Crib Sheet : Linear Kalman Smoothing

Gabriel A. Terejanu
Department of Computer Science and Engineering
University at Buffalo, Buffalo, NY 14260
terejanu@buffalo.edu

1 Introduction

Smoothing can be separated into three classes [6]:

1. **Fixed-interval smoothing**. The goal is to obtain the estimates $\mathbf{x}_k^s$, for $k = 0 \dots N$, given a fixed observation interval $\mathbf{Z}_N = \{\mathbf{z}_k | 0 \leq k \leq N\}$ [6].

2. **Fixed-point smoothing**. The goal is to obtain the estimate $\mathbf{x}_N^s$ (the time t_N is fixed, $N < k$) given a set of observations $\mathbf{Z}_k = \{\mathbf{z}_k | 0 \leq N \leq k\}$ [6].

3. **Fixed-lag smoothing**. The goal is to obtain the estimate $\mathbf{x}_k^s$ given the observations $\mathbf{Z}_{k+N} = \{\mathbf{z}_k | 0 \leq k \leq k+N\}$ [6]. Where $k+N$ is the most current measurement and N is a constant lag.

“It is possible to employ a single smoothing scheme, based on fixed-interval smoothing to solve all three problems” [1].

“A state is said to be **smoothable** if an optimal smoother provides a state estimate superior to that obtained when the final optimal filter estimate is extrapolated backwards in time” [4].

“Only those states which are controllable by the noise driving the system state vector are smoothable” (Weiss 1970).

2 Fixed Interval Smoothing

The goal is to obtain the estimates $\mathbf{x}_k^s$, for $k = 0 \dots N$, given a fixed observation interval $\mathbf{Z}_N = \{\mathbf{z}_k | 0 \leq k \leq N\}$ [6].

Fraser and Potter (1969): two optimal linear filters [3, 2]

The smoothed estimate is expressed as a linear combination between the forward and backward filter state estimates. The optimal weighting between the two is also known as "*Millman's theorem* which is also an exact analog to maximum likelihood of a scalar with independent measurements" [2].

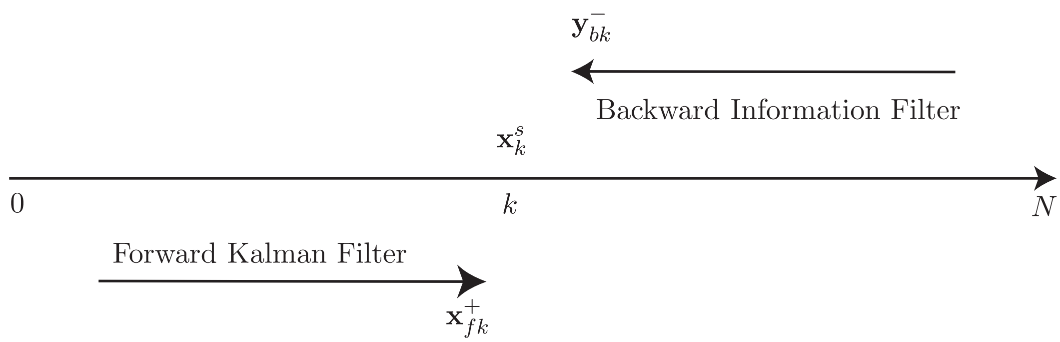

Figure 1: Fixed Interval Smoothing : two-filter approach

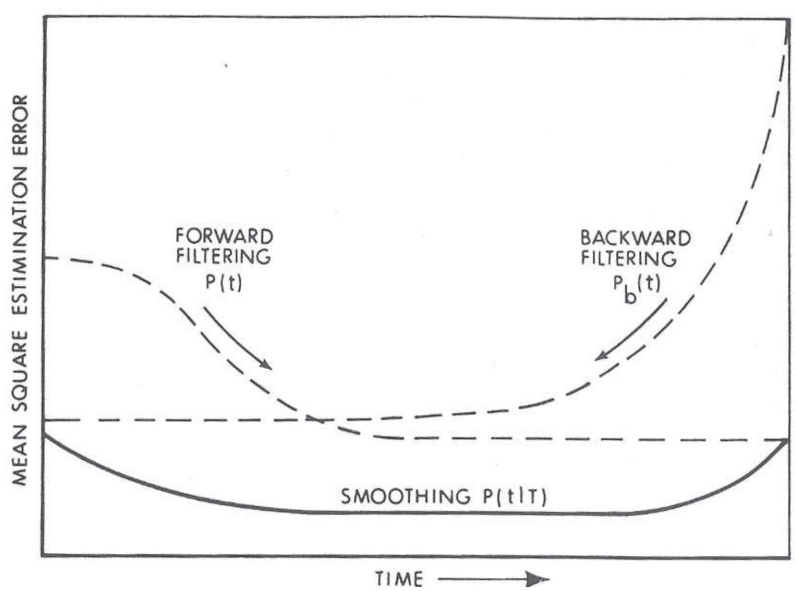

Figure 2: Advantage of performing optimal smoothing [4]

Model and Observation	
Model and Observation	$\mathbf{x}_k = \mathbf{A}_{k-1}\mathbf{x}_{k-1} + \mathbf{B}_{k-1}\mathbf{u}_{k-1} + \mathbf{w}_{k-1}$ $\mathbf{z}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k$
Forward Filter	
Initialization	$\mathbf{x}^+_{f0} = \mu_0$ with error covariance $\mathbf{P}^+_{f0}$
Model Forecast Step/Predictor	$\mathbf{x}^-_{fk} = \mathbf{A}_{k-1}\mathbf{x}^+_{fk-1} + \mathbf{B}_{k-1}\mathbf{u}_{k-1}$ $\mathbf{P}^-_{fk} = \mathbf{A}_{k-1}\mathbf{P}^+_{fk-1}\mathbf{A}^T_{k-1} + \mathbf{Q}_{k-1}$
Data Assimilation Step/Corrector	$\mathbf{x}^+_{fk} = \mathbf{x}^-_{fk} + \mathbf{K}_{fk}(\mathbf{z}_k - \mathbf{H}_k\mathbf{x}^-_{fk})$ $\mathbf{K}_{fk} = \mathbf{P}^-_{fk}\mathbf{H}^T_k(\mathbf{H}_k\mathbf{P}^-_{fk}\mathbf{H}^T_k + \mathbf{R}_k)^{-1}$ $\mathbf{P}^+_{fk} = (\mathbf{I} - \mathbf{K}_{fk}\mathbf{H}_k)\mathbf{P}^-_{fk}$
Backward Information Filter	
Initialization	$\mathbf{y}^-_{bN} = 0$ $\mathbf{Y}^-_{bN} = 0$
Backward Update	$\mathbf{y}^+_{bk} = \mathbf{y}^-_{bk} + \mathbf{H}^T_k\mathbf{R}^{-1}_k\mathbf{z}_k$ $\mathbf{Y}^+_{bk} = \mathbf{Y}^-_{bk} + \mathbf{H}^T_k\mathbf{R}^{-1}_k\mathbf{H}_k$
Backard Propagation	$\mathbf{K}_{bk} = \mathbf{Y}^+_{bk+1}\left(\mathbf{Y}^+_{bk+1} + \mathbf{Q}^{-1}_k\right)^{-1}$ $\mathbf{y}^-_{bk} = \mathbf{A}^T_k\left(\mathbf{I} - \mathbf{K}_{bk}\right)\left(\mathbf{y}^+_{bk+1} - \mathbf{Y}^+_{bk+1}\mathbf{B}_k\mathbf{u}_k\right)$ $\mathbf{Y}^-_{bk} = \mathbf{A}^T_k\left(\mathbf{I} - \mathbf{K}_{bk}\right)\mathbf{Y}^+_{bk+1}\mathbf{A}_k$
Smoother	
Estimate	$\mathbf{x}^s_k = \left(\mathbf{I} - \mathbf{K}^s_k\right)\mathbf{x}^+_{fk} + \mathbf{P}^s_k\mathbf{y}^-_{bk}$ $\mathbf{K}^s_k = \mathbf{P}^+_{fk}\mathbf{Y}^-_{bk}\left(\mathbf{I} + \mathbf{P}^+_{fk}\mathbf{Y}^-_{bk}\right)^{-1}$ $\mathbf{P}^s_k = \left(\mathbf{I} - \mathbf{K}^s_k\right)\mathbf{P}^+_{fk}$

Table 1: Fraser and Potter: two optimal linear filters.

Notes

- The infinite value of $\mathbf{P}^-_{bN}$ and the boundary condition to specify $\mathbf{x}^-_{bN}$ are difficult to apply [3]. To avoid these problems the information form approach is used to represent the backward filter.
- The smoother error covariance $\mathbf{P}^s_k = \left((\mathbf{P}^+_{fk})^{-1} + (\mathbf{P}^-_{bk})^{-1}\right)^{-1}$ [2]. At $k = N$ the smoother estimate and covariance have to be the same as the forward Kalman Filter. Thus, $\mathbf{P}^s_N = \mathbf{P}^+_{fk}$ yields $(\mathbf{P}^-_{bk})^{-1} = 0$. Since we use the information form for the backward filter: $\mathbf{Y}^-_{bk} = (\mathbf{P}^-_{bk})^{-1} = 0$ and $\mathbf{y}^-_{bk} = \mathbf{Y}^-_{bk}\mathbf{x}^-_{bk} = 0$ (are zero for $k = N$).
- The smoothed covariance is equivalent to Joseph's stabilized form.
- Forward error estimate and backward "forecast" error are uncorrelated.

Rauch, Tung, and Striebel (1965): correction to the Kalman Filter [5, 2]

Combines the backward filter and the smoother in one single recursive step. **Notes**

Model and Observation	
Model and Observation	$\mathbf{x}_k = \mathbf{A}_{k-1}\mathbf{x}_{k-1} + \mathbf{B}_{k-1}\mathbf{u}_{k-1} + \mathbf{w}_{k-1}$ $\mathbf{z}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k$
Forward Filter	
Initialization	$\mathbf{x}_{f0}^{+} = \mu_0$ with error covariance $\mathbf{P}_{f0}^{+}$
Model Forecast Step/Predictor	$\mathbf{x}_{fk}^{-} = \mathbf{A}_{k-1}\mathbf{x}_{fk-1}^{+} + \mathbf{B}_{k-1}\mathbf{u}_{k-1}$ $\mathbf{P}_{fk}^{-} = \mathbf{A}_{k-1}\mathbf{P}_{fk-1}^{+}\mathbf{A}_{k-1}^{T} + \mathbf{Q}_{k-1}$
Data Assimilation Step/Corrector	$\mathbf{x}_{fk}^{+} = \mathbf{x}_{fk}^{-} + \mathbf{K}_{fk}(\mathbf{z}_k - \mathbf{H}_k\mathbf{x}_{fk}^{-})$ $\mathbf{K}_{fk} = \mathbf{P}_{fk}^{-}\mathbf{H}_k^{T}(\mathbf{H}_k\mathbf{P}_{fk}^{-}\mathbf{H}_k^{T} + \mathbf{R}_k)^{-1}$ $\mathbf{P}_{fk}^{+} = (\mathbf{I} - \mathbf{K}_{fk}\mathbf{H}_k)\mathbf{P}_{fk}^{-}$
Smoother	
Initialization	$\mathbf{x}_N^{s} = \mathbf{x}_{fk}^{+}$ $\mathbf{P}_N^{s} = \mathbf{P}_{fk}^{+}$
Update	$\mathbf{K}_k^{s} = \mathbf{P}_{fk}^{+}\mathbf{A}_k^{T}(\mathbf{P}_{fk+1}^{-})^{-1}$ $\mathbf{P}_k^{s} = \mathbf{P}_{fk}^{+} - \mathbf{K}_k^{s}\left(\mathbf{P}_{fk+1}^{-} - \mathbf{P}_{k+1}^{s}\right)(\mathbf{K}_k^{s})^{T}$ $\mathbf{x}_k^{s} = \mathbf{x}_{fk}^{+} + \mathbf{K}_k^{s}\left(\mathbf{x}_{k+1}^{s} - \mathbf{x}_{fk+1}^{-}\right)$

Table 2: Rauch, Tung, and Striebel: correction to the Kalman Filter.

- The smoother does not depend on either backward covariance or backward estimate. Its form reveals just a correction of the current Kalman Filter using only the data provided by the forward filter.
- The smoothed estimate does not depend on the smoothed covariance.
- In order to obtain the smoothed estimate only the forward state estimate and the smoothed gain have to be stored.
- It is completely derived from optimal control theory [2]

$$\begin{aligned}
&\text{Minimize} \\
J(\mathbf{w}_k) &= \frac{1}{2}\sum_{k=1}^{N}(\mathbf{z}_k - \mathbf{H}_k\mathbf{x}_k)^T\mathbf{R}_k^{-1}(\mathbf{z}_k - \mathbf{H}_k\mathbf{x}_k) + \mathbf{w}_k^t\mathbf{Q}_k^{-1}\mathbf{w}_k \\
&+\frac{1}{2}\left(\mathbf{x}_{f0}^{+} - \mathbf{x}_0\right)^T(\mathbf{P}_{f0}^{+})^{-1}\left(\mathbf{x}_{f0}^{+} - \mathbf{x}_0\right) \\
&\text{subject to} \\
\mathbf{x}_k &= \mathbf{A}_{k-1}\mathbf{x}_{k-1} + \mathbf{B}_{k-1}\mathbf{u}_{k-1} + \mathbf{w}_{k-1} \\
\mathbf{z}_k &= \mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k
\end{aligned}$$

this yields a Two-Point-Boundary-Value-Problem (TPBVP).

- It can be derived as MAP estimate of $\mathbf{x}_k$, given the data set $\mathbf{Z}_N$, $N > k$, where $\mathbf{x}_k$ maximizes the conditional density $p(\mathbf{x}_k|\mathbf{Z}_N)$ [6, p.367].

3 Fixed Point Smoothing

The goal is to obtain the estimate $\mathbf{x}^s_N$ (the time t_N is fixed, $N < k$) given a set of observations $\mathbf{Z}_k = \{\mathbf{z}_k | 0 \leq N \leq k\}$ [6].

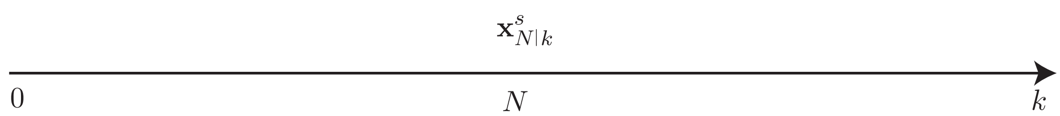

Figure 3: Fixed Point Smoothing

Meditch (1967): extended RTS for fixed-point smoothing [6, 2]

Model and Observation	
Model and Observation	$\mathbf{x}_k = \mathbf{A}_{k-1}\mathbf{x}_{k-1} + \mathbf{B}_{k-1}\mathbf{u}_{k-1} + \mathbf{w}_{k-1}$ $\mathbf{z}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k$
Forward Filter	
Initialization	$\mathbf{x}^+_{f0} = \mu_0$ with error covariance $\mathbf{P}^+_{f0}$
Model Forecast Step/Predictor	$\mathbf{x}^-_{fk} = \mathbf{A}_{k-1}\mathbf{x}^+_{fk-1} + \mathbf{B}_{k-1}\mathbf{u}_{k-1}$ $\mathbf{P}^-_{fk} = \mathbf{A}_{k-1}\mathbf{P}^+_{fk-1}\mathbf{A}^T_{k-1} + \mathbf{Q}_{k-1}$
Data Assimilation Step/Corrector	$\mathbf{x}^+_{fk} = \mathbf{x}^-_{fk} + \mathbf{K}_{fk}(\mathbf{z}_k - \mathbf{H}_k\mathbf{x}^-_{fk})$ $\mathbf{K}_{fk} = \mathbf{P}^-_{fk}\mathbf{H}^T_k(\mathbf{H}_k\mathbf{P}^-_{fk}\mathbf{H}^T_k + \mathbf{R}_k)^{-1}$ $\mathbf{P}^+_{fk} = (\mathbf{I} - \mathbf{K}_{fk}\mathbf{H}_k)\mathbf{P}^-_{fk}$
Smoother	
Initialization	$\mathbf{x}^s_{N\|N} = \mathbf{x}^+_{fN}$ $\mathbf{P}^s_{N\|N} = \mathbf{P}^+_{fN}$ $\mathbf{M}_{N\|N} = \mathbf{I}$
Update	$\mathbf{M}_{N\|k} = \mathbf{M}_{N\|k-1}\mathbf{K}^s_{k-1}$ $\mathbf{K}^s_k = \mathbf{P}^+_{fk}\mathbf{A}^T_k(\mathbf{P}^-_{fk+1})^{-1}$ $\mathbf{P}^s_{N\|k} = \mathbf{P}^s_{N\|k-1} - \mathbf{M}_{N\|k}\left(\mathbf{P}^+_{fk} - \mathbf{P}^-_{fk}\right)\mathbf{M}^T_{N\|k}$ $\mathbf{x}^s_{N\|k} = \mathbf{x}^s_{N\|k-1} - \mathbf{M}_{N\|k}\left(\mathbf{x}^+_{fk} - \mathbf{x}^-_{fk}\right)$

Table 3: Meditch (1967): extended RTS for fixed-point smoothing

4 Fixed Lag Smoothing

The goal is to obtain the estimate $\mathbf{x}_k^s$ given the observations $\mathbf{Z}_{k+N} = \{\mathbf{z}_k | 0 \leq k \leq k+N\}$ [6]. Where $k+N$ is the most current measurement and N is a constant lag.

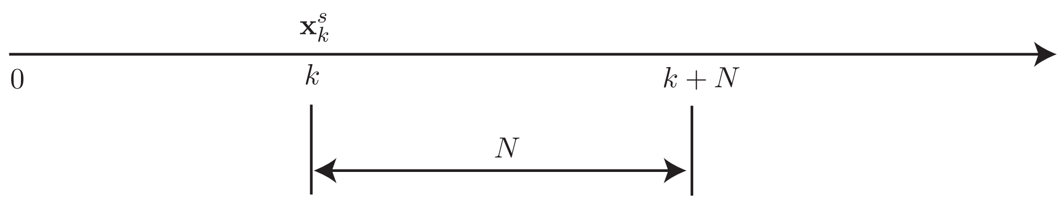

Figure 4: Fixed Lag Smoothing

Meditch (1967): extended RTS for fixed-lag smoothing [6, 2]
"The fixed-lag-smoothing algorithm for discrete processed is obtained by combinng both the fixed-interval and fixed-lag algorithms" [6].

Model and Observation	
Model and Observation	$\mathbf{x}_k = \mathbf{A}_{k-1}\mathbf{x}_{k-1} + \mathbf{B}_{k-1}\mathbf{u}_{k-1} + \mathbf{w}_{k-1}$ $\mathbf{z}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k$
Forward Filter	
Initialization	$\mathbf{x}_{f0}^+ = \mu_0$ with error covariance $\mathbf{P}_{f0}^+$
Model Forecast Step/Predictor	$\mathbf{x}_{fk}^- = \mathbf{A}_{k-1}\mathbf{x}_{fk-1}^+ + \mathbf{B}_{k-1}\mathbf{u}_{k-1}$ $\mathbf{P}_{fk}^- = \mathbf{A}_{k-1}\mathbf{P}_{fk-1}^+\mathbf{A}_{k-1}^T + \mathbf{Q}_{k-1}$
Data Assimilation Step/Corrector	$\mathbf{x}_{fk}^+ = \mathbf{x}_{fk}^- + \mathbf{K}_{fk}(\mathbf{z}_k - \mathbf{H}_k\mathbf{x}_{fk}^-)$ $\mathbf{K}_{fk} = \mathbf{P}_{fk}^-\mathbf{H}_k^T(\mathbf{H}_k\mathbf{P}_{fk}^-\mathbf{H}_k^T + \mathbf{R}_k)^{-1}$ $\mathbf{P}_{fk}^+ = (\mathbf{I} - \mathbf{K}_{fk}\mathbf{H}_k)\mathbf{P}_{fk}^-$
Smoother	
Initialization	$\mathbf{x}_{k\|k+N}^s = \mathbf{x}_{fN}^+$ $\mathbf{P}_{k\|k+N}^s = \mathbf{P}_{fN}^+$ $\mathbf{M}_{k\|k+N} = \mathbf{I}$
Update	$\mathbf{M}_{k+1\|k+N+1} = \mathbf{M}_{k\|k+N}\mathbf{K}_{k+N}^s$ $\mathbf{K}_{k+N}^s = \mathbf{P}_{fk+N}^+\mathbf{A}_{k+N}^T(\mathbf{P}_{fk+N+1}^-)^{-1}$ $\mathbf{P}_{k+1\|k+N+1}^s = \mathbf{P}_{fk+1}^- - (\mathbf{K}_k^s)^{-1}\left(\mathbf{P}_{fk+}^+ - \mathbf{P}_{k\|k+N}^s\right)(\mathbf{K}_k^s)^{-T}$ $-\mathbf{M}_{k+1\|k+N+1}\mathbf{K}_{fk+N+1}\mathbf{H}_{k+N+1}\mathbf{P}_{fk+N+1}^-\mathbf{M}_{k+1\|k+N+1}^T$ $\mathbf{x}_{k+1\|k+N+1}^s = \mathbf{A}_k\mathbf{x}_{k\|k+N}^s + \mathbf{B}_k\mathbf{u}_k$ $+\mathbf{Q}_k\mathbf{A}_k^{-T}(\mathbf{P}_{fk}^+)^{-1}\left(\mathbf{x}_{k\|k+N}^s - \mathbf{x}_{fk}^+\right)$ $+\mathbf{M}_{k+1\|k+N+1}\mathbf{K}_{fk+N+1}(\mathbf{z}_{k+N+1} - \mathbf{H}_{k+N+1}\mathbf{x}_{fk+N+1}^-)$

Table 4: Meditch (1967): extended RTS for fixed-lag smoothing

References

[1] Doucet A. & Maskell S. Briers, M. Smoothing algorithms for state-space models. *Submission IEEE Transactions on Signal Processing*, 2004.

[2] John Crassidis and John Junkins. *Optimal Estimation of Dynamic Systems*. CRC Press, 2004.

[3] D. Fraser and J. Potter. The optimum linear smoother as a combination of two optimum linear filters. *Automatic Control, IEEE Transactions on*, 14(4):387–390, Aug 1969.

[4] Arthur Gelb. *Applied Optimal Estimation*. The M.I.T. Press, 1974.

[5] H. E. Rauch, F. Tung, and C. T. Striebel. Maximum likelihood estimates of linear dynamic systems. *J. Amer. Inst. Aeronautics and Astronautics*, 3 (8):1445–1450, 1965.

[6] Andrew Sage and James Melsa. *Estimation Theory with Applications to Communications and Control*. McGraw-Hill Book Company, 1971.