A Kalman-like algorithm with no requirements for noise and initial conditions

Plenary Lecture

YURIY S. SHMALIY
Guanajuato University
Department of Electronics, DICIS
Ctra. Salamanca-Valle, 3.5+1.8km, Palo Blanco, Salamanca, 36855, Gto.
MEXICO
shmaliy@salamanca.ugto.mx

Abstract: We address and study a new Kalman-like estimator for solving universally the problems of filtering \((p = 0)\), prediction \((p > 0)\), and smoothing \((p < 0)\) of discrete time-varying state-space models with no requirements for noise and initial conditions. A solution is first found in a batch form and then represented in the computationally efficient iterative Kalman-like one. It is shown that the estimator proposed overperforms the Kalman one when 1) noise covariances and initial conditions are not known exactly, 2) noise constituents are not white sequences, and 3) both the system and measurement noise components need to be filtered out and the deterministic state estimated. Otherwise, the Kalman-like and Kalman filters produce similar errors. A numerical comparison with the Kalman filter is provided for the two-state space models.

Key–Words: Kalman filter, unbiased FIR filter, Kalman-like algorithm

1 Introduction

In signal processing, the terms “Kalman-like” and “Kalman-type” are commonly used whenever the standard linear Kalman filtering algorithm [1] is modified to estimate state of the nonlinear model, under unknown initial conditions, in the presence of nonwhite or multiplicative noise sources, etc. In such improper applications for the Kalman filter, the Kalman-like one is designed to save the recursive structure, while connecting the algorithm components with the model in different ways. Because there can be found an infinity of the Kalman-like solutions depending on applications, we meet a number of propositions suggesting some new qualities while saving (or not deteriorating substantially) the advantages of the Kalman filter: accuracy, fast computation, and small memory. There were also proposed iterative Kalman-like forms [2, 3] for the time-invariant finite impulse response (FIR) filters, having certain advantages against the Kalman filter such as better stability and robustness and lower sensitivity to noise and initial conditions.

In this paper, we employ the approach developed in [3] and address a \(p\)-shift time-variant Kalman-like algorithm intended for filtering \((p = 0)\), prediction \((p > 0)\), and smoothing \((p < 0)\) of discrete time-varying state-space models with no requirements for noise and initial conditions.

2 Signal Model and Problem Formulation

Consider a discrete time-variant linear model represented with the state and observation equations, respectively,

\[
\begin{align*}
\mathbf{x}_n &= \mathbf{A}_n \mathbf{x}_{n-1} + \mathbf{B}_n \mathbf{w}_n, \\
\mathbf{y}_n &= \mathbf{C}_n \mathbf{x}_n + \mathbf{D}_n \mathbf{v}_n,
\end{align*}
\]

where the \(K \times 1\) state vector and \(M \times 1\) observation vector are given by \(\mathbf{x}_n = [x_{1n} \ x_{2n} \ldots x_{Kn}]^T\) and \(\mathbf{y}_n = [y_{1n} \ y_{2n} \ldots y_{Mn}]^T\), respectively; \(\mathbf{A}_n\) is the \(K \times K\) system matrix and \(\mathbf{C}_n\) is the \(M \times K\) measurement matrix. Most generally, \(\mathbf{B}_n\) and \(\mathbf{D}_n\) have \(K \times K\) and \(M \times M\) dimensions, respectively. The \(K \times 1\) input noise vector \(\mathbf{w}_n = [w_{1n} \ w_{2n} \ldots w_{Kn}]^T\) and \(M \times 1\) measurement noise vector \(\mathbf{v}_n = [v_{1n} \ v_{2n} \ldots v_{Mn}]^T\) have zero-mean components, \(E\{\mathbf{w}_n\} = 0\) and \(E\{\mathbf{v}_n\} = 0\). It is implied that \(\mathbf{w}_n\) and \(\mathbf{v}_n\) are mutually uncorrelated and independent processes, \(E\{\mathbf{w}_i \mathbf{v}_j^T\} = 0\), having arbitrary covariances, \(\mathbf{Q}_w(i, j) = E\{\mathbf{w}_i \mathbf{w}_j^T\}\) and \(\mathbf{Q}_v(i, j) = E\{\mathbf{v}_i \mathbf{v}_j^T\}\), respectively, for all \(i\) and \(j\).

The strategies of the recursive Kalman algorithm [1] and iterative Kalman-like FIR one [3] are illustrated in Fig. 1. The Kalman filter starts at some initial
In order to provide FIR estimation, (1) and (2) can be expanded on an averaging interval of \( N \) points from \( m = n - N + 1 \) to \( n \) as in [3], respectively,

\[
\begin{align*}
X_{n,m} &= A_{n,m}x_m + B_{n,m}W_{n,m}, \\
Y_{n,m} &= C_{n,m}x_m + G_{n,m}W_{n,m} + D_{n,m}v_{n,m},
\end{align*}
\]

(3)

(4)

where the \( KN \times 1 \) state vector \( X_{n,m} \), the \( MN \times 1 \) observation vector \( Y_{n,m} \), the \( KN \times 1 \) state noise vector \( W_{n,m} \), and the \( MN \times 1 \) observation noise vector \( V_{n,m} \) are specified from \( m \) to \( n \) by, respectively,

\[
\begin{align*}
X_{n,m} &= [x_n^T x_{n-1}^T \ldots x_m^T]^T, \\
Y_{n,m} &= [y_n^T y_{n-1}^T \ldots y_m^T]^T, \\
W_{n,m} &= [w_n^T w_{n-1}^T \ldots w_m^T]^T, \\
V_{n,m} &= [v_n^T v_{n-1}^T \ldots v_m^T]^T.
\end{align*}
\]

(5)

(6)

(7)

(8)

The \( KN \times K \) system noise matrix \( A_{n,m} \), the \( MN \times K \) observation matrix \( C_{n,m} \), and the \( MN \times MN \) measurement noise matrix \( D_{n,m} \) are given with, respectively,

\[
\begin{align*}
A_{n,m} &= \begin{bmatrix}
n-m-1 \\
\prod_{i=0}^{n-m-1} A_{n-i} \\
\vdots \\
A_{m+1} \\
I
\end{bmatrix}, \\
C_{n,m} &= \begin{bmatrix}
n-m-1 \\
\prod_{i=0}^{n-m-1} A_{n-i} \\
\vdots \\
C_{m+1}A_{m+1} \\
C_m
\end{bmatrix}, \\
D_{n,m} &= \text{diag}(D_n D_{n-1} \ldots D_m).
\end{align*}
\]

(9)

(10)

(11)

The \( KN \times KN \) system noise matrix \( B_{n,m} \) and the \( MN \times KN \) matrix \( G_{n,m} \) are not further involved and omitted.

Now, assign some \( K \times MN \) time-variant gain matrix \( H_n(m,p) \) such that the estimate \( \tilde{x}_{n+p|m} \) of \( x_n \) is

\[
\begin{align*}
\tilde{x}_{n+p|m} &= H_n(m,p)Y_{n,m} \quad (12a) \\
&= H_n(m,p)[C_{n,m}x_m + G_{n,m}W_{n,m} + D_{n,m}v_{n,m}].
\end{align*}
\]

(12b)

By the unbiasedness condition, \( E\{x_{n+p}|n\} = E\{\tilde{x}_{n+p|m}\} \), and \( x_{n+p} \) specified via (3) with a shift \( p \) induced as

\[
\begin{align*}
\bar{x}_{n+p|m} &= \prod_{i=0}^{p-m-1} A_{n+p-i}x_{m},
\end{align*}
\]

(13)

where \( \bar{x}_{n+p|m} \) is obtained at \( n + p \) via measurement from the past to \( n \).
we arrive at the unbiasedness constraint
\[
\prod_{i=0}^{n+p-m-1} A_{n+p-i} = \mathbf{H}_n(m, p) C_{n,m},
\]
(14)
where \( \mathbf{H}_n(m, p) \) is the time-varying unbiased FIR estimator gain. This constraint gives us the unbiased FIR estimator gain
\[
\mathbf{H}_n(m, p) = \prod_{i=0}^{n+p-m-1} A_{n+p-i}(C^T_{n,m} C_{n,m})^{-1} C^T_{n,m}.
\]
(15)
It can easily be verified that (15) becomes \([3, \text{eq. (33)}]\) derived for time-invariant models. Provided (15), the time-varying unbiased FIR estimate is thus specified in the batch form as
\[
\mathbf{x}_{n+p|n} = \mathbf{H}_n(m, p) \mathbf{Y}_{n,m}
\]
(16a)
\[
= \prod_{i=0}^{n+p-m-1} A_{n+p-i}(C^T_{n,m} C_{n,m})^{-1}
\times C^T_{n,m} \mathbf{Y}_{n,m},
\]
(16b)
where \( C_{n,m} \) is given by (10) and \( \mathbf{Y}_{n,m} \) by (6).

### 3.1 Estimate Errors
The MSE in (16b) can be ascertained at \( n + p \) by
\[
\mathbf{J}(p) = E\{ (\mathbf{x}_{n+p} - \bar{\mathbf{x}}_{n+p|n})(\mathbf{x}_{n+p} - \bar{\mathbf{x}}_{n+p|n})^T \},
\]
(17)
if we substitute \( \mathbf{x}_{n+p|n} \) with (13). Following [3] one can arrive at
\[
\mathbf{J}(p) = \mathbf{H}_n(m, p)(\tilde{\mathbf{z}}_p + \bar{\mathbf{z}}_p)\mathbf{H}_n^T(m, p)
\]
(18)
where \( \tilde{\mathbf{z}}_p = \mathbf{G}_{n,m} E\{ \mathbf{W}_{n,m} \mathbf{W}^T_{n,m}\} \mathbf{C}_{n,m}^T \) and \( \bar{\mathbf{z}}_p = \mathbf{D}_{n,m} E\{ \mathbf{V}_{n,m} \mathbf{V}^T_{n,m}\} \mathbf{D}^T_{n,m} \) are noise functions, and deduce that, by zero noise matrices \( \tilde{\mathbf{z}}_p = 0 \) and \( \bar{\mathbf{z}}_p = 0 \), the MSE becomes identically zero. In the white Gaussian approximation \([6]\), the estimate errors can also be evaluated employing the noise power gain (NPG) computed via (15) as \([3]\)
\[
\mathbf{K}_n(m, p) = \mathbf{H}_n(m, p)\mathbf{H}_n^T(m, p).
\]
(19)

### 4 Time-Varying Kalman-Like Estimator
For fast computation of (16b), the following theorem suggests an iterative Kalman-like algorithm, which proof is obtained similarly to that given in [3] for time-invariant models.

**Theorem 1** Given the unbiased FIR estimator (16b), then its iterative Kalman-like algorithm is the following:
\[
\Xi_l = A_l^T C^T_l C_l A_l, \hspace{1cm} (20)
\]
\[
\mathbf{P} = (C^T_{m,s} C_{m,s})^{-1}, \hspace{1cm} (21)
\]
\[
\mathbf{F}_s = \prod_{i=0}^{K-2} A_{s-i} \left( \prod_{i=0}^{K-2} A_{s-i} \right)^T, \hspace{1cm} (22)
\]
\[
\mathbf{F}_{l|p} = A_{l+p}\mathbf{F}_{l+p|l+p-1}^{-1} A_l^T, \hspace{1cm} (24)
\]
\[
\mathbf{x}_{l+p|l} = A_{l+p}\mathbf{x}_{l+p|l+p-1} + A_{l+p}\Phi_{l+p|l+p-1}^{-1} \mathbf{F}_{l|p} C_l^T \times [\mathbf{Y}_l - C_l\Phi_{l+p|l+p-1}^{-1} \mathbf{x}_{l+p|l+p-1}] , \hspace{1cm} (25)
\]
where \( \Phi_{l+p} \) takes values depending on \( p \) as
\[
\Phi_{l+p} = \begin{cases} \prod_{i=0}^{p} A_{l-i} , & p \leq -1 \quad \text{(smoothing)} \\ A_l , & p = 0 \quad \text{(filtering)} \\ \mathbf{I} , & p = 1 \quad \text{(prediction)} \end{cases} \quad (26)
\]
s = \( m - K + 1 \), \( m = n - N + 1 \), and an iterative variable \( l \) ranges from \( m + K \) to \( n \) to produce a true value at \( l = n \).

As can be seen, (25) is the Kalman-like estimate, which gain does not depend on noise and initial conditions. The algorithm has two batch forms, (22) and (23), that can be computed fast for a typically small \( K \). Note that the minimum value for \( l \) is chosen to be \( m + K \), because \( C_l^T C_l \) is singular otherwise.

### 5 Examples of Applications
The purpose of this section is to show that the Kalman-like filter (theorem 1) is able to produce errors similar or even lower than in the Kalman one.

#### 5.1 Polynomial two-state model
In the first experiment, a two-state signal model was chosen with \( p = 0 \), \( \mathbf{B}_n = \mathbf{I} \), \( \mathbf{D}_n = \mathbf{I} \), \( \mathbf{C}_n = [1 \ 0] \), and
\[
\mathbf{A}_n = \begin{bmatrix} 1 & (1 + d_n) \tau \\ 0 & 1 \end{bmatrix}, \hspace{1cm} (27)
\]
where \( d_n = 20 \) for \( 120 \leq n \leq 160 \) and \( d_n = 0 \) otherwise. The variances of zero mean white sequences
in $x_{1n}$ and $x_{2n}$ were allowed to be $\sigma_1^2 = 10^{-4}$ and $\sigma_2^2 = 4 \times 10^{-2}/s^2$, respectively. Measurement of the first state was organized in the presence of noise $w_n$ having $\sigma_w^2 = 50^2$. The process was simulated at 340 subsequent points and the Kalman and Kalman-like algorithms applied in two options. The time-varying filtering was organized with $d_n$ nonconstant and the time-invariant one with $d_n = 0$ over all measurement. For the Kalman-like filter, the optimum $N$ were found to be $N = 54$ and $N = 12$ beyond and within the uncertainty region, respectively, provided minimum MSEs in the estimates.

Figure 2 sketches typical measurement and estimates obtained with the time-varying and time-invariant filters. As can be seen in Fig. 2a, both time-invariant, the Kalman and Kalman-like estimates of the first state produce large errors in the region of variations. In contrast, the difference between the time-varying estimates is poorly distinguishable in this figure and we go to the estimate error shown in Fig. 2b and Fig. 2c. Here, the Kalman-like filter demonstrates an important peculiarity that still has not been discussed in the literature. Because the model properties are changed by time variations, the optimal horizon length also changes. In fact, with $N = 54$, errors in the Kalman-like filter appear to be larger in the region of variations (Fig. 2b). However, both filters produce similar errors if to let $N = 12$ in this region. It is also seen that $N$ can be unaltered for short-time variations. In fact, there is no substantial discrepancy between the estimates from $n = 120$ to $130$ in Fig. 2b. Another difference between the time-invariant estimates is in the transients that are inherently finite in the Kalman-like filter (as associated with length $N$) and infinite in the Kalman one (Fig. 2c). Note that transients in the Kalman-like filter exhibit a bit larger excursions.

In addition, Fig. 2d shows filtering errors in the second state. The horizon length here does not affect substantially the estimates and we watch for similar behaviors, except for the inherently longer transient in the Kalman filter.

### 5.2 Harmonic two-state model

In the second experiment, we compare the Kalman and Kalman-like estimates obtained for the two-state harmonic model, (1) and (2), with the following vectors and matrices: $w_n = [w_n]$, $v_n = [v_n]$, $B_n = [1\ 1]^T$, $D_n = [1]$, $C_n = [1\ 0]$, and

$$A_n = \begin{bmatrix} \cos \left( \frac{\pi}{2n} \right) + \delta_n & (\sin \left( \frac{\pi}{2n} \right) \left( \sin \left( \frac{\pi}{2n} \right) - \frac{\pi}{2n} \right) \cos \left( \frac{\pi}{2n} \right) + \delta_n \end{bmatrix}. \tag{28}$$

The process was generated with $x_{10} = 100$, $x_{20} = 10^{-2}$, $\sigma_w^2 = 1$, and $\sigma_v^2 = 100$. Uncertainty

![Figure 2: Kalman and Kalman-like filtering of the time-variant two-state polynomial model: (a) measurement and estimates of the first state, (b) filtering error of the first state with a constant $N = 54$, (c) filtering error of the first state with a variable $N$, and (d) filtering errors of the second state.](image)
Kalman-like filter. We finally employ the full-horizon Kalman-like algorithm with \( N = n - 1 \). For the case of filtering (\( p = 0 \)), this algorithm needs only \( K \) and simplifies to:

\[
P_{K-1} = \left( C_{K-1,0}^T C_{K-1,0} \right)^{-1},
\]

\[
F_{K-1} = \prod_{i=0}^{K-2} A_{K-1-i} P_{K-1} \left( \prod_{i=0}^{K-2} A_{K-1-i} \right)^T,
\]

\[
\bar{x}_{K-1|K-1} = \prod_{i=0}^{K-2} A_{K-1-i} P_{K-1} C_{K-1,0}^T Y_{K-1,0},
\]

\[
F_n = \left[ C_n^T C_n + (A_n F_{n-1} A_n^T)^{-1} \right]^{-1},
\]

\[
\bar{x}_{n|n} = A_n \bar{x}_{n-1|n-1} + F_n C_n^T (y_n - C_n A_n \bar{x}_{n-1|n-1}).
\]

It was shown in [3] that the full-horizon FIR algorithm allows for filtering out both the measurement and system noise components. In this sense, the algorithm (29)–(33) provides us with the best fit for the noisy signal measured with noise, by changing \( p \) for a fixed \( n \). Figure 4 illustrates a typical example of filtering of a harmonic signal without uncertainties. It is neatly seen in Fig. 4a that the full-horizon Kalman-like filter tracks the noise-free (actual) behavior of the signal, whereas the Kalman one tracks the mean-value of the measurement. Accordingly, the filtering errors appear to be lower in the Kalman-like estimates (Fig. 4b). Three typical regions can be recognized observing Fig. 4b. In the first region of small averaging horizons, the Kalman filter produces lower errors than in the Kalman-like one. In the intermediate region (from 30 to 50), both filters produce similar errors. An averaging horizon taken as \( 30 < N < 50 \) can thus be said to be optimal for the fixed-horizon state filtering in the sense of the same MSE as in the optimal Kalman filter. In the third region of larger horizons, filtering out of both the measurement and system noise components is much better provided by the Kalman-like filter. We notice that this effect is reproducible disregarding the model.

5.2.1 Filtering with uncertainties

Figure 3 sketches the filtering estimates and errors for the time-varying signal. One can observe that, in the case of time-invariant filtering (Fig. 3a), both filters produce negligible errors beyond the uncertainty region. Then the Kalman-like filter reaches a new state of the signal much faster than the Kalman one. The relevant errors associated with this case are shown in Fig. 3b. In the time-varying filtering, both filters exhibit similar errors (Fig. 3c), although a bit better performance has the Kalman-like filter.

5.2.2 Full-horizon filtering of a distinct Model

We finally employ the full-horizon Kalman-like algorithm with \( N = n - 1 \). For the case of filtering (\( p = 0 \)), this algorithm needs only \( K \) and simplifies to:

\[
P_{K-1} = \left( C_{K-1,0}^T C_{K-1,0} \right)^{-1},
\]

\[
F_{K-1} = \prod_{i=0}^{K-2} A_{K-1-i} P_{K-1} \left( \prod_{i=0}^{K-2} A_{K-1-i} \right)^T,
\]

\[
\bar{x}_{K-1|K-1} = \prod_{i=0}^{K-2} A_{K-1-i} P_{K-1} C_{K-1,0}^T Y_{K-1,0},
\]

\[
F_n = \left[ C_n^T C_n + (A_n F_{n-1} A_n^T)^{-1} \right]^{-1},
\]

\[
\bar{x}_{n|n} = A_n \bar{x}_{n-1|n-1} + F_n C_n^T (y_n - C_n A_n \bar{x}_{n-1|n-1}).
\]

It was shown in [3] that the full-horizon FIR algorithm allows for filtering out both the measurement and system noise components. In this sense, the algorithm (29)–(33) provides us with the best fit for the noisy signal measured with noise, by changing \( p \) for a fixed \( n \). Figure 4 illustrates a typical example of filtering of a harmonic signal without uncertainties. It is neatly seen in Fig. 4a that the full-horizon Kalman-like filter tracks the noise-free (actual) behavior of the signal, whereas the Kalman one tracks the mean-value of the measurement. Accordingly, the filtering errors appear to be lower in the Kalman-like estimates (Fig. 4b). Three typical regions can be recognized observing Fig. 4b. In the first region of small averaging horizons, the Kalman filter produces lower errors than in the Kalman-like one. In the intermediate region (from 30 to 50), both filters produce similar errors. An averaging horizon taken as \( 30 < N < 50 \) can thus be said to be optimal for the fixed-horizon state filtering in the sense of the same MSE as in the optimal Kalman filter. In the third region of larger horizons, filtering out of both the measurement and system noise components is much better provided by the Kalman-like filter. We notice that this effect is reproducible disregarding the model.
have shown that this estimator is able to overperform the Kalman filter in the following special cases when

- Noise covariances and initial conditions are not known exactly,
- Noise components are not white,
- Both the system and measurement noise parts need to be filtered out.

Otherwise, the Kalman-like and Kalman filters produce similar errors.

References:


