\begin{aligned}
\hat{x}_m &= P_{bm}^- (P_{fm}^+ + P_{bm}^-)^{-1} \hat{x}_{fm}^+ + \left[I - P_{bm}^- (P_{fm}^+ + P_{bm}^-)^{-1} \right] \hat{x}_{bm}^- \\
&= P_{bm}^- (P_{fm}^+ + P_{bm}^-)^{-1} \hat{x}_{fm}^+ + \left[(P_{fm}^+ + P_{bm}^-) - P_{bm}^- \right] (P_{fm}^+ + P_{bm}^-)^{-1} \hat{x}_{bm}^- \\
&= P_{bm}^- (P_{fm}^+ + P_{bm}^-)^{-1} \hat{x}_{fm}^+ + P_{fm}^+ (P_{fm}^+ + P_{bm}^-)^{-1} \hat{x}_{bm}^-
\end{aligned} \tag{9.79}$$

Using the matrix inversion lemma on the rightmost inverse in the above equation and performing some other manipulations gives

$$\begin{aligned}
\hat{x}_m &= [(P_{bm}^- + P_{fm}^+) - P_{fm}^+] (P_{fm}^+ + P_{bm}^-)^{-1} \hat{x}_{fm}^+ + \\
&\quad P_{fm}^+ [\mathcal{I}_{bm}^- - \mathcal{I}_{bm}^- (\mathcal{I}_{fm}^+ + \mathcal{I}_{bm}^-)^{-1} \mathcal{I}_{bm}^-] \hat{x}_{bm}^- \\
&= \left[I - P_{fm}^+ (P_{fm}^+ + P_{bm}^-)^{-1} \right] \hat{x}_{fm}^+ + \\
&\quad P_{fm}^+ \left[I - \mathcal{I}_{bm}^- (\mathcal{I}_{fm}^+ + \mathcal{I}_{bm}^-)^{-1} \right] \mathcal{I}_{bm}^- \hat{x}_{bm}^- \\
&= \left[I - P_{fm}^+ \mathcal{I}_{bm}^- (I + P_{fm}^+ \mathcal{I}_{bm}^-)^{-1} \right] \hat{x}_{fm}^+ + \\
&\quad P_{fm}^+ \left[I - \mathcal{I}_{bm}^- (I + P_{fm}^+ \mathcal{I}_{bm}^-)^{-1} P_{fm}^+ \right] \mathcal{I}_{bm}^- \hat{x}_{bm}^- \\
&= P_{fm}^+ \left[I - \mathcal{I}_{bm}^- (I + P_{fm}^+ \mathcal{I}_{bm}^-)^{-1} P_{fm}^+ \right] \mathcal{I}_{fm}^+ \hat{x}_{fm}^+ + \\
&\quad P_{fm}^+ \left[I - \mathcal{I}_{bm}^- (I + P_{fm}^+ \mathcal{I}_{bm}^-)^{-1} P_{fm}^+ \right] \mathcal{I}_{bm}^- \hat{x}_{bm}^-
\end{aligned} \tag{9.80}$$

where we have relied on the identity $(A + B)^{-1} = B^{-1}(AB^{-1} + I)^{-1}$ (see Problem 9.2). The coefficients of \hat{x}_{fm}^+ and \hat{x}_{bm}^- in the above equation both have a common factor which can be written as follows:

$$\begin{aligned}
& P_{fm}^+ \left[I - \mathcal{I}_{bm}^- (I + P_{fm}^+ \mathcal{I}_{bm}^-)^{-1} P_{fm}^+ \right] \\
&= P_{fm}^+ - P_{fm}^+ \mathcal{I}_{bm}^- (I + P_{fm}^+ \mathcal{I}_{bm}^-)^{-1} P_{fm}^+ \\
&= P_{fm}^+ - P_{fm}^+ (\mathcal{I}_{fm}^+ P_{bm}^- + I)^{-1} \\
&= \left[P_{fm}^+ (\mathcal{I}_{fm}^+ P_{bm}^- + I) - P_{fm}^+ \right] (\mathcal{I}_{fm}^+ P_{bm}^- + I)^{-1} \\
&= P_{bm}^- (\mathcal{I}_{fm}^+ P_{bm}^- + I)^{-1} \\
&= (\mathcal{I}_{fm}^+ + \mathcal{I}_{bm}^-)^{-1}
\end{aligned} \tag{9.81}$$

Therefore, using Equation (9.78), we can write Equation (9.80) as

$$\begin{aligned}
\hat{x}_m &= P_m \mathcal{I}_{fm}^+ \hat{x}_{fm}^+ + P_m \mathcal{I}_{bm}^- \hat{x}_{bm}^- \\
&= P_m (\mathcal{I}_{fm}^+ \hat{x}_{fm}^+ + \mathcal{I}_{bm}^- \hat{x}_{bm}^-)
\end{aligned} \tag{9.82}$$

Figure 9.9 illustrates how the forward-backward smoother works.

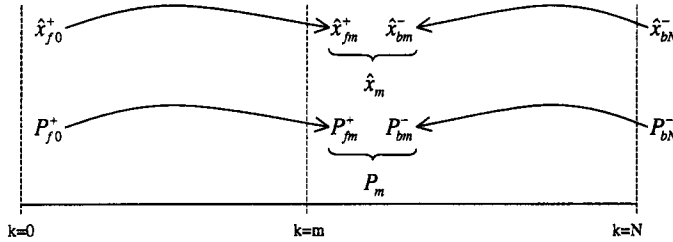


Figure 9.9 This figure illustrates the concept of the forward-backward smoother. The forward filter is run to obtain *a posteriori* estimates and covariances up to time m . Then the backward filter is run to obtain *a priori* estimates and covariances back to time m (i.e., *a priori* from a reversed time perspective). Then the forward and backward estimates and covariances at time m are combined to obtain the final estimate \hat{x}_m and covariance P_m .

■ EXAMPLE 9.3

In this example we consider the same problem given in Example 9.1. Suppose that we want to estimate the position and velocity of the vehicle at $t = 5$ seconds. We have measurements every 0.1 seconds for a total of 10 seconds. The standard deviation of the measurement noise is 10, and the standard deviation of the acceleration noise is 10. Figure 9.10 shows the trace of the covariance of the estimation of the forward filter as it runs from $t = 0$ to $t = 5$, the backward filter as it runs from $t = 10$ back to $t = 5$, and the smoothed estimate at $t = 5$. The forward and backward filters both converge to the same steady-state value, even though the forward filter was initialized to a covariance of 20 for both the position and velocity estimation errors, and the backward filter was initialized to an infinite covariance. The smoothed filter has a covariance of about 7.6, which shows the dramatic improvement that can be obtained in estimation accuracy when smoothing is used.

▽▽▽

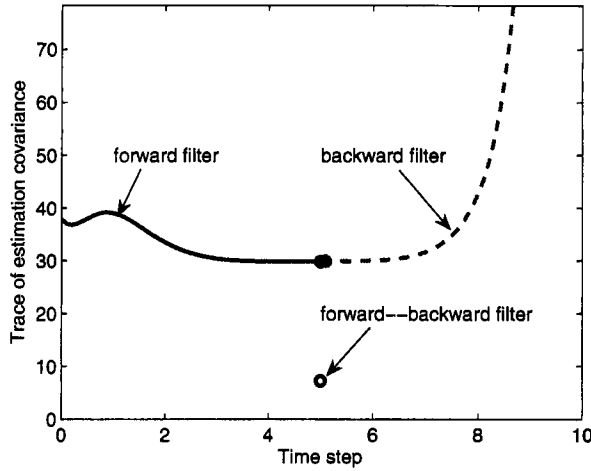


Figure 9.10 This shows the trace of the estimation-error covariance for Example 9.3. The forward filter runs from $t = 0$ to $t = 5$, the backward filter runs from $t = 10$ to $t = 5$, and the trace of the covariance of the smoothed estimate is shown at $t = 5$.

9.4.2 RTS smoothing

Several other forms of the fixed-interval smoother have been obtained. One of the most common is the smoother that was presented by Rauch, Tung, and Striebel, usually called the RTS smoother [Rau65]. The RTS smoother is more computationally efficient than the smoother presented in the previous section because we do not need to directly compute the backward estimate or covariance in order to get the smoothed estimate and covariance. In order to obtain the RTS smoother, we will first look at the smoothed covariance given in Equation (9.78) and obtain an equivalent expression that does not use P_{bm} . Then we will look at the smoothed estimate given in Equation (9.78), which uses the gain K_f , which depends on P_{bm} , and obtain an equivalent expression that does not use P_{bm} or \hat{x}_{bm} .

9.4.2.1 RTS covariance update First consider the smoothed covariance given in Equation (9.78). This can be written as

$$\begin{aligned} P_m &= \left[(P_{fm}^+)^{-1} + (P_{bm}^-)^{-1} \right]^{-1} \\ &= P_{fm}^+ - P_{fm}^+ (P_{fm}^+ + P_{bm}^-)^{-1} P_{fm}^+ \end{aligned} \quad (9.83)$$

where the second expression comes from an application of the matrix inversion lemma to the first expression (see Problem 9.3). From Equation (9.72) we see that

$$P_{bm}^- = F_m^{-1} \left[P_{b,m+1}^+ + Q_m \right] F_m^{-T} \quad (9.84)$$

Substituting this into the expression $(P_{fm}^+ + P_{bm}^-)^{-1}$ gives the following:

$$\begin{aligned}
(P_{fm}^+ + P_{bm}^-)^{-1} &= \left[P_{fm}^+ + F_m^{-1}(P_{b,m+1}^+ + Q_m)F_m^{-T} \right]^{-1} \\
&= \left[F_m^{-1}F_m P_{fm}^+ F_m^T F_m^{-T} + F_m^{-1}(P_{b,m+1}^+ + Q_m)F_m^{-T} \right]^{-1} \\
&= \left[F_m^{-1}(F_m P_{fm}^+ F_m^T + P_{b,m+1}^+ + Q_m)F_m^{-T} \right]^{-1} \\
&= F_m^T (F_m P_{fm}^+ F_m^T + P_{b,m+1}^+ + Q_m)^{-1} F_m \\
&= F_m^T (P_{f,m+1}^- + P_{b,m+1}^+)^{-1} F_m
\end{aligned} \tag{9.85}$$

From Equations (6.26) and (9.76) recall that

$$\begin{aligned}
\mathcal{I}_{fm}^+ &= \mathcal{I}_{fm}^- + H_m^T R_m^{-1} H_m \\
\mathcal{I}_{bm}^+ &= \mathcal{I}_{bm}^- + H_m^T R_m^{-1} H_m
\end{aligned} \tag{9.86}$$

We can combine these two equations to obtain

$$\mathcal{I}_{b,m+1}^+ = \mathcal{I}_{b,m+1}^- + \mathcal{I}_{f,m+1}^+ - \mathcal{I}_{f,m+1}^- \tag{9.87}$$

Substituting this into Equation (9.78) gives

$$\begin{aligned}
P_{m+1} &= \left[\mathcal{I}_{f,m+1}^+ + \mathcal{I}_{b,m+1}^- \right]^{-1} \\
&= \left[\mathcal{I}_{b,m+1}^+ + \mathcal{I}_{f,m+1}^- \right]^{-1} \\
P_{m+1}^{-1} &= \mathcal{I}_{b,m+1}^+ + \mathcal{I}_{f,m+1}^- \\
P_{b,m+1}^+ &= \left[P_{m+1}^{-1} - \mathcal{I}_{f,m+1}^- \right]^{-1}
\end{aligned} \tag{9.88}$$

Substituting this into Equation (9.85) gives

$$\begin{aligned}
(P_{fm}^+ + P_{bm}^-)^{-1} &= F_m^T \left[P_{f,m+1}^- + \left(P_{m+1}^{-1} - \mathcal{I}_{f,m+1}^- \right)^{-1} \right]^{-1} F_m \\
&= F_m^T \mathcal{I}_{f,m+1}^- \left[\mathcal{I}_{f,m+1}^- + \mathcal{I}_{f,m+1}^- \left(P_{m+1}^{-1} - \mathcal{I}_{f,m+1}^- \right)^{-1} \mathcal{I}_{f,m+1}^- \right]^{-1} \mathcal{I}_{f,m+1}^- F_m \\
&= F_m^T \mathcal{I}_{f,m+1}^- \left(P_{f,m+1}^- - P_{m+1} \right) \mathcal{I}_{f,m+1}^- F_m
\end{aligned} \tag{9.89}$$

where the last equality comes from an application of the matrix inversion lemma. Substituting this expression into Equation (9.83) gives

$$P_m = P_{fm}^+ - K_m (P_{f,m+1}^- - P_{m+1}) K_m^T \tag{9.90}$$

where the smoother gain K_m is given as

$$K_m = P_{fm}^+ F_m^T \mathcal{I}_{f,m+1}^- \tag{9.91}$$

The covariance update equation for P_m is not a function of the backward covariance. The smoother covariance P_m can be solved by using only the forward covariance P_{fm} , which reduces the computational effort (compared to the algorithm presented in Section 9.4.1).

9.4.2.2 RTS state estimate update Next we consider the smoothed estimate \hat{x}_m given in Equation (9.78). We will find an equivalent expression that does not use P_{bm} or \hat{x}_{bm} . In order to do this we will first need to establish a few lemmas.

Lemma 1

$$F_{k-1}^{-1}Q_{k-1}F_{k-1}^{-T} = F_{k-1}^{-1}P_{fk}^{-}F_{k-1}^{-T} - P_{f,k-1}^{+} \quad (9.92)$$

Proof: From Equation (9.67) we see that

$$P_{fk}^{-} = F_{k-1}P_{f,k-1}^{+}F_{k-1}^T + Q_{k-1} \quad (9.93)$$

Rearranging this equation gives

$$Q_{k-1} = P_{fk}^{-} - F_{k-1}P_{f,k-1}^{+}F_{k-1}^T \quad (9.94)$$

Premultiplying both sides by F_{k-1}^{-1} and postmultiplying both sides by F_{k-1}^{-T} gives the desired result.

QED

Lemma 2 The a posteriori covariance P_{bk}^{+} of the backward filter satisfies the equation

$$P_{bk}^{+} = (P_{fk}^{-} + P_{bk}^{+})\mathcal{I}_{fk}^{-}P_k \quad (9.95)$$

Proof: From Equation (9.78) we obtain

$$\begin{aligned} I &= (\mathcal{I}_{bk}^{+} + \mathcal{I}_{fk}^{-})P_k \\ P_{bk}^{+} &= (I + P_{bk}^{+}\mathcal{I}_{fk}^{-})P_k \\ &= P_k + P_{bk}^{+}\mathcal{I}_{fk}^{-}P_k \\ &= P_{fk}^{-}\mathcal{I}_{fk}^{-}P_k + P_{bk}^{+}\mathcal{I}_{fk}^{-}P_k \\ &= (P_{fk}^{-} + P_{bk}^{+})\mathcal{I}_{fk}^{-}P_k \end{aligned} \quad (9.96)$$

QED

Lemma 3 The covariances of the forward and backward filters satisfy the equation

$$P_{fk}^{-} + P_{bk}^{+} = F_{k-1}(P_{f,k-1}^{+} + P_{b,k-1}^{-})F_{k-1}^T \quad (9.97)$$

Proof: From Equation (9.67) and (9.72) we see that

$$\begin{aligned} P_{f,k-1}^{+} &= F_{k-1}^{-1}P_{fk}^{-}F_{k-1}^{-T} - F_{k-1}Q_{k-1}F_{k-1}^{-T} \\ P_{b,k-1}^{-} &= F_{k-1}^{-1}P_{bk}^{+}F_{k-1}^{-T} + F_{k-1}Q_{k-1}F_{k-1}^{-T} \end{aligned} \quad (9.98)$$

Adding these two equations and rearranging gives

$$\begin{aligned} P_{f,k-1}^{+} + P_{b,k-1}^{-} &= F_{k-1}^{-1}(P_{fk}^{-} + P_{bk}^{+})F_{k-1}^{-T} \\ P_{fk}^{-} + P_{bk}^{+} &= F_{k-1}(P_{f,k-1}^{+} + P_{b,k-1}^{-})F_{k-1}^T \end{aligned} \quad (9.99)$$

QED

Lemma 4 The smoothed estimate \hat{x}_k can be written as

$$\hat{x}_k = P_k\mathcal{I}_{fk}^{+}\hat{x}_{fk}^{-} - P_kH_k^TR_k^{-1}H_k\hat{x}_{fk}^{-} + P_k s_k^{+} \quad (9.100)$$

Proof: From Equations (9.69) and (9.82) we have

$$\begin{aligned}\hat{x}_k &= P_k \mathcal{I}_{fk}^+ \hat{x}_{fk}^+ + P_k \mathcal{I}_{bk}^- \hat{x}_{bk}^- \\ &= P_k \mathcal{I}_{fk}^+ \hat{x}_{fk}^+ + P_k s_k^- \end{aligned} \quad (9.101)$$

From Equation (9.76) we see that

$$s_k^- = s_k^+ - H_k^T R_k^{-1} y_k \quad (9.102)$$

Substitute this expression for s_k^- , and the expression for \hat{x}_{fk}^+ from Equation (9.67), into Equation (9.101) to obtain

$$\hat{x}_k = P_k \mathcal{I}_{fk}^+ \hat{x}_{fk}^- + P_k \mathcal{I}_{fk}^+ K_{fk} (y_k - H_k \hat{x}_{fk}^-) + P_k s_k^+ - P_k H_k^T R_k^{-1} y_k \quad (9.103)$$

Now substitute $P_{fk}^+ H_k^T R_k^{-1}$ for K_{fk} [from Equation (9.67)] in the above equation to obtain

$$\begin{aligned}\hat{x}_k &= P_k \mathcal{I}_{fk}^+ \hat{x}_{fk}^- + P_k \mathcal{I}_{fk}^+ P_{fk}^+ H_k^T R_k^{-1} (y_k - H_k \hat{x}_{fk}^-) + P_k s_k^+ - P_k H_k^T R_k^{-1} y_k \\ &= P_k \mathcal{I}_{fk}^+ \hat{x}_{fk}^- + P_k H_k^T R_k^{-1} (y_k - H_k \hat{x}_{fk}^-) + P_k s_k^+ - P_k H_k^T R_k^{-1} y_k \\ &= P_k \mathcal{I}_{fk}^+ \hat{x}_{fk}^- - P_k H_k^T R_k^{-1} H_k \hat{x}_{fk}^- + P_k s_k^+ \end{aligned} \quad (9.104)$$

QED

Lemma 5

$$(P_{f,k-1}^+ + P_{b,k-1}^-)^{-1} = F_{k-1}^T \mathcal{I}_{fk}^- (P_{fk}^- - P_k) \mathcal{I}_{fk}^- F_{k-1} \quad (9.105)$$

Proof: Recall from Equations (6.26) and (9.72) that

$$\begin{aligned}\mathcal{I}_{fk}^+ &= \mathcal{I}_{fk}^- + H_k^T R_k^{-1} H_k \\ \mathcal{I}_{bk}^+ &= \mathcal{I}_{bk}^- + H_k^T R_k^{-1} H_k \end{aligned} \quad (9.106)$$

Combining these two equations gives

$$\begin{aligned}\mathcal{I}_{bk}^+ &= \mathcal{I}_{bk}^- + \mathcal{I}_{fk}^+ - \mathcal{I}_{fk}^- \\ &= \left[\left(\mathcal{I}_{bk}^- + \mathcal{I}_{fk}^+ \right)^{-1} \right]^{-1} - \mathcal{I}_{fk}^- \\ &= P_k^{-1} - \mathcal{I}_{fk}^- \\ P_{bk}^+ &= \left(\mathcal{I}_k - \mathcal{I}_{fk}^- \right)^{-1} \end{aligned} \quad (9.107)$$

where we have used Equation (9.78) in the above derivation. Substitute this expression for P_{bk}^+ into Equation (9.97) to obtain

$$\begin{aligned}F_{k-1} (P_{f,k-1}^+ + P_{b,k-1}^-) F_{k-1}^T &= P_{fk}^- + P_{bk}^+ \\ &= P_{fk}^- + \left(\mathcal{I}_k - \mathcal{I}_{fk}^- \right)^{-1} \end{aligned} \quad (9.108)$$

Invert both sides to obtain

$$\begin{aligned}
F_{k-1}^{-T}(P_{f,k-1}^+ + P_{b,k-1}^-)^{-1}F_{k-1}^{-1} &= \left[P_{fk}^- + (\mathcal{I}_k - \mathcal{I}_{fk}^-)^{-1} \right]^{-1} \\
(P_{f,k-1}^+ + P_{b,k-1}^-)^{-1} &= F_{k-1}^T \left[P_{fk}^- + (\mathcal{I}_k - \mathcal{I}_{fk}^-)^{-1} \right]^{-1} F_{k-1} \\
&= F_{k-1}^T \left[P_{fk}^- \mathcal{I}_{fk}^- P_{fk}^- + P_{fk}^- \mathcal{I}_{fk}^- (\mathcal{I}_k - \mathcal{I}_{fk}^-)^{-1} \mathcal{I}_{fk}^- P_{fk}^- \right]^{-1} F_{k-1} \\
&= F_{k-1}^T \mathcal{I}_{fk}^- \left[\mathcal{I}_{fk}^- + \mathcal{I}_{fk}^- (\mathcal{I}_k - \mathcal{I}_{fk}^-)^{-1} \mathcal{I}_{fk}^- \right]^{-1} \mathcal{I}_{fk}^- F_{k-1} \quad (9.109)
\end{aligned}$$

Now apply the matrix inversion lemma to the term $(\mathcal{I}_k - \mathcal{I}_{fk}^-)^{-1}$ in the above equation. This results in

$$\begin{aligned}
(P_{f,k-1}^+ + P_{b,k-1}^-)^{-1} &= F_{k-1}^T \mathcal{I}_{fk}^- \left[\mathcal{I}_{fk}^- + \mathcal{I}_{fk}^- \left(-P_{fk}^- - P_{fk}^- (-P_{fk}^- + P_k)^{-1} P_{fk}^- \right) \mathcal{I}_{fk}^- \right]^{-1} \mathcal{I}_{fk}^- F_{k-1} \\
&= F_{k-1}^T \mathcal{I}_{fk}^- \left[\mathcal{I}_{fk}^- + \left(-I - (-P_{fk}^- + P_k)^{-1} P_{fk}^- \right) \mathcal{I}_{fk}^- \right]^{-1} \mathcal{I}_{fk}^- F_{k-1} \\
&= F_{k-1}^T \mathcal{I}_{fk}^- \left[\mathcal{I}_{fk}^- - \mathcal{I}_{fk}^- - (P_{fk}^- + P_k)^{-1} \right]^{-1} \mathcal{I}_{fk}^- F_{k-1} \\
&= F_{k-1}^T \mathcal{I}_{fk}^- (P_{fk}^- - P_k) \mathcal{I}_{fk}^- F_{k-1} \quad (9.110)
\end{aligned}$$

QED

With the above lemmas we now have the tools that we need to obtain an alternate expression for the smoothed estimate. Starting with the expression for s_{k-1}^- in Equation (9.76), and substituting the expression for $\mathcal{I}_{b,k-1}^-$ from Equation (9.72) gives

$$\begin{aligned}
s_{k-1}^- &= \mathcal{I}_{bk}^- F_{k-1}^{-1} P_{bk}^+ s_k^+ \\
&= F_{k-1}^T \left[Q_{k-1}^{-1} - Q_{k-1}^{-1} (\mathcal{I}_{bk}^+ + Q_{k-1}^{-1})^{-1} Q_{k-1}^{-1} \right] F_{k-1} F_{k-1}^{-1} P_{bk}^+ s_k^+ \\
&= F_{k-1}^T Q_{k-1}^{-1} \left[I - (\mathcal{I}_{bk}^+ + Q_{k-1}^{-1})^{-1} Q_{k-1}^{-1} \right] P_{bk}^+ s_k^+ \\
&= F_{k-1}^T Q_{k-1}^{-1} (\mathcal{I}_{bk}^+ + Q_{k-1}^{-1}) (\mathcal{I}_{bk}^+ + Q_{k-1}^{-1} - Q_{k-1}^{-1}) P_{bk}^+ s_k^+ \\
&= F_{k-1}^T Q_{k-1}^{-1} (\mathcal{I}_{bk}^+ + Q_{k-1}^{-1})^{-1} s_k^+ \\
&= F_{k-1}^T (I + \mathcal{I}_{bk}^+ Q_{k-1})^{-1} s_k^+ \quad (9.111)
\end{aligned}$$

Rearranging this equation gives

$$(I + \mathcal{I}_{bk}^+ Q_{k-1}) F_{k-1}^{-T} s_{k-1}^- = s_k^+ \quad (9.112)$$

Multiplying out this equation, and premultiplying both sides by $F_{k-1}^{-1} P_{bk}^+$, gives

$$F_{k-1}^{-1} P_{bk}^+ F_{k-1}^{-T} s_{k-1}^- + F_{k-1}^{-1} Q_{k-1} F_{k-1}^{-T} s_{k-1}^- = F_{k-1}^{-1} P_{bk}^+ s_k^+ \quad (9.113)$$

Substituting for $F_{k-1}^{-1}Q_{k-1}F_{k-1}^{-T}$ from Equation (9.92) gives

$$\begin{aligned} F_{k-1}^{-1}P_{bk}^+F_{k-1}^{-T}s_{k-1}^- + F_{k-1}^{-1}P_{fk}^-F_{k-1}^{-T}s_{k-1}^- - P_{f,k-1}^+s_{k-1}^- &= F_{k-1}^{-1}P_{bk}^+s_k^+ \\ \left[F_{k-1}^{-1}(P_{fk}^- + P_{bk}^+)F_{k-1}^{-T} - P_{f,k-1}^+ \right] s_{k-1}^- &= F_{k-1}^{-1}P_{bk}^+s_k^+ \\ \left[(P_{fk}^- + P_{bk}^+)F_{k-1}^{-T} - F_{k-1}P_{f,k-1}^+ \right] s_{k-1}^- &= P_{bk}^+s_k^+ \end{aligned} \quad (9.114)$$

Substituting in this expression for P_{bk}^+ from Equation (9.95) gives

$$\left[(P_{fk}^- + P_{bk}^+)F_{k-1}^{-T} - F_{k-1}P_{f,k-1}^+ \right] s_{k-1}^- = (P_{fk}^- + P_{bk}^+)\mathcal{I}_{fk}^-P_k s_k^+ \quad (9.115)$$

Substituting for $(P_{fk}^- + P_{bk}^+)$ from Equation (9.97) on both sides of this expression gives

$$\left[F_{k-1}(P_{f,k-1}^+ + P_{b,k-1}^-) - F_{k-1}P_{f,k-1}^+ \right] s_{k-1}^- = F_{k-1}(P_{f,k-1}^+ + P_{b,k-1}^-)F_{k-1}^T\mathcal{I}_{fk}^-P_k s_k^+ \quad (9.116)$$

Premultiplying both sides by $(P_{f,k-1}^+ + P_{b,k-1}^-)^{-1}F_{k-1}^{-1}$ gives

$$\left[I - (P_{f,k-1}^+ + P_{b,k-1}^-)^{-1}P_{f,k-1}^+ \right] s_{k-1}^- = F_{k-1}^T\mathcal{I}_{fk}^-P_k s_k^+ \quad (9.117)$$

Substituting Equation (9.105) for $(P_{f,k-1}^+ + P_{b,k-1}^-)^{-1}$ gives

$$s_{k-1}^- - F_{k-1}^T\mathcal{I}_{fk}^-(P_{fk}^- - P_k)\mathcal{I}_{fk}^-F_{k-1}P_{f,k-1}^+s_{k-1}^- = F_{k-1}^T\mathcal{I}_{fk}^-P_k s_k^+ \quad (9.118)$$

Now from Equation (9.105) we see that

$$-(P_{f,k-1}^+ + P_{b,k-1}^-)^{-1}F_{k-1}^{-1}\hat{x}_{fk}^- = F_{k-1}^T\mathcal{I}_{fk}^-(P_k - P_{fk}^-)\mathcal{I}_{fk}^-\hat{x}_{fk}^- \quad (9.119)$$

So we can add the two sides of this equation to the two sides of Equation (9.118) to get

$$\begin{aligned} s_{k-1}^- - F_{k-1}^T\mathcal{I}_{fk}^-(P_{fk}^- - P_k)\mathcal{I}_{fk}^-F_{k-1}P_{f,k-1}^+s_{k-1}^- - (P_{f,k-1}^+ + P_{b,k-1}^-)^{-1}F_{k-1}^{-1}\hat{x}_{fk}^- \\ = F_{k-1}^T\mathcal{I}_{fk}^-P_k s_k^+ + F_{k-1}^T\mathcal{I}_{fk}^-(P_k - P_{fk}^-)\mathcal{I}_{fk}^-\hat{x}_{fk}^- \end{aligned} \quad (9.120)$$

Now use Equation (9.100) to substitute for $P_k s_k^+$ in the above equation and obtain

$$\begin{aligned} s_{k-1}^- - F_{k-1}^T\mathcal{I}_{fk}^-(P_{fk}^- - P_k)\mathcal{I}_{fk}^-F_{k-1}P_{f,k-1}^+s_{k-1}^- - (P_{f,k-1}^+ + P_{b,k-1}^-)^{-1}F_{k-1}^{-1}\hat{x}_{fk}^- \\ = F_{k-1}^T\mathcal{I}_{fk}^-\hat{x}_k - F_{k-1}^T\mathcal{I}_{fk}^-P_k\mathcal{I}_{fk}^+\hat{x}_{fk} + \\ F_{k-1}^T\mathcal{I}_{fk}^-P_kH_k^TR_k^{-1}H_k\hat{x}_{fk} + F_{k-1}^T\mathcal{I}_{fk}^-(P_k - P_{fk}^-)\mathcal{I}_{fk}^-\hat{x}_{fk} \end{aligned} \quad (9.121)$$

Rearrange this equation to obtain

$$\begin{aligned} s_{k-1}^- - F_{k-1}^T\mathcal{I}_{fk}^-(P_{fk}^- - P_k)\mathcal{I}_{fk}^-F_{k-1}P_{f,k-1}^+s_{k-1}^- + \\ \left[-(P_{f,k-1}^+ + P_{b,k-1}^-)^{-1}F_{k-1}^{-1} + F_{k-1}^T\mathcal{I}_{fk}^-P_k\mathcal{I}_{fk}^+ - F_{k-1}^T\mathcal{I}_{fk}^-P_kH_k^TR_k^{-1}H_k - \right. \\ \left. F_{k-1}^T\mathcal{I}_{fk}^-P_k\mathcal{I}_{fk}^- \right] \hat{x}_{fk}^- = F_{k-1}^T\mathcal{I}_{fk}^-(\hat{x}_k - \hat{x}_{fk}^-) \end{aligned} \quad (9.122)$$

From Equation (9.106) we see that $\mathcal{I}_{fk}^+ - \mathcal{I}_{fk}^- = H_k^T R_k^{-1} H_k$. Also note that part of the coefficient of \hat{x}_{fk}^- on the left side of the above equation can be expressed as

$$(P_{f,k-1}^+ + P_{b,k-1}^-)^{-1} F_{k-1}^{-1} = \mathcal{I}_{b,k-1}^- (I + P_{f,k-1}^+ \mathcal{I}_{b,k-1}^-)^{-1} F_{k-1}^{-1} \quad (9.123)$$

From Equation (9.67) we see that $F_{k-1}^{-1} \hat{x}_{fk}^- = \hat{x}_{f,k-1}^+$. Therefore Equation (9.122) can be written as

$$\begin{aligned} s_{k-1}^- - F_{k-1}^T \mathcal{I}_{fk}^- (P_{fk}^- - P_k) \mathcal{I}_{fk}^- F_{k-1} P_{f,k-1}^+ s_{k-1}^- - \\ \mathcal{I}_{b,k-1}^- (I + P_{f,k-1}^+ \mathcal{I}_{b,k-1}^-)^{-1} \hat{x}_{f,k-1}^+ = F_{k-1}^T \mathcal{I}_{fk}^- (\hat{x}_k - \hat{x}_{fk}^-) \end{aligned} \quad (9.124)$$

Now substitute for P_k from Equation (9.90) and use Equation (9.91) in the above equation to obtain

$$\begin{aligned} s_{k-1}^- - F_{k-1}^T \mathcal{I}_{fk}^- (P_{fk}^- - P_{fk}^+ + K_k P_{f,k+1}^- K_k^T - K_k P_{k+1} K_k^T) K_{k-1}^T s_{k-1}^- - \\ \mathcal{I}_{b,k-1}^- (I + P_{f,k-1}^+ \mathcal{I}_{b,k-1}^-)^{-1} \hat{x}_{f,k-1}^+ = F_{k-1}^T \mathcal{I}_{fk}^- (\hat{x}_k - \hat{x}_{fk}^-) \end{aligned} \quad (9.125)$$

Premultiplying both sides by $P_{f,k-1}^+$ gives

$$\begin{aligned} \left[P_{f,k-1}^+ - P_{f,k-1}^+ F_{k-1}^T \mathcal{I}_{fk}^- (P_{fk}^- - P_{fk}^+ + K_k P_{f,k+1}^- K_k^T - K_k P_{k+1} K_k^T) K_{k-1}^T \right] s_{k-1}^- \\ - P_{f,k-1}^+ \mathcal{I}_{b,k-1}^- (I + P_{f,k-1}^+ \mathcal{I}_{b,k-1}^-)^{-1} \hat{x}_{f,k-1}^+ = K_{k-1} (\hat{x}_k - \hat{x}_{fk}^-) \end{aligned} \quad (9.126)$$

Now use Equation (9.91) to notice that the coefficient of s_{k-1}^- on the left side of the above equation can be written as

$$P_{f,k-1}^+ - K_{k-1} (P_{fk}^- - P_{fk}^+ + K_k P_{f,k+1}^- K_k^T - K_k P_{k+1} K_k^T) K_{k-1}^T \quad (9.127)$$

Using Equation (9.90) to substitute for $K_k P_{k+1} K_k^T$ allows us to write the above expression as

$$\begin{aligned} P_{f,k-1}^+ - K_{k-1} (P_{fk}^- - P_{fk}^+ + K_k P_{f,k+1}^- K_k^T - P_k + P_{fk}^+ - K_k P_{f,k+1}^- K_k^T) K_{k-1}^T \\ = P_{f,k-1}^+ - K_{k-1} P_{fk}^- K_{k-1}^T + K_{k-1} P_k K_{k-1}^T \\ = P_{f,k-1}^+ - K_{k-1} P_{fk}^- K_{k-1}^T + P_{k-1} - P_{f,k-1}^+ + K_{k-1} P_{fk}^- K_{k-1}^T \\ = P_{k-1} \end{aligned} \quad (9.128)$$

Since this is the coefficient of s_{k-1}^- in Equation (9.126), we can write that equation as

$$P_{k-1} s_{k-1}^- - P_{f,k-1}^+ \mathcal{I}_{b,k-1}^- (I + P_{f,k-1}^+ \mathcal{I}_{b,k-1}^-)^{-1} \hat{x}_{f,k-1}^+ = K_{k-1} (\hat{x}_k - \hat{x}_{fk}^-) \quad (9.129)$$

Now from Equations (9.78) and (9.82) we see that

$$\begin{aligned} \hat{x}_k &= (\mathcal{I}_{fk}^+ + \mathcal{I}_{bk}^-)^{-1} \mathcal{I}_{fk}^+ \hat{x}_{fk}^+ + P_k \mathcal{I}_{bk}^- \hat{x}_{bk}^- \\ &= (I + P_{fk}^+ \mathcal{I}_{bk}^-)^{-1} \hat{x}_{fk}^+ + P_k s_k^- \end{aligned} \quad (9.130)$$

From this we see that

$$\begin{aligned} \hat{x}_k - \hat{x}_{fk}^+ &= \left[(I + P_{fk}^+ \mathcal{I}_{bk}^-)^{-1} - I \right] \hat{x}_{fk}^+ + P_k s_k^- \\ &= \left[I - (I + P_{fk}^+ \mathcal{I}_{bk}^-) \right] (I + P_{fk}^+ \mathcal{I}_{bk}^-)^{-1} \hat{x}_{fk}^+ + P_k s_k^- \\ &= -P_{fk}^+ \mathcal{I}_{bk}^- (I + P_{fk}^+ \mathcal{I}_{bk}^-)^{-1} \hat{x}_{fk}^+ + P_k s_k^- \end{aligned} \quad (9.131)$$

Rewriting the above equation with the time subscripts $(k-1)$ and then substituting for the left side of Equation (9.129) gives

$$\hat{x}_{k-1} - \hat{x}_{f,k-1}^+ = K_{k-1}(\hat{x}_k - \hat{x}_{fk}^-) \quad (9.132)$$

from which we can write

$$\hat{x}_k = \hat{x}_{fk}^+ + K_k(\hat{x}_{k+1} - \hat{x}_{f,k+1}^-) \quad (9.133)$$

This gives the smoothed estimate \hat{x}_k without needing to explicitly calculate the backward estimate. The RTS smoother is implemented by first running the standard Kalman filter of Equation (9.67) forward in time to the final time, and then implementing Equations (9.90), (9.91), and (9.133) backward in time. The RTS smoother can be summarized as follows.

The RTS smoother

1. The system model is given as follows:

$$\begin{aligned} x_k &= F_{k-1}x_{k-1} + G_{k-1}u_{k-1} + w_{k-1} \\ y_k &= H_kx_k + v_k \\ w_k &\sim (0, Q_k) \\ v_k &\sim (0, R_k) \end{aligned} \quad (9.134)$$

2. Initialize the forward filter as follows:

$$\begin{aligned} \hat{x}_{f0} &= E(x_0) \\ P_{f0}^+ &= E[(x_0 - \hat{x}_{f0})(x_0 - \hat{x}_{f0})^T] \end{aligned} \quad (9.135)$$

3. For $k = 1, \dots, N$ (where N is the final time), execute the standard forward Kalman filter:

$$\begin{aligned} P_{fk}^- &= F_{k-1}P_{f,k-1}^+F_{k-1}^T + Q_{k-1} \\ K_{fk} &= P_{fk}^-H_k^T(H_kP_{fk}^-H_k^T + R_k)^{-1} \\ &= P_{fk}^+H_k^TR_k^{-1} \\ \hat{x}_{fk}^- &= F_{k-1}\hat{x}_{f,k-1}^+ + G_{k-1}u_{k-1} \\ \hat{x}_{fk}^+ &= \hat{x}_{fk}^- + K_{fk}(y_k - H_k\hat{x}_{fk}^-) \\ P_{fk}^+ &= (I - K_{fk}H_k)P_{fk}^-(I - K_{fk}H_k)^T + K_{fk}R_kK_{fk}^T \\ &= \left[(P_{fk}^-)^{-1} + H_k^TR_k^{-1}H_k \right]^{-1} \\ &= (I - K_{fk}H_k)P_{fk}^- \end{aligned} \quad (9.136)$$

4. Initialize the RTS smoother as follows:

$$\begin{aligned} \hat{x}_N &= \hat{x}_{fN}^+ \\ P_N &= P_{fN}^+ \end{aligned} \quad (9.137)$$

5. For $k = N - 1, \dots, 1, 0$, execute the following RTS smoother equations:

$$\begin{aligned}
 \mathcal{I}_{f,k+1}^- &= \left(P_{f,k+1}^-\right)^{-1} \\
 K_k &= P_{fk}^+ F_k^T \mathcal{I}_{f,k+1}^- \\
 P_k &= P_{fk}^+ - K_k (P_{f,k+1}^- - P_{k+1}) K_k^T \\
 \hat{x}_k &= \hat{x}_{fk}^+ + K_k (\hat{x}_{k+1} - \hat{x}_{f,k+1}^-)
 \end{aligned} \tag{9.138}$$

9.5 SUMMARY

In this chapter we derived the optimal smoothing filters. These filters, sometimes called retrodiction filters [Bar01], include the following variants.

- $\hat{x}_{j,k} = E(x_j | y_1, \dots, y_{k-1})$ ($k \geq j$) is the output of the fixed-point smoother. In this filter we find the estimate of the state at the fixed time j when measurements continue to arrive at the filter at times greater than j . The time index j is fixed while k continues to increase as we obtain more measurements.
- $\hat{x}_{k-N,k} = E(x_{k-N} | y_1, \dots, y_k)$ for a fixed N is the output of the fixed-lag smoother. In this filter we find the estimate of the state at each time k while using measurements up to and including time $(k + N)$. The time index k varies while N remains fixed.
- $\hat{x}_{k,N} = E(x_k | y_1, \dots, y_N)$ for a fixed N is the output of the fixed-interval smoother. In this filter we find the estimate of the state at each time k while using measurements up to and including time N . The time index k varies while the total number of measurements N is fixed. The two formulas we derived for this type of smoothing included the forward-backward smoother and the RTS smoother.

Just as steady-state filters can be used for standard filtering, we can also derive steady-state smoothers to save computational effort [Gel74]. An early survey of smoothing algorithms is given in [Med73].

PROBLEMS

Written exercises

9.1 Prove or disprove the following conjecture: The trace of the inverse of a matrix is equal to the inverse of the trace of the matrix.

9.2 Show that $(A + B)^{-1} = B^{-1}(AB^{-1} + I)^{-1}$.

9.3 Derive Equation (9.83).

9.4 Consider a scalar system with $F = 1$, $H = 1$, and $R = 2Q$.

- a) What is the steady-state value of the *a priori* estimation-error covariance P_k^- ?

- b) Suppose that after the Kalman filter has reached steady state, the fixed-point smoother begins to operate. Find a closed-form solution to the covariance of the smoothed estimate Π_k as a function of the time index k . What is the limiting value of Π_k as $k \rightarrow \infty$?

9.5 Repeat Problem 9.4 for the case $R = 12Q$. What is the percent improvement in the estimation-error covariance due to smoothing? Explain why the percent improvement due to smoothing for this case differs in the way that it does from the results of Problem 9.4.

9.6 Consider a scalar system with $F = 1$, $H = 1$, and $R = 2Q$. Suppose that the fixed-lag smoother for this system is in steady state so that $P_{k+1}^- = P_k^-$, $L_{k+1,i} = L_{k,i}$, $P_{k+1}^{i,i} = P_k^{i,i}$, and $P_{k+1}^{0,i} = P_k^{0,i}$, for $i = 1, \dots, N+1$. Find closed-form expressions for P_k^- , $L_{k,i}$, $P_k^{i,i}$, and $P_k^{0,i}$ as functions of i . What is the limit as $i \rightarrow \infty$ of $L_{k,i}$, $P_k^{i,i}$, and $P_k^{0,i}$?

9.7 Suppose you have a fixed-lag smoother as shown in Equation (9.43) that is in steady state. How do the eigenvalues of the fixed-lag smoother relate to the eigenvalues of the standard Kalman filter? What do you conclude about the stability of the fixed-lag smoother?

9.8 Solve Equation (9.10) for $(y_k - H_k \hat{x}_k^-)$ [assuming that $\rho(L_k) = r$, where r is the number of measurements in the system]. Substitute the resulting expression for $(y_k - H_k \hat{x}_k^-)$ in the fixed-lag smoother equation for $\hat{x}_{k+1-i,k}$ to show that the smoothed state estimate can be driven by the state estimates without any input from the measurements [And79].

9.9 Suppose that \hat{x}_f and \hat{x}_b are unbiased estimates of x , and $\hat{x} = K_f \hat{x}_f + K_b \hat{x}_b$. Show that if \hat{x} is an unbiased estimate of x , then we must have $K_f + K_b = I$.

9.10 Consider a scalar system with $F = 1$, $H = 1$, and $R = 2Q$. Use the forward-backward smoother of Section 9.4.1 to find the steady-state value of the covariance of the smoothed state estimate.

9.11 Consider a scalar system with $F = 1$, $H = 1$, and $R = 2Q$. Use the RTS smoother of Section 9.4.2 to find the steady-state value of the covariance of the smoothed state estimate.

9.12 Consider a scalar system with $F = 1$, $H = 1$, and $R = 2Q$. Suppose that the forward filter has reached steady state. Use the RTS smoother of Section 9.4.2 to find the covariance of the smoothed state estimate for $k = N, N-1, N-2, N-3$, and $N-4$. At what point does the covariance of the smoothed state estimate get within 1% of its steady-state value?

9.13 Repeat Problem 9.12 for $R = 12Q$. How do you intuitively explain the quicker convergence of P_k to steady state?

9.14 Use the RTS smoother equations to show that constant states are not smoothable. That is, if $F = I$ and $Q = 0$, then $P_k = P_{fN}^+$ for all k .

Computer exercises

9.15 Consider the second-order system

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\omega^2 & -2\zeta\omega \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} w(t)$$

where $\omega = 6$ rad/s is the natural frequency of the system, and $\zeta = 0.16$ is the damping ratio. The input $w(t)$ is continuous-time white noise with a variance of 0.01. Measurements of the first state are taken every 0.5 s:

$$y(t_k) = \begin{bmatrix} 1 & 0 \end{bmatrix} x(t_k) + v(t_k)$$

where $v(t_k)$ is discrete-time white noise with a variance of 10^{-4} . The initial state, estimate, and covariance are

$$\begin{aligned} x(0) &= \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ \hat{x}(0) &= x(0) \\ P(0) &= \begin{bmatrix} 10^{-5} & 0 \\ 0 & 10^{-2} \end{bmatrix} \end{aligned}$$

- a) Discretize the system equation.
- b) Implement the discrete-time Kalman filter and the RTS smoother for 10 s (20 time steps). Plot the variance of the estimation error of the first state for the forward filter and for the RTS smoother on a single plot. Do the same for the second state. Why is the second state more smoothable than the first state?

9.16 Repeat Problem 9.15 with the continuous-time process noise $w(t)$ having a variance of 1. How does this change the smoothability of the states?

9.17 Design a fixed-interval smoother for the system described in Problem 5.11 to estimate the state at each time on the basis of measurements at all 10 time steps.

- a) Plot the *a posteriori* covariance of the forward state estimate and the covariance of the smoothed state estimate as a function of time for both states.
- b) What are the percent improvements in the estimation-error variances due to smoothing for the two states at the initial time? Why is there so much more improvement for one state than for the other state?
- c) Simulate the system and smoother a hundred times or so, each simulation with a different noise history. On the basis of your simulations, derive a numerical estimate of the smoother estimation-error variances of the two states at the initial time. How do your numerical variances compare with the theoretical variances obtained in part (b)?