Technion—Israel Intitute of Technology
The Erna and Andrew Viterbi Department of Electrical Engineering

Estimation and Identification in Dynamical Systems (048825)
Lecture Notes: Prof. Nahum Shimkin, Spring 2016

7 Optimal Smoothing

The basic smoothing problem is:

$$\text{Compute } \hat{x}_{j|k} \simeq E(x_j|Z_k), \text{ with } \underline{j < k}.$$

The smoothed estimate improves the standard one, by adding "future" measurements. It is relevant for off-line estimation problems, or for on-line problems which can tolerate some delay.

In a recursive context, we can distinguish three main types of smoothing problems:

1. Fixed-lag smoothing: Estimate $\hat{x}_{j|j+\ell_0}$, for $j = 0, 1, \dots$
2. Fixed-point smoothing: Estimate $\hat{x}_{j_0|k}$, for $k = j_0, j_0 + 1, \dots$
3. Fixed-interval smoothing: Estimate $\hat{x}_{j|k_0}$, for $j = 0, \dots, k_0$

The third problem is in fact a "batch" (off-line) problem, and is the most common.

7.1 Fixed-point smoothing: Basic equations

We consider the standard (Gaussian or 2nd order) model:

$$\begin{aligned} x_{k+1} &= F_k x_k + w_k \\ z_k &= H_k x_k + v_k \end{aligned}$$

with $x_0 \sim (\overline{x}_0, P_0)$, $E(w_k w_l^T) = Q_k \delta_{kl}$, $E(v_k v_l^T) = R_k \delta_{kl}$, $w \perp v \perp x_0$.

We wish to determine, for $k \geq j$,

$$\begin{aligned} \hat{x}_{j|k} &= E(x_j | Z_k) \\ P_{j|k} &= E\{(x_j - \hat{x}_{j|k})(x_j - \hat{x}_{j|k})^T\} \end{aligned}$$

Solution 1: The problem can be tackled by reducing it to a standard (Kalman) filtering. To this end, define the augmented state:

$$X_k = \begin{bmatrix} x_k \\ x_j \end{bmatrix}, \quad k \geq j$$

and compute $\hat{X}_{k|k}$ for $k = j, j+1, \ldots$. The resulting equations can be simplified to give the equations below.

Solution 2: Here we proceed with a direct derivation, using the innovations approach. Recall that

$$\begin{aligned} E(x_j | Z_k) &= E(x_j | \tilde{Z}_k) = E(x_j | \tilde{Z}_{k-1}) + [E(x_j | \tilde{z}_k) - E(x_j)] \\ &= \hat{x}_{j|k-1} + \mathrm{cov}(x_j, \tilde{z}_k) \mathrm{cov}(\tilde{z}_k)^{-1} \tilde{z}_k \end{aligned}$$

where $\tilde{z}_k = z_k - H_k \hat{x}_{k|k-1}$. It remains to compute the two covariances.

We already know that

$$\mathrm{cov}(\tilde{z}_k) = (H_k P_k^- H_k^T + R_k) \triangleq M_k \,.$$

Further, using $\tilde{z}_k = (H_k x_k + v_k) - H_k \hat{x}_{k|k-1} = H_k \tilde{x}_{k|k-1} + v_k$,

$$\mathrm{cov}(x_j, \tilde{z}_k) = \mathrm{cov}(x_j, \tilde{x}_{k|k-1}) H_k^T \,.$$

Denote

$$\Sigma_{k,j} \triangleq \mathrm{cov}(x_j, \tilde{x}_{k|k-1}) = \mathrm{cov}(\tilde{x}_{j|j-1}, \tilde{x}_{k|k-1}) \,.$$

(the last equality holds for $k \geq j$). Then $\Sigma_{k,j}$ can computed recursively via

$$
\begin{aligned}
\Sigma_{k+1,j} &= \operatorname{cov}(x_j, \tilde{x}_{k+1|k}) \\
&= \operatorname{cov}(x_j, (F_k - K_k H_k)\tilde{x}_{k|k-1} + (w_k - K_k v_k)) \\
&= \Sigma_{k,j} \overline{F}_k^T
\end{aligned}
$$

where

$$\overline{F}_k = F_k - K_k H_k \,.$$

<u>Summarizing</u>, we obtain the following recursive update:

$$\hat{x}_{j|k} = \hat{x}_{j|k-1} + L_{k,j} \tilde{z}_k \,, \quad k = j, j+1, \ldots \tag{1}$$

where

$$L_{k,j} = \Sigma_{k,j} H_k^T M_k^{-1} \,, \tag{2}$$

$$\Sigma_{k+1,j} = \Sigma_{k,j} \overline{F}_k^T \,. \tag{3}$$

The initial conditions are $\hat{x}_{j|j-1}$ and $\Sigma_{j,j} = P_{j|j-1} \equiv P_j^-$.

Note that computing $\tilde{z}_k$ requires the "standard" estimate $\hat{x}_{k|k-1}$, hence it must be also be computed (using the standard Kalman filter).

The *covariance* of the smoothed estimator is obtained recursively by substituting (1):

$$P_{j|k} \triangleq \operatorname{cov}(x_j - \hat{x}_{j|k}) = P_{j|k-1} - \Sigma_{k,j} H_k^T L_{k,j}^T \tag{4}$$

Remarks (error improvement due to smoothing):

1. The covariance can be expressed more explicitly as

$$\begin{aligned} P_{j|k} &= P_{j|k-1} - \Sigma_{k,j} H_k^T M_k^{-1} H_k \Sigma_{k,j}^T = \dots \\ &= P_{j|j-1} - \sum_{l=j}^{k} \Sigma_{l,j} H_l^T M_l^{-1} H_l \Sigma_{l,j}^T \,. \end{aligned}$$

The l-th term gives the improvement due to z_l $(l \geq j)$ over the non-smoothed estimator.

2. Consider a stationary setting. From $\Sigma_{k+1,j} = \Sigma_{k,j}\overline{F}^T$ it follows that $\Sigma_{k,j}$ (and hence the gain $L_{k,j}$) decays exponentially in k–j according to the time constant (spectral radius) of $\overline{F}$. The improvement in $P_{j|k}$ decays accordingly, and becomes marginal when k–j increases beyond 2 or 3 times that time constant.

7.2 Fixed-interval smoothing: the two-pass smoother

Recall that here we wish to compute $\hat{x}_{k|n}$ for $k = 0, \ldots, n$. There are two main recursive algorithms for that purpose:

1. The two-pass smoother: Also known as the RTS (Rauch-Tung-Striebel) smoother. Here the standard Kalman estimate and covariance are computed in a forward pass, and the smoothed quantities are then computed in a backward pass.

2. The forward-backward algorithm: Here the smoother combines two estimates of the state, one given the past and the other given the future (see next section).

The two pass smoother is defined as follows:

- *Forward pass:* Compute the standard filtered quantities $\hat{x}_{k|k-1}$, $\hat{x}_{k|k}$, $P_{k|k-1}$, $P_{k|k}$ for $k = 0 \ldots n$, and store in memory.

- *Backward pass:* Compute $\hat{x}_{k|n}$, via

$$\hat{x}_{k|n} = \hat{x}_{k|k} + A_k(\hat{x}_{k+1|n} - \hat{x}_{k+1|k}), \quad k = n-1, \ldots, 0 \tag{5}$$

where

$$A_k = P_{k|k-1}\overline{F}_k^T P_{k+1|k}^{-1}. \tag{6}$$

The error covariance is

$$P_{k|n} = P_{k|k} + A_k(P_{k+1|n} - P_{k+1|k})A_k^T. \tag{7}$$

To prove these backward-pass relations, observe from (1) that

$$\hat{x}_{k|n} = \hat{x}_{k|n-1} + L_{n,k}\tilde{z}_n = \cdots = \hat{x}_{k|k} + \sum_{l=k+1}^{n} L_{l,k}\tilde{z}_l$$

Similarly,

$$\hat{x}_{k+1|n} = \hat{x}_{k+1|k} + \sum_{l=k+1}^{n} L_{l,k+1}\tilde{z}_l$$

It is easily seen that $L_{l,k} = A_k L_{l,k+1}$, with A_k as above. Therefore, using the last two equations,

$$\begin{aligned}
\hat{x}_{k|n} &= \hat{x}_{k|k} + A_k \sum_{l=k+1}^{n} L_{l,k+1} \tilde{z}_l \\
&= \hat{x}_{k|k} + A_k (\hat{x}_{k+1|n} - \hat{x}_{k+1|k})
\end{aligned}$$

as claimed. The equation for $P_{k|n}$ follows from (5) after some straightforward calculation. $\square$

We note that $P_{k|k-1} \overline{F}_k^T = P_{k|k} F_k^T$, so that (6) is equivalent to $A_k = P_{k|k} F_k^T P_{k+1|k}^{-1}$.

7.3 Fixed-Interval smoothing: the forward-backward filter

This smoother is usually less convenient computationally than the RTS smoother, hence we only outline the approach which is itself of some interest.

In this filter, the smoothed estimate $\hat{x}_{k|n}$ is obtained by combining the (standard) forward estimate:

$$\hat{x}^f \triangleq \hat{x}_{k|k}$$

with the *backward estimate:*

$$\hat{x}_k^b \triangleq E(x_k|z_{k+1}\dots z_n)$$

Define the respective covariances:

$$P_k^f \triangleq P_{k|k}, \quad P_k^b \triangleq \mathrm{cov}(x_k - \hat{x}_k^b)$$

The backward estimate can essentially be computed backwards from time n by considering the reverse-time system:

$$x_{k-1} = (F_{k-1})^{-1}x_k - (F_{k-1})^{-1}w_{k-1}$$

(assuming that F_{k-1} is invertible), with "initial conditions" $P_n = \infty$.

It may be shown that

$$\hat{x}_{k|n} = P_k^*[(P_k^f)^{-1}\hat{x}_k^f + (P_k^b)^{-1}\hat{x}_k^b]$$

where

$$P_k^* = [(P_k^f)^{-1} + (P_k^b)^{-1}]^{-1} = \mathrm{cov}(x_k - \hat{x}_{k|n})\,.$$

The (non-trivial) derivation is omitted here.

7.4 Fixed-lag smoothing

The problem here is to compute recursively $\hat{x}_{k|k+N}$, for $k = 0, 1, 2, \ldots$.

One way to obtain the fixed-lag smoother is to define an augmented super-state:

$$X_k = \begin{bmatrix} x_k \\ x_{k-1} \\ \vdots \\ x_{k-N} \end{bmatrix}, \quad k \geq 0$$

(with $x_i \triangleq 0$ for $i < 0$). After writing state equations for X_k we can estimate $\hat{X}_{k|k}$ using the Kalman filter for this system, and the last component of that vector is the required $x_{k-N|k}$. The obtained filter for X_k is obviously of high dimension, but the equations may be simplified.

A more direct solution builds on the innovations approach. Here we can simply combine the relations already obtained above. From (1) we have (after some index change)

$$\hat{x}_{k+1|k+1+N} = \hat{x}_{k+1|k+N} + (L_{k+1+N,k+1})\tilde{z}_{k+1+N}$$

and from (5),

$$\hat{x}_{k+1|k+N} = \hat{x}_{k+1|k} + A_k^{-1}(\hat{x}_{k|k+N} - \hat{x}_{k|k})\,.$$

Combining these two equations yields a recursive relation between $\hat{x}_{k+1|k+1+N}$ and $\hat{x}_{k|k+N}$.