A. Germani, C. Manes, P. Palumbo

A MINIMUM VARIANCE FILTER
FOR DISCRETE-TIME LINEAR SYSTEMS
PERTURBED BY UNKNOWN NONLINEARITIES

R. 575  Settembre 2002

Alfredo Germani  – Dipartimento di Ingegneria Elettrica, Università degli Studi dell’Aquila,
67040 Monteluco (L’Aquila), Italy and Istituto di Analisi dei Sistemi ed Informatica del
CNR, Viale Manzoni 30, 00185 Roma, Italy. Email: germani@ing.univaq.it.

Costanzo Manes  – Dipartimento di Ingegneria Elettrica, Università degli Studi dell’Aquila,
67040 Monteluco (L’Aquila), Italy and Istituto di Analisi dei Sistemi ed Informatica del
CNR, Viale Manzoni 30, 00185 Roma, Italy. Email: manes@ing.univaq.it.

Pasquale Palumbo  – Istituto di Analisi dei Sistemi ed Informatica del CNR, Viale Manzoni
30, 00185 Roma, Italy. Email: palumbo@iasi.rm.cnr.it.

ISSN: 1128–3378
Abstract

This paper investigates the problem of state estimation for discrete-time stochastic systems with linear dynamics perturbed by unknown nonlinearities. The Extended Kalman Filter (EKF) cannot be applied in this framework, because the lack of knowledge on the nonlinear terms forbids a reliable linear approximation of the perturbed system. Following the idea to compensate this lack of knowledge suitably exploiting the information brought by the measured output, a recursive linear filter is developed according to the minimum error variance criterion. Differently from what happens for the EKF, the gain of the proposed filter can be computed off-line. Numerical simulations show the effectiveness of the filtering algorithm.

Key words: State estimation, nonlinear systems, uncertain systems, Extended Kalman Filter.
1. Introduction

This paper is concerned with the filtering problem for systems that have a nominal linear behavior perturbed by nonlinear terms that are unknown to the filter designer. In addition, random noises affect both the state and output equations. A short version of this paper was presented in [Bangkok]. Many works in literature deal with the problem of state estimation when the system model and/or the noise model are uncertain. The \( H_\infty \) approach provides a solution for the filtering problem of linear systems when the noise energy is bounded [7], or when the uncertainties on the system matrices are bounded [8]. In [9] an \( H_\infty \) filter is presented for uncertain linear systems perturbed by known nonlinear terms satisfying a Lipschitz condition. A classical filtering approach in the presence of known nonlinearities is the so-called Extended Kalman Filter (EKF), whose performance is not guaranteed, but is easy to implement and in many cases gives satisfactory results. A robust version of the EKF filter, in the case of known nonlinear terms, is presented in [2]. In [3] a robust \( H_\infty \) filter is presented for continuous time linear systems in the presence of uncertain nonlinear perturbations that satisfy a known bound on the norm. All the cited papers do not handle a stochastic noise model. Moreover, a partial knowledge on the system nonlinearities (the norm bound or the Lipschitz constant) is required and appears in the filter equations or plays a role in the filter design.

This paper develops a filter that does not require any knowledge on the form of the nonlinear perturbation terms, and that is optimal, according to the minimum error variance criterion, in a specific class of estimators, following the approach in [4], [5], [6]. The key-point in the filter derivation is a clever use of the measured output vector in such a way to compensate the lack of knowledge on the nonlinear terms.

2. Linear systems with nonlinear perturbations

This paper considers stochastic systems of the type

\[
\begin{align*}
    x(0) &= x_0, \\
    x(k+1) &= A(k)x(k) + B(k)u(k) + R(k)h(k, x(k)) + N_f(k), \quad k \geq 0, \\
    y(k) &= C(k)x(k) + N_g(k),
\end{align*}
\]

(2.1)

where \( x(k) \in \mathbb{R}^n \) is the system state, \( u(k) \in \mathbb{R}^p \) is a known input, \( y(k) \in \mathbb{R}^q \) is the measured output, \( h(\cdot, \cdot) : \mathbb{R}^+ \times \mathbb{R}^n \mapsto \mathbb{R}^m \) is a nonlinear map, such that it guarantees the finite second order moment for the state. No more assumptions on its shape will be assumed in the following. \( \{N_f(k)\}, \{N_g(k)\} \) are the state and output noise sequences, and are assumed white, mutually uncorrelated and not necessarily Gaussian, with known covariances \( \{Q_f(k)\} \) and \( \{Q_g(k)\} \). The initial state \( x_0 \) is a random variable, with mean value \( \bar{x}_0 \) and covariance matrix \( \Psi_0 \), uncorrelated with the noise sequences. \( A(k), B(k), R(k), C(k) \) are matrices of suitable dimensions.

The novelty of this approach is to compensate the lack of any knowledge on the nonlinear map \( h(\cdot, \cdot) \) by a suitable processing of the output measurements \( y(k) \). The following Lemma illustrates how this can be done.

Lemma 2.1. Define the extended state \( X_e(k) \in \mathbb{R}^{n+m} \) as follows:

\[
    X_e(0) = \begin{bmatrix} x_0 \\ 0 \end{bmatrix}, \quad X_e(k) = \begin{bmatrix} x(k) \\ h(k-1, x(k-1)) \end{bmatrix}, \quad k > 0.
\]

(2.2)
Suppose that for all \( k > 0 \) the matrix sequence
\[
H(k) = \begin{bmatrix}
I_n & -R(k-1) \\
C(k) & O_{q \times m}
\end{bmatrix}, \quad k > 0,
\]
has full column rank \( n + m \), and define a sequence of left-inverses \( \{H^\dagger(k)\} \) (i.e., matrices such that \( H^\dagger(k)H(k) = I_{n+m} \)). Then the evolution of \( X_e(k) \) obeys the following (non-strictly causal) model:
\[
\begin{align*}
X_e(k+1) &= A_e(k)X_e(k) + B_e(k)u(k) + D_e(k)y(k+1) + F_e(k), \\
y(k) &= C_e(k)X_e(k) + G_e(k),
\end{align*}
\]
where
\[
\begin{align*}
A_e(k) &= H^\dagger(k+1) \begin{bmatrix} A(k) & O_{n \times m} \\ O_{q \times n} & O_{q \times m} \end{bmatrix}, \\
B_e(k) &= H^\dagger(k+1) \begin{bmatrix} B(k) \\ O_{q \times p} \end{bmatrix}, \\
D_e(k) &= H^\dagger(k+1) \begin{bmatrix} O_{n \times q} \\ I_q \end{bmatrix}, \\
C_e(k) &= [C(k) \ O_{q \times m}],
\end{align*}
\]
and with \( F_e(k) \) and \( G_e(k) \) extended noise sequences defined as:
\[
\begin{align*}
F_e(k) &= H^\dagger(k+1) \begin{bmatrix} N_f(k) \\ -N_g(k+1) \end{bmatrix}, \\
G_e(k) &= N_g(k).
\end{align*}
\]

Proof. Taking into account the matrices in (2.3), (2.5), the extended state in (2.2) and the extended state noise sequence in (2.6), the right hand side of the state equation in (2.4) is:
\[
\begin{align*}
H(k+1)^\dagger \begin{bmatrix} A(k)x(k) \\ 0 \end{bmatrix} + H(k+1)^\dagger \begin{bmatrix} B(k)u(k) \\ 0 \end{bmatrix} \\
+ H(k+1)^\dagger \begin{bmatrix} 0 \\ y(k+1) \end{bmatrix} + H(k+1)^\dagger \begin{bmatrix} N_f(k) \\ -N_g(k+1) \end{bmatrix}
\end{align*}
\]
\[
= H(k+1)^\dagger \begin{bmatrix} A(k)x(k) + B(k)u(k) + N_f(k) \\ y(k+1) - N_g(k+1) \end{bmatrix}
= H(k+1)^\dagger \begin{bmatrix} x(k+1) - R(k)h(k, x(k)) \\ C(k+1)x(k+1) \end{bmatrix}
= H(k+1)^\dagger \begin{bmatrix} I_n \\ C(k+1) \end{bmatrix} \begin{bmatrix} -R(k) \\ O_{q \times m} \end{bmatrix} \begin{bmatrix} x(k+1) \\ h(k, x(k)) \end{bmatrix}
= H(k+1)^\dagger H(k+1)X_e(k+1) = X_e(k+1).
\]
The measurement equation in (2.4) easily comes through direct substitution. ■

Remark 2.2. Since all matrices of the sequence \( \{H(k)\} \) have dimensions \((n + q) \times (n + m)\), the full column-rank condition assumed in Lemma 2.1 requires that \( q \geq m \): this means that the number of output components must be at least equal to the number of independent nonlinear perturbation terms. This is a reasonable necessary condition, in that the measurements have been used to replace the nonlinear terms in a state-space equation. The important result is that in equation (2.4) the unknown nonlinear term \( h(\cdot, \cdot) \) does not appear. On the other hand, the strict causality of the original system is lost, because the linear recursive state equation requires \( y(k+1) \) to compute \( X_e(k+1) \).

Remark 2.3. The extended state noise \( \{F_e(k)\} \) is still a white sequence, but is no more uncorrelated with the output noise \( \{G_e(k)\} \): from definition (2.6) it is evident a one-step correlation.
3. The filtering algorithm

The class of estimators considered in this paper is given by:

**Definition 3.1.** A state estimator for the class of linear stochastic systems perturbed by unknown nonlinear terms of the type (2.1) is said to be insensitive to the Nonlinear Perturbation (shortly, NLP-insensitive) if its structure does not depend explicitly on the nonlinear terms.

The filtering algorithm here proposed is based on the estimation of the extended state of system (2.4). The same ideas developed in [6] are used here to deal with the non-strict causality of system (2.4). The following decomposition is suggested:

**Proposition 3.2.** The extended state (2.2) and the output of system (2.4) can be decomposed as:

\[ X_e(k) = X_d(k) + X_s(k), \]
\[ y(k) = y_s(k) + C_e X_d(k), \]

where the dynamics of \(X_d(k)\) and \(X_s(k)\) are described by the two systems

\[
\begin{align*}
X_d(0) &= \mathbb{E}[X_e(0)] = (\bar{x}_0^T & 0^T)^T, \\
X_d(k+1) &= A_e(k)X_d(k) + B_e(k)u(k) + D_e(k)y(k+1),
\end{align*}
\]
\[
\begin{align*}
X_s(0) &= X_e(0) - \mathbb{E}[X_e(0)] = (\bar{x}_0^T - \bar{x}_0^T & 0^T)^T, \\
X_s(k+1) &= A_e(k)X_s(k) + \mathcal{F}_e(k), \\
y_s(k) &= C_e X_s(k) + \mathcal{G}_e(k).
\end{align*}
\]

According to Proposition 3.2, the extended state is split into two components: one, named \(X_d(k)\), is not directly affected by the noises and can be exactly computed as a linear function of the measured output up to the current time \(k\); the other, named \(X_s(k)\), is generated by a strictly causal stochastic system, whose output is \(y_s(k)\). Hence, when facing the problem to estimate \(X_e(k)\), it is clear that only the component \(X_s(k)\) needs to be estimated. When an estimate of \(X_s(k)\) is available, the estimate of \(x(k)\), the state of the original system, is obtained by picking up the first \(n\) components.

**Definition 3.3.** The class \(\mathcal{P}\) of estimators of \(x(k)\) in (2.1) is defined as the set of all the NLP-insensitive estimators of the form:

\[
\tilde{X}_e(k) = X_d(k) + \tilde{X}_s(k), \\
\tilde{x}(k) = S \tilde{X}_e(k),
\]

where \(S = [I_n \enspace O_{n \times m}]\) and \(\tilde{X}_s(k)\) is any estimate of the process \(X_s(k)\) among all the Borel functions of the measurements \(\{y_s(\tau), \tau \leq k\}\).

It is well known that the optimal choice for \(\tilde{X}_s(k)\) is given by the conditional expectation \(\tilde{X}_s(k) = \mathbb{E}[X_s(k)|Y_s^k]\), whose computation in general can not be obtained through algorithms of finite dimensions. From an applicative point of view, it is useful to develop finite-dimensional approximations of the optimal filter, for instance the optimal linear filter:

\[
\bar{X}_s(k) = \Pi[X_s(k)|L_t(Y_s^k)],
\]
that is the projection of \( X_s(k) \) onto the Hilbert space of finite second order moments random vectors, generated by all the linear transformations of the measurements \( \{ y_s(\tau), \tau \leq k \} \). The \( \mathcal{P} \)-estimator (3.4) with \( \tilde{X}_s(k) \) computed as in (3.5), will be denoted the optimal \( \mathcal{P} \)-linear estimator of \( x(k) \).

**Remark 3.4.** Since the component \( X_d(k) \) can be exactly computed, the error covariance matrix for a \( \mathcal{P} \)-linear estimate is:

\[
\text{Cov}(x(k) - \hat{x}(k)) = S \text{Cov}(X_e(k) - \bar{X}_e(k)) S^T = S \text{Cov}(X_s(k) - \bar{X}_s(k)) S^T, \tag{3.6}
\]

and therefore it depends only on the filtering error on the component \( X_s(k) \).

**Theorem 3.5.** The optimal \( \mathcal{P} \)-linear estimate \( \hat{x}(k) \) of the state of system (2.1) is achieved by the following filter:

\[
\begin{align*}
\hat{x}(0) &= \bar{x}_0, \\
\hat{x}(0) &= \hat{x}(0) - 1 + K_G(0)[y(0) - C(0)\hat{x}(0) - 1], \\
\hat{x}(k + 1|k) &= A(k)\hat{x}(k) + B(k)u(k), \\
\hat{x}(k + 1) &= \hat{x}(k + 1|k) + D(k)g(k + 1) \\
&\quad + K_G(k + 1)[y(k + 1) - C(k + 1)D(k + 1)g(k + 1) - C(k + 1)\hat{x}(k + 1|k)],
\end{align*}
\]

where:

\[
A(k) = SH^T(k + 1) \begin{bmatrix} I_n \\ O_{q \times n} \end{bmatrix}, \quad B(k) = SB_e(k), \quad D(k) = SD_e(k), \tag{3.8}
\]

with \( S = [I_n \quad O_{n \times m}] \) and the filter gain \( K_G(k) \) recursively computed as:

\[
\begin{align*}
K_G(0) &= \Psi_0 C(0)^T \left( C(0)\Psi_0 C(0)^T + Q_g(0) \right)^{\dagger}, \\
P(0) &= \left[ I_n - K_G(0)C(0) \right] \Psi_0, \\
M_P(k + 1) &= \begin{bmatrix} A(k)P(k)A(k)^T + Q_f(k) & O_{n \times q} \\ O_{q \times n} & Q_g(k + 1) \end{bmatrix}, \\
K_G(k + 1) &= -SH^T(k)M_P(k + 1)L(k)^T[L(k)M_P(k + 1)L(k)^T]^\dagger, \\
P(k + 1) &= \begin{bmatrix} SH^T(k) + K_G(k + 1)L(k) \end{bmatrix} M_P(k + 1)H(k)^T S^T,
\end{align*}
\]

with

\[
L(k) = \begin{bmatrix} O_{q \times n} & I_q \end{bmatrix} (I_{n + q} - H(k)H^T(k)). \tag{3.10}
\]

\( P(k) \) is the sequence of covariance matrices of the error \( \hat{x}(k) - x(k) \).

**Proof.** From a mathematical point of view, the optimal \( \mathcal{P} \)-linear filtering problem for \( X_e(k) \) is similar to the one solved in [4], whose solution is given by:

\[
\begin{align*}
\bar{X}_e(0|0) &= \bar{x}_0^T, \\
\bar{X}_e(0) &= \bar{X}_e(0|0) + K(0)[y(0) - C_e(0)\bar{X}_e(0|0)], \\
\bar{X}_e(k + 1) &= \bar{X}_e(k + 1|k) + D_e(k)g(k + 1) + K_E(k + 1)[y(k + 1) \\
&\quad - C_e(k + 1)D_e(k)g(k + 1) - C_e(k + 1)\bar{X}_e(k + 1|k)], \\
\bar{X}_e(k + 1|k) &= A_e(k)\bar{X}_e(k) + B_e(k)u(k),
\end{align*}
\]

with the Kalman gain $K_E(k)$ given by the following Riccati equations:

\[
\begin{align*}
K_E(0) &= \Sigma_0 S^T C(0)^T [C(0)S\Sigma_0 S^T C(0)^T + Q_g(0)]^+, \\
P_E(0) &= \left[I_{n+m} - K_E(0)C(0)S\right]\Sigma_0, \\
M_P(k + 1) &= \left[A(k)SP_E(k)S^T A(k)^T + Q_f(k) \begin{bmatrix} O_{n \times q} \\ Q_g(k + 1) \end{bmatrix}\right], \\
K_E(k + 1) &= -H(k)^T M_P(k + 1)L(k)^T [L(k)M_P(k + 1)L(k)^T]^+, \\
P_E(k + 1) &= [H(k)^T + K_E(k + 1)L(k)]M_P(k + 1)H(k)^T,
\end{align*}
\]

(3.13)

and $L(k)$ as in (3.10). Note that $\Sigma_0$, the covariance matrix of the extended initial state $X_e(0)$, is given by:

\[
\Sigma_0 = \text{Cov}(X_e(0)) = \text{Cov}\left(\begin{bmatrix} x_0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} \Sigma_0 & O_{n \times m} \\ O_{m \times n} & \Sigma_{e0} \end{bmatrix}.
\]

(3.14)

According to definition 3.3, it comes that the state estimate $\hat{x}(k)$ is given by the first $n$ components of the estimated extended state, so that the state equations for the filtered state are:

\[
\begin{align*}
\hat{x}(k + 1) &= S\hat{X}_e(k + 1) = S\hat{X}_e(k + 1|k) + SD_e(k)y(k + 1) + SK_E(k + 1)[y(k + 1) \\
&\quad - C(k + 1)SD_e(k)y(k + 1) - C(k + 1)S\hat{X}_e(k + 1|k)] \\
&= \hat{x}(k + 1|k) + D(k)y(k + 1) + K_G(k + 1)[y(k + 1) \\
&\quad - C(k + 1)D(k)y(k + 1) - C(k + 1)\hat{x}(k + 1|k)], \\
\hat{x}(k + 1|k) &= S\hat{X}_e(k + 1|k) = SA_e(k)\hat{X}_e(k) + SB_e(k)u(k) \\
&= SH(k)^T \begin{bmatrix} I_n \\ O_{q \times n} \end{bmatrix} A(k)S\hat{X}_e(k) + SB_e(k)u(k) = A(k)\hat{x}(k) + B(k)u(k),
\end{align*}
\]

(3.15)

(3.16)

where $K_G = SK_E$ and $A(k), B(k), D(k)$ are as in (3.8). The initial state estimate

\[
\begin{align*}
\hat{x}(0) &= S\hat{X}_e(0|1) + SK_E(0)[y(0) - C(0)S\hat{X}_e(0|1)] \\
&= \hat{x}(0|1) + K_E(0)[y(0) - C(0)\hat{x}(0|1 - 1)], \\
\hat{x}(0|1) &= S\hat{X}_e(0|1) = S\hat{x}_0.
\end{align*}
\]

(3.17)

Taking into account the Riccati equations (3.13), the Kalman gain $K_G(k)$ is given by:

\[
K_G(k) = -SH(k)^T M_P(k + 1)L(k)^T [L(k)M_P(k + 1)L(k)^T]^+. 
\]

(3.18)

According to [4], the error covariance matrices of the extended state is given by the sequence $\{P_E(k)\}$. Name $\{P(k)\}$ the sequence of the error covariance matrices for $\hat{x}(k)$:

\[
P(k) = E\left(\left(\hat{x}(k) - \hat{x}(k)\right)^T \left(\hat{x}(k) - \hat{x}(k)\right)\right) = SP_E(k)S^T.
\]

(3.19)

Then:

\[
P(k + 1) = \left[S\hat{H}(k)^T + SK_E(k + 1)L(k)\right]M_P(k + 1)H(k)^T S^T
\]

\[
= \left[S\hat{H}(k)^T + K_G(k + 1)L(k)\right]M_P(k + 1)H(k)^T S^T,
\]

(3.20)
with
\[
M_P(k+1) = \begin{bmatrix}
A(k)S \bar{P}E(k)S^T A(k)^T + Q_f(k) & O_{n \times q} \\
O_{q \times n} & Q_g(k+1)
\end{bmatrix}
= \begin{bmatrix}
A(k)P(k)A(k)^T + Q_f(k) & O_{n \times q} \\
O_{q \times n} & Q_g(k+1)
\end{bmatrix}.
\]

At last, the initial conditions of (3.13) become:
\[
K_G(0) = SK_E(0) = S \Psi_0 S^T C(0)^T [C(0)S \Psi_0 S^T C(0)^T + Q_g(0)]^\dagger
= \Psi_0 C(0)^T [C(0)\Psi_0 C(0)^T + Q_g(0)]^\dagger,
\]
and
\[
P(0) = SP_E(0)S^T = S[I_n+m - K_E(0)C(0)S] \bar{\Psi}_0 S^T = [I_n - K_G(0)C(0)] \Psi_0.
\]

**Remark 3.6.** The performance of the proposed filter could be improved by extending the class of the estimator to the wider class of polynomial transformations of the output, following the approach described in [1].

**Remark 3.7.** Note that the correction gain $K_G(k)$ used in the proposed filtering algorithm can be computed off-line, differently from what happens for the Extended Kalman Filter, where the correction gain is a function of the current state estimate, and needs to be computed on-line.

### 4. Simulation results

As it has been said in the introduction, the proposed filtering algorithm is useful in all cases in which the linear part of the system dynamics is well-known, while the nonlinear part is practically unknown. In these circumstances the Extended Kalman Filter can not use the linear approximation (around the previous estimate) of the nonlinear terms. An interesting point is that the gain of the proposed filter is not forced by the measurements, and can be computed off-line.

Let system (2.1) be described by the time-invariant matrices reported below
\[
A = \begin{bmatrix} 0.3 & 0.2 & 0 \\ 0 & 0.7 & 0 \\ 0.2 & 0 & 0.3 \end{bmatrix}, \quad B = \begin{bmatrix} 2 \\ 0.4 \\ 0.8 \end{bmatrix}, \quad R = \begin{bmatrix} -2 \\ 2 \\ 0 \end{bmatrix}, \quad C = \begin{bmatrix} -1 & 1 & 1 \\ 1 & 2 & -1 \end{bmatrix},
\]
and the nonlinear term $h(x(k)) = x_2(k) \cos(x_1(k))$. The state and output noises are discrete zero-mean asymmetric distribution, whose covariance matrices are:
\[
Q_f = \begin{bmatrix} 0.024 & 0.024 & 0 \\ 0.024 & 0.064 & -0.080 \\ 0 & -0.080 & 0.160 \end{bmatrix}, \quad Q_g = \begin{bmatrix} 0.4532 & -0.5636 \\ -0.5636 & 6.2228 \end{bmatrix}.
\]

There is a deterministic drift given by the unitary step $u(k) = \delta_1(k)$. The initial state is given by $x_0 = (8 \ -6.5 \ 10)^T$, far from the initialization of the filter: without any information
concerning the statistics of $x_0$, the \textit{a priori} initial estimate $\hat{x}(0|−1)$ is chosen null with the identity as covariance matrix.

As it can be easily verified, matrix $A$ is an asymptotically stable matrix, so that the convergence of the filter is guaranteed.

![Fig. 4.1 – True and estimated state: first component](image1)

![Fig. 4.2 – True and estimated state: second component.](image2)
5. Conclusions

This work presents a new approach in filtering a stochastic linear system perturbed by uncertain nonlinearities. A clever use of the measurements allows the construction of a recursive filter that does not need any knowledge on the nonlinear perturbation terms.

References


