Extension of minimum variance estimation for systems with unknown inputs

Mohamed Darouach, Michel Zasadzinski, Mohamed Boutayeb

Extension of minimum variance estimation for systems with unknown inputs

M. Darouach,[*] **M. Zasadzinski**[a][†] **M. Boutayeb**[b][‡]

[a] *CRAN – CNRS*
IUT de Longwy, Université Henri Poincaré – Nancy I
186 rue de Lorraine, 54400 Cosnes et Romain, FRANCE
[b] *LSIIT – CNRS*
Université Louis Pasteur – Strasbourg I
Pôle API, boulevard Sébastien Brant, BP 10413, 67412 Illkirch Cedex, FRANCE

Abstract

In this paper, we address the problem of minimum variance estimation for discrete-time time-varying stochastic systems with unknown inputs. The objective is to construct an optimal filter in the general case where the unknown inputs affect both the stochastic model and the outputs. It extends the results of Darouach and Zasadzinski (1997) where the unknown inputs are only present in the model. The main difficulty in treating this problem lies in the fact that the estimation error is correlated with the systems noises, this fact leads generally to suboptimal filters. Necessary and sufficient conditions for the unbiasedness of this filter are established. Then conditions under which the estimation error and the system noises are uncorrelated are presented, and an optimal estimator and a predictor filters are derived. Sufficient conditions for the existence of these filters are given and sufficient conditions for their stability are obtained for the time-invariant case. A numerical example is given in order to illustrate the proposed method.

Keywords : Minimum variance filter; Unknown inputs; Stability; Convergence; Linear time-varying systems.

1 Introduction

State estimation for linear systems in the presence of unknown inputs has received considerable attention in the last two decades. In the deterministic context this problem has been extensively studied and the conditions for the stability of the obtained observers were developed (see e.g. (Hou and Müller, 1992; Hou and Müller, 1994; Darouach *et al.*, 1994)). On the other hand, stochastic models constitute a large class of models which can represent systems with noises, disturbances and uncertainties. For this set of systems, it is well known that the state estimation can be obtained from the Kalman filter (see e.g. (Anderson and Moore, 1979; Middleton and Goodwin, 1990; Kailath *et al.*, 2000).

However few results were obtained for stochastic systems with unknown inputs. The most common one is to treat the unknown input as a stochastic process with known wide-sense description (known mean and covariance for example, see the recent reference (Park *et al.*, 2000) and references therein). Generally, there may not be any knowledge concerning the model of these inputs. As shown by Kitanidis (1987), the problem is of great importance in geophysical and environmental applications. Other applications of the state estimation of systems with unknown inputs are in the fault detection and isolation (FDI) problems (Patton *et al.*, 1989; Basseville and Nikiforov, 1994).

A new method for constructing an innovation filter has been developed by Nikoukhah (1994) for stochastic systems with unknown inputs. This filter (or residual) generates from the observed quantities (known

---

[*]Corresponding author : Mohamed.Darouach@iut-longwy.uhp-nancy.fr

[†]Michel.Zasadzinski@iut-longwy.uhp-nancy.fr

[‡]Mohamed.Boutayeb@ipst-ulp.u-strasbg.fr

inputs and outputs) an output decoupled from the unknown inputs. This output is a zero-mean and white process in the absence of failure (normal behavior of the system) and a non-zero-mean when a failure has occurred. This work considers only the time-invariant continuous case, which can be easily extended to the discrete-time case. It uses a frequency domain approach and polynomial matrix factorizations. The optimal and unbiased state estimation problem is not considered. In our paper, minimum variance unbiased state estimators are considered and the solutions are given in the time domain.

This paper deals with optimal unbiased filtering for stochastic discrete time-varying systems with unknown inputs. We extend the results obtained by Kitanidis (1987), Darouach and Zasadzinski (1997), and Hsieh (2000) to the general case where the unknown inputs affect both the model and the outputs. Therefore in that case, the major problem to deal with is that the state estimation error is correlated with the model and the measurement noises. The unbiasedness of this error leads to the minimization of the variance of the estimation error under algebraic constraints. As in Darouach and Zasadzinski (1997), we parametrize all solutions of the algebraic constraints by two parameter matrices Z and α. The first one plays a role of a gain matrix and can be chosen such that the estimation error variance is minimum. The second parameter matrix is used to guarantee the rank conditions of some used matrices.

As shown in Anderson and Moore (1979), pages 296-300, in the general case where the estimation error is correlated with the model and measurement noises, only a suboptimal filter can be obtained. Necessary and sufficient conditions for the existence of this filter are given.

The conditions for the estimation error to be uncorrelated with the system noises are presented, these conditions lead to two filters. The first one is an estimator and has as inputs the known inputs of the systems at time instant t and the outputs at time instant $t + 1$. The second one is a predictor and has as inputs the known inputs of the systems and the outputs at time instant t. All the obtained results are expressed in terms of the initial system matrices.

In the time-invariant case, the conditions for the stability of the obtained filters are derived. When the measurements are free of unknown inputs, these conditions become necessary and sufficient and correspond to those given by Darouach and Zasadzinski (1997).

2   Optimal filter design

Let us consider the linear discrete-time stochastic time-varying system

$$x(t + 1) = \Phi(t)x(t) + \Gamma(t)u(t) + F(t)d(t) + w(t) \tag{1a}$$
$$y(t) = H(t)x(t) + G(t)d(t) + v(t) \tag{1b}$$

where $x(t) \in \mathbb{R}^n$ is the state, $u(t) \in \mathbb{R}^r$ is the known input, $d(t) \in \mathbb{R}^q$ is the unknown input, $y(t) \in \mathbb{R}^p$ is the output. $w(t)$ and $v(t)$ are uncorrelated white noises sequences of zero-mean and covariance matrices $Q(t) \geqslant 0$ and $R(t) \geqslant 0$ respectively. The initial state x_0 is of mean $\overline{x}_0$ and covariance P_0 and is independent of $w(t)$ and $v(t)$. We assume that $p \geqslant q$ and, without loss of generality, that rank $[\, H(t)\ G(t) \,] = p$.

The problem is to design an unbiased minimum variance linear estimator of $x(t)$ given measurements up to time instant t (with $t \in \mathbb{N}$) and the known input $u(t)$ without any information concerning the disturbance $d(t)$.

Let $\widehat{x}(t)$ denote the estimate of the state $x(t)$ at time instant t. The estimation error is defined as

$$e(t) = x(t) - \widehat{x}(t).$$

Definition 1. *An estimator $\widehat{x}(t)$ is unbiased if its expected value satisfies $\mathcal{E}\left\{\widehat{x}(t)\right\} = \mathcal{E}\left\{x(t)\right\}$, $\forall t \geqslant 0$, i.e.*

$$\mathcal{E}\left\{e(t)\right\} = 0,\ \forall t \geqslant 0 \tag{2}$$

where $\mathcal{E}\left\{\textbf{.}\right\}$ denotes the expectation operator.

The estimation error covariance $P(t)$ is defined by

$$P(t) = \mathcal{E}\left\{e(t)e^T(t)\right\}. \tag{3}$$

Consider the following filter

$$\xi(t + 1) = N(t)\xi(t) + J(t)y(t) + E(t)u(t) \tag{4a}$$
$$\widehat{x}(t) = \xi(t) + L(t)y(t) \tag{4b}$$

where the matrices $N(t) \in \mathbb{R}^{n \times n}$, $J(t) \in \mathbb{R}^{n \times p}$, $E(t) \in \mathbb{R}^{n \times r}$ and $L(t) \in \mathbb{R}^{n \times p}$ must be determined such that

$$\mathcal{J} = \text{tr} \left(P(t) \right) \tag{5}$$

is minimized under the constraint (2).

From equations (1b) and (4b), the estimation error can be written as

$$e(t) = \Psi(t)x(t) - \xi(t) - L(t)G(t)d(t) - L(t)v(t) \tag{6}$$

where $\Psi(t) = I - L(t)H(t)$. The dynamics of this error is given by

$$
\begin{aligned}
e(t+1) = & N(t)e(t) + (\Psi(t+1)\Phi(t) - N(t)\Psi(t) - J(t)H(t)) \, x(t) \\
& + (\Psi(t+1)\Gamma(t) - E(t)) \, u(t) + (\Psi(t+1)F(t) + N(t)L(t)G(t) - J(t)G(t)) \, d(t) \\
& + \Psi(t+1)w(t) - L(t+1)G(t+1)d(t+1) + (N(t)L(t) - J(t)) \, v(t) - L(t+1)v(t+1) \quad (7)
\end{aligned}
$$

where we have used (1), (4) and (6). By using (7), the constraint (2) yields to the following equations

$$0 = \Psi(t+1)\Phi(t) - N(t)\Psi(t) - J(t)H(t) \tag{8a}$$
$$0 = \Psi(t+1)F(t) + N(t)L(t)G(t) - J(t)G(t) \tag{8b}$$
$$0 = L(t+1)G(t+1) \tag{8c}$$
$$0 = \Psi(t+1)\Gamma(t) - E(t). \tag{8d}$$

Now from the definition of $\Psi(t+1)$, equation (8a) can be rewritten as

$$N(t) = \Psi(t+1)\Phi(t) - K(t)H(t) \tag{9}$$

where

$$K(t) = J(t) - N(t)L(t). \tag{10}$$

Define the following matrices $\mathcal{L}(t+1) = \begin{bmatrix} L(t+1) & K(t) \end{bmatrix}$, $\Sigma(t+1) = \begin{bmatrix} H(t+1)F(t) & G(t+1) \\ G(t) & 0 \end{bmatrix}$ and $\mathcal{F}(t) = \begin{bmatrix} F(t) & 0 \end{bmatrix}$ where the unknown matrices are $L(t+1)$ and $K(t)$.

With these notations, the constraints (8b) and (8c) can be written as

$$\mathcal{L}(t+1)\Sigma(t+1) = \mathcal{F}(t). \tag{11}$$

The unbiasedness of the estimate (4) is guaranteed when a solution to (11) exists, and we have the following lemma.

Lemma 1. *A necessary and sufficient condition for the estimate (4b) to be unbiased is*

$$\text{rank} \, \Sigma(t+1) = \text{rank} \begin{bmatrix} F(t) \\ G(t) \end{bmatrix} + \text{rank} \, G(t+1). \tag{12}$$

Proof. From Lancaster and Tismenetsky (1985), a necessary and sufficient condition for the existence of the solution to (11) is

$$\mathcal{F}(t) \left(\Sigma^{\dagger}(t+1)\Sigma(t+1) - I \right) = 0 \tag{13}$$

where $A^{\dagger}$ denotes any generalized inverse of matrix A. Condition (13) can be also written as

$$\text{rank} \begin{bmatrix} F(t) & 0 \\ H(t+1)F(t) & G(t+1) \\ G(t) & 0 \end{bmatrix} = \text{rank} \begin{bmatrix} H(t+1)F(t) & G(t+1) \\ G(t) & 0 \end{bmatrix}$$

which yields to

$$\text{rank} \, \Sigma(t+1) = \text{rank} \begin{bmatrix} H(t+1) & I & 0 \\ I & 0 & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} -F(t) & 0 \\ H(t+1)F(t) & G(t+1) \\ G(t) & 0 \end{bmatrix} = \text{rank} \begin{bmatrix} F(t) \\ G(t) \end{bmatrix} + \text{rank} \, G(t+1).$$

$\square$

Under (12), the general solution to (11) is given by

$$\mathcal{L}(t+1) = \overline{F}(t+1) + Z(t)\overline{G}(t+1) \tag{14}$$

where $\overline{F}(t+1) = \mathcal{F}(t)\Sigma^\dagger(t+1)$,
$\overline{G}(t+1) = \alpha(t)\left(I - \Sigma(t+1)\Sigma^\dagger(t+1)\right)$, and $Z(t)$ is an arbitrary matrix. The choice of the matrix parameter $\alpha(t)$ will be made in remark 1. Now, from the definitions of $\Psi(t)$, $\mathcal{L}(t+1)$ and using (14), (9) can be written as

$$N(t) = A(t) - Z(t)B(t) \tag{15}$$

where $A(t) = \Phi(t) - \overline{F}(t+1)\mathcal{H}(t+1)$, $B(t) = \overline{G}(t+1)\mathcal{H}(t+1)$ and $\mathcal{H}(t+1) = \begin{bmatrix} H(t+1)\Phi(t) \\ H(t) \end{bmatrix}$.

At this stage, filter (4) has been parametrized as a function of matrix parameter $Z(t)$. Thus the problem of designing an unbiased minimum variance filter (4) is reduced to the determination of the gain matrix $Z(t)$ such that (5) is minimized. From the above results and under assumption (12), the filter (4) can be written as

$$\widehat{x}(t+1) = (A(t) - Z(t)B(t))\,\widehat{x}(t) + E(t)u(t) + \left(\overline{F}(t+1) + Z(t)\overline{G}(t+1)\right)\overline{y}(t+1) \tag{16}$$

where $\overline{y}(t+1) = \begin{bmatrix} y(t+1) \\ y(t) \end{bmatrix}$ and the dynamic error (7) as

$$e(t+1) = (A(t) - Z(t)B(t))\,e(t) + \overline{H}(t+1)\eta(t+1) \tag{17}$$

where $\overline{H}(t+1) = \left[\begin{bmatrix} I - \mathcal{L}(t+1) \begin{bmatrix} H(t+1) \\ 0 \end{bmatrix} \end{bmatrix} \middle| -\mathcal{L}(t+1) \right]$ and $\eta(t+1) = \begin{bmatrix} w(t) \\ v(t+1) \\ v(t) \end{bmatrix}$

2.1 Suboptimal filter design

In this section, we propose a suboptimal solution to the design of the filter (4).

From equation (17) it is easy to see that $\eta(t+1)$ and $e(t)$ are correlated, and $\eta(t+1)$ is a colored noise of covariance

$$\Pi(t+1) = \begin{bmatrix} Q(t) & 0 & 0 \\ 0 & R(t+1) & 0 \\ 0 & 0 & R(t) \end{bmatrix}.$$

On the other hand, following Anderson and Moore (1979), pages 296-300, we have

$$\eta(t+2) = M\eta(t+1) + \overline{\eta}(t+1) \tag{18}$$

where $M = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & I & 0 \end{bmatrix}$ and $\overline{\eta}(t+1) = \begin{bmatrix} w(t+1) \\ v(t+2) \\ 0 \end{bmatrix}$. Here $\overline{\eta}(t+1)$ is a zero-mean white sequence of covariance

$\Xi(t+1) = \begin{bmatrix} Q(t+1) & 0 & 0 \\ 0 & R(t+2) & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

From (17) and (18), we obtain

$$\begin{bmatrix} e(t+1) \\ \eta(t+2) \end{bmatrix} = \begin{bmatrix} A(t) - Z(t)B(t) & \overline{H}(t+1) \\ 0 & M \end{bmatrix} \begin{bmatrix} e(t) \\ \eta(t+1) \end{bmatrix} + \begin{bmatrix} 0 \\ \overline{\eta}(t+1) \end{bmatrix}. \tag{19}$$

This is a linear system with a white noise excitation, and

$$\mathcal{E}\left\{ \begin{bmatrix} e(t+1) \\ \eta(t+2) \end{bmatrix} \begin{bmatrix} e(t+1)^T & \eta(t+2)^T \end{bmatrix} \right\} = \begin{bmatrix} P(t+1) & \Delta(t+1) \\ \Delta^T(t+1) & \Pi(t+2) \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 \\ 0 & \Xi(t+1) \end{bmatrix} + \begin{bmatrix} A(t) - Z(t)B(t) & \overline{H}(t+1) \\ 0 & M \end{bmatrix} \begin{bmatrix} P(t) & \Delta(t) \\ \Delta^T(t) & \Pi(t+1) \end{bmatrix} \begin{bmatrix} A(t) - Z(t)B(t) & \overline{H}(t+1) \\ 0 & M \end{bmatrix}^T. \tag{20}$$

The determination of $Z(t)$ can be obtained by the minimization of (5).

We notice (see Anderson and Moore (1979), page 300) that only a suboptimal solution can be obtained. It is the presence of non zero $\Delta(t+1)$ which causes this suboptimality. This suboptimal solution which minimizes the trace of $P(t+1)$ can be obtained along the line of the work of Anderson and Moore (1979), pages 296-300.

2.2 Optimal filter design

This section is devoted to the optimal filter design. Let us come back to equation (17), by using the definition of $\overline{H}(t)$ and $\mathcal{L}(t)$, one can see that a necessary and sufficient condition for $N(t)e(t)$ and $\overline{H}(t+1)\eta(t+1)$ to be uncorrelated is

$$\mathcal{E}\left\{N(t)e(t)\eta^T(t+1)\overline{H}^T(t+1)\right\} = N(t)\overline{H}(t)\begin{bmatrix}0 & 0 & 0 \\ 0 & 0 & R(t) \\ 0 & 0 & 0\end{bmatrix}\overline{H}^T(t+1) = N(t)L(t)R(t)K^T(t) = 0. \quad (21)$$

Sufficient conditions for this equation to be satisfied for all $N(t)$ and $R(t)$ are

$$K(t) = 0 \text{ or } L(t) = 0.$$

Then the decorrelation constraint (21) leads to two filters. In the case where $K(t) = 0$, the filter (4) is an estimator one since the estimate $\widehat{x}(t+1)$ uses the known inputs of the system u at time instant t and the outputs y at time instant $t+1$ (see section 2.2.1). In the case where $L(t) = 0$, the filter (4) is a predictor one since the estimate $\widehat{x}(t+1)$ is a function of the known inputs u and the outputs y at time instant t (see section 2.2.2).

2.2.1 Optimal estimator filter design

This section considers the case where $K(t) = 0$, we show that the optimal filter is an estimator one.
If $K(t) = 0$, then from equations (9), (10) and (11), it is easy to see that the filter (4) can be written as the following estimator

$$\widehat{x}(t+1) = N(t)\widehat{x}(t) + E(t)u(t) + L(t+1)y(t+1) \quad (22)$$

with

$$N(t) = \Psi(t+1)\Phi(t) \text{ and } J(t) = N(t)L(t) \quad (23a)$$
$$\mathcal{F}(t) = L(t+1)\overline{\Sigma}(t+1) \quad (23b)$$

where $\overline{\Sigma}(t+1) = \begin{bmatrix}H(t+1)F(t) & G(t+1)\end{bmatrix}$. Then the condition of lemma 1 becomes.

Lemma 2. *A sufficient condition for the unbiasedness of the estimator (22) is*

$$\operatorname{rank}\overline{\Sigma}(t+1) = \operatorname{rank}F(t) + \operatorname{rank}G(t+1). \quad (24)$$

Proof. The proof is similar to that of lemma 1 and is omitted. $\qquad\square$
Under condition $K(t) = 0$, equation (17) becomes

$$e(t+1) = N(t)e(t) + \begin{bmatrix}(I - L(t+1)H(t+1)) & -L(t+1)\end{bmatrix}\begin{bmatrix}w(t) \\ v(t+1)\end{bmatrix}. \quad (25)$$

The problem is reduced to determine matrix $L(t)$ such that the criterion (5) is minimum. Assume that (24) is satisfied, then as in (14) the solution of (23b) is given by

$$L(t+1) = \widehat{F}(t+1) + \overline{Z}(t)\widehat{G}(t+1) \quad (26)$$

where $\overline{Z}(t)$ is an arbitrary matrix, $\widehat{F}(t+1) = \mathcal{F}(t)\overline{\Sigma}^\dagger(t+1)$ and $\widehat{G}(t+1) = \overline{\alpha}(t)(I - \overline{\Sigma}(t+1)\overline{\Sigma}^\dagger(t+1))$. The choice of the matrix parameter $\overline{\alpha}(t)$ is given in remark 1.
From (26), the matrix $N(t)$ in (23a) can be written as

$$N(t) = \overline{A}(t) - \overline{Z}(t)\overline{B}(t) \quad (27)$$

where $\overline{A}(t) = \Phi(t) - \widehat{F}(t+1)H(t+1)\Phi(t)$ and $\overline{B}(t) = \widehat{G}(t+1)H(t+1)\Phi(t)$. Then filter (22) can be written as

$$\widehat{x}(t+1) = \overline{A}(t)\widehat{x}(t) + E(t)u(t) + \widehat{F}(t+1)y(t+1) + \overline{Z}(t)\left(\widehat{G}(t+1)y(t+1) - \overline{B}(t)\widehat{x}(t)\right) \quad (28)$$

and equation (25) becomes

$$e(t+1) = \left(\overline{A}(t) - \overline{Z}(t)\overline{B}(t)\right)e(t) + \left(I - \widehat{F}(t+1)H(t+1) - \overline{Z}(t)\widehat{G}(t+1)H(t+1)\right)w(t)$$
$$- \left(\widehat{F}(t+1) + \overline{Z}(t)\widehat{G}(t+1)\right)v(t+1). \quad (29)$$

Then the propagation of the error covariance is given by

$$P(t+1) = \left(\overline{A}(t) - \overline{Z}(t)\overline{B}(t)\right) P(t) \left(\overline{A}(t) - \overline{Z}(t)\overline{B}(t)\right)^T$$
$$+ \overline{Q}(t) - \overline{Z}(t)S^T(t) - S(t)\overline{Z}^T(t) + \overline{Z}(t)\theta(t)\overline{Z}^T(t) \quad (30)$$

where

$$\overline{Q}(t) = \left(I - \widehat{F}(t+1)H(t+1)\right)Q(t)\left(I - \widehat{F}(t+1)H(t+1)\right)^T + \widehat{F}(t+1)R(t+1)\widehat{F}^T(t+1) \quad (31a)$$

$$S(t) = \left(I - \widehat{F}(t+1)H(t+1)\right)Q(t)H^T(t+1)\widehat{G}^T(t+1) - \widehat{F}(t+1)R(t+1)\widehat{G}^T(t+1) \quad (31b)$$

$$\theta(t) = \widehat{G}(t+1)(R(t+1) + H(t+1)Q(t)H^T(t+1))\widehat{G}^T(t+1) \quad (31c)$$

The minimum variance estimation problem in the presence of the unknown input $d(t)$ is reduced to find the matrix $\overline{Z}(t)$ which minimizes the trace of $P(t)$. The result is given by (see for example (Kailath *et al.*, 2000))

$$\overline{Z}(t) = \left(\overline{A}(t)P(t)\overline{B}^T(t) + S(t)\right)\left(\overline{B}(t)P(t)\overline{B}^T(t) + \theta(t)\right)^{-1}. \quad (32)$$

The condition for the existence of $\overline{Z}(t)$ is that $(\overline{B}(t)P(t)\overline{B}^T(t) + \theta(t)) > 0$.

Assumption 1. *In the sequel, we assume that $\theta(t) > 0$.*

Remark 1. From (31c), a necessary condition for $\theta(t)$ to be nonsingular is that $\widehat{G}(t+1)$ must be of full row rank, therefore the matrix parameter $\overline{\alpha}(t)$ must be chosen such that this constraint is satisfied. Another constraint on the choice of $\overline{\alpha}(t)$ is that the matrix $\begin{bmatrix} \overline{\Sigma}(t+1)\overline{\Sigma}^\dagger(t+1) \\ \widehat{G}(t+1) \end{bmatrix}$ is of full column rank. This condition is required for the main results of the paper (see the proof of lemma 4). One choice of this matrix parameter is the following. Let

$$\overline{\Sigma}(t+1) = U^T(t+1) \begin{bmatrix} \Upsilon(t+1) & 0 \\ 0 & 0 \end{bmatrix} V(t+1)$$

be the singular value decomposition of matrix $\overline{\Sigma}(t+1)$, where $\Upsilon(t+1) \in \mathbb{R}^{r_{\overline{\Sigma}} \times r_{\overline{\Sigma}}}$ is a nonsingular matrix and $r_{\overline{\Sigma}}$ is the rank of matrix $\overline{\Sigma}(t+1)$, then

$$\begin{bmatrix} \overline{\Sigma}(t+1)\overline{\Sigma}^\dagger(t+1) \\ \widehat{G}(t+1) \end{bmatrix} = \begin{bmatrix} U^T(t+1)\begin{bmatrix} I_{r_{\overline{\Sigma}}} & 0 \\ 0 & 0 \end{bmatrix}U(t+1) \\ \overline{\alpha}(t)U^T(t+1)\begin{bmatrix} 0 & 0 \\ 0 & I_{p-r_{\overline{\Sigma}}} \end{bmatrix}U(t+1) \end{bmatrix}.$$

We can see that one choice of $\overline{\alpha}(t)$ is

$$\overline{\alpha}(t) = \begin{bmatrix} 0 & I_{p-r_{\overline{\Sigma}}} \end{bmatrix} U(t+1). \quad (33)$$

In this case $\widehat{G}(t+1) = \overline{\alpha}(t)$ and

$$\begin{bmatrix} \overline{\Sigma}(t+1)\overline{\Sigma}^\dagger(t+1) \\ \widehat{G}(t+1) \end{bmatrix} = \begin{bmatrix} U^T(t+1) & 0 \\ 0 & I_{p-r_{\overline{\Sigma}}} \end{bmatrix} \begin{bmatrix} I_{r_{\overline{\Sigma}}} & 0 \\ 0 & 0 \\ 0 & I_{p-r_{\overline{\Sigma}}} \end{bmatrix} U(t+1)$$

□

Remark 2. If matrix $\overline{\Sigma}(t+1)$ is of full row rank, then $\overline{\Sigma}(t+1)\overline{\Sigma}^\dagger(t+1) = I_p$ (and $r_{\overline{\Sigma}} = p$), and we have $\widehat{G}(t+1) = 0$, $\overline{B}(t) = 0$, $S(t) = 0$ and $\theta(t) = 0$. In this case, the optimal estimator filter (28) becomes

$$\widehat{x}(t+1) = \overline{A}(t)\widehat{x}(t) + E(t)u(t) + \widehat{F}(t+1)y(t+1)$$

and the propagation of the error covariance is given by the following discrete-time Lyapunov equation

$$P(t+1) = \overline{A}(t)P(t)\overline{A}^T(t) + \widehat{F}(t+1)R(t+1)\widehat{F}^T(t+1)$$
$$+ \left(I - \widehat{F}(t+1)H(t+1)\right)Q(t)\left(I - \widehat{F}(t+1)H(t+1)\right)^T.$$

□

2.2.2 Optimal predictor filter design

This section considers the case where $L(t) = 0$, we show that the optimal filter is a predictor one. If $L(t) = 0$, then from equations (9), (10) and (11), it is easy to see that the filter (4) can be written as the following predictor

$$\widehat{x}(t+1) = N(t)\widehat{x}(t) + E(t)u(t) + K(t)y(t) \tag{34}$$

with

$$N(t) = \Psi(t+1)\Phi(t) - K(t)H(t) \tag{35a}$$
$$K(t)G(t) = F(t). \tag{35b}$$

The matrix $K(t)$ exists if and only if $\text{Im}(F^T(t)) \subset \text{Im}(G^T(t))$ and lemma 1 becomes.

Lemma 3. *A sufficient condition for the unbiasedness of the predictor (34) is*

$$\text{rank} \begin{bmatrix} F(t) \\ G(t) \end{bmatrix} = \text{rank}\, G(t). \tag{36}$$

Under condition (36), as in section 2.2.1, the solution of (35b) is parametrized by a matrix $Y(t)$ as follows

$$K(t) = J(t) = \widetilde{F}(t) + Y(t)\widetilde{G}(t) \tag{37}$$

where $\widetilde{F}(t) = F(t)G^\dagger(t)$, $\widetilde{G}(t) = \beta(t)\left(I - G(t)G^\dagger(t)\right)$ and $Y(t)$ is an arbitrary matrix. The choice of the matrix parameter $\beta(t)$ is similar to that of $\overline{\alpha}(t)$ (see remark 1). From relations (9) and (37), we have

$$N(t) = \widetilde{A}(t) - Y(t)\widetilde{B}(t) \tag{38}$$

where $\widetilde{A}(t) = \Phi(t) - \widetilde{F}(t)H(t)$ and $\widetilde{B}(t) = \widetilde{G}(t)H(t)$. Then filter (34) becomes

$$\widehat{x}(t+1) = \widetilde{A}(t)\widehat{x}(t) + \Gamma(t)u(t) + \widetilde{F}(t)y(t) + Y(t)\left(\widetilde{G}(t)y(t) - \widetilde{B}(t)\widehat{x}(t)\right) \tag{39}$$

and equation (7) can be rewritten as

$$e(t+1) = \left(\widetilde{A}(t) - Y(t)\widetilde{B}(t)\right)e(t) + w(t) - \left(\widetilde{F}(t) + Y(t)\widetilde{G}(t)\right)v(t). \tag{40}$$

Then the propagation of the error covariance is given by

$$P(t+1) = \left(\widetilde{A}(t) - Y(t)\widetilde{B}(t)\right) P(t) \left(\widetilde{A}(t) - Y(t)\widetilde{B}(t)\right)^T$$
$$+ \widetilde{Q}(t) - Y(t)\widetilde{S}^T(t) - \widetilde{S}(t)Y^T(t) + Y(t)\widetilde{\theta}(t)Y(t) \tag{41}$$

where $\widetilde{Q}(t) = Q(t) + \widetilde{F}(t)R(t)\widetilde{F}^T(t)$, $\widetilde{S}(t) = -\widetilde{F}(t)R(t)\widetilde{G}^T(t)$ and $\widetilde{\theta}(t) = \widetilde{G}(t)R(t)\widetilde{G}^T(t)$.

The minimization of $\text{tr}\,(P(t+1))$ yields the following gain matrix $Y(t)$

$$Y(t) = \left(\widetilde{A}(t)P(t)\widetilde{B}^T(t) + \widetilde{S}(t)\right)\left(\widetilde{B}(t)P(t)\widetilde{B}^T(t) + \widetilde{\theta}(t)\right)^{-1}. \tag{42}$$

Assumption 2. *In the sequel, we assume that $\widetilde{\theta}(t) > 0$.*

Remark 3. As in remark 2, if matrix $G(t)$ is of full row rank, then $G(t)G^\dagger(t) = I$ and we have $\widetilde{G}(t) = 0$, $\widetilde{B}(t) = 0$, $\widetilde{S}(t) = 0$ and $\widetilde{\theta}(t) = 0$. In this case, the optimal predictor filter (39) becomes

$$\widehat{x}(t+1) = \widetilde{A}(t)\widehat{x}(t) + \Gamma(t)u(t) + \widetilde{F}(t+1)y(t+1)$$

and the propagation of the error covariance is given by the following discrete-time Lyapunov equation

$$P(t+1) = \widetilde{A}(t)P(t)\widetilde{A}^T(t) + \widetilde{F}(t+1)R(t+1)\widetilde{F}^T(t+1) + Q(t).$$

$\square$

Remark 4.

(i) For $F(t) = 0$ and $G(t) = 0$, conditions (24) and (36) are verified. Then the obtained filter (22) is the Kalman estimator one, while the obtained filter (34) is the Kalman predictor one (Anderson and Moore, 1979). In this case we obtain
 - $\widehat{F}(t) = 0$ and $\widehat{G}(t+1) = \overline{\alpha}(t)$ with $\det(\overline{\alpha}(t)) \neq 0$,
 - $\widetilde{F}(t) = 0$ and $\widetilde{G}(t) = \beta(t)$ with $\det(\beta(t)) \neq 0$.

(ii) Condition (24) generalizes the one obtained by (Kitanidis, 1987; Darouach and Zasadzinski, 1997) in the case where $G(t) = 0$. In this case, these conditions become necessary and sufficient.

(iii) It is easy to see that the obtained estimator and predictor filters are independent from the matrix parameters $\overline{\alpha}(t)$ and $\beta(t)$.

$\square$

3 Time-invariant case

In this section, we study the asymptotic properties of the optimal filters of section 2. We consider the following time-invariant system

$$x(t+1) = \Phi x(t) + \Gamma u(t) + Fd(t) + w(t) \tag{43a}$$
$$y(t) = Hx(t) + Gd(t) + v(t) \tag{43b}$$

where $w(t)$ and $v(t)$ are uncorrelated white noises sequences of zero-mean and covariance matrices $Q \geqslant 0$ and $R \geqslant 0$ respectively. The initial state x_0 is of mean $\overline{x}_0$ and covariance P_0, and is independent of $w(t)$ and $v(t)$.

As in section 2.2, we shall consider the optimal estimator filter and the extension to the optimal predictor filter is treated in remark 7.

Assume that (24) is satisfied, i.e. that the following assumption holds.

Assumption 3. *We assume that $\theta > 0$ and, with $\overline{\Sigma} = [HF \quad G]$,*

$$\operatorname{rank} \overline{\Sigma} = \operatorname{rank} F + \operatorname{rank} G \tag{44}$$

In this case, from the results of section 2.2.1, the minimum variance estimator filter becomes

$$\widehat{x}(t+1) = \left(\overline{A} - \overline{Z}(t)\overline{B}\right)\widehat{x}(t) + \overline{E}(t)u(t) + \left(\widehat{F} + \overline{Z}(t)\widehat{G}\right)y(t+1) \tag{45}$$

with $\overline{A} = \Phi - \widehat{F}H\Phi$, $\overline{B} = \widehat{G}H\Phi$, $\widehat{G} = \overline{\alpha}(I - \overline{\Sigma}\,\overline{\Sigma}^\dagger)$, $\overline{E}(t) = (I - (\widehat{F} + \overline{Z}(t)\widehat{G})H)\Gamma$, $\widehat{F} = \mathcal{F}\overline{\Sigma}^\dagger$, $\mathcal{F} = [F \quad 0]$, $\overline{Z}(t) = (\overline{A}P(t)\overline{B}^T + S)(\overline{B}P(t)\overline{B}^T + \theta)^{-1}$, $\overline{Q} = (I - \widehat{F}H)Q(I - \widehat{F}H)^T + \widehat{F}R\widehat{F}^T$, $\theta = \widehat{G}\left(R + HQH^T\right)\widehat{G}^T$, $S = (I - \widehat{F}H)QH^T\widehat{G}^T - \widehat{F}R\widehat{G}^T$ and

$$P(t+1) = \left(\overline{A} - \overline{Z}(t)\overline{B}\right)P(t)\left(\overline{A} - \overline{Z}(t)\overline{B}\right)^T + \overline{Q} - \overline{Z}(t)S^T - S\overline{Z}^T(t) + \overline{Z}(t)\theta\overline{Z}^T(t) \tag{46}$$

where the matrix parameter $\overline{\alpha}$ is chosen as in remark 1.

Now define the following matrices $\overline{A}_s = \overline{A} - S\theta^{-1}\overline{B}$, $\overline{Q}_s = \overline{Q} - S\theta^{-1}S^T$ and $\mathcal{G} = [0 \quad G]$. Before giving the asymptotic properties of the filter (45), we need to prove the following lemmas.

Lemma 4. *Under assumption 3, the pair $(\overline{A}_s, \overline{B})$ is detectable if and only if*

$$\operatorname{rank}\begin{bmatrix} zI - \Phi & -\mathcal{F} \\ H & \mathcal{G} \end{bmatrix} = n + \operatorname{rank}\begin{bmatrix} \mathcal{F} \\ \mathcal{G} \end{bmatrix} = n + \operatorname{rank} F + \operatorname{rank} G, \ \forall z \in \mathbb{C}, |z| \geqslant 1. \tag{47}$$

Proof. First, it is easy to see that the pair $(\overline{A}_s, \overline{B})$ is detectable if and only if the pair $(\overline{A}, \overline{B})$ is detectable. On the other hand, since matrix $\begin{bmatrix} \overline{\Sigma}\,\overline{\Sigma}^\dagger \\ \widehat{G} \end{bmatrix}$ is of full column rank, we have,

$$\operatorname{rank}\begin{bmatrix} zI - \Phi & -\mathcal{F} \\ H & \mathcal{G} \end{bmatrix} = \operatorname{rank}\begin{bmatrix} I & 0 \\ z^{-1}H & z^{-1}I \end{bmatrix}\begin{bmatrix} zI - \Phi & -F & 0 \\ H\Phi & HF & G \end{bmatrix}\begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & zI \end{bmatrix}$$

$$= \operatorname{rank}\begin{bmatrix} I & \widehat{F} \\ 0 & \widehat{G} \\ 0 & \overline{\Sigma}\,\overline{\Sigma}^\dagger \end{bmatrix}\begin{bmatrix} zI - \Phi & -F & 0 \\ H\Phi & HF & G \end{bmatrix}\begin{bmatrix} I & 0 \\ -\overline{\Sigma}^\dagger H\Phi & I \end{bmatrix}$$

$$= \operatorname{rank}\begin{bmatrix} zI - \overline{A} \\ \overline{B} \end{bmatrix} + \operatorname{rank} \overline{\Sigma}, \ \forall z \in \mathbb{C}, |z| \geqslant 1.$$

Using lemma 2, lemma 4 is proved. $\square$

Lemma 5. *Under assumption 3, the existence of an unreacheable mode λ of $(\overline{A}_s, \overline{Q}_s^{1/2})$ on the stability boundary is equivalent to the existence of a row vector $\mu \neq 0$ such that*

$$\mu\begin{bmatrix} \Phi - \lambda I & -F & Q^{1/2} & 0 \\ -\lambda H & G & 0 & R^{1/2} \end{bmatrix} = 0. \tag{48}$$

Proof. Let λ be an unreachable mode of $(\overline{A}_s,\ \overline{Q}_s^{1/2})$, then there exists a row vector $\varpi \neq 0$ such that

$$\varpi\overline{A}_s = \lambda\varpi \tag{49a}$$
$$\varpi\overline{Q}_s^{1/2} = 0 \tag{49b}$$

where $\overline{Q}_s^{1/2} = \begin{bmatrix} I & -S\theta^{-1} \end{bmatrix} \begin{bmatrix} I - \widehat{F}H & -\widehat{F} \\ \widehat{G}H & \widehat{G} \end{bmatrix} \begin{bmatrix} Q^{1/2} & 0 \\ 0 & R^{1/2} \end{bmatrix}$.

From (44) in assumption 3 and by using (13) and the definitions of $\overline{A}_s$, $\widehat{F}$, $\widehat{G}$, equations (49a) and (49b) can be written as

$$\mu \begin{bmatrix} \Phi - \lambda I & -F & Q^{1/2} & 0 \\ -\lambda H & G & 0 & R^{1/2} \end{bmatrix} = 0$$

where $\mu = \varpi \begin{bmatrix} I & -S\theta^{-1} \end{bmatrix} \begin{bmatrix} I - \widehat{F}H & -\widehat{F} \\ \widehat{G}H & \widehat{G} \end{bmatrix}$, which is the desired result. $\qquad\square$

The above Riccati difference equation (RDE) (46) can be written as

$$P(t+1) = \left(\overline{A} - \widehat{Z}(t)\overline{B}\right)P(t)\left(\overline{A} - \widehat{Z}(t)\overline{B}\right)^T + \widehat{Z}(t)\theta\widehat{Z}^T(t) + \overline{Q}_s \tag{50}$$

where $\widehat{Z}(t) = \overline{Z}(t) - S\theta^{-1}$, and its associated algebraic Riccati equation (ARE) is

$$P = \left(\overline{A} - \widehat{Z}\,\overline{B}\right)P\left(\overline{A} - \widehat{Z}\,\overline{B}\right)^T + \widehat{Z}\theta\widehat{Z}^T + \overline{Q}_s \tag{51}$$

where $\widehat{Z} = \overline{Z} - S\theta^{-1}$.

Then, from the above results and from the well-known results on the stability and convergence of the standard Kalman filter (Anderson and Moore, 1979; de Souza *et al.*, 1986), we can give the convergence and the stability for the unknown input estimator filter (45) for the time-invariant system (43).

Theorem 1. *Under assumption 3, with $P_0 > 0$, the sequence $P(t)$ generated by (50) converges exponentially to the stabilizing solution P of the algebraic Riccati equation (51) if and only if*

$$\text{rank}\begin{bmatrix} zI - \Phi & -F & 0 \\ H & 0 & G \end{bmatrix} = n + \text{rank}\, F + \text{rank}\, G,\ \forall z \in \mathbb{C},\ |z| \geqslant 1, \tag{52}$$

$$\text{rank}\begin{bmatrix} \Phi - e^{j\omega}I & -F & Q^{1/2} & 0 \\ -e^{j\omega}H & G & 0 & R^{1/2} \end{bmatrix} = n + p,\ \forall\omega \in [0,\ 2\pi]. \tag{53}$$

Remark 5. Conditions (52) and (53) generalize those given by Darouach and Zasadzinski (1997). Indeed, for $G = 0$, condition (52) becomes

$$\text{rank}\begin{bmatrix} zI - \Phi & -F \\ H & 0 \end{bmatrix} = n + \text{rank}\, F,\ \forall z \in \mathbb{C},\ |z| \geqslant 1,$$

and condition (53) can be written as

$$\text{rank}\begin{bmatrix} \Phi - e^{j\omega}I & F & Q^{1/2} & 0 \\ e^{j\omega}H & 0 & 0 & R^{1/2} \end{bmatrix} = n + p,\ \forall\omega \in [0,\ 2\pi].$$

$\qquad\square$

Remark 6. As shown in remark 2, if matrix $\overline{\Sigma}$ is of full row rank, then $\overline{\Sigma}\,\overline{\Sigma}^{\dagger} = I$ and we have $\widehat{G} = 0$, $\overline{B} = 0$, $S = 0$ and $\theta = 0$. In this case, the optimal estimator filter (45) becomes

$$\widehat{x}(t+1) = \overline{A}\widehat{x}(t) + Eu(t) + \widehat{F}y(t+1)$$

and the propagation of the error covariance is given by the following discrete-time Lyapunov equation

$$P(t+1) = \overline{A}P(t)\overline{A}^T + \widehat{F}R\widehat{F}^T + \left(I - \widehat{F}H\right)Q\left(I - \widehat{F}H\right)^T.$$

From lemma 4, $\overline{A}$ is a stability matrix. $\qquad\square$

Remark 7 (Optimal predictor). Assume that $\widetilde{\theta} > 0$ and rank $\left[\begin{smallmatrix} F \\ G \end{smallmatrix}\right] = \text{rank } G$. Then, from the results of section 2.2.2, the optimal predictor filter for system (43) is given by

$$\widehat{x}(t+1) = \widetilde{A}\widehat{x}(t) + \Gamma u(t) + \widetilde{F}y(t) + Y(t)\left(\widetilde{G}y(t) - \widetilde{B}\widehat{x}(t)\right) \tag{54}$$

with $\widetilde{A} = \Phi - \widetilde{F}H$, $\widetilde{B} = \widetilde{G}H$, $\widetilde{F} = FG^\dagger$, $\widetilde{G} = \beta(I - GG^\dagger)$, $Y(t) = (\widetilde{A}P(t)\widetilde{B}^T + \widetilde{S})(\widetilde{B}P(t)\widetilde{B}^T + \widetilde{\theta})^{-1}$, $\widetilde{Q} = Q + \widetilde{F}R\widetilde{F}^T$, $\widetilde{S} = -\widetilde{F}R\widetilde{G}^T$, $\widetilde{\theta} = \widetilde{G}R\widetilde{G}^T$,

$$P(t+1) = \left(\widetilde{A} - Y(t)\widetilde{B}\right)P(t)\left(\widetilde{A} - Y(t)\widetilde{B}\right)^T + \widetilde{Q} - Y(t)\widetilde{S}^T - \widetilde{S}Y^T(t) + Y(t)\widetilde{\theta}Y^T(t) \tag{55}$$

where the matrix parameter β is chosen as in remark 1.
Define the algebraic Riccati equation (ARE)

$$P = \left(\widetilde{A} - Y\widetilde{B}\right)P\left(\widetilde{A} - Y\widetilde{B}\right)^T + \widetilde{Q} - Y\widetilde{S}^T - \widetilde{S}Y^T + Y\widetilde{\theta}Y^T \tag{56}$$

associated with the Riccati difference equation (55), where $Y = (\widetilde{A}P\widetilde{B}^T + \widetilde{S})(\widetilde{B}P\widetilde{B}^T + \widetilde{\theta})^{-1}$.
The stability and the convergence of the optimal predictor filter are guaranteed if and only if

$$\text{rank}\begin{bmatrix} zI - \Phi & -F \\ H & G \end{bmatrix} = n + \text{rank}\begin{bmatrix} F \\ G \end{bmatrix}, \ \forall z \in \mathbb{C}, \ |z| \geqslant 1,$$

$$\text{rank}\begin{bmatrix} \Phi - e^{j\omega}I & -F & Q^{1/2} & 0 \\ -H & G & 0 & R^{1/2} \end{bmatrix} = n + p, \ \forall \omega \in [0, \ 2\pi].$$

$\square$

4　Numerical example

In this section, we present a numerical example in order to illustrate the effectiveness of the obtained results. Consider the estimation problem of the discretized DC motor system given in (Park *et al.*, 2000). This system is described by the linear discrete-time stochastic time-invariant system (43) where

$$\Phi = \begin{bmatrix} -0.0005 & -0.0084 \\ 0.0517 & 0.8069 \end{bmatrix}, \Gamma = \begin{bmatrix} 0.1815 \\ 1.7902 \end{bmatrix}, F = \begin{bmatrix} 0.0129 & 0 \\ -1.2504 & 0 \end{bmatrix},$$

$$Q = \begin{bmatrix} 0.0036 & 0.0342 \\ 0.0342 & 0.3249 \end{bmatrix}, H = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, G = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \text{ and } R = \begin{bmatrix} 0.01 & 0 \\ 0 & 0.16 \end{bmatrix}.$$

Assumption 3 and theorem 1 hold since

$$\text{rank }\overline{\Sigma} = \text{rank } F + \text{rank } G = 2$$

$$\text{rank}\begin{bmatrix} zI - \Phi & -F & 0 \\ H & 0 & G \end{bmatrix} = 2 + \text{rank } F + \text{rank } G, \ \forall z \in \mathbb{C}, \ \text{where } z \neq 0 \text{ and } z \neq -0.0073$$

$$\text{rank}\begin{bmatrix} \Phi - e^{j\omega}I & -F & Q^{1/2} & 0 \\ -e^{j\omega}H & G & 0 & R^{1/2} \end{bmatrix} = 4, \ \forall \omega \in [0, \ 2\pi].$$

Since matrix $\overline{\Sigma}$ is of full row rank, we can choose the parameter matrix $\overline{\alpha}$ as $\overline{\alpha} = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$ and we have $N = \overline{A}$ and $\overline{B} = 0$ as shown in remark 6.
We have taken as initial condition $P_0 = \left[\begin{smallmatrix} 10 & 0 \\ 0 & 200 \end{smallmatrix}\right]$ for the discrete-time Lyapunov equation given in remark 6 and the discrete-time Lyapunov equation converges to the solution P of the algebraic Lyapunov equation given by

$$P = \begin{bmatrix} 0.01 & -0.9693 \\ -0.9693 & 134.7406 \end{bmatrix}.$$

The results of section 3 applied to this problem give the following filter matrices

$$N = \begin{bmatrix} 0 & 0 \\ 0.0032 & -0.0073 \end{bmatrix}, J = \begin{bmatrix} 0 & 0 \\ 0.7122 & 0 \end{bmatrix}, E = \begin{bmatrix} 0 \\ 19.3890 \end{bmatrix} \text{ and } L = \begin{bmatrix} 1 & 0 \\ -96.9302 & 0 \end{bmatrix}.$$

In accordance with lemma 4, the filter matrix N has two fixed modes given by $\overline{z}_1 = 0$ and $\overline{z}_2 = -0.0073$ which satisfy the following equality

$$\text{rank}\begin{bmatrix} zI - \Phi & -F & 0 \\ H & 0 & G \end{bmatrix} = 1 + \text{rank } F + \text{rank } G, \ \text{for } z = \overline{z}_1 = 0 \text{ and } z = \overline{z}_2 = -0.0073.$$

Figure 1 presents the known input $u(t)$ and the unknown input $d(t)$. In figures 2 and 3, we have plotted the actual state $x(t)$ and the estimated state $\widehat{x}(t)$. Convergence of the trace of $P(t)$ is shown in figure 4.

5 Conclusion

In this paper, we have extended the minimum variance estimation for linear stochastic discrete time-varying systems with unknown inputs to the general case where the unknown inputs affect both the model and the measurements. The necessary and sufficient conditions for the existence of the filter are derived. Unlike existing approaches where the unknown inputs are assumed to be stochastic processes with known wide-sense description, the optimal filtering presented does not assume any knowledge concerning the model of these inputs. Based on the independence of the estimation error and the system and measurement noises, optimal estimator and predictor filters are proposed. Sufficient conditions for their stability are derived for the time-invariant case.

References

Anderson, B. and J.B. Moore (1979). *Optimal Filtering*. Prentice Hall. Englewood Cliffs, New Jersey.

Basseville, M. and I. Nikiforov (1994). *Detection of Abrupt Changes : Theory and Application*. Prentice Hall. Englewood Cliffs, New Jersey.

Darouach, M. and M. Zasadzinski (1997). 'Unbiased minimum variance estimation for systems with unknown exogenous inputs'. *Automatica* **33**, 717–719.

Darouach, M., M. Zasadzinski and S.J. Xu (1994). 'Full-order observers for linear systems with unknown inputs'. *IEEE Trans. Aut. Contr.* **39**, 606–609.

de Souza, C., M.R. Gevers and G.C. Goodwin (1986). 'Riccati equations in optimal filtering of nonstabilizable systems having singular state transition matrices'. *IEEE Trans. Aut. Contr.* **31**, 831–838.

Hou, M. and P.C. Müller (1992). 'Design of observers for linear systems with unknown inputs'. *IEEE Trans. Aut. Contr.* **37**, 871–875.

Hou, M. and P.C. Müller (1994). 'Disturbance decoupled observer design : a unified point of view'. *IEEE Trans. Aut. Contr.* **39**, 1338–1341.

Hsieh, C. (2000). 'Robust two-stage Kalman filters for systems with unknown inputs'. *IEEE Trans. Aut. Contr.* **45**, 2374–2378.

Kailath, T., A.H. Sayed and B. Hassibi (2000). *Linear Estimation*. Prentice Hall. Englewood Cliffs, New Jersey.

Kitanidis, P. (1987). 'Unbiased minimum-variance linear state estimation'. *Automatica* **23**, 775–778.

Lancaster, P. and M. Tismenetsky (1985). *The Theory of Matrices*. 2nd edn. Academic Press. Orlando, USA.

Middleton, R. and G.C. Goodwin (1990). *Digital Control and Estimation*. Prentice Hall. Englewood Cliffs, New Jersey.

Nikoukhah, R. (1994). 'Innovations generation in the presence of unknown inputs : application to robust failure detection'. *Automatica* **30**, 1851–1867.

Park, S., P.S. Kim, O.H. Kwon and W.H. Kwon (2000). 'Estimation and detection of unknown inputs using optimal FIR filter'. *Automatica* **36**, 1481–1488.

Patton, R., R.N. Clark and P.M. Frank (1989). *Fault Diagnosis in Dynamic Systems*. Prentice Hall. Englewood Cliffs, New Jersey.
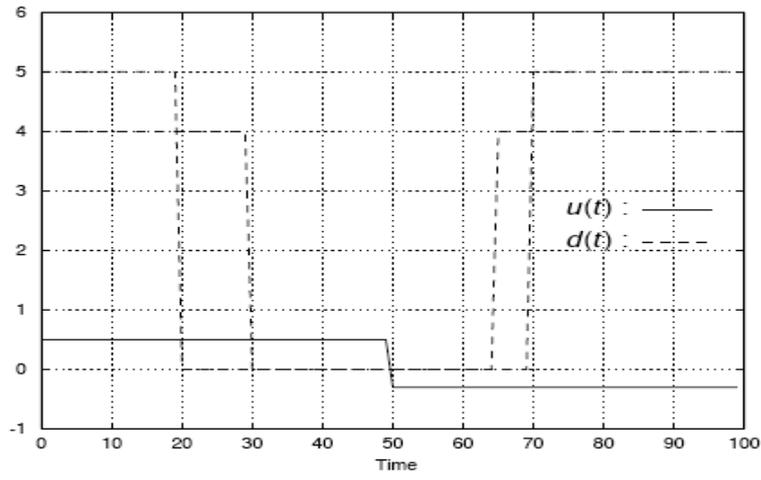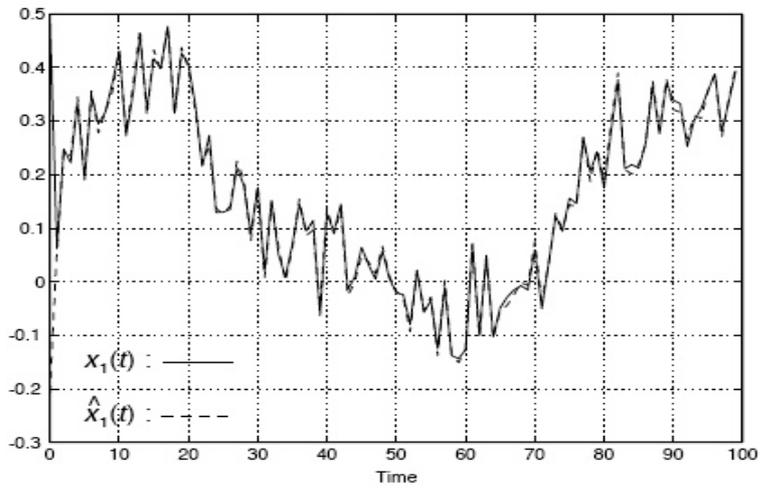
Figure 1: Known input $u(t)$ and unknown input $d(t)$.

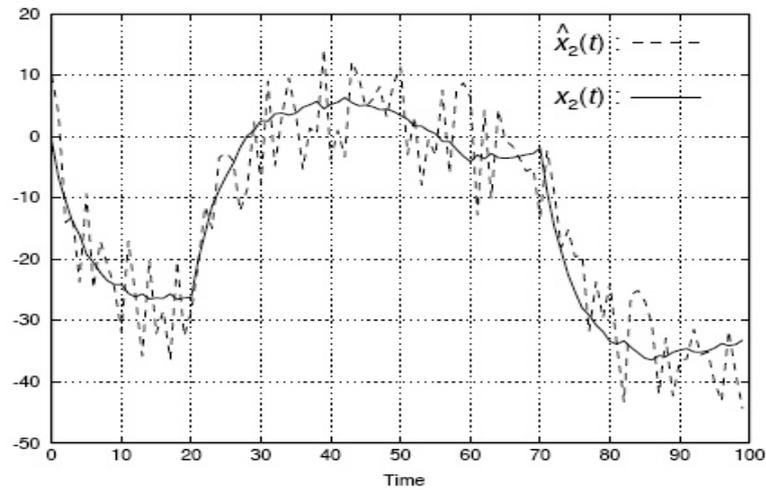Figure 2: Actual state $x_1(t)$ and estimated state $\widehat{x}_1(t)$.

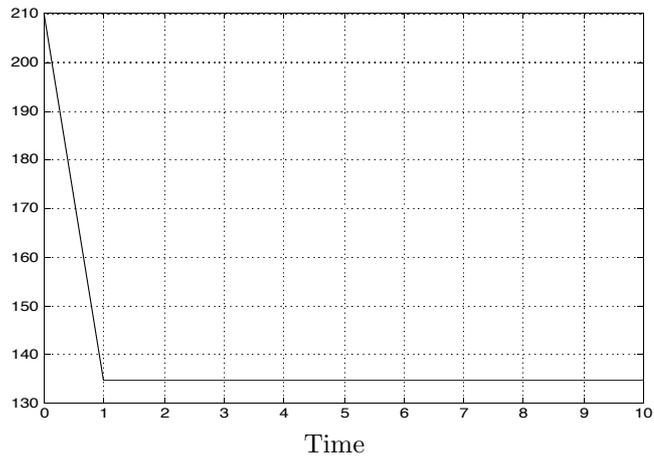Figure 3: Actual state $x_2(t)$ and estimated state $\widehat{x}_2(t)$.

Figure 4: Trace of the estimation covariance $P(t)$.