International Mathematical Forum, Vol. 7, 2012, no. 18, 887 - 903

A Solution to the Filtering Problem for Stochastic Systems with Multi-Sensor Uncertain Observations

I. García-Garrido

Dpto. Estadística e I. O., Universidad de Granada
Campus Fuentenueva, s/n, 18071 Granada, Spain
irenegarciag@ugr.es

J. Linares-Pérez

Dpto. Estadística e I. O., Universidad de Granada
Campus Fuentenueva, s/n, 18071 Granada, Spain
jlinares@ugr.es

R. Caballero-Águila

Dpto. Estadística e I. O., Universidad de Jaén
Campus Las Lagunillas, s/n, 23071 Jaén, Spain
raguila@ujaen.es

A. Hermoso-Carazo

Dpto. Estadística e I. O., Universidad de Granada
Campus Fuentenueva, s/n, 18071 Granada, Spain
ahermoso@ugr.es

Abstract

In this paper, the least-squares linear and quadratic filtering problems are studied in discrete-time linear stochastic systems with uncertain observations coming from multiple sensors, when the variables describing the uncertainty in the observations are correlated at instants that differ two units of time. The least-squares linear filter is obtained by using an approach based on innovations. The least-squares quadratic estimation problem is solved by defining an appropriate augmented system, whose state linear filtering estimate provides the quadratic filtering estimate of the original state vector. A numerical simulation example shows the effectiveness of the proposed estimation algorithms.

Mathematics Subject Classification: 62M20, 60G35

Keywords: uncertain observations, multiple sensors, least-squares estimation

1 Introduction

The state estimation problem in discrete-time linear systems using observations which are linear functions of the state and are perturbed by additive Gaussian white noise has been widely studied. Although different techniques have been used to address this problem, one of the major contributions has been the well-known Kalman filter, a recursive solution to the least-squares (LS) optimal filtering problem in Gaussian linear systems. However, in many practical situations, the widely used assumption of Gaussian noises cannot be accepted as a realistic statistical description of the processes involved (see for instance [11]). In these non-Gaussian systems, the Kalman filter only provides the LS linear estimator and, generally, determining the optimal estimator in these cases involves severe computational difficulties, thus being necessary to look for suboptimal estimators which are easier to obtain, such as linear estimators or, even, polynomial estimators which improve the commonly used linear ones.

Signal estimation problems in the presence of non-Gaussian noises have been widely studied. De Santis et al. [3] were the first who proposed a quadratic estimator, that is more accurate than the commonly used linear one and retains the features of easy computability and recursivity. A generalization of this study is proposed in [2], where a recursive algorithm for arbitrary degree polynomial estimators is proposed.

On the other hand, there exists a large number of situations where some observations may not contain information about the system state, thus being only noise (*uncertain observations*). In these cases, the observation equation includes not only additive noises, but also a multiplicative noise, which is modelled by a sequence of Bernoulli random variables, whose values –one or zero– indicate the presence or absence of the state in the observation, respectively. Even if the state and the additive noises are gaussian processes, this multiplicative noise component makes the joint distribution of the system state and observations be non-gaussian and, consequently, the optimal estimator of the signal is not easily obtainable.

The LS linear and polynomial estimation problems in systems with uncertain observations have been studied by several authors under different hypotheses on the variables modelling the uncertainty of the observations. Among others, in [10] the linear estimation problem is addressed when the uncertainty is modelled by independent variables. A more general model is considered in [8], where polynomial estimation algorithms are proposed using observations affected by non-independent uncertainty; this model covers those systems with different transmission channels, where the signal is randomly transmitted by one of them and there exists a different probability of uncertainty at each channel. More recently, other models have been studied to cover situations where the maximum number of consecutive observations without information on the signal is bounded; for example, in [9] the polynomial filtering and smoothing problems are addressed assuming that the variables describing the uncertainty are correlated at consecutive sampling times, and this model is applied to situations where the signal cannot be missing in two consecutive observations.

In all the mentioned papers it is assumed that the observations available for the estimation come either from a single sensor or from multiple sensors with identical uncertainty characteristics. In the last years, motivated by the increasing development of sensor networks for data acquisition and signal processing, several authors have generalized this study to cases in which there are multiple sensors featuring different uncertainty characteristics (see e.g. [4] for independent variables modelling the uncertainty in the observations, and [5],[1] for situations where the uncertainty in each sensor is modelled by variables correlated at consecutive sampling times).

In this paper, the LS linear and quadratic filtering problems with uncertain observations coming from multiple sensors is addressed, when the Bernoulli variables modelling the uncertainty in the observations are correlated in instants that differ two units of time. Hence, the current study provides an extension of the results established in [1] where Bernoulli variables correlated at consecutive instants are considered. The aim of this extension is to cover more general practical situations, for example, signal transmission problems where no more than two consecutive observations without information of the signal can be transmitted by the same sensor (this occurs e.g. in sensor networks where sensor failures may occur and a failed sensor is replaced not immediately, but two sampling times after having failed). Finally, the behaviour of the proposed estimators is illustrated by a numerical simulation example where a signal generated by a first-order autoregressive model is estimated from uncertain observations coming from two sensors. The linear and quadratic estimators are compared in terms of their error covariance matrices and the performance of both estimators is analyzed for different values of the uncertainty probabilities.

2 Model description

Consider a discrete-time linear stochastic system with uncertain observations coming from multiple sensors, whose mathematical modelling is accomplished by the following equations.

The state equation is given by

$$x_k = F_{k-1}x_{k-1} + w_{k-1}, \quad k \geq 1, \tag{1}$$

where $\{x_k;\ k \geq 0\}$ is an n-dimensional stochastic process representing the system state, $\{w_k;\ k \geq 0\}$ is a white noise process and F_k, for $k \geq 1$, are known deterministic matrices.

We consider scalar uncertain observations $\{y_k^i;\ k \geq 1\}$, $i = 1, \dots, m$, coming from m sensors and perturbed by noises whose statistical properties are not necessarily the same for all the sensors. Specifically, we assume that, at each time k, the observation y_k^i from the i-th sensor is perturbed by an additive noise vector v_k^i and by a multiplicative Bernoulli variable θ_k^i; that is

$$y_k^i = \theta_k^i H_k^i x_k + v_k^i, \quad k \geq 1, \quad i = 1, \dots, m, \tag{2}$$

where H_k^i, for $k \geq 1$ and $i = 1, \dots, m$, are known deterministic matrices of appropriate dimensions. When $\theta_k^i = 1$, which occurs with known probability $\overline{\theta}_k^i$, the state x_k is present in the observation y_k^i coming from the i-th sensor at time k, whereas if $\theta_k^i = 0$ such observation only contains additive noise v_k^i with probability $1 - \overline{\theta}_k^i$.

Denoting $y_k = (y_k^1, \dots, y_k^m)^T$, $v_k = (v_k^1, \dots, v_k^m)^T$, $H_k = (H_k^{1T}, \dots, H_k^{mT})^T$ and $\Theta_k = Diag(\theta_k^1, \dots, \theta_k^m)$, equation (2) is equivalent to the following stacked observation equation

$$y_k = \Theta_k H_k x_k + v_k, \quad k \geq 1. \tag{3}$$

The LS linear and quadratic estimation problems of the state x_k given by (1) from the observations $y_1, \dots, y_k$ given by (3) have been addressed in [1] assuming correlation at consecutive time instants between the Bernoulli variables that model the presence or absence of the state in the observations coming from each sensor. Our aim in this paper is to generalize the results obtained in [1] to the case where these variables are correlated in instants that differ two units of time. Specifically, the following hypothesis is assumed:

Hypothesis 1. For $i = 1, \dots, m$, $\{\theta_k^i;\ k \geq 1\}$ is a sequence of Bernoulli random variables with $P[\theta_k^i = 1] = \overline{\theta}_k^i$. The variables θ_k^i and θ_s^j are independent for $|k - s| \neq 0, 2$, $i, j = 1, \dots, m$, and $Cov[\theta_k^i, \theta_s^j]$ are known for $|k - s| = 0, 2$. Defining $\theta_k = (\theta_k^1, \dots, \theta_k^m)^T$, the covariance matrices of θ_k and θ_s will be denoted by $K_{k,s}^\theta$.

On the other hand, the LS linear estimator of x_k based on $y_1, \dots, y_k$, i.e. the linear filter of x_k, is the orthogonal projection of x_k onto the space of n-dimensional random variables obtained as linear transformations of the observations $y_1, \dots, y_k$, which requires the existence of the second-order moments of such observations. The LS quadratic estimator of x_k based on $y_1, \dots, y_k$ is obtained as the orthogonal projection of x_k onto the space of n-dimensional random variables obtained by linear transformations of $y_1, \dots, y_k$ and their second-order powers, $y_1^{[2]}, \dots, y_k^{[2]}$, defined by the Kronecker product, $y_i^{[2]} = y_i \otimes y_i$. Hence, the existence of the second-order moments of the vectors $y_1^{[2]}, \dots, y_k^{[2]}$ is required to address the LS quadratic estimation problem.

To assure the existence of such moments, the following hypotheses are assumed:

Hypothesis 2. The initial state x_0 is a random vector with zero mean and $Cov[x_0] = P_0$, $Cov[x_0, x_0^{[2]}] = P_0^{(3)}$, $Cov[x_0^{[2]}] = P_0^{(4)}$.

Hypothesis 3. The process $\{w_k;\ k \geq 0\}$ is a zero-mean white noise with $Cov[w_k] = Q_k$, $Cov[w_k, w_k^{[2]}] = Q_k^{(3)}$ and $Cov[w_k^{[2]}] = Q_k^{(4)}$.

Hypothesis 4. The noise $\{v_k;\ k \geq 1\}$ is a zero-mean white process with $Cov[v_k] = R_k$, $Cov[v_k, v_k^{[2]}] = R_k^{(3)}$ and $Cov[v_k^{[2]}] = R_k^{(4)}$.

Finally, we assume the following hypothesis on the independence of the initial state and noises:

Hypothesis 5. The initial state x_0 and the noise processes $\{w_k;\ k \geq 0\}$, $\{v_k;\ k \geq 1\}$ and $\{\theta_k;\ k \geq 1\}$ are mutually independent.

3 LS linear estimation problem

Our aim in this section is to obtain the LS linear filter of x_k from a recursive algorithm. For this purpose, an approach based on innovations is used. This approach consists of transforming the observation process $\{y_k;\ k \geq 1\}$ into an equivalent one (innovation process) of orthogonal vectors $\{\nu_k;\ k \geq 1\}$, equivalent in the sense that each set $\{\nu_1, \dots, \nu_k\}$ generates the same linear subspace as $\{y_1, \dots, y_k\}$ (see [6]).

The innovation at time k is defined as $\nu_k = y_k - \widehat{y}_{k/k-1}$, where $\widehat{y}_{k/k-1}$ is the one-stage LS linear predictor of y_k. Using the orthogonal projection lemma, the predictor can be expressed as a linear combination of innovations

$$\widehat{y}_{k/k-1} = \sum_{i=1}^{k-1} T_{k,i} \Pi_i^{-1} \nu_i, \quad k \geq 2; \quad \widehat{y}_{1/0} = 0, \tag{4}$$

where $T_{k,i} = E[y_k \nu_i^T]$ and $\Pi_i = E[\nu_i \nu_i^T]$ is the covariance of ν_i.

Similarly, since the innovation is a white process, the state linear filter,

$\widehat{x}_{k/k}$, and the state one-stage linear predictor, $\widehat{x}_{k/k-1}$, satisfy

$$\widehat{x}_{k/L} = \sum_{i=1}^{L} S_{k,i}\Pi_i^{-1}\nu_i, \quad k \geq 1, \quad L = k,\ k-1, \tag{5}$$

where $S_{k,i} = E[x_k \nu_i^T]$.

So, the following expression for the filter in terms of the predictor is immediate:

$$\widehat{x}_{k/k} = \widehat{x}_{k/k-1} + S_{k,k}\Pi_k^{-1}\nu_k, \quad k \geq 1; \quad \widehat{x}_{0/0} = 0. \tag{6}$$

To derive a recursive linear filtering algorithm, it is necessary to give an expression for each element of equation (6). Next, an equation for the state predictor $\widehat{x}_{k/k-1}$ in terms of the filter $\widehat{x}_{k/k}$ and expressions for the innovation ν_k, its covariance matrix Π_k and the matrix $S_{k,k}$ are obtained.

State predictor $\boldsymbol{\widehat{x}_{k/k-1}}$. From hypotheses 3 and 5, it is immediate that the filter of the noise w_{k-1} is $\widehat{w}_{k-1/k-1} = E[w_{k-1}] = 0$ and hence, taking into account equation (1), we have

$$\widehat{x}_{k/k-1} = F_{k-1}\widehat{x}_{k-1/k-1}, \quad k \geq 1. \tag{7}$$

Innovation $\boldsymbol{\nu_k}$. Since the innovation process is given by $\nu_k = y_k - \widehat{y}_{k/k-1}$, it is enough to get an expression for the one-stage predictor of the observation, $\widehat{y}_{k/k-1}$.

From the model hypotheses, it follows that $T_{k,i} = \Theta_k^p H_k S_{k,i}$, for $k \geq 3$, $i < k-2$, where $\Theta_k^p = E[\Theta_k]$. From (4) and after some manipulations, we obtain

$$\begin{aligned}\widehat{y}_{k/k-1} = \sum_{i=1}^{k-1} \Theta_k^p H_k S_{k,i}\Pi_i^{-1}\nu_i + \left(T_{k,k-2} - \Theta_k^p H_k S_{k,k-2}\right)\Pi_{k-2}^{-1}\nu_{k-2} \\ + \left(T_{k,k-1} - \Theta_k^p H_k S_{k,k-1}\right)\Pi_{k-1}^{-1}\nu_{k-1}.\end{aligned} \tag{8}$$

Denoting $\Psi_{k,k-2} = T_{k,k-2} - \Theta_k^p H_k S_{k,k-2}$ and taking into account equation (3), we obtain:

$$\begin{aligned}\Psi_{k,k-2} &= E[(\Theta_k - \Theta_k^p)\, H_k x_k \nu_{k-2}^T] \\ &= E[(\Theta_k - \Theta_k^p)\, H_k x_k y_{k-2}^T] - E[(\Theta_k - \Theta_k^p)\, H_k x_k \widehat{y}_{k-2/k-3}^T].\end{aligned}$$

From hypotheses 1 and 5, Θ_k is independent of the innovation ν_i, for $i \leq k-3$; hence, $E[(\Theta_k - \Theta_k^p)\, H_k x_k \widehat{y}_{k-2/k-3}^T] = 0$, and using again (3) for y_{k-2}, we obtain

$$\Psi_{k,k-2} = E[(\Theta_k - \Theta_k^p)\, H_k x_k x_{k-2}^T H_{k-2}^T \Theta_{k-2}]. \tag{9}$$

To obtain the above expectation, we use the following property [1]:

Property 1. The random matrices Θ_k satisfy the equality $E[\Theta_k G_{m\times m}\Theta_s] = E[\theta_k\theta_s^T]\circ E[G_{m\times m}]$ for any random matrix $G_{m\times m}$ independent of $\{\Theta_k;\ k\geq 1\}$, where $\circ$ denotes the Hadamard product ($[A\circ B]_{ij}=A_{ij}B_{ij}$).

Then, from (9), we have

$$\Psi_{k,k-2} = K^{\theta}_{k,k-2}\circ\left(H_k E[x_k x_{k-2}^T]H_{k-2}^T\right).$$

Now, if we denote $E_k = E[x_k x_k^T]$, from equation (1) it is clear that

$$\Psi_{k,k-2} = K^{\theta}_{k,k-2}\circ\left(H_k F_{k-1}F_{k-2}E_{k-2}H_{k-2}^T\right) \tag{10}$$

and E_k can be obtained recursively by

$$E_k = F_{k-1}E_{k-1}F_{k-1}^T + Q_{k-1},\quad k\geq 1;\quad E_0 = P_0. \tag{11}$$

Analogously,

$$\begin{aligned}T_{k,k-1}-\Theta^p_k H_k S_{k,k-1} &= E[(\Theta_k-\Theta^p_k)\,H_k x_k\nu_{k-1}^T]\\ &= K^{\theta}_{k,k-1}\circ\left(H_k E[x_k x_{k-1}^T]H_{k-1}^T\right) - E[(\Theta_k-\Theta^p_k)\,H_k x_k\widehat{y}^T_{k-1/k-2}].\end{aligned}$$

Since $K^{\theta}_{k,k-1}=0$ and $\widehat{y}_{k-1/k-2}=\sum_{i=1}^{k-3}T_{k-1,i}\Pi_i^{-1}\nu_i + T_{k-1,k-2}\Pi^{-1}_{k-2}\nu_{k-2}$, we have:

$$\begin{aligned}T_{k,k-1}-\Theta^p_k H_k S_{k,k-1} &= -E[(\Theta_k-\Theta^p_k)\,H_k x_k\nu_{k-2}^T]\Pi^{-1}_{k-2}T^T_{k-1,k-2}\\ &= -\Psi_{k,k-2}\Pi^{-1}_{k-2}T^T_{k-1,k-2}.\end{aligned} \tag{12}$$

Next, substituting (10) and (12) in (8) and using (5) for $L=k-1$, it is concluded that

$$\begin{aligned}&\widehat{y}_{k/k-1} = \Theta^p_k H_k\widehat{x}_{k/k-1} - \Psi_{k,k-2}\Pi^{-1}_{k-2}\left(\nu_{k-2}-T^T_{k-1,k-2}\Pi^{-1}_{k-1}\nu_{k-1}\right),\quad k\geq 3,\\ &\widehat{y}_{2/1} = T_{2,1}\Pi_1^{-1}y_1,\\ &\widehat{y}_{1/0}=0\end{aligned} \tag{13}$$

where $\Psi_{k,k-2}$ is given in (10) with E_k recursively obtained from (11).

Finally, using (12) and taking into account that, from (1), $S_{k,k-1}=F_{k-1}S_{k-1,k-1}$, the matrices $T_{k,k-1}$ in (13) are recursively obtained from

$$\begin{aligned}&T_{k,k-1} = \Theta^p_k H_k F_{k-1}S_{k-1,k-1} - \Psi_{k,k-2}\Pi^{-1}_{k-2}T^T_{k-1,k-2},\quad k\geq 3,\\ &T_{2,1} = \Theta^p_2 H_2 F_1 E_1 H_1^T\Theta^p_1.\end{aligned} \tag{14}$$

Matrix $\boldsymbol{S_{k,k}}$. Since $\nu_k = y_k - \widehat{y}_{k/k-1}$, we have $S_{k,k}=E[x_k\nu_k^T]=E[x_k y_k^T] - E[x_k\widehat{y}^T_{k/k-1}]$. Next, we calculate these expectations.

On the one hand, from (3) we have that $E[x_k y_k^T] = E_k H_k^T \Theta_k^p, \quad \forall k \geq 1$.

On the other, from (13) we obtain

$$E[x_k \widehat{y}_{k/k-1}^T] = E[x_k \widehat{x}_{k/k-1}^T] H_k^T \Theta_k^p + \left(S_{k,k-2} - S_{k,k-1}\Pi_{k-1}^{-1}T_{k-1,k-2}\right)\Pi_{k-2}^{-1}\Psi_{k,k-2}^T, \quad k \geq 3,$$
$$E[x_2 \widehat{y}_{2/1}^T] = F_1 E_1 H_1^T \Theta_1^p \Pi_1^{-1} T_{2,1}^T.$$

Now, the orthogonal projection lemma assures that $E[x_k \widehat{x}_{k/k-1}^T] = E_k - P_{k/k-1}$, where $P_{k/k-1} = E[(x_k - \widehat{x}_{k/k-1})(x_k - \widehat{x}_{k/k-1})^T]$ is the prediction error covariance matrix.

Then, using again that $S_{k,k-1} = F_{k-1}S_{k-1,k-1}$ and $S_{k,k-2} = F_{k-1}F_{k-2}S_{k-2,k-2}$, the following expression for $S_{k,k}$ is derived

$$\begin{aligned} S_{k,k} &= P_{k/k-1}H_k^T\Theta_k^p \\ &\quad - F_{k-1}\left(F_{k-2}S_{k-2,k-2} - S_{k-1,k-1}\Pi_{k-1}^{-1}T_{k-1,k-2}\right)\Pi_{k-2}^{-1}\Psi_{k,k-2}^T, \quad k \geq 3, \\ S_{2,2} &= E_2 H_2^T \Theta_2^p - F_1 E_1 H_1^T \Theta_1^p \Pi_1^{-1} T_{2,1}^T, \\ S_{1,1} &= E_1 H_1^T \Theta_1^p. \end{aligned} \tag{15}$$

Finally, from (1), the prediction error covariance matrix $P_{k/k-1}$ is recursively calculated from

$$P_{k/k-1} = F_{k-1}P_{k-1/k-1}F_{k-1}^T + Q_{k-1}, \quad k \geq 1 \tag{16}$$

with $P_{k/k} = E[(x_k - \widehat{x}_{k/k})(x_k - \widehat{x}_{k/k})^T]$, the filtering error covariance matrix, verifying

$$P_{k/k} = P_{k/k-1} - S_{k,k}\Pi_k^{-1}S_{k,k}^T, \quad k \geq 1; \quad P_{0/0} = P_0. \tag{17}$$

Covariance matrix of the innovation $\boldsymbol{\Pi_k = E[\nu_k \nu_k^T]}$. From the orthogonal projection lemma, the covariance matrix of the innovation is obtained as $\Pi_k = E[y_k y_k^T] - E[\widehat{y}_{k/k-1}\widehat{y}_{k/k-1}^T]$. From (3) and using Property 1, we have that

$$E[y_k y_k^T] = E[\theta_k \theta_k^T] \circ \left(H_k E_k H_k^T\right) + R_k, \quad k \geq 1.$$

On the other hand, taking into account (13), a similar reasoning to that used to derive the matrix $S_{k,k}$ leads to the following expression, for $k \geq 3$,

$$\begin{aligned} E[\widehat{y}_{k/k-1}\widehat{y}_{k/k-1}^T] &= \overline{\theta}_k\overline{\theta}_k^T \circ \left(H_k\left(E_k - P_{k/k-1}\right)H_k^T\right) + \Psi_{k,k-2}\Pi_{k-2}^{-1}\Psi_{k,k-2}^T \\ &\quad + \Psi_{k,k-2}\Pi_{k-2}^{-1}T_{k-1,k-2}\Pi_{k-2}^{-1}\Psi_{k,k-2}^T \\ &\quad + \Theta_k^p H_k F_{k-1}\left(F_{k-2}S_{k-2,k-2} + S_{k-1,k-1}\Pi_{k-1}^{-1}T_{k-1,k-2}\right)\Pi_{k-2}^{-1}\Psi_{k,k-2}^T \\ &\quad + \left[\Theta_k^p H_k F_{k-1}\left(F_{k-2}S_{k-2,k-2} + S_{k-1,k-1}\Pi_{k-1}^{-1}T_{k-1,k-2}\right)\Pi_{k-2}^{-1}\Psi_{k,k-2}^T\right]^T. \end{aligned}$$

Next, from (15) we have that, in the above expression,

$$F_{k-1}\left(F_{k-2}S_{k-2,k-2}+S_{k-1,k-1}\Pi_{k-1}^{-1}T_{k-1,k-2}\right)\Pi_{k-2}^{-1}\Psi_{k,k-2}^{T} \\ =-S_{k,k}+P_{k/k-1}H_k^T\Theta_k^p, \qquad k\geq 3.$$

Therefore, the innovation covariance matrix verifies

$$\begin{aligned} \Pi_k &= K_{k,k}^{\theta}\circ(H_kE_kH_k^T)+R_k-\Theta_k^pH_kP_{k/k-1}H_k^T\Theta_k^p+\Theta_k^pH_kS_{k,k} \\ &\quad +S_{k,k}^TH_k^T\Theta_k^p-\Psi_{k,k-2}\Pi_{k-2}^{-1}\left(I+T_{k-1,k-2}^T\Pi_{k-1}^{-1}T_{k-1,k-2}\Pi_{k-2}^{-1}\right)\Psi_{k,k-2}^T,\ k\geq 3, \\ \Pi_2 &= E[\theta_2\theta_2^T]\circ(H_2E_2H_2^T)+R_2-T_{2,1}\Pi_1^{-1}T_{2,1}^T, \\ \Pi_1 &= E[\theta_1\theta_1^T]\circ(H_1E_1H_1^T)+R_1. \end{aligned} \tag{18}$$

where I denotes the identity matrix of appropriate dimensions.

Finally, the proposed linear filtering algorithm is constituted by equations (6), (7) and (13)-(18).

4 LS quadratic estimation problem

This section is concerned with the problem of deriving a recursive algorithm for the LS quadratic estimator of x_k based on the observations until time k, that is, the quadratic filter $\widehat{x}_{k/k}^q$.

To obtain this estimator, the following *augmented state* and *observation* vectors are defined

$$\mathcal{X}_k=\begin{pmatrix} x_k \\ x_k^{[2]}\end{pmatrix}\in\mathbb{R}^{n+n^2}, \qquad \mathcal{Y}_k=\begin{pmatrix} y_k \\ y_k^{[2]}\end{pmatrix}\in\mathbb{R}^{m+m^2}.$$

Clearly, the n-dimensional space of linear transformations of $\mathcal{Y}_1,\dots,\mathcal{Y}_k$ is equal to the n-dimensional space of linear transformations of $y_1,\dots,y_k$ and $y_1^{[2]},\dots,y_k^{[2]}$. Therefore, as mentioned in Section 2, the LS quadratic estimator of x_k is the LS linear estimator of x_k based on $\mathcal{Y}_1,\dots,\mathcal{Y}_k$, which is obtained by extracting the first n entries of the LS linear estimator of $\mathcal{X}_k$ based on $\mathcal{Y}_1,\dots,\mathcal{Y}_k$.

To address this problem, we consider the centered vectors $X_k=\mathcal{X}_k-E[\mathcal{X}_k]$ and $Y_k=\mathcal{Y}_k-E[\mathcal{Y}_k]$, which verify the following system [1]:

$$\begin{aligned} X_k &= \mathcal{F}_{k-1}X_{k-1}+W_{k-1}, \quad k\geq 1, \\ Y_k &= D_k^{\Theta}\mathcal{H}_kX_k+V_k, \quad k\geq 1 \end{aligned} \tag{19}$$

with

$$\mathcal{F}_k=\begin{pmatrix} F_k & 0 \\ 0 & F_k^{[2]}\end{pmatrix}, \quad \mathcal{H}_k=\begin{pmatrix} H_k & 0 \\ 0 & H_k^{[2]}\end{pmatrix}, \quad D_k^{\Theta}=\begin{pmatrix} \Theta_k & 0 \\ 0 & \Theta_k^{[2]}\end{pmatrix},$$

$$W_k = \begin{pmatrix} w_k \\ (I+K)((F_k x_k)\otimes w_k) + w_k^{[2]} - vec(Q_k) \end{pmatrix},$$

$$V_k = \begin{pmatrix} v_k \\ (I+K)((\Theta_k H_k x_k)\otimes v_k) + v_k^{[2]} - vec(R_k) \end{pmatrix} + (D_k^{\Theta} - D_k^{p})\mathcal{H}_k E[\mathcal{X}_k],$$

where $D_k^p = E[D_k^{\Theta}]$, $vec(\cdot)$ denotes the 'vec' or 'stack' operator, which vectorizes a matrix (i.e. given a matrix $A = ((a_{ij}))_{\substack{i=1,\dots,n,\\ j=1\dots,m}}$, $vec(A) = (a_{11},\dots,a_{n1},\dots,a_{1m},\dots,a_{nm})^T$) and K denotes the commutation matrix of appropriate dimensions ($K(v\otimes u) = u\otimes v$, for any vectors u, v). See [7] for details and properties of the 'vec' operator and commutation matrix.

Next, some relevant statistical properties of the initial condition, X_0, and the noises involved in the augmented system (19) are presented; these properties are necessary to derive the linear filtering algorithm for this system.

Proposition 1. The initial state X_0 is a zero-mean random vector with covariance matrix given by

$$P_0^* = \begin{pmatrix} P_0 & P_0^{(3)} \\ P_0^{(3)} & P_0^{(4)} \end{pmatrix}.$$

Proposition 2. The noise $\{W_k;\ k\geq 0\}$ is a zero-mean white noise process with

$$E[W_k W_k^T] = Q_k^W = \begin{pmatrix} Q_k & Q_k^{(3)} \\ Q_k^{(3)T} & Q_k^{22} \end{pmatrix},$$

where $Q_k^{22} = (I+K)((F_k D_k F_k^T)\otimes Q_k)(I+K) + Q_k^{(4)}$.

The reader is referred to [1] for complete proof of propositions 1 and 2.

Proposition 3. The noise $\{V_k;\ k\geq 1\}$ is a zero-mean process with

i) $E[V_k V_s^T] = 0, \quad |k-s|\neq 0, 2.$

ii) $E[V_k V_k^T] = R_{k,k}^V = \overline{R}_k + Cov[C_k^{\Theta}]\circ\left(\mathcal{H}_k E[\mathcal{X}_k]E[\mathcal{X}_k^T]\mathcal{H}_k^T\right).$

iii) $E[V_k V_{k-2}^T] = R_{k,k-2}^V = Cov[C_k^{\Theta}, C_{k-2}^{\Theta}]\circ\left(\mathcal{H}_k E[\mathcal{X}_k]E[\mathcal{X}_{k-2}^T]\mathcal{H}_{k-2}^T\right),$

where

$$\overline{R}_k = \begin{pmatrix} R_k & R_k^{(3)} \\ R_k^{(3)} & R_k^{22} \end{pmatrix}, \qquad C_k^{\Theta} = \left(\theta_k^T, \theta_k^{[2]T}\right)^T$$

with $R_k^{22} = (I+K)\left(\left(E[\theta_k\theta_k^T] \circ \left(H_k D_k H_k^T\right)\right) \otimes R_k\right)(I+K) + R_k^{(4)}$.

Proof. Obviously, $E[V_k] = 0$, $\forall k \geq 0$. Next, denoting

$$\mathcal{V}_k = \begin{pmatrix} v_k \\ (I+K)((\Theta_k H_k x_k) \otimes v_k) + v_k^{[2]} - vec(R_k) \end{pmatrix},$$

we can write $V_k = \mathcal{V}_k + (D_k^\Theta - D_k^p)\mathcal{H}_k E[\mathcal{X}_k]$.

Taking into account Hypothesis 5, we have $E[\mathcal{V}_k\left((D_s^\Theta - D_s^p)\mathcal{H}_s E[\mathcal{X}_s]\right)^T] = 0$, $\forall k, s$ and therefore

$$E[V_k V_s^T] = E[\mathcal{V}_k \mathcal{V}_s^T] + E[\left((D_k^\Theta - D_k^p)\mathcal{H}_k E[\mathcal{X}_k]\right)\left((D_s^\Theta - D_s^p)\mathcal{H}_s E[\mathcal{X}_s]\right)^T].$$

Considering hypotheses 4 and 5, Property 1 and the Kronecker product properties [7], we obtain that $E[\mathcal{V}_k\mathcal{V}_s^T] = \overline{R}_k\delta_{k,s}$, where δ denotes the Kronecker delta function and $\overline{R}_k$ is a matrix whose blocks are given by:

$$\begin{aligned}
\overline{R}_k^{11} &= E[v_k v_k^T] = R_k \\
\overline{R}_k^{12} &= E[(\Theta_k H_k x_k)^T \otimes v_k v_k^T](I+K) + E[v_k v_k^{[2]T}] \\
&= \left(\Theta_k^p E[(H_k x_k)^T] \otimes E[v_k v_k^T]\right)(I+K) + R_k^{(3)} = R_k^{(3)} \\
\overline{R}_k^{22} &= (I+K)\left(E[\Theta_k H_k x_k x_k^T H_k^T \Theta_k] \otimes E[v_k v_k^T]\right)(I+K) \\
&\quad + E[(v_k^{[2]} - vec(R_k))(v_k^{[2]} - vec(R_k))^T] \\
&\quad + (I+K)\left(E[\Theta_x H_k x_k] \otimes E[v_k(v_k^{[2]} - vec(R_k))^T]\right) \\
&\quad + E[(\Theta_x H_k x_k)^T] \otimes E[(v_k^{[2]} - vec(R_k))v_k^T](I+K) \\
&= (I+K)\left(E[\theta_k\theta_k^T] \circ E[H_k x_k x_k^T H_k^T] \otimes E[v_k v_k^T]\right)(I+K) + R_k^{(4)} = R_k^{22}.
\end{aligned}$$

Finally, from Property 1, we have that

$$\begin{aligned}
&E[\left((D_k^\Theta - D_k^p)\mathcal{H}_k E[\mathcal{X}_k]\right)\left((D_s^\Theta - D_s^p)\mathcal{H}_s E[\mathcal{X}_s]\right)^T] \\
&\quad = Cov[C_k^\Theta, C_s^\Theta] \circ \left(\mathcal{H}_k E[\mathcal{X}_k] E[\mathcal{X}_s^T]\mathcal{H}_s^T\right),
\end{aligned}$$

where $Cov[C_k^\Theta, C_s^\Theta] = 0$ for $|k-s| \neq 0, 2$.

So, the proposition is proven. □

Proposition 4.
a) The initial state X_0 and the noises $\{W_k;\ k \geq 0\}$ and $\{V_k;\ k \geq 1\}$ are uncorrelated.
b) The matrix D_k^Θ is independent of $(X_0, \{W_k;\ k \geq 0\}, V_1, \ldots, V_{k-3}, V_{k-1})$.

Proof. Taking into account Hypothesis 5 and the Kronecker product properties, the uncorrelation between X_0, $\{W_k;\ k \geq 0\}$ and $\{V_k;\ k \geq 1\}$ is obtained. Moreover, since

$$\begin{aligned}&(X_0, \{W_k;\ k \geq 0\}, V_1, \dots, V_{k-3}, V_{k-1}) \ \text{is a function of}\\ &(x_0, \{w_k;\ k \geq 0\}, \{v_j;\ j \leq k-3,\ j = k-1\}, \{\theta_j;\ j \leq k-3,\ j = k-1\}),\end{aligned}$$

and D_k^Θ is a function of θ_k, the model independence assumptions guarantees that the matrix D_k^Θ is independent of $(X_0, \{W_k;\ k \geq 0\}, V_1, \dots, V_{k-3}, V_{k-1})$, and the proof is concluded. □

Linear filtering algorithm for the augmented system. From the properties established in the previous propositions and reasoning similarly to Section 3, the following recursive algorithm for the LS linear estimator $\widehat{X}_{k/k}$ is derived.

The filter state is given by the following relation

$$\widehat{X}_{k/k} = \widehat{X}_{k/k-1} + \mathcal{G}_{k,k}\Lambda_k^{-1}\mathcal{I}_k, \quad k \geq 1; \quad \widehat{X}_{0/0} = 0,$$

where the predictor, $\widehat{X}_{k/k-1}$, is given by

$$\widehat{X}_{k/k-1} = \mathcal{F}_{k-1}\widehat{X}_{k-1/k-1}, \quad k \geq 1.$$

The innovation process satisfies

$$\begin{aligned}\mathcal{I}_k &= Y_k - D_k^p\mathcal{H}_k\widehat{X}_{k/k-1} - \Upsilon_{k,k-2}\Lambda_{k-2}^{-1}\left(\mathcal{I}_{k-2} - \mathcal{T}_{k-1,k-2}^T\Lambda_{k-1}^{-1}\mathcal{I}_{k-1}\right), \quad k \geq 3,\\ \mathcal{I}_2 &= Y_2 - \mathcal{T}_{2,1}\Lambda_1^{-1}Y_1,\\ \mathcal{I}_1 &= Y_1,\end{aligned}$$

where

$$\begin{aligned}&\Upsilon_{k,k-2} = Cov[C_k^\Theta, C_{k-2}^\Theta] \circ \left(\mathcal{H}_k\mathcal{F}_{k-1}\mathcal{F}_{k-2}\mathcal{E}_{k-2}\mathcal{H}_{k-2}^T\right) + R_{k,k-2}^V, \quad k \geq 3,\\ &\mathcal{E}_k = \mathcal{F}_{k-1}\mathcal{E}_{k-1}\mathcal{F}_{k-1}^T + Q_{k-1}^W, \quad k \geq 1; \quad \mathcal{E}_0 = P_0^*,\\ &\mathcal{T}_{k,k-1} = D_k^p\mathcal{H}_k\mathcal{F}_{k-1}\mathcal{G}_{k-1,k-1} - \Upsilon_{k,k-2}\Lambda_{k-2}^{-1}\mathcal{T}_{k-1,k-2}^T, \quad k \geq 3,\\ &\mathcal{T}_{2,1} = D_2^p\mathcal{H}_2\mathcal{F}_1\mathcal{E}_1\mathcal{H}_1^T D_1^p.\end{aligned}$$

The innovation covariance matrix is specified by

$$\begin{aligned}\Lambda_k &= Cov[C_k^\Theta] \circ \left(\mathcal{H}_k\mathcal{E}_k\mathcal{H}_k^T\right) + R_{k,k}^V - D_k^p\mathcal{H}_k\Sigma_{k/k-1}\mathcal{H}_k^T D_k^p + D_k^p\mathcal{H}_k\mathcal{G}_{k,k}\\ &\quad + \mathcal{G}_{k,k}^T\mathcal{H}_k^T D_k^p - \Upsilon_{k,k-2}\Lambda_{k-2}^{-1}\left(I + \mathcal{T}_{k-1,k-2}^T\Lambda_{k-1}^{-1}\mathcal{T}_{k-1,k-2}\Lambda_{k-2}^{-1}\right)\Upsilon_{k,k-2}^T, \ k \geq 3,\\ \Lambda_2 &= E[C_2^\Theta C_2^{\Theta^T}] \circ \left(\mathcal{H}_2\mathcal{E}_2\mathcal{H}_2^T\right) + R_{2,2}^V - \mathcal{T}_{2,1}\Lambda_1^{-1}\mathcal{T}_{2,1}^T,\\ \Lambda_1 &= E[C_1^\Theta C_1^{\Theta^T}] \circ \left(\mathcal{H}_1\mathcal{E}_1\mathcal{H}_1^T\right) + R_{1,1}^V.\end{aligned}$$

The matrix $\mathcal{G}_{k,k}$ is determined by

$$\begin{aligned}\mathcal{G}_{k,k} &= \Sigma_{k/k-1}\mathcal{H}_k^T D_k^p \\ &\quad - \mathcal{F}_{k-1}\left(\mathcal{F}_{k-2}\mathcal{G}_{k-2,k-2} - \mathcal{G}_{k-1,k-1}\Lambda_{k-1}^{-1}\mathcal{T}_{k-1,k-2}\right)\Lambda_{k-2}^{-1}\Upsilon_{k,k-2}^T, \quad k \geq 3, \\ \mathcal{G}_{2,2} &= \mathcal{E}_2\mathcal{H}_2^T D_2^p - \mathcal{F}_1\mathcal{E}_1\mathcal{H}_1^T D_1^p \Lambda_1^{-1}\mathcal{T}_{2,1}^T, \\ \mathcal{G}_{1,1} &= \mathcal{E}_1\mathcal{H}_1^T D_1^p\end{aligned}$$

with

$$\begin{aligned}\Sigma_{k/k-1} &= \mathcal{F}_{k-1}\Sigma_{k-1/k-1}\mathcal{F}_{k-1}^T + Q_{k-1}^W, \quad k \geq 1, \\ \Sigma_{k/k} &= \Sigma_{k/k-1} - \mathcal{G}_{k,k}\Lambda_k^{-1}\mathcal{G}_{k,k}^T, \quad k \geq 1, \quad \Sigma_{0/0} = P_0^*.\end{aligned}$$

5 Numerical simulation example

To illustrate the effectiveness of the proposed estimation algorithms, we consider the same system as in [1] but, according to the current theoretical study, now the variables modelling the uncertainty of the observations are assumed to be correlated at sampling times that differ two units of time.

This system consists of a scalar first-order autoregressive model,

$$x_k = 0.95x_{k-1} + w_{k-1}, \quad k \geq 1,$$

where x_0 is a zero-mean Gaussian variable with variance $P_0 = 1$ and the process $\{w_k;\ k \geq 0\}$ is a zero-mean white Gaussian noise with variances $Q_k = 0.1$, $\forall k \geq 0$. The scalar uncertain observations come from two sensors and are perturbed by zero-mean white noise processes $\{v_k^i;\ k \geq 1\}$, $i = 1, 2$,

$$y_k^i = \theta_k^i x_k + v_k^i, \quad k \geq 1, \quad i = 1, 2.$$

These noises have the following probability distributions:

$$\begin{aligned}&P[v_k^1 = -8] = \frac{1}{8}, \quad P[v_k^1 = \frac{8}{7}] = \frac{7}{8}, \quad \forall k \geq 1, \\ &P[v_k^2 = 1] = \frac{15}{18}, \quad P[v_k^2 = -3] = \frac{2}{18}, \quad P[v_k^2 = -9] = \frac{1}{18}, \quad \forall k \geq 1,\end{aligned}$$

and variances given by $R_k^1 = 64/7$ and $R_k^2 = 19/3$, $\forall k \geq 1$, respectively.

The variables modelling the uncertainty of each sensor, θ_k^i, are defined from two independent sequences of independent Bernoulli random variables, $\{\gamma_k^i;\ k \geq 0\}$, $i = 1, 2$ with $P[\gamma_k^i = 1] = \gamma_i$. Specifically, the uncertainty variables are defined by the following relation

$$\theta_k^i = 1 - \gamma_{k+2}^i(1 - \gamma_k^i), \quad i = 1, 2.$$

Note that if $\theta_k^i = 0$, then $\gamma_{k+2}^i = 1$ and $\gamma_k^i = 0$, and hence, $\theta_{k+2}^i = 1$. This fact guarantees that, at each sensor, no more than two consecutive observations consisting of noise only can be transmitted.

Since the variables γ_k^i and γ_s^i are independent, θ_k^i and θ_s^i are also independent for $|k-s| \neq 0, 2$. The mean of these variables is given by $\overline{\theta}^i = 1-\gamma_i(1-\gamma_i)$ and

$$E[(\theta_k^i - \overline{\theta}^i)(\theta_s^i - \overline{\theta}^i)] = \begin{cases} 0 & \text{si } |k-s| \neq 0, 2 \\ -(1-\overline{\theta}^i)^2 & \text{si } |k-s| = 2 \\ \overline{\theta}^i(1-\overline{\theta}^i) & \text{si } |k-s| = 0 \end{cases}$$

To illustrate the effectiveness of the proposed linear and quadratic filtering estimators, both algorithms have been implemented in Matlab using different values of the probabilities γ_1 and γ_2, which provide different values of the probabilities $\overline{\theta}^1$ and $\overline{\theta}^2$. Since the values of $\overline{\theta}^i$ are the same if $1-\gamma_i$ is used instead of γ_i, only the case $\gamma_i \leq 0.5$ has been considered (note that, in such case, the probability of transmitting only noise at the ith sensor, $1-\overline{\theta}^i$, is an increasing function of γ_i).

In all the cases considered, the filtering error variances generally present insignificant variation from a certain iteration onwards; therefore, only the error variances at a fixed iteration (namely, $k = 50$) are displayed here in order to show more clearly the error variance evolution with respect to the values of γ_1 and γ_2.

Figure **??** shows the linear and quadratic filtering error variances at $k = 50$ versus γ_1 (for constant values of γ_2) and Figure **??** presents these variances versus γ_2 (for constant values of γ_1). Both figures show that, as the values γ_1 and γ_2 increase (which means that the probability of transmitting only noise increases), the filtering error variances also increase and, therefore, the filtering estimates are worse. In addition, we also observe that the error variances of the quadratic filter are always significantly smaller than those of the linear filter, confirming the higher accuracy of the quadratic estimator over the linear one.

6 Conclusion

Linear and quadratic state estimation problems have been addressed for uncertain observations coming from multiple sensors and featuring correlation in the uncertainty at instants that differ two units of time. The observation

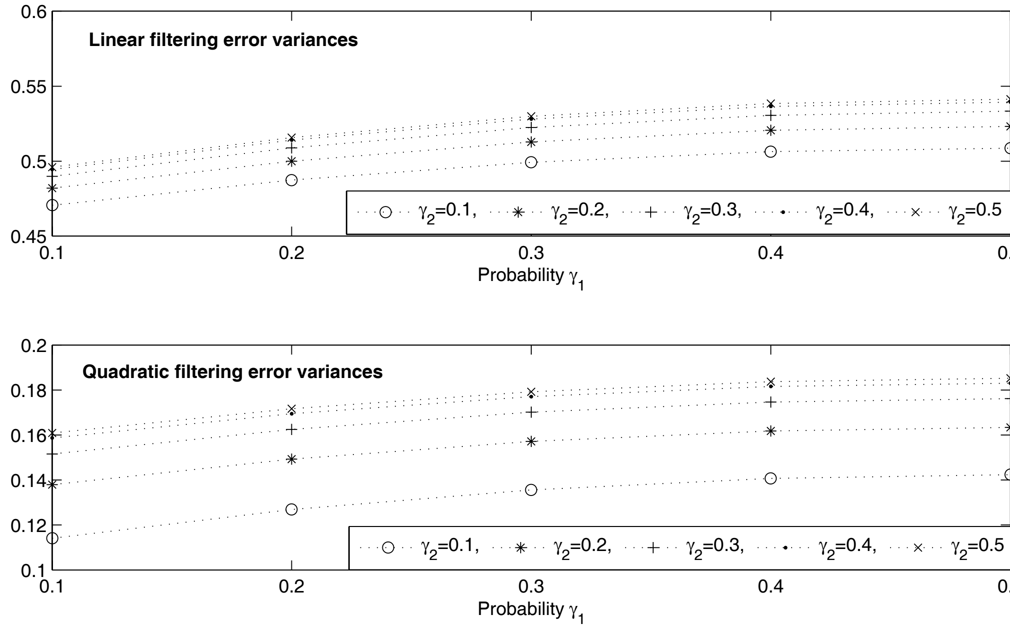

Figure 1: Linear and quadratic filtering error variances at $k = 50$ versus γ_1 with γ_2 varying from 0.1 to 0.5

model considered covers those situations where sensors can fail and transmit only noise, but no more than two consecutive observations without information on the system state can be transmitted by the same sensor. Simulation results confirm that, as expected, if the probability of transmitting only noise at one of the sensors increases, worse estimations are obtained. It must be noticed that the current correlation model of the variables describing the uncertainty can be generalized by considering variables correlated at instants that differ m units of time ($m \geq 2$).

ACKNOWLEDGEMENTS. This research is supported by Ministerio de Educación y Ciencia (grant No. MTM2008-05567) and Junta de Andalucía (grant No. P07-FQM-02701).

References

[1] R. Caballero-Águila, A. Hermoso-Carazo and J. Linares-Pérez, Linear and quadratic estimation using uncertain observations from multiple sensors

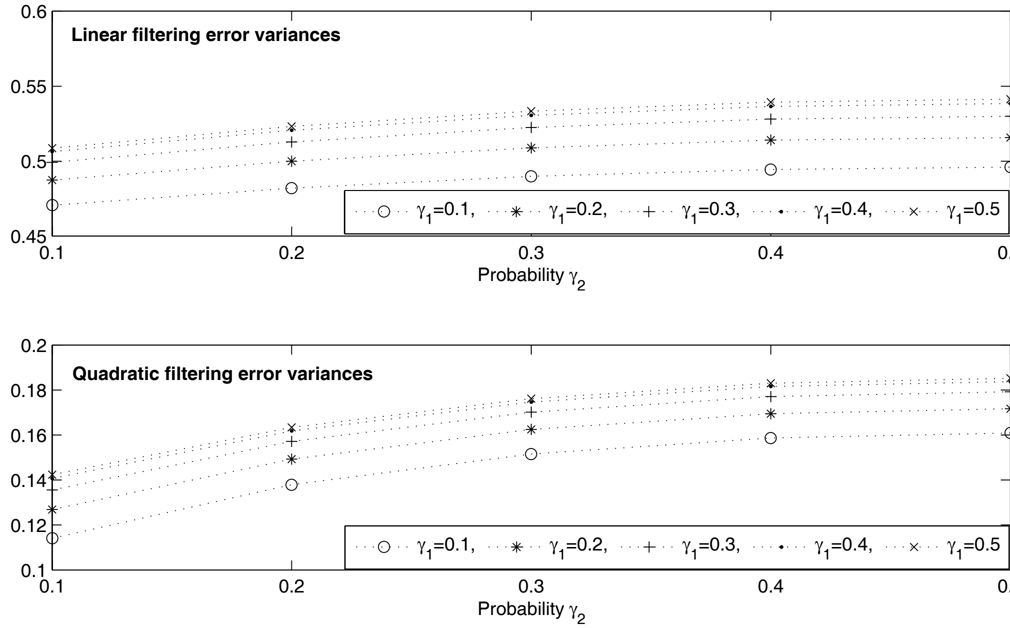

Figure 2: Linear and quadratic filtering error variances at $k = 50$ versus γ_2 with γ_1 varying from 0.1 to 0.5

with correlated uncertainty, *Signal Processing,* **91** (2011), 330 - 337.

[2] F. Carravetta, A. Germani and M. Raimondi, Polynomial filtering for linear discrete time non-Gaussian systems, *SIAM J. Control and Optimization,* **34** (1996), 1666 - 1690.

[3] A. De Santis, A. Germani and M. Raimondi, Optimal quadratic filtering of linear discrete-time non-Gaussian systems, *IEEE Transactions on Automatic Control,* **40** (1995), 1274 - 1278.

[4] F.O. Hounkpevi and E.E. Yaz, Robust minimum variance linear state estimators for multiple sensors with different failure rates, *Automatica,* **43** (2007), 1274 - 1280.

[5] J.D. Jiménez-López, J. Linares-Pérez, S. Nakamori, R. Caballero-Águila, A. Hermoso-Carazo, Signal estimation based on covariance information from observations featuring correlated uncertainty and coming from multiple sensors, *Signal Processing,* **88** (2008), 2998 - 3006.

[6] T. Kailath, A.H. Sayed and B. Hassibi, *Linear Estimation,* Prentice Hall, New Jersey, 2000.

[7] J.R. Magnus and H.H. Neudecker, *Matrix differential calculus with applications in statistics and econometrics,* revised ed, Wiley, New York, 1999.

[8] S. Nakamori, R. Caballero-Águila, A. Hermoso-Carazo, J.D. Jiménez-López and J. Linares-Pérez, Least-squares ν th-order polynomial estimation of signals from observations affected by non-independent uncertainty, *Applied Mathematics and Computation,* **176** (2006), 642 - 653.

[9] S. Nakamori, R. Caballero-Águila, A. Hermoso-Carazo, J.D. Jiménez-López and J. Linares-Pérez, Signal polynomial smoothing from correlated interrupted observations based on covariances, *Mathematical Methods in the Applied Sciences,* **30** (2007), 1645 - 1665.

[10] W. NaNacara and E.E. Yaz, Recursive estimator for linear and nonlinear systems with uncertain observations, *Signal Processing,* **62** (1997), 215 - 228.

[11] E. Yaz, Relationship between several novel control schemes proposed for a class of nonlinear stochastic system, *International Journal of Control,* **45** (1987), 1447 - 1454.

Received: October, 2011