Notes on Fixed-Lag Smoothing

CDS 112

1 Two-Step Kalman Filter Equations

Consider linear discrete time dynamical systems of the form:

$$x_{k+1} = A_k \, x_k + G_k \, \eta_k; \tag{1}$$
$$y_{k+1} = H_{k+1} x_{k+1} + \omega_{k+1}; \tag{2}$$

where $x \in \mathbb{R}^p$ and $y \in \mathbb{R}^m$. The process noise, η_k, and measurement noise, ω_k are assumed to be zero mean, white, Gaussian random processes:

$$\eta_k \sim N(0, Q_k) \qquad \omega_k \sim N(0, R_k) \; . \tag{3}$$

The 2-step Kalman filter equations for this system are:

Dynamic Update Step

$$\hat{x}_{k+1|k} = A_k \, \hat{x}_{k|k} \tag{4}$$
$$\Sigma_{k+1|k} = A_k \, \Sigma_{k|k} \, A_k^T + G_k \, Q_k \, G_k^T \tag{5}$$

where $\hat{x}_{k+1|k}$ is the state estimate at time t_{k+1} given measurements up until time t_k, and $\Sigma_{k+1|k}$ is the covariance of the state estimate.

Measurement Update Step

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_k \, (y_{k+1} - H_{k+1} \, \hat{x}_{k+1|k}) \tag{6}$$
$$K_{k+1} = \Sigma_{k+1|k} \, H_{k+1}^T \, (H_{k+1} \Sigma_{k+1|k} \, H_{k+1}^T + R_k)^{-1} \tag{7}$$
$$\Sigma_{k+1|k+1} = \Sigma_{k+1|k} - \Sigma_{k+1|k} \, H_{k+1}^T \, (H_{k+1} \Sigma_{k+1|k} \, H_{k+1}^T + R_k)^{-1} H_{k+1} \Sigma_{k+1|k} \tag{8}$$
$$= (I - K_{k+1} \, H_{k+1}) \Sigma_{k+1|k}$$
$$= \Sigma_{k+1|k} \, (I - H_{k+1}^T K_{k+1}^T) \tag{9}$$

where K_{k+1} is the Kalman filter gain matrix at time t_{k+1}

2 One-Step Kalman Filter Equations

The two Kalman filtering steps can be practically merged into one step as follows. First, the dynamic propagation and measurement update equations can be merged:

$$\hat{x}_{k+1|k} = A_k \, \hat{x}_{k|k} = A_k [\hat{x}_{k|k-1} + K_k \, (y_k - H_k \, \hat{x}_{k|k-1})]$$
$$= (A_k - L_k \, H_k) \hat{x}_{k|k-1} + L_k \, y_k \tag{10}$$

where we have introduced the symbol L_k as shorthand notation for:

$$L_k = A_k K_k \tag{11}$$

where K_k is the Kalman gain matrix at time t_k

Next, the covariance update across the two steps can be similarly merged by substituting Equation (9) into Equation (5):

$$\begin{aligned}\Sigma_{k+1|k} &= A_k \Sigma_{k|k} A_k + G_k Q_k G_k^T \\ &= A_k \Sigma_{k|k-1} [I - H_k^T K_k^T] A_k^T + G_k Q_k G_k^T \\ &= A_k \Sigma_{k|k-1} [A_k - L_k H_k]^T + G_k Q_k G_k^T \end{aligned} \tag{12}$$

3 A Fixed-Lag Smoother: Background

Our goal is to compute the estimate $\hat{x}_{k-N|k}$ for some fixed "time lag" N. That is, we want to estimate the state at time t_{k-N} given measurements until t_k. There are *many* fixed lag smoothing techniques. This note develops an approach which will produce smoothed estimates for all states in the fixed smoothing window: $\hat{x}_{k-N|k}$, $\hat{x}_{k-N+1|k}$, $\cdots$, $\hat{x}_{k|k}$.

The key idea is to defined an *augmented* state vector, and an associated linear discrete time *augmented dynamical system*. The augmented state vector, $\vec{X}^a \in \mathbb{R}^{p(N+2)}$ is defined and governed by the following discrete time state equation:

$$\vec{X}^a_{k+1} = \begin{bmatrix} x^{(0)}_{k+1} \\ x^{(1)}_{k+1} \\ x^{(2)}_{k+1} \\ \vdots \\ x^{(N+1)}_{k+1} \end{bmatrix} = \begin{bmatrix} A_k & 0 & 0 & \cdots & 0 \\ I & 0 & 0 & \cdots & 0 \\ 0 & I & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & I & 0 \end{bmatrix} \begin{bmatrix} x^{(0)}_{k} \\ x^{(1)}_{k} \\ x^{(2)}_{k} \\ \vdots \\ x^{(N+1)}_{k} \end{bmatrix} + \begin{bmatrix} G_k \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \eta_k \tag{13}$$

while the *augmented measurement equation* takes the form:

$$y_k = \begin{bmatrix} H_k & 0 & 0 & \cdots & 0 \end{bmatrix} \vec{X}^a_k + \omega_k \tag{14}$$

where the terms A_k, G_k, H_k, η_k, and ω_k are the same as those in the dynamic equations (1).

If $x^{(0)}_k$ is identified with the state of the dynamic system (1) at time t_k, then the components of the augmented vector $\vec{X}^a_{k+1}$ are essentially sequentially delayed versions of the state vector of system (1). From Equation (13) we can see that the first row of the augmented dynamic equations implements the dynamic system (1), while the remaining rows implement successive time delays:

$$
\begin{array}{rcl}
x_{k+1}^{(0)} & = & A_k x_k^{(0)} \; + \; G_k \; \eta_k \; = \; A_k x_k \; + \; G_k \; \eta_k \; = \; \text{state of system (1) at } t_{k+1} \\
x_{k+1}^{(1)} & = & x_k^{(0)} \; = \; x_k \; = \; \text{state of system (1) at } t_k \\
x_{k+1}^{(2)} & = & x_k^{(1)} \; = \; x_{k-1}^{(0)} \; = \; x_{k-1} \; = \; \text{state of system (1) at } t_{k-1} \\
 & \vdots & \\
x_{k+1}^{(i)} & = & x_k^{(i-1)} \; = \; \cdots \; = \; x_{k-i+1} \; = \; \text{state of system (1) at } t_{k-i+1} \\
 & \vdots & \\
x_{k+1}^{(N+1)} & = & x_k^{(N)} \; = \; x_{k-1}^{(N-1)} \; = \; \cdots \; = \; x_{k-N} \; = \; \text{state of system (1) at } t_{k-N}
\end{array}
$$

The last component of the augmented state vector at time t_{k+1} is the actual state for which we wish to provided a smoothed estimate. Hence, if we can find an estimate for $\hat{\vec{X}}^a_{k+1|k}$, then the last component of this estimate is the estimate $\hat{x}_{k-N|k}$ that we seek.

Because the augmented system satisfies all of the requirements needed to apply Kalman Filtering theory (the dynamics are discrete time and linear, and we have previously proved that all of the states will be Gaussian random variables), we can construct the smoother by constructing the Kalman filter for the augmented dynamical system. This approach has two benefits: (1) one need only understand the Kalman filter to construct the smoother; (2) the smoother will construct $\hat{x}_{l|k}$ and its covariance for $l = k-N, k-N+1, \ldots, k$ (i.e., all of the states in the smoothing window). This property is an advantage for some applications.

4 Derivation of the Fixed-Lag Smoother

To construct the fixed-lag smoothing equations, substitute the augmented dynamic equation (13) into the one-step Kalman filter equations (10) and (12) of Section 2, and then isolate the terms of interest.

Notation: Let's introduce some additional notation to simplify the bookkeeping process during the derivations. Let the augmented system (13) and (14) have the symbolic forms:

$$
\begin{array}{rcl}
\vec{X}^a_{k+1} & = & A^a_k \; \vec{X}^a_k \; + \; G^a_k \; \eta_k \\
y_{k+1} & = & H^a_{k+1} \; \vec{X}^a_{k+1} \; + \; \omega_{k+1}
\end{array} \qquad (15)
$$

where the matrices A^a_k, G^a_k, and H^a_k have the obvious definitions via Equations (13) and (14).

$$
A^a_k = \begin{bmatrix} A_k & 0 & 0 & \cdots & 0 \\ I & 0 & 0 & \cdots & 0 \\ 0 & I & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & I & 0 \end{bmatrix}; \qquad G^a_k = \begin{bmatrix} G_k \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}; \qquad H^a_k = \begin{bmatrix} H_k & 0 & 0 & \cdots & 0 \end{bmatrix}. \qquad (16)
$$

Let the covariance of the augmented state estimate and the augmented Kalman gain matrix have the forms:

$$\Sigma^a_{k+1|k} = \begin{bmatrix} \Sigma^{(0,0)}_{k+1|k} & \Sigma^{(0,1)}_{k+1|k} & \cdots & \Sigma^{(0,N+1)}_{k+1|k} \\ \Sigma^{(1,0)}_{k+1|k} & \Sigma^{(1,1)}_{k+1|k} & \cdots & \Sigma^{(1,N+1)}_{k+1|k} \\ \vdots & \vdots & \cdots & \vdots \\ \Sigma^{(N+1,0)}_{k+1|k} & \Sigma^{(N+1,1)}_{k+1|k} & \cdots & \Sigma^{(N+1,N+1)}_{k+1|k} \end{bmatrix} \qquad L^a_k = \begin{bmatrix} L^{(0)}_k \\ L^{(1)}_k \\ \vdots \\ L^{(N+1)}_k \end{bmatrix} \tag{17}$$

where $\Sigma^a_{k+1|k} = E[(\vec{X}^a_{k+1} - \hat{\vec{X}}^a_{k+1|k})(\vec{X}^a_{k+1} - \hat{\vec{X}}^a_{k+1|k})^T]$, and thus the superscripts (i,j) refer to the i^{th} and j^{th} components of the augmented state vectors. Since this covariance matrix is necessarily symmetric, $\Sigma^{(i,j)}_{k+1|k} = (\Sigma^{(j,i)}_{k+1|k})^T$.

Note that:

$$\begin{aligned} \Sigma^{(0,0)}_{k+1|k} &= \Sigma_{k+1|k}, \text{ the covariance associated to } \hat{x}_{k+1|k} \text{ of system (1)} \\ \Sigma^{(i+1,i+1)}_{k+1|k} &= \Sigma_{k-i|k}, \text{ the covariance associated to } \hat{x}_{k-i|k} \text{ of system (1)} \\ \Sigma^{(N+1,N+1)}_{k+1|k} &= \Sigma_{k-N|k}, \text{ the covariance associated to } \hat{x}_{k-N|k} \text{ of system (1)} \end{aligned}$$

Dynamic Update Equation: Substituting the notations introduced above into the one-step state update equation (10)

$$\hat{\vec{X}}^a_{k+1|k} = A^a_k \, \hat{\vec{X}}_{k|k-1} + L^a_k \, [y_k - H^a_k \hat{\vec{X}}^a_{k|k-1}] \tag{18}$$

yields:

$$\begin{bmatrix} \hat{x}^{(0)}_{k+1|k} \\ \hat{x}^{(1)}_{k+1|k} \\ \vdots \\ \hat{x}^{(N+1)}_{k+1|k} \end{bmatrix} = \begin{bmatrix} A_k & 0 & \cdots & 0 & 0 \\ I & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & 0 & 0 \\ 0 & 0 & \cdots & I & 0 \end{bmatrix} \begin{bmatrix} \hat{x}^{(0)}_{k|k-1} \\ \hat{x}^{(1)}_{k|k-1} \\ \vdots \\ \hat{x}^{(N+1)}_{k|k-1} \end{bmatrix} + \begin{bmatrix} L^{(0)}_k \\ L^{(1)}_k \\ \vdots \\ L^{(N+1)}_k \end{bmatrix} \left(y_k - \begin{bmatrix} H_k & 0 & 0 & \cdots & 0 \end{bmatrix} \hat{\vec{X}}^a_{k|k-1}\right) \tag{19}$$

The rows of this equation define one-step updates for each of the augmented state components:

$$\begin{aligned} \hat{x}^{(0)}_{k+1|k} &= A_k \hat{x}^{(0)}_{k|k-1} + L^{(0)}_k \, (y_k - H_k \, \hat{x}^{(0)}_{k|k-1}) \\ &\vdots \\ \hat{x}^{(i+1)}_{k+1|k} &= \hat{x}^{(i)}_{k|k-1} + L^{(i+1)}_k \, (y_k - H_k \, \hat{x}^{(0)}_{k|k-1}) \\ &\vdots \\ \hat{x}^{(N+1)}_{k+1|k} &= \hat{x}^{(N)}_{k|k-1} + L^{(N+1)}_k \, (y_k - H_k \, \hat{x}^{(0)}_{k|k-1}) \end{aligned}$$

which is equivalent to the following filtering operations on the states of the original dynamical system:

$$\begin{aligned}
\hat{x}_{k+1|k} &= A_k \hat{x}_{k|k-1} + L_k^{(0)} (y_k - H_k \hat{x}_{k|k-1}) \\
&\vdots \\
\hat{x}_{k-i|k} &= \hat{x}_{k-i|k-1} + L_k^{(i+1)} (y_k - H_k \hat{x}_{k|k-1}) \\
&\vdots \\
\hat{x}_{k-N|k} &= \hat{x}_{k-N|k-1} + L_k^{(N+1)} (y_k - H_k \hat{x}_{k|k-1})
\end{aligned}$$

The first equation above is simply the one-step Kalman filter state estimate update for (1), while the subsequent expressions show how the smoothing filter uses the measurment at t_k to update the state estimates at earlier times in the smoothing window.

Filter Gains: Next let's compute the gain matrices, $\{L_k^{(i)}\}$ using the definition of $L_k^a = A_k^a K_k^a$ and the Kalman gain definition from (7):

$$L_k^a = A_k^a \Sigma_{k|k-1}^a H_k^{a^T} \left(H_k^a \Sigma_{k|k-1}^a H_k^{a^T} + R_k\right)^{-1} . \tag{20}$$

Substituting the appropriate notation into Equation (20) yields:

$$\begin{bmatrix} L_k^{(0)} \\ L_k^{(1)} \\ \vdots \\ L_k^{(N+1)} \end{bmatrix} = \begin{bmatrix} A_k & 0 & \cdots & 0 & 0 \\ I & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & 0 & 0 \\ 0 & 0 & \cdots & I & 0 \end{bmatrix} \begin{bmatrix} \Sigma_{k|k-1}^{(0,0)} & \cdots & \Sigma_{k|k-1}^{(0,N+1)} \\ \vdots & \cdots & \vdots \\ \Sigma_{k|k-1}^{(N+1,0)} & \cdots & \Sigma_{k|k-1}^{(N+1,N+1)} \end{bmatrix} \begin{bmatrix} H_k^T \\ 0 \\ \vdots \\ 0 \end{bmatrix} * \tag{21}$$

$$\left(\begin{bmatrix} H_k & \cdots & 0 \end{bmatrix} \Sigma_{k|k-1}^a \begin{bmatrix} H_k^T \\ \vdots \\ 0 \end{bmatrix} + R_k \right)^{-1} \tag{22}$$

which results in:

$$\begin{aligned}
L_k^{(0)} &= A_k \Sigma_{k|k-1}^{(0,0)} H_k^T (H_k \Sigma_{k|k-1}^{(0,0)} H_k^T + R_k)^{-1} \\
&= A_k \Sigma_{k|k-1} H_k^T (H_k \Sigma_{k|k-1} H_k^T + R_k)^{-1} \\
&= A_k K_k
\end{aligned} \tag{23}$$

where K_k is the usual Kalman gain matrix of the two-step Kalman filter for system (1). For $i = 0, \ldots, N$ the smoothing filter gains are:

$$\begin{aligned}
L_k^{(i+1)} &= \Sigma_{k|k-1}^{(i,0)} H_k^T (H_k \Sigma_{k|k-1}^{(0,0)} H_k^T + R_k)^{-1} \\
&= \Sigma_{k|k-1}^{(i,0)} H_k^T (H_k \Sigma_{k|k-1} H_k^T + R_k)^{-1}
\end{aligned} \tag{24}$$

Covariance Update: Finally, the covariance of the smoothed state estimates can be updated following Equation (12):

$$\Sigma_{k+1|k}^a = A_k^a \Sigma_{k|k-1}^a [A_k^a - L_k^a H_k^a]^T + G_k^a Q_k^a G_k^{a^T}$$

Expanding out this equation in terms of the notation introduced above yields:

$$\begin{bmatrix} \Sigma^{(0,0)}_{k+1|k} & \cdots & \Sigma^{(0,N+1)}_{k+1|k} \\ \vdots & \cdots & \vdots \\ \Sigma^{(N+1,0)}_{k+1|k} & \cdots & \Sigma^{(N+1,N+1)}_{k+1|k} \end{bmatrix} = \begin{bmatrix} A_k & 0 & \cdots & 0 \\ I & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & 0 \\ 0 & 0 & I & 0 \end{bmatrix} \begin{bmatrix} \Sigma^{(0,0)}_{k|k-1} & \cdots & \Sigma^{(0,N+1)}_{k|k-1} \\ \vdots & \cdots & \vdots \\ \Sigma^{(N+1,0)}_{k|k-1} & \cdots & \Sigma^{(N+1,N+1)}_{k|k-1} \end{bmatrix} *$$

$$\left(\begin{bmatrix} A_k^T & I & 0 & \cdots & 0 \\ 0 & 0 & I & \cdots & \vdots \\ \vdots & \vdots & \vdots & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix} - \begin{bmatrix} H_k^T \\ \vdots \\ 0 \end{bmatrix} \begin{bmatrix} L_k^{(0)^T} & L_k^{(1)^T} & \cdots & L_k^{(N+1)^T} \end{bmatrix} \right) + \begin{bmatrix} G_k Q_k G_k^T & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

From this expanded equation one can then read off the results that the first column of the covariance matrix can be updated as:

$$\begin{aligned} \Sigma^{(0,0)}_{k+1|k} &= A_k\, \Sigma^{(0,0)}_{k|k-1}[A_k \; - \; L_k^{(0)} \; H_k]^T + G_k \; Q_k \; G_k^T \\ \Sigma^{(1,0)}_{k+1|k} &= \Sigma^{(0,0)}_{k|k-1}[A_k \; - \; L_k^{(0)} \; H_k]^T \\ \Sigma^{(2,0)}_{k+1|k} &= \Sigma^{(1,0)}_{k|k-1}[A_k \; - \; L_k^{(0)} \; H_k]^T \\ \vdots &= \vdots \\ \Sigma^{(N+1,0)}_{k+1|k} &= \Sigma^{(N,0)}_{k|k-1}[A_k \; - \; L_k^{(0)} \; H_k]^T \end{aligned}$$

Careful scrutiny of all of the smoother equations above shows that in fact only the elements of the first column of the covariance matrix need be updated, as these are the only terms used in the smoothing equations. However, for application purposes one may also wish to update the diagonal terms covariance terms, as they indicate the quality of the estimates of all the states in the smoothing window:

$$\begin{aligned} \Sigma^{(1,1)}_{k+1|k} &= \Sigma^{(0,0)}_{k|k-1} - \Sigma^{(0,0)}_{k|k-1} \; H_k^T \; L_k^{(1)^T} \\ \Sigma^{(2,2)}_{k+1|k} &= \Sigma^{(1,1)}_{k|k-1} - \Sigma^{(1,0)}_{k|k-1} H_k^T \; L_k^{(2)^T} \\ \vdots &= \vdots \\ \Sigma^{(i,i)}_{k+1|k} &= \Sigma^{(i-1,i-1)}_{k|k-1} \; - \; \Sigma^{(i-1,0)}_{k|k-1} H_k^T \; L_k^{(i)^T} \end{aligned} \tag{25}$$

Summary: Based on these derivations, the fixed-lag smoother for system (1) is organized as follows:

1. *Initialize the states of the Fixed-Lag smoother.* This step is somewhat problematic, since a correct initialization of the complete filter requires not only an initialization of all components of the augmented state vector $\vec{X}^a$, but also of the covariance matrix $\Sigma^a_{k|k-1}$. Naively, one could run the standard one-step Kalman Filter for system (1) until

time k (where $k \geq N$). This approach will produce the correct estimate $\hat{x}^{(0)}_{k+1|k}$, as well as the correct estimate covariance component, $\Sigma^{(0,0)}_{k+1|k} = \Sigma_{k+1|k}$, and the Kalman gain K_k calculated by the filter will correctly equal $L^{(0)}_k$. However, the one-step Kalman filter estimates derived for times within the desired smoothing window do not have the correct conditioning. That is, the one-step Kalman filter will produce estimates $\hat{x}_{i|i-1}$ for $i = 1, N$. However, we need to initialize the augmented state vector of the filter with $\hat{x}_{i|k-1}$. Moreover, the covariances produced by this approach also do not have the correct conditioning, nor does this approach produce the correlations between different state delays which are needed in the update equations.

Thus, one correct approach to initialize the filter is to initialize $\hat{x}_{0|-1}$ to the initial best estimate of the state, $\overline{x}_0$, and to set $\Sigma^{(0,0)}_{0|-1} = P_0$, where P_0 is the covariance of the initial state location. All of the remaining delayed components of the augmented state vector and the augmented covariance matrix are then zeroed. With this approach, the augmented state components and augmented covariance components will not all assume their true values until $k > N$, where N is the length of the smoothing window.

2. *Update the Fixed-Lag Smoother.* For each successive k, sequentially compute the equations that were derived above for $i = 1, \ldots, N+1$:

$$
\begin{aligned}
L^{(i)}_k &= \Sigma^{(0,i-1)}_{k|k-1} \, H^T_k (H_k \, \Sigma^{(0,0)}_{k|k-1} \, H^T_k + R_k)^{-1} \\
\Sigma^{(0,i)}_{k+1|k} &= \Sigma^{(0,i-1)}_{k|k-1} [A_k \; - \; L^{(0)}_k H_k]^T \\
\hat{x}_{k+1|k} &= \hat{x}^{(i-1)}_{k|k-1} + L^{(i)}_k (y_k \; - \; H_k \, \hat{x}^{(0)}_{k|k-1})
\end{aligned}
$$

If desired, compute the diagonal covariance terms according to (25).