Automatica 49 (2013) 1892–1899

Contents lists available at SciVerse ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Brief paper

Unified forms for Kalman and finite impulse response filtering and smoothing[☆]

Dan Simon[a,1], Yuriy S. Shmaliy[b]

[a] Cleveland State University, Cleveland, OH, USA

[b] Universidad de Guanajuato, Salamanca, Mexico

ARTICLE INFO

Article history:
Received 25 July 2012
Received in revised form
30 October 2012
Accepted 12 February 2013
Available online 26 March 2013

Keywords:
Kalman filters
Unbiased FIR filters
Smoothing
Filtering
Robust estimation

ABSTRACT

The Kalman filter and smoother are optimal state estimators under certain conditions. The Kalman filter is typically presented in a predictor/corrector format, but the Kalman smoother has never been derived in that format. We derive the Kalman smoother in a predictor/corrector format, thus providing a unified form for the Kalman filter and smoother. We also discuss unbiased finite impulse response (UFIR) filters and smoothers, which can provide a suboptimal but robust alternative to Kalman estimators. We derive two unified forms for UFIR filters and smoothers, and we derive lower and upper bounds for their estimation error covariances.

1. Introduction

We assume that we have a linear system described as

$$
\begin{aligned}
x_i &= F_i x_{i-1} + w_i \\
y_i &= H_i x_i + v_i
\end{aligned}
\qquad (1)
$$

where the time index $i \geq 1$, x_i is the K-dimensional state vector, y_i is the M-dimensional measurement, $\{w_i\}$ is a process noise sequence, $\{v_i\}$ is a measurement noise sequence, and system matrices F_i and H_i are known. Our objective is to estimate x_i based on the measurements and our knowledge of the system dynamics.

We use the term *estimator* to refer to the class of algorithms that includes filtering, prediction, and smoothing. A *filter* estimates x_i based on measurements up to and including time i. A *predictor* estimates x_i based on measurements prior to time i. A *smoother* estimates x_i based on measurements prior to time i, at time i, and later than time i.

Kalman estimation

The Kalman smoother can be written in fixed-lag form, fixed-interval form, or fixed-point form. These algorithms can be described as follows (Anderson & Moore, 2005) and (Simon, 2006, ch. 9).

- A fixed-lag smoother estimates x_i for $i \geq 1$ using measurements up to and including time $i + q$ for a fixed value of $q > 0$.
- A fixed-interval smoother estimates x_i for $i \in [1, N]$ using measurements up to and including time N.
- A fixed-point smoother estimates x_i using measurements up to and including time $i+q$ for a fixed value of i and for $q = 1, 2, \ldots$.

As we will see in Section 2, the form of the Kalman smoother is much different than that of the Kalman filter. Section 2.1 derives a Kalman smoother that is in the same form as the predictor/corrector form of the Kalman filter.

The Kalman filter is an infinite impulse response (IIR) filter; that is, each measurement y_m affects each estimate $\hat{x}_i$ for all $m \leq i$. The IIR nature of the Kalman filter makes it sensitive to modeling errors (Heffes, 1966; Nishimura, 1966; Soong, 1965). Over the past few decades, researchers have proposed many methods of making the Kalman estimator more robust (Peña & Guttman, 1988). Kalman estimation with uncertainties in the system matrices has been considered by many authors (Kosanam & Simon, 2004; Theodor & Shaked, 1996; Xie, Lu, Zhang, & Zhang, 2004; Zhang, Heemink, & Van Eijkeren, 1995); this is often called adaptive or robust Kalman estimation (Hide, Moore, & Smith, 2003). Methods for identifying noise covariances are presented in Alspach (1974); Mehra (1972) and Myers and Tapley (1976).

[☆] The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Tongwen Chen under the direction of Editor Ian R. Petersen.

E-mail addresses: d.j.simon@csuohio.edu (D. Simon), shmaliy@ugto.mx (Y.S. Shmaliy).

[1] Tel.: +1 216 687 5407; fax: +1 216 687 5405.

0005-1098/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.automatica.2013.02.026

Finite impulse response estimation

Whereas the research efforts mentioned above aimed to improve the Kalman estimator in the presence of mismodeling, we propose instead to use a finite impulse response (FIR) estimator. The advantages of transversal FIR estimators over Kalman estimators were recognized as far back as the 1960s, particularly in the areas of stability and robustness (Jazwinski, 1970). In spite of their history, FIR filters are not commonly used for state estimation. This is probably due to their analytical complexity and large computational burden. FIR smoothers can be used for polynomial models (Wang, 1991; Zhou & Wang, 2004). Order-recursive FIR smoothers were proposed for state space (Yuan & Stuller, 1994). General receding horizon FIR smoother theory has been developed (Ahn & Kim, 2008; Han & Kwon, 2007, 2008; Kwon, Han, Kwon, & Kwon, 2007). More recently, unbiased FIR (UFIR) smoothing of polynomial state space models has been considered (Shmaliy & Morales-Mendoza, 2010), and FIR smoothing was developed from the general p-shift estimator (Shmaliy, 2010, 2011; Shmaliy & Ibarra-Manzano, 2012). Iterative UFIR algorithms have also been developed (Shmaliy, 2010, 2011). These algorithms have the same predictor/corrector structure as the Kalman filter, often ignore the statistics of the noise and initial estimation errors, and become virtually optimal as the length of the FIR window increases.

Overview of the paper

Section 2 gives a brief review of Kalman filtering and smoothing, and derives a unified form for the two algorithms. Section 3 gives a review of UFIR filtering and smoothing, and derives two distinct but mathematically equivalent unified forms for the two algorithms. It also derives upper and lower bounds for the estimation error covariance. Section 4 presents some simulation results.

2. Kalman filtering and smoothing

If our estimate of x_i is based on measurements up to and including time t, we denote the estimate as $\hat{x}_{i|t}$. If $t = i$ then we have $\hat{x}_{i|i}$, which is called the *a posteriori* state estimate. If $t = i - 1$ then we have $\hat{x}_{i|i-1}$, which is called the *a priori* state estimate. If $t > i$, then we have a non-causal smoothed estimate. Suppose the following conditions hold:

(1) $\{w_i\}$ and $\{v_i\}$ are zero-mean, Gaussian, white, and uncorrelated, with known covariances Q_i and R_i respectively;
(2) We have an initial state estimate before any measurements are processed that we denote as $\hat{x}_{0|0}$;
(3) $(x_0 - \hat{x}_{0|0}) \sim N(0, P_{0|0})$, which means that the initial estimation error is Gaussian and zero-mean with covariance $P_{0|0}$.

Then the Kalman filter output is the mean of the state conditioned on measurements up to and including the current time:

$$\hat{x}_{i|i} = E(x_i | y_1, y_2, \ldots, y_i) \tag{2}$$

for $i \geq 1$. Furthermore, the Kalman filter estimate is the one that minimizes the trace of the covariance of the estimation error. The Kalman filter algorithm can be described as shown in Fig. 1, although there are also other equivalent formulations of the Kalman filter (Simon, 2006).

In the case of smoothing, we use future measurements to obtain the state estimate. One well-known smoothing algorithm is called the Rauch–Tung–Striebel (RTS) smoother, which is a type of fixed-interval smoother (Rauch, Tung, & Striebel, 1965) and (Simon, 2006, Section 9.4.2). Given measurements y_i for $i \in [1, N]$, the RTS smoother outputs $\hat{x}_{i|N}$ for all $i \in [0, N]$. The RTS smoother algorithm is summarized in Fig. 2.

$$\begin{aligned}
&\hat{x}_{0|0} = E(x_0)\\
&P_{0|0} = E\left[(x_0 - \hat{x}_{0|0})(x_0 - \hat{x}_{0|0})^T\right]\\
&\text{For } i = 1, 2, \cdots\\
&\qquad \hat{x}_{i|i-1} = F_i \hat{x}_{i-1|i-1}\\
&\qquad P_{i|i-1} = F_i P_{i-1|i-1} F_i^T + Q_i\\
&\qquad K_i = P_{i|i-1} H_i^T (H_i P_{i|i-1} H_i^T + R_i)^{-1}\\
&\qquad \hat{x}_{i|i} = \hat{x}_{i|i-1} + K_i (y_i - H_i \hat{x}_{i|i-1})\\
&\qquad P_{i|i} = (I - K_i H_i) P_{i|i-1}\\
&\text{Next } i
\end{aligned}$$

Fig. 1. The Kalman filter. K_i is the Kalman gain, $P_{i|i}$ is the *a posteriori* estimation error covariance, and $P_{i|i-1}$ is the *a priori* estimation error covariance.

$$\begin{aligned}
&\text{Execute the Kalman filter for } i = 1, 2, \cdots, n \text{ (See Fig. 1)}\\
&\text{Initialize } P_n^s = P_{n|n}\\
&\text{For } q = 1, 2, \cdots, n\\
&\qquad i = n - q\\
&\qquad K_i^s = P_{i|i} F_{i+1}^T (P_{i+1|i})^{-1}\\
&\qquad P_i^s = P_{i|i} - K_i^s (P_{i+1|i} - P_{i+1}^s)(K_i^s)^T\\
&\qquad \hat{x}_{i|n} = \hat{x}_{i|i} + K_i^s (\hat{x}_{i+1|n} - \hat{x}_{i+1|i})\\
&\text{Next } i
\end{aligned}$$

Fig. 2. The RTS smoother. K_i^s is the Kalman smoother gain, and P_i^s is the covariance of the error of the smoothed estimate at time i.

2.1. Unified Kalman filtering and smoothing

Fig. 1 shows that the Kalman filter estimate can be written in the form

$$\hat{x}_{i|i} = \gamma_i \hat{x}_{i-1|i-1} + K_i y_i$$
$$\text{where } \gamma_i = (I - K_i H_i) F_i \tag{3}$$

for $i \geq 1$. This is called a predictor/corrector form. However, the smoothed estimate in Fig. 2 does not have this form. We would like to find a similar form for the smoothed estimate:

$$\hat{x}_{n-q|n} = \gamma_{n,q} \hat{x}_{n-1|n-1} + \sum_{m=n-q+1}^{n} \beta_{n,q,m} y_m \tag{4}$$

where the smoother lag $q > 0$. Such a form could serve at least two purposes.

First, we find it mathematically attractive to obtain unified forms for different algorithms. We see this in many areas of science and engineering (Fonseca & Fleming, 1998; Guerreiro & Trigueiros, 2010; Miller & Boxer, 1999), so the parallel form of (3) and (4) is intuitively appealing.

Second, the smoother form of (4) may have practical benefits because it directly shows the additional sensitivity of the smoothed estimate to each measurement, beyond the sensitivity already incorporated in $\hat{x}_{n-1|n-1}$. $\beta_{n,q,m}$ is the sensitivity of $\hat{x}_{n-q|n}$ to y_m for $m \in [n-q+1, n]$ beyond the sensitivity that is implicit in $\hat{x}_{n-1|n-1}$. These sensitivities could be used to process measurements in order of decreasing sensitivity so that the most important measurements are processed first, in case the timeliness of the smoothed estimate is important.

Note that all of the measurements up to and including time $n-1$ are incorporated in the filtered estimate $\hat{x}_{n-1|n-1}$ in (4). However, the additional contribution of those measurements to obtain the smoothed estimate $\hat{x}_{n-q|n}$ is determined by the $\beta_{n,q,m}$ coefficients. We suppose that the estimate $\hat{x}_{n-1|n-1}$ is available and that the user may want to process only a subset of the measurements to obtain the smoothed estimate.

To be more specific, (4) can be written algorithmically by computing

$$\mu(l) = \text{value of } m \text{ in the } l\text{-th largest value of } \beta_{n,q,m} \tag{5}$$

for $m \in [n-q+1, n]$ and $l \in [1, q]$. When we say "l-th largest value of $\beta_{n,q,m}$", we implicitly assume some matrix or vector norm. After