EXTRACTA MATHEMATICAE Vol. **18**, Núm. 1, 119 – 128 (2003)

Filtering and Fixed-Point Smoothing from an Innovation Approach in Systems with Uncertainty

R. Caballero(*), A. Hermoso(**), J. Jiménez(*), J. Linares(**)

(*) *Dpto. Estadística e I.O., Universidad de Jaén, Paraje Las Lagunillas, s/n,* 23071-*Jaén, Spain, e-mail: raguila@ujaen.es, jdomingo@ujaen.es*

(**) *Dpto. Estadística e I.O., Universidad de Granada, Campus Fuentenueva, s/n,* 18071-*Granada, Spain, e-mail: ahermoso@ugr.es, jlinares@ugr.es*

(Presented by M. Molina)

AMS *Subject Class.* (2000): 62M20, 60G35 *Received January* 21, 2002

1. Introduction

The least mean-squared error linear estimation problem of signals in systems with uncertain observations has been widely studied, since many practical situations (for example, in Communication Theory) are modelled by this kind of systems. The systems with uncertain observations are characterized by the fact that the observation equation contains, besides additive noise, a multiplicative noise component, defined by a sequence of Bernoulli random variables, which determines the presence of the signal in the observations.

Under the hypothesis that the state-space model is known, Nahi [6] and, subsequently, Monzingo [5] obtained the solution of this problem assuming independence of the variables that characterize the uncertainty and, also, that the additive noises of the signal and the observation are uncorrelated. Later on, these results were generalized by Hermoso and Linares ([1], [2]) to the case in which the additive noises are correlated.

In the papers mentioned above, the algorithms are obtained by starting from the general expression of the estimators as a linear function of the observations. In contrast, the technique based on innovations, consisting of converting the observation process into the innovation process, and expressing the estimators as a linear function of the innovations, allows to make a simpler derivation of the estimation algorithms. This technique is justified by two main ideas: on the one hand, the set of observations and the set of innovations can be determined one from the other by a causal and causally invertible operation and, on the other hand, the Wiener-Hopf equation is easier to solve since the innovation process is white, that is, their variables are independent.

The innovation approach has been used by several authors. For example, Kailath ([3], [4]) obtains the Kalman-Bucy filtering and smoothing algorithms in both continuous and discrete-time systems by using this approach. Later on, Nakamori ([7], [8], [9]) deals with the estimation problem in continuous-time systems without requiring the whole knowledge of the state-space model, but only the covariance information.

In this paper the least mean-squared error linear filtering and fixed-point smoothing problems in systems with uncertain observations are treated, assuming that the state-space model is not available. It is supposed that the variables describing the uncertainty are independent and the covariance matrix of the signal is known and presents a factorization in a semidegenerate kernel form. By applying an innovation approach, recursive algorithms for the filtering and fixed-point smoothing estimates are obtained; also, formulas for the error covariance matrices of the proposed estimators are presented.

2. Problem formulation

Let $z(k)$ and $y(k)$ be $n \times 1$ vectors which describe the signal we wish to estimate and the observation of this signal at time k, respectively. Let us suppose, on the one hand, that the observation is perturbed by an additive noise $v(k)$ and, on the other hand, that in each instant of time the observation may not contain the signal (in such a case, it will consist only of noise). So, the observation equation is described by

$$y(k) = u(k)z(k) + v(k) \tag{1}$$

where $u(k)$ is a Bernoulli random variable.

In order to analyze the estimation problem of the signal $z(k)$, we assume the following hypotheses on the signal and the noise processes involved in equation (1):

(I) The signal process $\{z(k);\ k \geq 0\}$ has zero mean and its autocovariance function is expressed in a semidegenerate kernel form, that is,

$$K_z(k,s) = E[z(k)z^T(s)] = \begin{cases} A(k)B^T(s), & 0 \leq s \leq k \\ B(k)A^T(s), & 0 \leq k \leq s \end{cases} \tag{2}$$

where A and B are $n \times M'$ matrix functions.

(II) The noise process $\{v(k);\ k \geq 0\}$ is a zero-mean white sequence with autocovariance function $E[v(k)v^T(s)] = R(k)\delta_K(k-s)$, being δ_K the Kronecker delta function and R a known $n \times n$ matrix function.

(III) The multiplicative noise $\{u(k);\ k \geq 0\}$ is a sequence of independent Bernoulli random variables with $P\left[u(k) = 1\right] = p(k)$.

(IV) The processes $\{z(k);\ k \geq 0\}$, $\{u(k);\ k \geq 0\}$ and $\{v(k);\ k \geq 0\}$ are mutually independent.

Our aim is to obtain the least mean-squared error linear estimator, $\widehat{z}(k, L)$, of the signal $z(k)$ based on the observations $\{y(1), \dots, y(L)\}$.

At each instant of time k, the innovation is defined by $\nu(k) = y(k) - \widehat{y}(k, k-1)$, being $\widehat{y}(k, k-1)$ the least mean-squared error linear estimator of the observation $y(k)$ based on $\{y(1), \dots, y(k-1)\}$ and $\widehat{y}(1, 0) = 0$. It can be observed that $\nu(k)$ represents the new information to estimate the signal which is provided by the observation $y(k)$.

Since the set of innovations $\{\nu(k);\ k \leq L\}$ and the set of observations $\{y(k);\ k \leq L\}$ can be determined one from the other by a causal and causally invertible linear transformation (Kailath [3]), we can replace one set by the other with no loss of information.

Consequently, $\widehat{z}(k, L)$ is the least mean-squared error linear estimator of the signal $z(k)$ given the innovations $\{\nu(1), \dots, \nu(L)\}$. Thus, this linear estimator can be expressed as

$$\widehat{z}(k, L) = \sum_{i=1}^{L} g(k, i, L)\nu(i)$$

being $g(k, i, L),\ i = 1, \dots, L$, the impulse response function which verifies the Wiener-Hopf equation

$$E[z(k)\nu^T(s)] = \sum_{i=1}^{L} g(k, i, L)E[\nu(i)\nu^T(s)], \quad s \leq L.$$

Taking now into account that the innovation $\{\nu(k);\ k \geq 1\}$ is a white process, and denoting $\Pi(s) = E[\nu(s)\nu^T(s)]$, the above equation can be rewritten as

$$g(k, s, L) = E[z(k)\nu^T(s)]\Pi^{-1}(s).$$

As it can be observed the impulse response function $g(k,i,L)$ does not depend on L. So, the linear estimator can be write as

$$\widehat{z}(k,L) = \sum_{i=1}^{L} g(k,i)\nu(i). \tag{3}$$

where

$$g(k,i) = E[z(k)\nu^T(i)]\Pi^{-1}(i) \tag{4}$$

In order to determine $\widehat{z}(k,L)$, first of all, in Section 3, we will obtain the innovation process and its covariance matrix. Afterwards, in Section 4, we will derive a recursive algorithm for the filter $\widehat{z}(k,k)$ and the fixed-point smoother $\widehat{z}(k,L)$, $L > k$. Finally, in Section 5, we will study the covariance matrices of the filtering and fixed-point smoothing errors establishing recursive expressions to calculate them.

3. Innovation process

As we have indicated in the above section, the innovation $\nu(k)$, defined by $\nu(k) = y(k) - \widehat{y}(k,k-1)$, represents the new information contained in the observation $y(k)$ that cannot be obtained from the previous observations. In the following theorem we establish the expression to calculate the innovation from a recursive formula.

Theorem 1. *Let us consider the observation equation described in* (1), *satisfying* (I)-(IV).

(i) *The innovation process is a zero-mean white sequence which verifies*

$$\nu(k) = y(k) - p(k)A(k)O(k-1), \quad k \geq 1 \tag{5}$$

where the vector $O(k)$ *can be calculated from the recursive relation*

$$O(k) = O(k-1) + \Delta(k)\Pi^{-1}(k)\nu(k), \qquad O(0) = 0 \tag{6}$$

being

$$\Delta(k) = p(k)\left[B^T(k) - r(k-1)A^T(k)\right] \tag{7}$$

where the matrix function r *verifies*

$$r(k) = r(k-1) + \Delta(k)\Pi^{-1}(k)\Delta^T(k), \qquad r(0) = 0. \tag{8}$$

(ii) *The covariance matrix of the innovation,* $\Pi(k)$, *is given by*

$$\Pi(k) = R(k) + p(k)A(k)\left[B^T(k) - p(k)r(k-1)A^T(k)\right]. \tag{9}$$

Remark. Let us note that the non-singularity of $\Pi(k)$ is not guaranteed in general. If $\Pi(k)$ were singular, the Moore-Penrose pseudo-inverse could be used.

Proof. (i) *Innovation process.*

By using the hypotheses (I)-(IV) and the Orthogonal Projection Lemma (O.P.L.) (Kailath, [3]), it is immediately obtained that

$$\widehat{y}(k,k-1) = p(k)\widehat{z}(k,k-1), \quad k \geq 1 \tag{10}$$

being $\widehat{z}(k,k-1)$ the linear one-stage predictor of the signal. Consequently,

$$\nu(k) = y(k) - p(k)\widehat{z}(k,k-1), \quad k \geq 1. \tag{11}$$

From (3) and (4), $\widehat{z}(k,k-1)$ is given by

$$\widehat{z}(k,k-1) = \sum_{i=1}^{k-1} S(k,i)\Pi^{-1}(i)\nu(i), \quad \widehat{z}(1,0) = 0 \tag{12}$$

where

$$S(k,i) = E[z(k)\nu^T(i)]. \tag{13}$$

So, we only need to calculate $S(k,i)$ for $i \leq k-1$.

By using expression (11) for the innovation $\nu(i)$, $S(k,i)$ can be expressed as

$$S(k,i) = E[z(k)y^T(i)] - p(i)E[z(k)\widehat{z}^T(i,i-1)].$$

From the hypotheses on the model, the first expectation on the right-hand side term is

$$E[z(k)y^T(i)] = p(i)A(k)B^T(i), \quad i \leq k-1$$

and, taking into account expressions (12) and (13), the second expectation is given by

$$E[z(k)\widehat{z}^T(i,i-1)] = \sum_{j=1}^{i-1} S(k,j)\Pi^{-1}(j)S^T(i,j), \quad i \leq k-1.$$

Consequently, we have that

$$S(k,i) = p(i)\left[A(k)B^T(i) - \sum_{j=1}^{i-1} S(k,j)\Pi^{-1}(j)S^T(i,j)\right], \quad i \le k-1.$$

Then, by introducing a function, Δ, satisfying

$$\Delta(i) = p(i)\left[B^T(i) - \sum_{j=1}^{i-1} \Delta(j)\Pi^{-1}(j)S^T(i,j)\right] \tag{14}$$

the expression of $S(k,i)$ can be transformed into

$$S(k,i) = A(k)\Delta(i), \quad i \le k-1. \tag{15}$$

By defining

$$O(k) = \sum_{i=1}^{k} \Delta(i)\Pi^{-1}(i)\nu(i) \tag{16}$$

we conclude, from (12) and (15), that the one-stage predictor of the signal can be expressed as

$$\widehat{z}(k,k-1) = A(k)O(k-1) \tag{17}$$

and expression (5) for the innovation is immediately obtained from (11).

The recursive relation (6) for the vectors $O(k)$ is easily deduced from (16).

Next, we will derive expression (7) for $\Delta(k)$. By substituting i for k in (14) and taking into account (15), it is clear that

$$\Delta(k) = p(k)\left[B^T(k) - \sum_{i=1}^{k-1} \Delta(i)\Pi^{-1}(i)\Delta^T(i)A^T(k)\right]$$

and, if we define $r(k)$ as

$$r(k) = E\left[O(k)O^T(k)\right] = \sum_{i=1}^{k} \Delta(i)\Pi^{-1}(i)\Delta^T(i) \tag{18}$$

expression (7) is immediately obtained.

Finally, from (18), the relation (8) for the matrices $r(k)$ is clear.

(ii) *Innovation covariance matrix.*

From the O.P.L., it is easily obtained that the innovation covariance matrix is given by

$$\Pi(k) = E\left[y(k)y^T(k)\right] - E\left[\widehat{y}(k,k-1)\widehat{y}^T(k,k-1)\right].$$

From the hypotheses on the model, the first expectation on the right-hand side term is

$$E\left[y(k)y^T(k)\right] = R(k) + p(k)A(k)B^T(k)$$

and, by using expression (10), the second expectation is

$$E\left[\widehat{y}(k,k-1)\widehat{y}^T(k,k-1)\right] = p^2(k)E\left[\widehat{z}(k,k-1)\widehat{z}^T(k,k-1)\right].$$

Hence, we conclude that

$$\Pi(k) = R(k) + p(k)A(k)B^T(k) - p^2(k)E\left[\widehat{z}(k,k-1)\widehat{z}^T(k,k-1)\right]. \tag{19}$$

Finally, from (19), using (17) and (18), expression (9) is easily deduced. ∎

4. Recursive algorithm for the filter and fixed-point smoother

The solution to the fixed-point smoothing problem together with the recursive formulas to obtain the filtering estimate of the signal (which provides the initial condition of the recursive equation for the fixed-point smoothing estimate) are established in the following theorem.

Theorem 2. *If we consider the observation equation* (1), *under the hypotheses* (I)-(IV), *the fixed-point smoothing estimates of the signal* $z(k)$ *are given by*

$$\widehat{z}(k,L) = \widehat{z}(k,L-1) + g(k,L)\nu(L), \qquad L > k \tag{20}$$

where the smoothing gain, $g(k,L)$, *is expressed as*

$$g(k,L) = p(L)\left[B(k) - E(k,L-1)\right]A^T(L)\Pi^{-1}(L) \tag{21}$$

and the matrices $E(k,L)$ *are recursively calculated from*

$$\begin{aligned} E(k,L) &= E(k,L-1) + g(k,L)\Delta^T(L), \qquad L > k \\ E(k,k) &= A(k)r(k). \end{aligned} \tag{22}$$

The algorithm is completed with the relations for $\nu(L)$, $\Pi(L)$ *and* $\Delta(L)$ *given in Theorem* 1.

The filter, $\widehat{z}(k,k)$, *which provides the initial condition for* (20), *is given by* $\widehat{z}(k,k) = A(k)O(k)$, *where* $O(k)$ *has been obtained in Theorem* 1.

Proof. Expression (20) for the fixed-point smoothing estimates is immediately obtained from (3).

Since, from (4) and (13), the smoothing gain, $g(k,L)$, is expressed as

$$g(k,L) = S(k,L)\Pi^{-1}(L) \tag{23}$$

we must calculate a recursive relation for $S(k,L)$. So, by using (10) and (12) for $k = L$, and taking into account that $E[z(k)y^T(L)] = p(L)B(k)A^T(L)$, for $L > k$, we have that

$$S(k,L) = p(L)\left[B(k)A^T(L) - \sum_{i=1}^{L-1} S(k,i)\Pi^{-1}(i)S^T(L,i)\right].$$

Moreover, from (15), for $i \leq L-1$, we have that $S(L,i) = A(L)\Delta(i)$ and, consequently,

$$S(k,L) = p(L)\left[B(k) - \sum_{i=1}^{L-1} S(k,i)\Pi^{-1}(i)\Delta^T(i)\right]A^T(L).$$

Then, if we denote

$$E(k,L) = \sum_{i=1}^{L} S(k,i)\Pi^{-1}(i)\Delta^T(i) \tag{24}$$

it is clear that

$$S(k,L) = p(L)\left[B(k) - E(k,L-1)\right]A^T(L)$$

and substituting this expression in (23), expression (21) for the smoothing gain is obtained.

To complete the fixed-point smoothing algorithm, it is necessary to find a recursive relation for the matrices $E(k,L)$. This relation, given in (22), is immediately derived from (23) and (24). Again from (24) and taking into account that, for $i \leq k$, $S(k,i) = A(k)\Delta(i)$, its initial condition is obtained, by using (18).

Let us now deduce the relation for the filter. By using (3), the filtering estimate is given by

$$\widehat{z}(k,k) = \widehat{z}(k,k-1) + g(k,k)\nu(k).$$

From (4) for $i = k$, we obtain $g(k,k) = E\left[z(k)\nu^T(k)\right]\Pi^{-1}(k)$. Moreover, by using (11) and the O.P.L., it is clear that

$$E\left[z(k)\nu^T(k)\right] = E\left[z(k)y^T(k)\right] - p(k)E\left[\widehat{z}(k,k-1)\widehat{z}^T(k,k-1)\right].$$

Finally, taking into account the hypotheses on the model, $E\left[z(k)y^T(k)\right] = p(k)A(k)B^T(k)$ and, from (17) and (18), $E\left[\widehat{z}(k,k-1)\widehat{z}^T(k,k-1)\right] = A(k)r(k-1)A^T(k)$. Thus, by using (7), the filter gain is given by $g(k,k) = A(k)\Delta(k)\Pi^{-1}(k)$ and, as a consequence of (6) and (17), equation for the filter is deduced. ∎

5. Error covariance matrices

The performance of the filtering and fixed-point smoothing estimates can be measured by the filtering and smoothing error $\widetilde{z}(k,L) = z(k) - \widehat{z}(k,L)$ and, more specifically, by the covariance matrices of these errors,

$$P(k,L) = E\left[\widetilde{z}(k,L)\widetilde{z}(k,L)^T\right].$$

In the following theorem we present a recursive formula to obtain $P(k,L)$, a measure of the estimation accuracy for the filter and fixed-point smoother proposed in Theorem 2.

THEOREM 3. *The smoothing and filtering error covariance matrices are respectively obtained from*

$$\begin{aligned} &P(k,L) = P(k,L-1) - g(k,L)\Pi(L)g^T(k,L), \quad L > k \\ &P(k,k) = A(k)\left[B^T(k) - r(k)A^T(k)\right] \end{aligned} \tag{25}$$

where $r(k)$ *and* $\Pi(L)$ *have been recursively calculated in Theorem* 1 *and* $g(k,L)$, *in Theorem* 2.

Proof. From the O.P.L. and expression (2), the matrices $P(k,L)$ can be written as

$$P(k,L) = A(k)B^T(k) - F(k,L) \tag{26}$$

where $F(k,L) = E\left[\widehat{z}(k,L)\widehat{z}^T(k,L)\right]$ denotes the estimator covariance matrix. Then, we must calculate the matrices $F(k,L)$ associated with the smoothing estimate and the filter.

On the one hand, by using (20) and taking into account that $\nu(L)$ and $\widehat{z}(k, L-1)$ are uncorrelated, we obtain

$$F(k,L) = F(k,L-1) + g(k,L)\Pi(L)g^T(k,L), \quad L > k \tag{27}$$

and, from (27) and (26), expression (25) for the smoothing error covariance matrices is deduced.

On the other hand, since the filter is given by $\widehat{z}(k,k) = A(k)O(k)$, we have that

$$F(k,k) = A(k)r(k)A^T(k)$$

and, substituting this expression in (26) for $L = k$, the filtering error covariance matrix established in (25) is easily obtained. ∎

Acknowledgements

This work has been partially supported by the "Ministerio de Ciencia y Tecnología" under contract BFM2000-0602.

References

[1] Hermoso, A., Linares, J., Linear estimation for discrete-time systems in the presence of time-correlated disturbances and uncertain observations, *IEEE Transactions on Automatic Control,* **AC-39** (8) (1994), 1636 – 1638.

[2] Hermoso, A., Linares, J.,, Linear smoothing for discrete-time systems in the presence of correlated disturbances and uncertain observations, *IEEE Transactions on Automatic Control,* **AC-40** (8) (1995), 1486 – 1488.

[3] Kailath, T., An innovations approach to least-squares estimation. Part I: linear filtering in additive white noise, *IEEE Transactions on Automatic Control,* **AC-13** (6) (1968), 646 – 655.

[4] Kailath, T., An innovations approach to least-squares estimation. Part II: linear smoothing in additive white noise, *IEEE Transactions on Automatic Control,* **AC-13** (6) (1968), 655 – 660.

[5] Monzingo, R.A., Discrete linear recursive smoothing for systems with uncertain observations, *IEEE Transactions on Automatic Control,* **AC-26** (3) (1981), 754 – 757.

[6] Nahi, N.E., Optimal Recursive Estimation with Uncertain Observation, *IEEE Transactions on Information Theory,* **IT-15** (4) (1969), 457 – 462.

[7] Nakamori, S., Filtering algorithm based on innovations theory for white gaussian plus colored observation noise, *Journal of Information Science and Engineering,* **5** (1989), 157 – 166.

[8] Nakamori, S., New design of estimators using covariance information based on an innovations approach, *Signal Processing,* **24** (1991), 157 – 282.

[9] Nakamori, S., Design of a fixed-point smoother based on an innovations theory for white gaussian plus coloured observation noise, *International Journal of Systems Science,* **22** (12) (1991), 2573 – 2584.