

Fixed-lag maximum likelihood FIR smoother for state-space models

Choon Ki Ahn^{1a)} and Pyung Soo Kim^{2b)}

¹ Digital Printing R&D Center, Samsung Electronics, 416, Maetan-3dong, Yeongtong-gu, Suwon, 442–742, Korea

² Department of Electronics Engineering, Korea Polytechnic University, 2121, Jeongwang-dong, Siheung-si, Gyeonggi-do, 429–793, Korea

^{a)} hironaka@cisl.snu.ac.kr

^{b)} pskim@kpu.ac.kr

Abstract: In this paper, we propose a new fixed-lag maximum likelihood smoother with a finite impulse response (FIR) structure for discrete-time state-space models. This smoother is called a maximum likelihood FIR smoother (MLFS). The MLFS is linear with the most recent finite outputs and does not require *a priori* initial state information on a receding horizon. It is shown that the proposed MLFS possesses the unbiasedness property and the deadbeat property. Simulation study illustrates that the proposed MLFS is more robust against uncertainties and faster in convergence than the conventional fixed-lag Kalman smoother.

Keywords: fixed-lag smoother, maximum likelihood, FIR structure, unbiasedness property, deadbeat property

Classification: Science and engineering for electronics

References

- [1] B. Hassibi, A. H. Sayed, and T. Kailath, *Indefinite Quadratic Estimation and Control: A Unified Approach to \mathcal{H}_2 and \mathcal{H}_∞ Theories*, SIAM Studies in Applied and Numerical Mathematics, 1999.
- [2] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*, Prentice Hall, 2000.
- [3] A. M. Bruckstein and T. Kailath, “Recursive limited memory filtering and scattering theory,” *IEEE Trans. Inform. Theory*, vol. 31, pp. 440–443, 1985.
- [4] M. S. Grewal and A. P. Andrews, *Kalman Filtering - Theory and Practice*, Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [5] R. W. Eustace, B. A. Woodyatt, G. L. Merrington, and A. Runacres, “Fault signatures obtained from fault implant tests on an F404 engine,” *ASME Trans. J. of Engine, Gas Turbines, and Power*, vol. 116, no 1, pp. 178–183, 1994.

1 Introduction

The most common estimation tool for state-space models is the Kalman filter [1, 2]. Though this filter has proven its efficiency, the filtering operation brings a slight time delay in the estimation due to the causality of the filter. In order to solve this problem, one has to consider a smoothing operation. This has led to an interest in the Kalman smoother [1, 2], which is a maximum likelihood smoother. Its significance is increasing by the need of making more correct estimations in telecommunication and information systems.

In Kalman smoothing problems, *a priori* information on the initial state is assumed to be completely known. However, since the initial state is also an unknown signal, it is somewhat unreasonable to assume that *a priori* information on the initial state is completely known. The Kalman smoother is not linearly related to the outputs since the smoothing algorithm includes the initial state term. In addition, it is not unbiased unless the mean of the initial state is completely known. Furthermore, due to its infinite impulse response (IIR) structure, the Kalman smoother may diverge for systems with temporary modeling uncertainties and numerical errors [3, 4]. Therefore, it is desirable to obtain an alternative maximum likelihood smoother that can overcome the above shortcomings of the Kalman smoother. In this paper, we propose a new maximum likelihood FIR smoother (MLFS), which solves the weak points of the Kalman smoother.

2 Fixed-lag Maximum Likelihood FIR Smoother

Consider a linear discrete-time state-space signal model:

$$x_{k+1} = Ax_k + Gw_k, \quad (1)$$

$$y_k = Cx_k + v_k, \quad (2)$$

where $x_k \in \mathbb{R}^n$ and $y_k \in \mathbb{R}^q$ are the unknown state and the measured output, respectively. At the initial time k_0 of the model, the state x_{k_0} is a random variable with a mean \bar{x}_{k_0} and a covariance P_{k_0} . The system noise $w_k \in \mathbb{R}^p$ and the measurement noise $v_k \in \mathbb{R}^q$ are zero-mean white Gaussian and mutually uncorrelated. The covariances of w_k and v_k are denoted by Q and R , respectively, which are assumed to be positive definite matrices. These noises are uncorrelated with the initial state x_{k_0} .

On the receding horizon $[k - N, k]$ where N is a horizon size, outputs are expressed in terms of the state x_{k-N} at the time $k - N$ as follows:

$$Y_{k-1} = \tilde{C}_N x_{k-N} + \tilde{G}_N W_{k-1} + V_{k-1} \quad (3)$$

where

$$Y_{k-1} \triangleq [y_{k-N}^T \ y_{k-N+1}^T \ \cdots \ y_{k-1}^T]^T, \quad (4)$$

$$W_{k-1} \triangleq [w_{k-N}^T \ w_{k-N+1}^T \ \cdots \ w_{k-1}^T]^T, \quad (5)$$

$$V_{k-1} \triangleq [v_{k-N}^T \ v_{k-N+1}^T \ \cdots \ v_{k-1}^T]^T, \quad (6)$$

and \tilde{C}_N and \tilde{G}_N are obtained from

$$\tilde{C}_N \triangleq \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{N-1} \end{bmatrix} \in \Re^{Nq \times n}, \quad (7)$$

$$\tilde{G}_N \triangleq \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ CG & 0 & \cdots & 0 & 0 \\ CAG & CG & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ CA^{N-2}G & CA^{N-3}G & \cdots & CG & 0 \end{bmatrix} \in \Re^{Nq \times Np}. \quad (8)$$

In order to obtain the MLFS, it is necessary to relate x_{N-k} to x_{k-h} , where h is a time-lag satisfying $0 \leq h < N$. From (1), the state x_{k-h} at time $k-h$ is represented by

$$x_{k-h} = A^{N-h}x_{k-N} + M_N W_{k-1} \quad (9)$$

where

$$M_N \triangleq [A^{N-h-1}G \ \cdots \ G \ \underbrace{0 \ \cdots \ 0}_h] \in \Re^{n \times Np}. \quad (10)$$

Therefore, we have

$$x_{k-N} = A^{-(N-h)}x_{k-h} - A^{-(N-h)}M_N W_{k-1} \quad (11)$$

under assumption of nonsingularity of the matrix A . From (3) and (11), we have

$$Y_{k-1} = \bar{C}_N x_{k-h} + \bar{G}_N W_{k-1} + V_{k-1} \quad (12)$$

where

$$\bar{C}_N \triangleq \tilde{C}_N A^{-(N-h)}, \quad (13)$$

$$\bar{G}_N \triangleq \tilde{G}_N - \tilde{C}_N A^{-(N-h)}M_N. \quad (14)$$

The noise term $\bar{G}_N W_{k-1} + V_{k-1}$ in (12) is zero-mean white Gaussian with covariance Π_N given by

$$\Pi_N \triangleq \bar{G}_N [\text{diag}(\overbrace{Q \cdots Q}^N)] \bar{G}_N^T + [\text{diag}(\overbrace{R \cdots R}^N)].$$

This has the following multivariable Gaussian density function:

$$f(\bar{G}_N W_{k-1} + V_{k-1}) = \frac{1}{\sqrt{(2\pi)^N \|\Pi_N\|}} \times \exp \left[-\frac{1}{2} (\bar{G}_N W_{k-1} + V_{k-1})^T \Pi_N^{-1} (\bar{G}_N W_{k-1} + V_{k-1}) \right].$$

It is noted that linear transformation on, and linear combinations of, Gaussian random processes are themselves Gaussian random processes. For this

reason, it is clear that when $\bar{G}_N W_{k-1} + V_{k-1}$ is Gaussian, Y_{k-1} is as well. The multivariable Gaussian density function of Y_{k-1} is derived from a shifted version of $f(\bar{G}_N W_{k-1} + V_{k-1})$ as follows:

$$\begin{aligned} f(Y_{k-1}|x_{k-h}) &= f(Y_{k-1} - \bar{C}_N x_{k-h}) \\ &= \frac{1}{\sqrt{(2\pi)^N \|\Pi_N\|}} \exp \left[-\frac{1}{2} (Y_{k-1} - \bar{C}_N x_{k-h})^T \Pi_N^{-1} (Y_{k-1} - \bar{C}_N x_{k-h}) \right], \end{aligned} \quad (15)$$

which called the likelihood function. The MLFS is obtained by maximizing the likelihood function (15) with respect to x_{k-h} . This maximization problem is equivalent to the minimization problem of the following cost:

$$J = \frac{1}{2} (Y_{k-1} - \bar{C}_N x_{k-h})^T \Pi_N^{-1} (Y_{k-1} - \bar{C}_N x_{k-h}). \quad (16)$$

Differentiating the cost (16) gives the following maximum likelihood criterion:

$$\frac{\partial J}{\partial x_{k-h}} = \bar{C}_N^T \Pi_N^{-1} (Y_{k-1} - \bar{C}_N x_{k-h}) = 0, \quad (17)$$

which is called the likelihood equation. The optimal MLFS \hat{x}_{k-h} is given by the solution of the likelihood equation (17) in the following theorem.

Theorem 1 Assume that (A, C) is observable and $N \geq n$. The MLFS is expressed as

$$\hat{x}_{k-h} = (\bar{C}_N^T \Pi_N^{-1} \bar{C}_N)^{-1} \bar{C}_N^T \Pi_N^{-1} Y_{k-1}. \quad (18)$$

Note that the matrix $\bar{C}_N^T \Pi_N^{-1} \bar{C}_N$ is guaranteed to be nonsingular if (A, C) is observable for $N \geq n$, since the matrix Π_N is positive definite.

Now we investigate inherent properties of the proposed MLFS.

Theorem 2 The MLFS is an unbiased smoother for noisy systems. It is also a deadbeat estimator for noise-free systems.

Proof. Taking the expectation on both sides of (12), the following relation is obtained:

$$E\{Y_{k-1}\} = \bar{C}_N E\{x_{k-h}\}. \quad (19)$$

From (18) and (19), we have

$$\begin{aligned} E\{\hat{x}_{k-h}\} &= (\bar{C}_N^T \Pi_N^{-1} \bar{C}_N)^{-1} \bar{C}_N^T \Pi_N^{-1} E\{Y_{k-1}\} \\ &= (\bar{C}_N^T \Pi_N^{-1} \bar{C}_N)^{-1} \bar{C}_N^T \Pi_N^{-1} \bar{C}_N E\{x_{k-h}\} \\ &= E\{x_{k-h}\}. \end{aligned}$$

This shows the unbiasedness property of the MLFS for the noisy system (1) and (2).

When there are no noises on the horizon $[k - N, k]$, the equation (12) becomes

$$Y_{k-1} = \bar{C}_N x_{k-h}. \quad (20)$$

From (18) and (20), it is shown that

$$\begin{aligned}\hat{x}_{k-h} &= (\bar{C}_N^T \Pi_N^{-1} \bar{C}_N)^{-1} \bar{C}_N^T \Pi_N^{-1} Y_{k-1} \\ &= (\bar{C}_N^T \Pi_N^{-1} \bar{C}_N)^{-1} \bar{C}_N^T \Pi_N^{-1} \bar{C}_N x_{k-h} \\ &= x_{k-h}.\end{aligned}$$

This completes the proof of the deadbeat property of the MLFS. ■

Remark 1 *The proposed MLFS exhibits unbiasedness and deadbeat properties. Note that these desirable properties indicate finite convergence time and fast tracking ability of the MLFS. Thus, it can be expected that the MLFS would be appropriate for fast estimation and detection of signals with unknown times of occurrence, which arise in many problems, such as fault detection, maneuver detection, and target tracking of flying objects.*

3 Numerical Example

A numerical example is given to provide a comparison between the proposed MLFS and the conventional fixed-lag Kalman smoother [1, 2] for the following F-404 engine model [5]:

$$\begin{aligned}x_{k+1} &= \begin{bmatrix} 0.9305 + \delta_k & 0 & 0.1107 \\ 0.0077 & 0.9802 + \delta_k & -0.0173 \\ 0.0142 & 0 & 0.8953 + 0.1\delta_k \end{bmatrix} x_k + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} w_k, \\ y_k &= \begin{bmatrix} 1 + 0.1\delta_k & 0 & 0 \\ 0 & 1 + 0.1\delta_k & 0 \end{bmatrix} x_k + v_k,\end{aligned}$$

where a model uncertain parameter δ_k is given by

$$\delta_k = \begin{cases} 0.1, & 50 \leq k \leq 100 \\ 0, & \text{otherwise.} \end{cases} \quad (21)$$

Noise covariance matrices are given by $Q = R = 0.02$. The horizon size and the fixed-lag size are taken as $N = 10$ and $h = 2$, respectively. Fig. 1 compares the robustness of two smoothers for the second state related to turbine temperature. Due to FIR structure, the estimation error of the proposed MLFS is considerably smaller than that of the conventional Kalman smoother with IIR structure on the interval where modeling uncertainty exists. Moreover, it is shown that the estimation error of the MLFS converges more rapidly than that of the Kalman smoother after the temporary modeling uncertainty disappears.

4 Conclusion

This paper has proposed the MLFS, which is a new fixed-lag maximum likelihood smoother with FIR structure, for state-space signal models. The proposed smoother does not require any *a priori* information on the horizon initial state. The MLFS processes only the finite outputs on the recent horizon $[k - N, k]$ linearly. The proposed MLFS exhibits the unbiasedness property

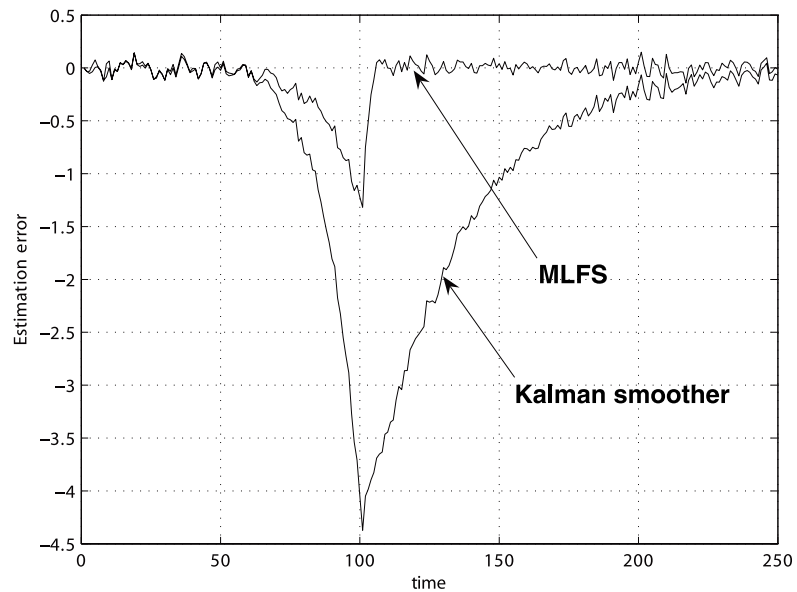


Fig. 1. Estimation errors for the MLFS and the Kalman smoother

and the deadbeat property. Furthermore, due to the FIR structure of the MLFS, the proposed smoother is believed to be robust against temporary modeling uncertainties or numerical errors, while the conventional Kalman smoother with the IIR structure may show poor robustness and even divergence phenomenon. Therefore, the MLFS can be a good substitute for the commonly used Kalman smoother in many signal processing problems where unknown signals are represented by state-space models.

Acknowledgments

The authors thank Prof. Soo Hee Han (Seoul National University) for his valuable discussions and Prof. Wook Hyun Kwon (Seoul National University) for his continuous encouragement.