# Chapter 12

# Smoothing

## 12.1 Overview

Problem: Estimate $\boldsymbol{x}(k)$ given $Z^j$ for some $j > k$.

3 classic types:

- Fixed point smoothing: Estimate $\boldsymbol{x}(k)$ for a fixed $k$ based on $Z^j$ with $j$ increasing.
- Fixed lag smoothing: Estimate $\boldsymbol{x}(k)$ based on $Z^j$ where $j - k$ is constant and $j > k$.
- Fixed interval smoothing: Estimate $\boldsymbol{x}(k), k = 1, 2, \ldots, N$ based on $Z^N$.

Our focus will be fixed interval smoothing. New notation:

$$\begin{aligned}
\boldsymbol{x}^*(k) &\triangleq \hat{\boldsymbol{x}}(k|N) \\
P^*(k) &\triangleq P(k|N) \\
\boldsymbol{v}^*(k) &\triangleq \hat{\boldsymbol{v}}(k|N)
\end{aligned}$$

Bar Shalom develops a smoother based on the traditional KF equations in Section 8.6. We will instead develop the square-root information smoother (SRIS), which is intuitive and more numerically stable.

Key observation: Smoother equations fall out of the MAP estimation approach.

## 12.2 Preliminaries

Suppose we have performed an SRIF forward (filtering) pass on measurements $Z^N$. From this will be left three groups of SRI equations:

1. Equations relating the state and process noise:

$$\begin{aligned}
\bar{\boldsymbol{z}}_v(0) &= \bar{R}_{vv}(0)v(0) + \bar{R}_{vx}(1)x(1) + \bar{\boldsymbol{w}}_v(0) \\
&\vdots \\
\bar{\boldsymbol{z}}_v(N-1) &= \bar{R}_{vv}(N-1)\boldsymbol{v}(N-1) + \bar{R}_{vx}(N)\boldsymbol{x}(N) + \bar{\boldsymbol{w}}_v(N-1)
\end{aligned}$$

where $\bar{\boldsymbol{w}}_v(k) \sim (0, I), \quad k = 0, \ldots, N-1.$

2. Equations for the residuals:

$$\boldsymbol{z}_r(1) = \boldsymbol{w}_r(1)$$
$$\vdots$$
$$\boldsymbol{z}_r(N) = \boldsymbol{w}_r(N)$$

where $\boldsymbol{w}_r(k) \sim (0, I)$ for $k = 1, \ldots, N$. These residuals equations can be discarded; they are not needed for the SRIS.

3. An equation for the terminal state:

$$\boldsymbol{z}_x(N) = R_{xx}(N)\boldsymbol{x}(N) + \boldsymbol{w}_x(N)$$

where $\boldsymbol{w}_x(k) \sim (0, I)$.

The SRIF applies the dynamics model

$$\boldsymbol{x}(k+1) = F(k)\boldsymbol{x}(k) + G(k)\boldsymbol{u}(k) + \Gamma(k)\boldsymbol{v}(k)$$

going *forward* through the data.

**Q:** Do you suppose the filtered estimates $\hat{\boldsymbol{x}}(k)$ actually conform to this dynamics model?
**A:** Interestingly, they do not. A time history of $\hat{\boldsymbol{x}}(k)$ doesn't necessarily look like it could have come from the dynamics model because the estimates get jostled around by the innovations $\boldsymbol{\nu}(k)$ (see [8]). ■

A smoother enforces the dynamics model going *backward* in time. After a smoother is applied, the time history $\boldsymbol{x}^*(k)$ actually behaves like the dynamics model would suggest (no discontinuous jumps due to innovations).

**Q:** Does a smoother need to take in the measurements $Z^N$?
**A:** No, the filter already extracted all possible information from the measurements. ■

## 12.3 SRIS Algorithm

1. Let

$$\boldsymbol{z}_x^*(N) = \boldsymbol{z}_x(N)$$
$$R_{xx}^*(N) = R_{xx}(N)$$
$$\boldsymbol{w}_x^*(N) = \boldsymbol{w}_x(N)$$

2. If a smoothed state estimate and covariance is needed (e.g., for control, etc.), then compute

$$\boldsymbol{x}^*(N) = R_{xx}^{*-1}(N)\boldsymbol{z}^*(N)$$
$$P^*(N) = R_{xx}^{*-1}(N)R_{xx}^{*-T}(N)$$

3. Set $k = N - 1$. Recall from the MAP derivation of the SRIF that the cost function $J_b\left[v(k), x(k+1), k\right]$ was written as

$$J_b\left[\boldsymbol{v}(k), \boldsymbol{x}(k+1), k\right] = \left\| \begin{bmatrix} \bar{R}_{vv}(k) & \bar{R}_{vx}(k+1) \\ \boldsymbol{0} & \bar{R}_{xx}(k+1) \end{bmatrix} \begin{bmatrix} \boldsymbol{v}(k) \\ \boldsymbol{x}(k+1) \end{bmatrix} - \begin{bmatrix} \bar{\boldsymbol{z}}_v(k) \\ \bar{\boldsymbol{z}}_x(k+1) \end{bmatrix} \right\|^2 + \left\| H_a(k+1)x(k+1) - \boldsymbol{z}_a(k+1) \right\|^2$$

Minimizing this cost function is equivalent to maximizing the *a posteriori* distribution $p[\boldsymbol{x}(k+1), \boldsymbol{v}(k)|Z^{k+1}]$. In the SRIF derivation, we solved for the minimizing $\boldsymbol{v}(k)$ in terms of $\boldsymbol{x}(k+1)$ and substituted this to obtain $J_c[\boldsymbol{x}(k+1), k+1]$. But we could just as well have kept working with $J_b\left[\boldsymbol{v}(k), \boldsymbol{x}(k+1), k\right]$ and have obtained, after the QR factorization involving the measurement, a cost of the form

$$J_b\left[\boldsymbol{v}(k), \boldsymbol{x}(k+1), k\right] = \left\| \begin{bmatrix} \bar{R}_{vv}(k) & \bar{R}_{vx}(k+1) \\ \mathbf{0} & R_{xx}(k+1) \end{bmatrix} \begin{bmatrix} \boldsymbol{v}(k) \\ \boldsymbol{x}(k+1) \end{bmatrix} - \begin{bmatrix} \bar{\boldsymbol{z}}_v(k) \\ \boldsymbol{z}_x(k+1) \end{bmatrix} \right\|^2 + \|\boldsymbol{z}_r(k+1)\|^2$$

The SRI equation for the *a posteriori* state estimate can be unpacked from the above cost as

$$\boldsymbol{z}_x(k+1) = R_{xx}(k+1)\boldsymbol{x}(k+1) + \boldsymbol{w}_x(k+1), \quad \boldsymbol{w}_x(k+1) \sim \mathcal{N}(0, I)$$

If instead of the *a posteriori* state estimate, we had the *smoothed* state estimate, expressed in SRI form as

$$\boldsymbol{z}_x^*(k+1) = R_{xx}^*(k+1)\boldsymbol{x}(k+1) + \boldsymbol{w}_x^*(k+1), \quad \boldsymbol{w}_x^*(k+1) \sim \mathcal{N}(0, I)$$

then, substituting this into $J_b\left[\boldsymbol{v}(k), \boldsymbol{x}(k+1), k\right]$ and discarding the irreductible component $\|\boldsymbol{z}_r(k+1)\|^2$, we could write a cost function for the smoothing problem at $k$:

$$J_b^*\left[\boldsymbol{v}(k), \boldsymbol{x}(k+1), k\right] = \left\| \begin{bmatrix} \bar{R}_{vv}(k) & \bar{R}_{vx}(k+1) \\ \mathbf{0} & R_{xx}^*(k+1) \end{bmatrix} \begin{bmatrix} \boldsymbol{v}(k) \\ \boldsymbol{x}(k+1) \end{bmatrix} - \begin{bmatrix} \bar{\boldsymbol{z}}_v(k) \\ \boldsymbol{z}_x^*(k+1) \end{bmatrix} \right\|^2$$

We seek to minimize this cost subject to the dynamics constraint.

4. Use the dynamics equation to eliminate $\boldsymbol{x}(k+1)$ in favor of $\boldsymbol{x}(k)$. Substituting for $\boldsymbol{x}(k+1)$ and stacking yields a cost of the form

$$J_a^*[\boldsymbol{v}(k), \boldsymbol{x}(k), k] = \left\| \begin{bmatrix} \quad \end{bmatrix} \begin{bmatrix} \boldsymbol{v}(k) \\ \boldsymbol{x}(k) \end{bmatrix} - \begin{bmatrix} \quad \end{bmatrix} \right\|^2$$

whose implied SRI equations are

$$\begin{bmatrix} \bar{\boldsymbol{z}}_v(k) - \bar{R}_{vx}(k+1)G(k)\boldsymbol{u}(k) \\ \boldsymbol{z}_x^*(k+1) - R_{xx}^*(k+1)G(k)\boldsymbol{u}(k) \end{bmatrix} = \begin{bmatrix} \bar{R}_{vv}(k) + \bar{R}_{vx}(k+1)\Gamma(k) & \bar{R}_{vx}(k+1)F(k) \\ R_{xx}^*(k+1)\Gamma(k) & R_{xx}^*(k+1)F(k) \end{bmatrix} \begin{bmatrix} \boldsymbol{v}(k) \\ \boldsymbol{x}(k) \end{bmatrix} + \begin{bmatrix} \bar{\boldsymbol{w}}_v(k) \\ \boldsymbol{w}_x^*(k+1) \end{bmatrix}, \qquad \begin{bmatrix} \bar{\boldsymbol{w}}_v(k) \\ \boldsymbol{w}_x^*(k+1) \end{bmatrix} \sim \mathcal{N}(0, I)$$

Multiplying both sides by $T_a(k) = Q_a^T(k)$ from QR factorization of the block matrix decouples the SRI equations as

$$\begin{bmatrix} \boldsymbol{z}_v^*(k) \\ \boldsymbol{z}_x^*(k) \end{bmatrix} = \begin{bmatrix} R_{vv}^*(k) & R_{vx}^*(k) \\ 0 & R_{xx}^*(k) \end{bmatrix} \begin{bmatrix} \boldsymbol{v}(k) \\ \boldsymbol{x}(k) \end{bmatrix} + \begin{bmatrix} \boldsymbol{w}_v^*(k) \\ \boldsymbol{w}_x^*(k) \end{bmatrix}, \qquad \begin{bmatrix} \boldsymbol{w}_v^*(k) \\ \boldsymbol{w}_x^*(k) \end{bmatrix} \sim \mathcal{N}(0, I)$$

This transformation does not change the associated cost:

$$J_a^*[\boldsymbol{v}(k), \boldsymbol{x}(k), k] = \|R_{vv}^*(k)\boldsymbol{v}(k) + R_{vx}^*(k)\boldsymbol{x}(k) - \boldsymbol{z}_v^*(k)\|^2 + \|R_{xx}^*(k)\boldsymbol{x}(k) - \boldsymbol{z}_x^*(k)\|^2$$

5. Choose $\boldsymbol{x}^*(k)$ and $\boldsymbol{v}^*(k)$ to minimize the cost:

$$\begin{aligned} \boldsymbol{x}^*(k) &= R_{xx}^{*-1}(k)\boldsymbol{z}_x^*(k) = \mathbb{E}\left[\boldsymbol{x}(k)|Z^N\right] \\ \boldsymbol{v}^*(k) &= R_{vv}^{*-1}(k)[\boldsymbol{z}_v^*(k) - R_{vx}^*(k)\boldsymbol{x}^*(k)] = \mathbb{E}\left[\boldsymbol{v}(k)|Z^N\right] \end{aligned}$$

The associated error covariance matrices are

$$
\begin{aligned}
P^*(k) &= R_{xx}^{*-1}(k) R_{xx}^{*-T}(k) \\
P_{vv}^*(k) &= R_{vv}^{*-1}(k)[I + R_{vx}^*(k) R_{xx}^{*-1}(k) R_{xx}^{*-T}(k) R_{vx}^{*T}(k)] R_{vv}^{*-T}(k) \\
P_{vx}^*(k) &= -R_{vv}^{*-1}(k) R_{vx}^*(k) R_{xx}^{*-1}(k) R_{xx}^{*-T}(k)
\end{aligned}
$$

Note that $J_a[\boldsymbol{v}^*(k), \boldsymbol{x}^*(k), k] = 0$ because in forming the smoother's cost function we discarded the irreducible components $\boldsymbol{z}_r(k),\ k = 1, \ldots, N$ from the SRIF.

6. If $k = 0$, stop; otherwise, decrement $k$ by 1 and go to step 4, now using the SRI equations

$$
\begin{aligned}
\bar{\boldsymbol{z}}_v(k) &= \bar{R}_{vv}(k)\boldsymbol{v}(k) + \bar{R}_{vx}(k+1)\boldsymbol{x}(k+1) + \bar{\boldsymbol{w}}_v(k) \\
\boldsymbol{z}_x^*(k+1) &= R_{xx}^*(k+1)\boldsymbol{x}(k+1) + \boldsymbol{w_x}^*(k+1)
\end{aligned}
$$