Istituto di Analisi dei Sistemi ed Informatica
CONSIGLIO NAZIONALE DELLE RICERCHE

A. Germani, C. Manes, P. Palumbo

STATE ESTIMATION OF A CLASS OF STOCHASTIC VARIABLE STRUCTURE SYSTEMS

R. 548 2001

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@iasi.rm.cnr.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

Collana dei Rapporti
dell'Istituto di Analisi dei Sistemi ed Informatica, CNR

viale Manzoni 30, 00185 ROMA, Italy

tel. ++39-06-77161
fax ++39-06-7716461
email: iasi@iasi.rm.cnr.it
URL: http://www.iasi.rm.cnr.it

Abstract

This paper considers the problem of state estimation for discrete-time systems whose dynamics switches within a finite set of linear stochastic behaviors. In recent years such systems are receiving a growing attention because of their importance from an applicative point of view, in that switching phenomena are normally present in many engineering problems. The solution of the filtering problem depends on the amount of the *a priori* information about the switching process.

In this paper a new approach in filtering the state of a stochastic variable structure system, driven by a Markovian jump is proposed. By using a state space realization for the Markov process and the formalism of bilinear systems, a linear filter achieves the best estimate among all the linear transformations of the measured output. The algorithm can be considered as a useful key tool to solve the problem of suboptimal estimates, when the system is affected by non-Gaussian noises, in the framework of polynomial filters.

Key words: State estimation, switching systems, hybrid systems, bilinear systems.

1. Introduction

Switching systems, also denoted hybrid systems or variable structure systems, are receiving a growing attention in recent years because of their importance from an applicative point of view, in that switching phenomena are normally present in many engineering problems (for a survey on hybrid systems control and applications see [5,2,10]). Many authors investigated the problem of state estimation for switching systems. Most papers in literature deal with systems with a stochastically driven switching sequence, modeled as a finite-state Markov Chain (see e.g. [1, 3, 8, 7, 4, 11, 6] for the discrete-time case and [9, 13, 12] for the continuous-time case).

In the framework of discrete-time systems, the problem of filtering the state of stochastic switching systems was first formulated in [1]. The authors pointed out the complexity of the exact solution of the problem and proposed an approximate solution. For the same problem in [7], where a partial observation of the switching process is assumed, an almost-recursive implementation of the exact solution is derived, whose complexity geometrically grows with time. In [4] a linear filter is implemented based on a clever use of the characteristic function associated to the Markovian jump. In [6] different approximate state estimators have been analyzed, without assuming observations on the switching process. All estimators proposed in [6] are iterative algorithms over a finite observation time, and do not allow a recursive implementation.

This paper proposes a linear algorithm in order to estimate the state of a discrete-time variable structure system driven by a Markovian jump. By using a bilinear approach, the optimal linear estimate, in the sense of minimum variance is here achieved. The paper develops the case of a two-state Markov chain, although with some more computations, a more general finite-state Markov chain could be investigated with the same methodology. The result is particularly useful to solve the problem of suboptimal estimates, when the system is affected by non-Gaussian noises, in the framework of polynomial filters.

2. The system to be filtered

The class of systems investigated in this paper is described by the following set of equations:

$$\begin{aligned} x(k+1) &= A_{\mu(k)}x(k) + B_{\mu(k)}u(k) + F_{\mu(k)}N_k, \qquad x(0)=x_0, \qquad k\geq 0,\\ y(k) &= C_{\mu(k)}x(k) + D_{\mu(k)}u(k) + G_{\mu(k)}N_k, \end{aligned} \tag{2.1}$$

where $x(k)$ is the stochastic state variable in $I\!R^n$, $u(k)$ is a deterministic known input in $I\!R^p$, $y(k)$ is the measured output in $I\!R^q$ and $N(k)$ is a zero-mean white noise standard sequence in $I\!R^b$, that is:

$$I\!E\big[N(k)N(h)^T\big] = \delta_{kh}I_b, \qquad \forall k,h\in I\!N \tag{2.2}$$

and, moreover, the state noise is supposed to be uncorrelated with the output noise, that is:

$$F_{\mu(k)}G^T_{\mu(h)} = O_{n\times q}, \qquad \forall k,h\in I\!N \tag{2.3}$$

The initial state x_0 is a random variable with mean and covariance matrix available, named χ and Ψ_0 respectively, and is independent of the noise sequence $\{N(k),\ k\in I\!N\}$.

The dynamic matrices of system (2.1) are forced to assume values on binary ranges, for instance: $\mathcal{R}\big(A_{\mu(k)}\big) = \{A_0, A_1\}$, all according to a two-state Markov chain $\mu(k)$, with range in $\{0,1\}$, transition probability matrix Π_μ and a given initial condition $P(\mu(0)=0)=p_0$.

Remark 2.1. By using the above notation the wide class of failure systems with a working plant, referred to the zero value of $\mu(k)$ for instance, and a failure plant driven by the Markovian jump, can be represented with a useful assignment of the Π_μ matrix. Multiple failure systems and hybrid systems could also be modeled by equations (2.1) with a finite state Markov chain and a few more computation efforts.

In this paper the filtering problem is solved by using a suitable state space model for the Markovian jump:

Lemma 2.2. *Let $\{v_0(k),\ k\in I\!N\}$, $\{v_1(k),\ k\in I\!N\}$ be white noise independent sequences, both assuming values in $\{0,1\}$, with distribution $P\big(v_0(k)=1\big)=\varepsilon_0$, $P\big(v_1(k)=1\big)=\varepsilon_1$ respectively; let $\{\mu(k),\ k\in I\!N\}$ be the following stochastic sequence, assuming values in $\{0,1\}$:*

$$\begin{aligned} \mu(k+1) &= \mu(k)+\big(1-\mu(k)\big)v_0(k)-\mu(k)v_1(k), \qquad && \mu(k)\in\{0,1\}, \qquad k\in I\!N \\ \mu(0) &= \mu_0, && P\big(\mu_0=0\big)=p_0. \end{aligned} \tag{2.4}$$

with μ_0 independent of $\{v_0(k),v_1(h),\quad k,h\in I\!N\}$. Then $\{\mu(k),\ k\in I\!N\}$ is a two-state Markov chain, with transition probability matrix Π_μ such that:

$$\Pi_\mu=\begin{bmatrix} P\big(\mu(k+1)=0|\mu(k)=0\big) & P\big(\mu(k+1)=0|\mu(k)=1\big) \\ P\big(\mu(k+1)=1|\mu(k)=0\big) & P\big(\mu(k+1)=1|\mu(k)=1\big)\end{bmatrix}=\begin{bmatrix}1-\varepsilon_0 & \varepsilon_1 \\ \varepsilon_0 & 1-\varepsilon_1\end{bmatrix} \tag{2.5}$$

Proof. The proof that $\{\mu(k),\ k\in I\!N\}$ is a Markov chain strictly follows from the recursive equations (2.4), in that for each $k,h\in I\!N$, it comes that the conditional random variable $\mu(k+h)|\mu(k)$ is a Borel function of $\{v_0(i),v_1(j),\ i,j=k,\dots,k+h-1\}$, independent of $\{\mu(l),\ l<k\}$. The coefficients of matrix Π_μ easily come from probability computation. ∎

Remark 2.3. The sequences $\{v_i(k),\ k\in I\!N\}$, $i=0,1$ are not zero-mean. It is useful for the sequel to rewrite the recursive Markov equations (2.4) introducing zero-mean sequences, so that they change in:

$$\mu(k+1)=(1-\varepsilon_0-\varepsilon_1)\mu(k)-\big(\mathcal{V}_0(k)+\mathcal{V}_1(k)\big)\mu(k)+\mathcal{V}_0(k)+\varepsilon_0, \qquad \mu(0)=\mu_0. \tag{2.6}$$

with $\mathcal{V}_i(k)=v_i(k)-I\!E\big[v_i(k)\big]=v_i(k)-\varepsilon_i$.

Remark 2.4. Suppose that system (2.1) models a failure system. Then, according to Remark 2.1, ε_0, ε_1 can be considered, respectively, as the probability of failure occuring from a working plant, and the probability of failure vanishing from a failure plant.

Remark 2.5. According to the Markov chain theory, a suitable choice for the initial distribution p_0 of μ_0 allows a constant value for the probability masses of $\mu(k)$. That means, if p_0 is such that:

$$\begin{pmatrix} p_0 \\ 1-p_0\end{pmatrix}=\begin{bmatrix}1-\varepsilon_0 & \varepsilon_1 \\ \varepsilon_0 & 1-\varepsilon_1\end{bmatrix}\begin{pmatrix} p_0 \\ 1-p_0\end{pmatrix} \qquad \Longrightarrow \qquad \begin{pmatrix} p_0 \\ 1-p_0\end{pmatrix}=\begin{pmatrix}\frac{\varepsilon_1}{\varepsilon_0+\varepsilon_1} \\ \\ \frac{\varepsilon_0}{\varepsilon_0+\varepsilon_1}\end{pmatrix} \tag{2.7}$$

$\mu(k)$ has a stationary distribution:

$$P\big(\mu(k)=0\big)=\frac{\varepsilon_1}{\varepsilon_0+\varepsilon_1} \tag{2.8}$$

Remark 2.6. The Markovian jump model (2.1) can be written by using the following bilinear formalism:

$$\begin{aligned} x(k+1) &= \Big(A_0 + (A_1 - A_0)\mu(k)\Big)x(k) + \Big(B_0 + (B_1 - B_0)\mu(k)\Big)u(k) \\ &\qquad + \Big(F_0 + (F_1 - F_0)\mu(k)\Big)N(k), \\ y(k) &= \Big(C_0 + (C_1 - C_0)\mu(k)\Big)x(k) + \Big(D_0 + (D_1 - D_0)\mu(k)\Big)u(k) \\ &\qquad + \Big(G_0 + (G_1 - G_0)\mu(k)\Big)N(k), \end{aligned} \tag{2.9}$$

with $\mu(k)$ as in (2.6).

3. The filtering algorithm

The filtering problem is solved by estimating the following extended state:

$$\xi(k) = \begin{pmatrix} \xi_1(k) \\ \xi_2(k) \\ \xi_3(k) \end{pmatrix} = \begin{pmatrix} x(k) \\ \mu(k)x(k) \\ \mu(k) \end{pmatrix} \in I\!R^{2n+1}, \qquad \xi_1(k), \xi_2(k) \in I\!R^n, \quad \xi_3(k) \in I\!R. \tag{3.1}$$

Theorem 3.1. *The evolution of $\xi(k)$ is given by a time-varying discrete-time bilinear system of the type*

$$\begin{aligned} \xi(k+1) &= \mathcal{A}(k)\xi(k) + \mathcal{B}(k, \xi(k), \eta(k)) + \mathcal{U}(k) + \mathcal{D}(k)\eta(k), \qquad && \xi(0) = \big(x_0^T,\ \mu_0 x_0^T,\ \mu_0\big)^T \\ y(k) &= \mathcal{C}(k)\xi(k) + \mathcal{M}(\xi(k), \eta(k)) + \mathcal{W}(k) + \mathcal{R}\eta(k), && k \geq 0 \end{aligned} \tag{3.2}$$

where $\eta(k)$ is a suitable zero-mean white noise sequence, $\mathcal{B}(k, \xi(k), \eta(k))$, $\mathcal{M}(\xi(k), \eta(k))$ are bilinear forms, $\mathcal{U}(k)$ and $\mathcal{W}(k)$ are deterministic terms and $\mathcal{A}(k)$, $\mathcal{D}(k)$, $\mathcal{C}(k)$ and $\mathcal{R}$ are suitable matrices.

Proof. The proof is constructive. The evolution of $\xi_1(k)$, $\xi_2(k)$ and $\xi_3(k)$ are given by:

$$\begin{aligned} \xi_1(k+1) &= A_0\xi_1(k) + (A_1 - A_0)\xi_2(k) + (B_1 - B_0)u(k)\xi_3(k) + B_0 u(k) \\ &\qquad + (F_1 - F_0)N(k)\xi_3(k) + F_0 N(k), \end{aligned} \tag{3.3}$$

$$\begin{aligned} \xi_2(k+1) &= \varepsilon_0 A_0 \xi_1(k) + \Big((1-\varepsilon_1)A_1 - \varepsilon_0 A_0\Big)\xi_2(k) + \Big((1-\varepsilon_1)B_1 - \varepsilon_0 B_0\Big)u(k)\xi_3(k) \\ &+ \varepsilon_0 B_0 u(k) - B_0 u(k)\xi_3(k)\mathcal{V}_0(k) - B_1 u(k)\xi_3(k)\mathcal{V}_1(k) + A_0\xi_1(k)\mathcal{V}_0(k) \\ &- A_0\xi_2(k)\mathcal{V}_0(k) - A_1\xi_2(k)\mathcal{V}_1(k) + (1-\varepsilon_1)F_1 N(k)\xi_3(k) - \varepsilon_0 F_0 N(k)\xi_3(k) \\ &- F_1 N(k)\mathcal{V}_1(k)\xi_3(k) - F_0 N(k)\mathcal{V}_0(k)\xi_3(k) + B_0 u(k)\mathcal{V}_0(k) + F_0 N(k)\mathcal{V}_0(k) + \varepsilon_0 F_0 N(k), \end{aligned} \tag{3.4}$$

$$\xi_3(k+1) = (1 - \varepsilon_0 - \varepsilon_1)\xi_3(k) - \big(\mathcal{V}_0(k) + \mathcal{V}_1(k)\big)\xi_3(k) + \mathcal{V}_0(k) + \varepsilon_0 \tag{3.5}$$

By defining the following stochastic sequence:

$$\eta(k) = \begin{pmatrix} \chi_1(k) \\ \chi_2(k) \\ \chi_3(k) \\ \chi_4(k) \\ \chi_5(k) \end{pmatrix} = \begin{pmatrix} N(k) \\ \mathcal{V}_0(k) \\ \mathcal{V}_1(k) \\ \mathcal{V}_0(k)N(k) \\ \mathcal{V}_1(k)N(k) \end{pmatrix} \in I\!R^m, \qquad \begin{array}{l} m = 3b+2, \\ \chi_i(k) \in I\!R^b, \quad i = 1,4,5, \\ \chi_i(k) \in I\!R, \quad i = 2,3, \end{array} \tag{3.6}$$

equation (3.2) is achieved with the time-variant matrix $\mathcal{A}(k)$ given by:

$$\mathcal{A}(k) = \begin{bmatrix} A_0 & A_1 - A_0 & (B_1 - B_0)u(k) \\ \varepsilon_0 A_0 & (1-\varepsilon_1)A_1 - \varepsilon_0 A_0 & \Big((1-\varepsilon_1)B_1 - \varepsilon_0 B_0\Big)u(k) \\ O_{1\times n} & O_{1\times n} & (1-\varepsilon_0-\varepsilon_1) \end{bmatrix}, \tag{3.7}$$

the time-variant bilinear form:

$$\mathcal{B}(k,\xi(k),\eta(k)) = \begin{pmatrix} \mathcal{B}_1(\xi(k)\eta(k)) \\ \mathcal{B}_2(k,\xi(k)\eta(k)) \\ \mathcal{B}_3(\xi(k)\eta(k)) \end{pmatrix}, \tag{3.8}$$

with

$$\begin{aligned} \mathcal{B}_1(\xi(k),\eta(k)) &= (F_1 - F_0)\chi_1(k)\xi_3(k) \\ \mathcal{B}_2(k,\xi(k),\eta(k)) &= -B_0u(k)\xi_3(k)\chi_2(k) - B_1u(k)\xi_3(k)\chi_3(k) + A_0\xi_1(k)\chi_2(k) \\ &\quad - A_0\xi_2(k)\chi_2(k) - A_1\xi_2(k)\chi_3(k) + \Big((1-\varepsilon_1)F_1 - \varepsilon_0F_0\Big)\chi_1(k)\xi_3(k) \\ &\quad - F_1\chi_5(k)\xi_3(k) - F_0\chi_4(k)\xi_3(k), \\ \mathcal{B}_3(\xi(k),\eta(k)) &= -\xi_3(k)\chi_2(k) - \xi_3(k)\chi_3(k) \end{aligned} \tag{3.9}$$

and

$$\mathcal{D}(k) = \begin{bmatrix} F_0 & O_{n\times 1} & O_{n\times 1} & O_{n\times b} & O_{n\times b} \\ \varepsilon_0F_0 & B_0u(k) & O_{n\times 1} & F_0 & O_{n\times b} \\ O_{1\times b} & 1 & 0 & O_{1\times b} & O_{1\times b} \end{bmatrix}, \qquad \mathcal{U}(k) = \begin{pmatrix} B_0u(k) \\ \varepsilon_0B_0u(k) \\ \varepsilon_0 \end{pmatrix} \tag{3.10}$$

The measurements equation in (3.2) comes by defining

$$\mathcal{C}(k) = \big[C_0 \quad C_1 - C_0 \quad (D_1 - D_0)u(k)\big], \qquad \mathcal{W}(k) = D_0u(k), \qquad \mathcal{R} = \big[G_0 \quad O_{q\times 2(b+1)}\big] \tag{3.11}$$

and

$$\mathcal{M}(\xi(k),\eta(k)) = (G_1 - G_0)\chi_1(k)\xi_3(k). \tag{3.12}$$

From Remark 2.3 it easily comes that $\eta(k)$ is a zero-mean sequence. In order to show the whiteness of $\eta(k)$, it has to be verified that

$$I\!E\big[\chi_i(k)\chi_j(h)^T\big] = O_{ij}, \qquad \forall i,j = 1,\dots 5, \qquad \forall k \neq h \tag{3.13}$$

where O_{ij} is a null matrix of suitable dimensions, according to (3.6). Some of (3.13) conditions are clearly satisfied by construction or definition ($i = 1,\dots,3$, $j = i,\dots,3$). All the others are easily achieved by using the whiteness properties of N or $\mathcal{V}_i$ as it is shown by the following computation: let $i = 4$, $j = 5$ and $k \neq h$:

$$\begin{aligned} E\big[\chi_4(k)\chi_5(h)^T\big] &= E\big[\mathcal{V}_0(k)N(k)N(h)^T\mathcal{V}_1(h)\big] = E\big[\mathcal{V}_0(k)\mathcal{V}_1(h)N(k)N(h)^T\big] \\ &= E\big[\mathcal{V}_0(k)\mathcal{V}_1(h)\big]E\big[N(k)N(h)^T\big] = O_{b\times b} \end{aligned} \tag{3.14}$$

∎

For the sequel, it is useful to rewrite the extended system (3.2) with respect to a zero-mean state variable

$$\zeta(k) = \xi(k) - \bar{\xi}(k), \tag{3.15}$$

with $\bar{\xi}(k) = E\big[\xi(k)\big]$.

Theorem 3.2. *The evolution of the zero-mean sequence $\zeta(k)$ defined in* (3.15) *is given by the following equations:*

$$\begin{aligned} \zeta(k+1) &= \mathcal{A}(k)\zeta(k) + \mathcal{F}(k), \qquad\qquad \zeta(0) = \zeta_0 = \xi(0) - \bar{\xi}(0), \\ y(k) &= \mathcal{C}(k)\zeta(k) + \mathcal{G}(k) + \mathcal{Z}(k) \end{aligned} \tag{3.16}$$

with

$$\begin{aligned} \mathcal{F}(k) &= \mathcal{B}\big(k, \zeta(k), \eta(k)\big) + \mathcal{B}\big(k, \bar{\xi}(k), \eta(k)\big) + \mathcal{D}(k)\eta(k) \\ \mathcal{G}(k) &= \mathcal{M}(\zeta(k), \eta(k)) + \mathcal{M}(\bar{\xi}(k), \eta(k)) + \mathcal{R}\eta(k) \end{aligned} \tag{3.17}$$

zero-mean white noise uncorrelated sequences. $\mathcal{Z}(k) = \mathcal{C}(k)\bar{\xi}(k) + \mathcal{W}(k)$ is a deterministic term.

Proof. Substitute $\xi(k) = \zeta(k) + \bar{\xi}(k)$ in (3.2). After some easy computations it comes:

$$\zeta(k+1) = \mathcal{A}(k)\zeta(k) + \mathcal{B}(k, \zeta(k), \eta(k)) + \mathcal{B}(k, \bar{\xi}(k), \eta(k)) + \mathcal{D}(k)\eta(k) - E\big[\mathcal{B}(k, \xi(k), \eta(k))\big] \tag{3.18}$$

Let us prove that the last term $E\big[\mathcal{B}(k, \xi(k), \eta(k))\big]$ is null. The first n components mean is:

$$E\big[\mathcal{B}_1(\xi(k), \eta(k))\big] = (F_1 - F_0)E\big[\mu(k)N(k)\big] = 0 \tag{3.19}$$

as it comes from the independence of the sequences μ and N. The second n components mean is given by the following eight terms:

$$\begin{aligned} E\big[\mathcal{B}_2(k, \xi(k), \eta(k))\big] &= -B_0u(k)E\big[\mu(k)\mathcal{V}_0(k)\big] - B_1u(k)E\big[\mu(k)\mathcal{V}_1(k)\big] \\ &\quad + A_0E\big[\mathcal{V}_0(k)x(k)\big] - A_0E\big[\mathcal{V}_0(k)\mu(k)x(k)\big] - A_1E\big[\mathcal{V}_1(k)\mu(k)x(k)\big] \\ &\quad + \Big((1-\varepsilon_1)F_1 - \varepsilon_0F_0\Big)E\big[\mu(k)N(k)\big] \\ &\quad - F_1E\big[\mu(k)\mathcal{V}_1(k)N(k)\big] - F_0E\big[\mu(k)\mathcal{V}_0(k)N(k)\big] \end{aligned} \tag{3.20}$$

As it can be seen by taking a look at equations (2.6) and (2.9), $\mu(k)$ and $x(k)$ are Borel functions of $\big\{\mu_0, \mathcal{V}_i(h),\ i = 0,1;\ h = 0,\dots,k-1\big\}$ and $\big\{x_0, \mu_0, \mathcal{V}_i(h), N(j),\ i = 0,1;\ h = 0,\dots,k-2;\ j = 0,\dots,k-1\big\}$ respectively, and moreover N is a stochastic sequence independent of $\mathcal{V}_i$, so that all the components of the eight terms in (3.20) are uncorrelated, and so (3.20) is null. The same can be repeated to prove that $E\big[\mathcal{B}_3(\xi(k), \eta(k))\big] = 0$.

To show that $\mathcal{F}(k)$ is a white noise sequence, note that it can be written as:

$$\mathcal{F}(k) = \mathcal{B}(k, \zeta(k), \eta(k)) + \widetilde{\mathcal{D}}(k)\eta(k) \tag{3.21}$$

where $\widetilde{\mathcal{D}}(k)$, the time-variant deterministic matrix such that $\widetilde{\mathcal{D}}(k)\eta(k) = \mathcal{B}(k, \bar{\xi}(k), \eta(k)) + \mathcal{D}(k)\eta(k)$, is given by:

$$\begin{bmatrix} \widetilde{\mathcal{D}}_{11}(k) & O_{n\times 1} & O_{n\times 1} & O_{n\times b} & O_{n\times b} \\ \widetilde{\mathcal{D}}_{21}(k) & \widetilde{\mathcal{D}}_{22}(k) & \widetilde{\mathcal{D}}_{23}(k) & \widetilde{\mathcal{D}}_{24}(k) & \widetilde{\mathcal{D}}_{25}(k) \\ O_{1\times b} & \widetilde{\mathcal{D}}_{32}(k) & \widetilde{\mathcal{D}}_{33}(k) & O_{1\times b} & O_{1\times b} \end{bmatrix} \tag{3.22}$$

with:

$$\begin{aligned} &\widetilde{\mathcal{D}}_{11}(k) = \big(1-\bar{\xi}_3(k)\big)F_0 + \bar{\xi}_3(k)F_1, && \widetilde{\mathcal{D}}_{24}(k) = \big(1-\bar{\xi}_3(k)\big)F_0, \\ &\widetilde{\mathcal{D}}_{21}(k) = \varepsilon_0\big(1-\bar{\xi}_3(k)\big)F_0 + (1-\varepsilon_1)\bar{\xi}_3(k)F_1, && \widetilde{\mathcal{D}}_{25}(k) = \bar{\xi}_3(k)F_1, \\ &\widetilde{\mathcal{D}}_{22}(k) = A_0\big(\bar{\xi}_1(k)-\bar{\xi}_2(k)\big) + B_0u(k) - B_0u(k)\bar{\xi}_3(k), && \widetilde{\mathcal{D}}_{32}(k) = 1-\bar{\xi}_3(k), \\ &\widetilde{\mathcal{D}}_{23}(k) = -A_1\bar{\xi}_2(k) - B_1u(k)\bar{\xi}_3(k), && \widetilde{\mathcal{D}}_{33}(k) = -\bar{\xi}_3(k). \end{aligned} \tag{3.22}$$

To show the whiteness of $\mathcal{F}(k)$ it has to be proved that the first sequence in (3.21) is white and, moreover, that it is uncorrelated with $\eta(k)$ at different instants as it has been proved in Theorem 3.1 that $\eta(k)$ is a white noise. Let $b_i(k) = \mathcal{B}_i(k, \zeta(k), \eta(k))$. b_1 is clearly a white noise sequence, as μ and N are independent sequences. b_2 is the sum of eight white noise sequences, uncorrelated at different instants. To show the effectiveness of the previous statement it is here reported the proof that the term $A_0\zeta_2(k)\chi_2(k)$, at time k, is uncorrelated with the term $F_0\chi_4(h)\zeta_3(h)$, at time $h < k$. Note that:

$$\begin{aligned} \zeta_2(k) &= \varphi_1\big(\mu_0, x_0, \mathcal{V}_i(j), N(j), \quad i = 0,1;, \quad j = 0,\dots k-1\big), \\ \zeta_3(h) &= \varphi_2\big(\mu_0, \mathcal{V}_i(j), \quad i = 0,1;, \quad j = 0,\dots h-1\big), \end{aligned} \tag{3.24}$$

where φ_1 and φ_2 are suitable Borel functions.

$$\begin{aligned} I\!\!E\big[\chi_2(k)\zeta_3(h)A_0\zeta_2(k)\chi_4(h)^T F_0^T\big] &= A_0 I\!\!E\big[\mathcal{V}_0(k)\mathcal{V}_0(h)\zeta_3(h)\zeta_2(k)N(h)^T\big]F_0^T \\ &= A_0 I\!\!E\big[\mathcal{V}_0(k)\big] I\!\!E\big[\mathcal{V}_0(h)\zeta_3(h)\zeta_2(k)N(h)^T\big]F_0^T = O_{n\times n} \end{aligned} \tag{3.25}$$

as $\mathcal{V}_0(k)$ is independent of the other components in (3.25). By using the same procedure it is easy to show that also b_3 is a white noise sequence and that b_1, b_2, b_3 and η are uncorrelated at different instants.

The measurement equation in (3.16) easily comes by substituting $\xi(k) = \zeta(k) + \bar{\xi}(k)$ in the output equation of (3.2). According to the state noise sequence (3.21), also the output noise sequence $\mathcal{G}(k)$ can be put in the form:

$$\mathcal{G}(k) = \mathcal{M}(\zeta(k), \eta(k)) + \widetilde{\mathcal{R}}(k)\eta(k), \tag{3.26}$$

where

$$\widetilde{\mathcal{R}}(k) = \big[(G_1 - G_0)\bar{\xi}_3(k) + G_0 \quad O_{q\times 2(b+1)}\big]. \tag{3.27}$$

By using the same procedures previously adopted, it is easy to prove that also $\mathcal{G}$ is a white noise sequence and, moreover, that $\mathcal{F}$ and $\mathcal{G}$ are uncorrelated sequences. ∎

Remark 3.3. Note that, according to Theorem 3.1:

$$\bar{\xi}(k+1) = \mathcal{A}(k)\bar{\xi}(k) + \mathcal{U}(k), \qquad \bar{\xi}(0) = \bar{\xi}_0 = \big(\bar{x}_0^T,\ (1-p_0)\bar{x}_0^T,\ 1-p_0\big)^T \tag{3.28}$$

Remark 3.4. Without a loss of generality, the bilinear term $\mathcal{B}\big(k, \zeta(k), \eta(k)\big)$ can be expressed in the form:

$$\mathcal{B}\big(k, \zeta(k), \eta(k)\big) = \sum_{i=1}^{m} \eta_i(k) B_i(k) \zeta(k), \tag{3.29}$$

where $\eta_i(k)$ stands for the i-th scalar component of $\eta(k)$ and $B_i(k)$ are time-variant matrices in $I\!R^{(2n+1)\times(2n+1)}$.

Definition 3.5. Consider the vector Y^k of all measurements y up to time k, defined as

$$Y^k = \begin{bmatrix} y(0) \\ \vdots \\ y(k) \end{bmatrix} \in I\!R^{q(k+1)}, \tag{3.30}$$

and let $L(Y^k)$ denote the Hilbert space of all the linear transformations of the extended measured output Y^k. Then the optimal linear estimate of system (2.1), driven by the Markovian jump defined in Lemma 2.2, is intended to be:

$$\hat{x}(k) = \begin{bmatrix} I_n & O_{n\times(n+1)} \end{bmatrix} \Big(\Pi\big[\zeta(k) | L(Y^k)\big] + \bar{\xi}(k) \Big). \tag{3.31}$$

where the first term of the sum is given by the projection of the zero-mean extended state $\zeta(k)$ onto $L(Y^c)$ and the second term is the mean value of $\xi(k)$.

Remark 3.6. Definition 3.5 considers only the case of filtering the state $x(k)$ of system (2.1). However, the methodology proposed is able to estimate also the Markovian jump, if the input suitably excites the system.

Theorem 3.7. *According to Definition* 3.5*, the optimal linear state estimate of system* (2.1) *is given by:*

$$\begin{aligned}
\hat{\zeta}(k+1) &= \hat{\zeta}(k+1|k) + K(k+1)\big[y(k+1) - \mathcal{C}(k+1)\hat{\zeta}(k+1|k) - \mathcal{Z}(k+1)\big], \\
\hat{\zeta}(k+1|k) &= \mathcal{A}(k)\hat{\zeta}(k), \qquad\qquad \hat{\zeta}(0|-1) = I\!E\big[\zeta_0\big] = 0 \\
\bar{\xi}(k+1) &= \mathcal{A}(k)\bar{\xi}(k) + \mathcal{U}(k), \qquad\qquad \bar{\xi}(0) = \xi_0, \\
\hat{x}(k) &= \begin{bmatrix} I_n & O_{n\times(n+1)} \end{bmatrix}\big(\hat{\zeta}(k) + \bar{\xi}(k)\big)
\end{aligned} \tag{3.32}$$

where $\hat{\zeta}(0|-1)$ is the minimum variance a priori prediction and the gain matrix $K(k)$ is given by the following Riccati equations:

$$\begin{aligned}
P_P(k+1) &= \mathcal{A}(k)P(k)\mathcal{A}(k)^T + \Psi_{\mathcal{F}}(k), \qquad\qquad k \geq 0 \\
K(k+1) &= P_P(k+1)\mathcal{C}(k+1)^T\big(\mathcal{C}(k+1)P_P(k+1)\mathcal{C}(k+1)^T + \Psi_{\mathcal{G}}(k+1)\big)^{-1} \\
P(k+1) &= \big(I_n - K(k+1)\mathcal{C}(k+1)\big)P_P(k+1) \\
P_P(0) &= \Psi_\zeta(0)
\end{aligned} \tag{3.33}$$

with $\Psi_{\mathcal{F}}(k)$ and $\Psi_{\mathcal{G}}(k)$ the covariance matrices of the noise sequences $\mathcal{F}$ and $\mathcal{G}$ given by:

$$\begin{aligned}
\Psi_\zeta(k+1) &= \mathcal{A}(k)\Psi_\zeta(k)\mathcal{A}(k)^T + \Psi_{\mathcal{F}}(k), \qquad\qquad \Psi_\zeta(0) = \mathrm{cov}(\xi_0) \\
\Psi_{\mathcal{F}}(k) &= \widetilde{\mathcal{D}}(k)\Psi_\eta\widetilde{\mathcal{D}}(k)^T + \sum_{i=1}^{m}\big(\Psi_\eta\big)_{ii}B_i(k)\Psi_\zeta(k)B_i(k)^T, \qquad \Psi_\eta = \mathrm{cov}(\eta(k)) \\
\Psi_{\mathcal{G}}(k) &= \widetilde{\mathcal{R}}(k)\Psi_\eta\widetilde{\mathcal{R}}(k)^T + (G_1 - G_0)(G_1 - G_0)^T\Psi_{\zeta_3}(k).
\end{aligned} \tag{3.34}$$

Proof. The filter and Riccati equations (3.32), (3.33) come out by applying the Kalman filter to system (3.16). According to (3.21) and (3.29), the extended state noise covariance matrix $\Psi_{\mathcal{F}}(k)$ is given by:

$$\begin{aligned}\Psi_{\mathcal{F}}(k) = &\sum_{i=1}^{m}\sum_{j=1}^{m} B_i(k)I\!E\big[\eta_i(k)\eta_j(k)\zeta(k)\zeta(k)^T\big]B_j(k)^T + \widetilde{\mathcal{D}}(k)I\!E\big[\eta(k)\eta(k)^T\big]\widetilde{\mathcal{D}}(k)^T \\ &+ \sum_{i=1}^{m} B_i(k)I\!E\big[\eta_i(k)\zeta(k)\eta(k)^T\big]\widetilde{\mathcal{D}}(k)^T + \sum_{i=1}^{m}\widetilde{\mathcal{D}}(k)I\!E\big[\eta_i(k)\eta(k)\zeta(k)^T\big]B_i(k)^T\end{aligned} \tag{3.35}$$

The last two terms in (3.35) are null, as $\eta(k)$ is uncorrelated to $\zeta(k)$ and, moreover, they are both zero-mean sequences. As far as $\Psi_\eta(k)$, it is easy to show that it is a diagonal matrix, given by:

$$\Psi_\eta(k) = \begin{bmatrix} I_b & \cdots & \cdots & \cdots & \cdots \\ \cdots & \varepsilon_0(1-\varepsilon_0) & \cdots & \cdots & \cdots \\ \cdots & \cdots & \varepsilon_1(1-\varepsilon_1) & \cdots & \cdots \\ \cdots & \cdots & \cdots & \varepsilon_0(1-\varepsilon_0)I_b & \cdots \\ \cdots & \cdots & \cdots & \cdots & \varepsilon_1(1-\varepsilon_1)I_b \end{bmatrix}. \tag{3.36}$$

∎

5. Simulations

In order to test the theory developed in this paper, numerical simulations have been produced. Without loss of generality, the system to be filtered has been supposed with no deterministic drift, that is equations (2.1) are considered with $B_{\mu(k)} = O_{n\times p}$, $\forall \mu(k) \in \{0,1\}$. Numerical data are the following:

$$\begin{aligned} A_0 &= \begin{bmatrix} 0.1 & 0 & 1 \\ 0 & 0.2 & 0.5 \\ -0.1 & 1 & 0.5 \end{bmatrix}, \quad C_0 = \begin{bmatrix} 0.5 & 1 & 0 \\ 0 & 1 & -2 \end{bmatrix}, \quad F_0 = \begin{bmatrix} 1 & 0 \\ -0.6 & 0 \\ 1 & 0 \end{bmatrix}, \quad G_0 = \begin{bmatrix} 0 & 1 \\ 0 & 0.5 \end{bmatrix} \\ A_1 &= \begin{bmatrix} -0.2 & 0 & 1 \\ 0 & 0.2 & 0.5 \\ -0.1 & 0.1 & -0.8 \end{bmatrix}, \quad C_1 = \begin{bmatrix} -0.5 & 1 & 0.2 \\ 0.5 & 1 & 2 \end{bmatrix}, \quad F_1 = \begin{bmatrix} 0.1 & 0 \\ -0.5 & 0 \\ -1.3 & 0 \end{bmatrix}, \quad G_1 = \begin{bmatrix} 0 & 1.2 \\ 0 & -0.5 \end{bmatrix} \end{aligned} \tag{3.37}$$

The Markovian jump transition matrix has been chosen as the following:

$$\Pi_\mu = \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix} \tag{3.38}$$

In order to show the efficiency of the algorithm, the noise sequence $N(k)$ is a zero-mean sequence, whose distribution is the following:

$$\begin{aligned} &P(N_1(k) = -2) = 0.2, &\quad &P\left(N_2(k) = -\frac{\sqrt{6}}{3}\right) = 0.6, \\ &P\left(N_1(k) = \frac{1}{2}\right) = 0.8, &\quad &P\left(N_2(k) = \frac{\sqrt{6}}{2}\right) = 0.4, \end{aligned} \tag{3.39}$$

The pictures below report the filtered state compared with the real one.

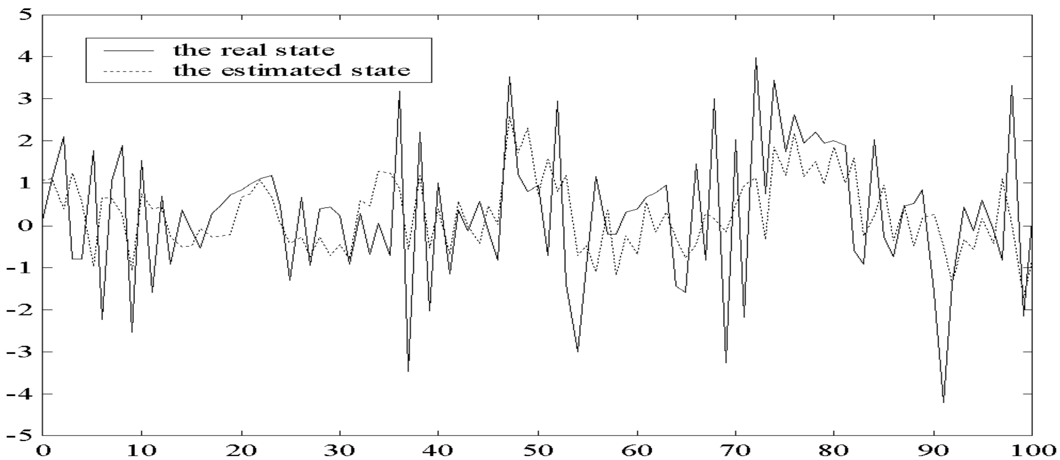

Fig. 4.1 – True and estimated state: the first component.

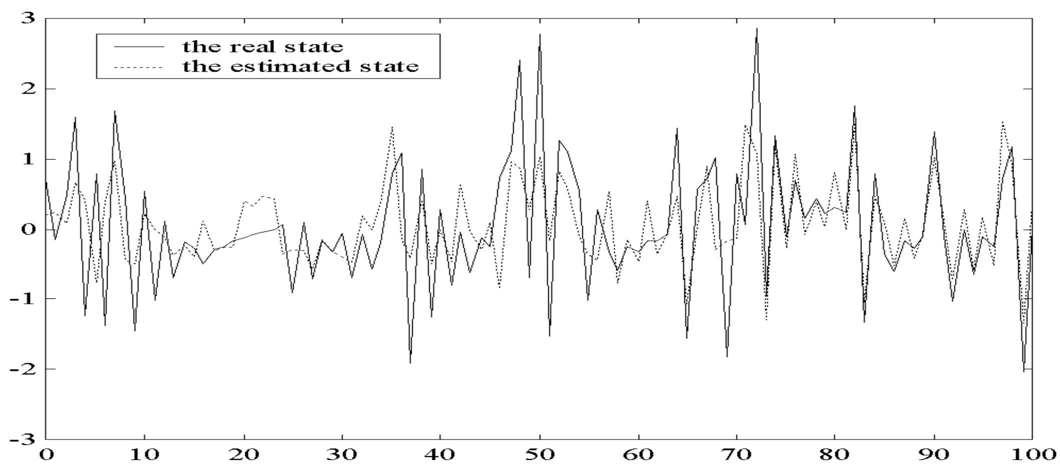

Fig. 4.2 – True and estimated state: the second component.

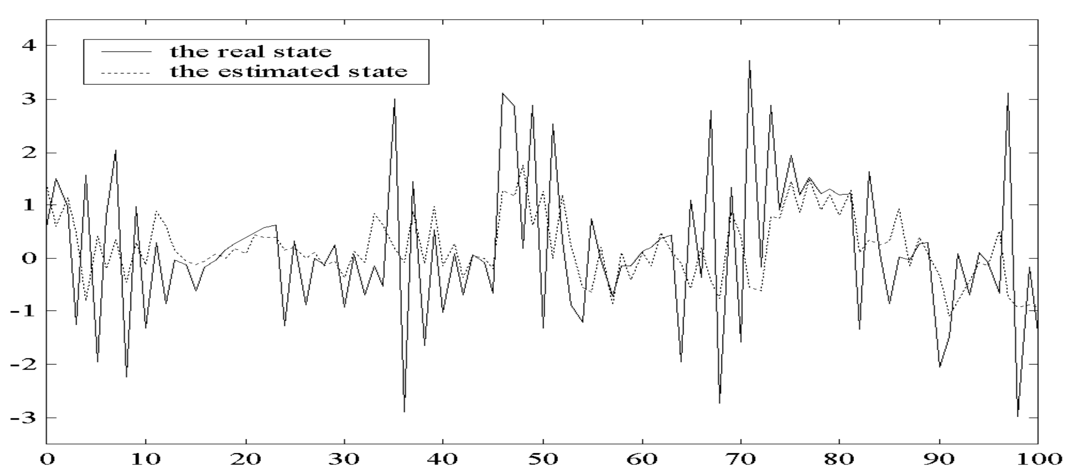

Fig. 4.3 – True and estimated state: the third component.

6. Conclusions

In this paper a new approach in filtering the state of a stochastic variable structure system, driven by a Markovian jump is proposed. By using a state space realization for the Markov process and the formalism of bilinear systems, the optimal linear filter, in the sense of minimum variance, achieves the best estimate among all the linear transformations of the measured output. The algorithm can be considered as a useful key tool to solve the problem of suboptimal estimates, when the system is affected by non-Gaussian noises, in the framework of polynomial filters. Numerical simulations show the goodness of the theoretical results.

References

[1] G. A. Ackerson, K. S. Fu, "On State Estimation in Switching Environments," *IEEE Trans. Automat. Contr.*, Vol. 15, No. 1, pp. 10–17, 1970.

[2] A. Balluchi, L. Benvenuti, M. D. Di Benedetto, C. Pinello, A. L. Sangiovanni-Vincentelli, "Automative Engine Control and Hybrid Systems: Challenge and Opportunities," *Proc. of IEEE*, Vol. 88, No. 7, pp. 888–912, 2000.

[3] H.A.P. Blom and Y. Bar-Shalom, "The interactive multiple model algorithm for systems with Markovian switching coefficients," *IEEE Trans. Automat. Contr.*, Vol. 33, No. 8, pp. 780–783, 1988.

[4] O. L. V. Costa, "Linear Minimum Mean Square Error Estimation for Discrete-Time Markovian Jump Linear Systems," *IEEE Trans. Automat. Contr.*, Vol. 39, No. 8, pp. 1685–1689, 1994.

[5] M. D. Di Benedetto, A. Sangiovanni-Vincentelli (Eds): Hybrid Systems: Computation and Control 2001, LNCS 2034, Springer-Verlag 2001.

[6] A. Doucet, A. Logothetis, V. Krishnamurthy, "Stochastic Sampling Algorithms for State Estimation of Jump Markov Linear Systems," *IEEE Trans. Automat. Contr.*, Vol. 45, No. 1, pp. 188–202, 2000.

[7] R. J. Elliot, F. Dufour, D. D. Sworder, "Exact Hybrid Filters in Discrete Time," *IEEE Trans. Automat. Contr.*, Vol. 41, No. 12, pp. 1807–1810, 1996.

[8] R.E. Helmick, W.D. Blair, S.A. Hoffman, "Fixed-interval smoothing for Markovian switching systems," *IEEE Trans. on Inf. Theory,* Vol. 41, No. 6, pp. 1845-1855, 1995.

[9] J. L. Hibey, C. D. Charalambous, "Conditional Densities for Contiuous-Time Nonlinear Hybrid Systems with Applications to Fault Detection," *IEEE Trans. Automat. Contr.*, Vol. 44, No. 11, pp. 2164–2169, 1999.

[10] K. H. Johansson, M. Egerstedt, J. Lygeros, S. Sastry, "On the regularization of Zeno hybrid automata. Hybrid control systems," *Systems Control Lett.*, Vol. 38, no. 3, 141–150, 1999.

[11] A. Logothetis, V. Krishnamurthy, "Expectation maximization algorithms for MAP estimation of Jump Markov Linear Systems," *IEEE Trans. n Signal Processing*, Vol. 47, No. 8, pp. 2139–2156, 1999.

[12] B. M. Miller, W. J. Runggaldier, "Kalman Filtering for Linear Systems with Coefficients Driven by a Hidden Markov Jump Process," *Syst. and Contr. Lett.*, No. 31, pp. 93–102, 1997.

[13] Q. Zhang, "Hybrid Filtering for Linear Systems with Non-Gaussian Disturbances," *IEEE Trans. Automat. Contr.*, Vol. 45, No. 1, pp. 50–61, 2000.