State Estimation of stochastic singular linear systems
Mohamed Darouach, Michel Zasadzinski, Driss Mehdi

To cite this version:
State Estimation of stochastic singular linear systems

M. DAROUACH*, M. ZASADZINSKI* and D. MEHDI**

* C.R.A.N., C.N.R.S., U.R.A. 821
Institut Universitaire de Technologie de Longwy
Université de NANCY I
route de Romain, 54400 Longwy, FRANCE

** Laboratoire L.A.I.I.
Ecole Nationale Supérieure de Physique de Strasbourg
7 rue de l'université, 67000 Strasbourg, FRANCE

Abstract : In this paper we present a simple algorithm for the state estimation of stochastic singular linear systems based on the least squares method.

1. Introduction

The analysis and design of linear singular systems have received great attention in the last few years as can be seen in the survey of Lewis (1986), and also in Verghese et al. (1981) and Cobb (1984). The control of these systems requires, as in the standard case, the knowledge of the state vector. In the deterministic case one can use the observer theory to estimate the state vector (El-Tohami et al. (1987), Fahmy and O'Reilly (1989)). Unfortunately, in the stochastic case there are few works which treat the problem of control (Bender and Laub (1987 a, b)) and estimation (Dai (1989a)). The state estimation problem is considered under the assumption of regularity \[ \text{det}(sE - A) \neq 0 \] and causality where matrices E and A are square and singular.

In a recent paper (Darouach and Zasadzinski (1990)), an extension to the deterministic system with uncertain measurements, where matrices E and A are rectangular, was presented. In this paper we shall consider a generalization to the stochastic systems case.

The paper is organized as follows. First the concept of estimability is introduced and the uniqueness conditions of problem solution are given in section 2. Then in section 3 an algorithm for the state estimation is derived from the least squares method. A numerical example is used to illustrate this algorithm. Section 4 contains conclusion and remarks.
2. Formulation of the problem

Let us consider the discrete singular linear stochastic system described by:

\[ \begin{align*}
E x_{k+1} &= A x_k + w_k \\
z_k &= H x_k + v_k
\end{align*} \tag{1} \tag{2} \]

where \( x_k \) is the \( n \)-dimensional state vector and \( z_k \) is the \( m \)-dimensional output vector. \( E \) and \( A \) are \( p \times n \) constant matrices and \( H \) is an \( m \times n \) constant matrix. \( w_k \) and \( v_k \) are \( p \times 1 \) and \( m \times 1 \) vectors of zero mean white sequences whose covariance matrices are given by:

\[ \begin{align*}
E\{w_k w_i^T\} &= \begin{cases} W > 0 & \text{if } i = k \\ 0 & \text{if } i \neq k \end{cases} \\
E\{v_k v_i^T\} &= \begin{cases} V > 0 & \text{if } i = k \\ 0 & \text{if } i \neq k \end{cases} \\
E\{w_k v_i^T\} &= 0 \text{ for all } k \text{ and } i.
\end{align*} \]

When \( E \) and \( A \) are square matrices (\( p = n \)), it is well known (Dai (1989b)) that the knowledge of the structure, especially the observability and controllability properties, is important in estimation and in optimal control. In this contribution, we introduce the notion of estimability in the general case where \( p \) may be different from \( n \).

**Definition** :

System (1)-(2) is said to be estimable if for \( w_k = v_k = 0 \) and for some \( N > 0 \), the knowledge of output \( z_k \) where \( k \in [0, N] \) and model equation (1) is sufficient to determine uniquely \( x_k \) (\( k \in [0, N] \)).

System (1)-(2) can be written as follows:

\[
\Phi_N \begin{bmatrix} x_0 \\ \vdots \\ x_N \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ z_0 \\ \vdots \\ z_N \end{bmatrix} \tag{3}
\]

where

\[
\Phi_N = \begin{bmatrix} -A & E & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -A & E \\ H & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & H \end{bmatrix}
\]

\]

2
System (1)-(2) is estimable if and only if $\Phi_N$ has a full column rank.

Our aim is to find the estimate of $x_k$, denoted $\hat{x}_{k/k}$, based on the observations $z$ over the time interval $[0, k]$.

If the initial state $x_0$ is assumed to be gaussian with known mean $\bar{x}_0$, and variance $P_0 > 0$, independent of $v_k$ and $w_k$, the problem in the least squares sense can be formulated:

$$\min_{\hat{x}_{i/k}} J_k = \frac{1}{2} (\|\hat{x}_{0/k} - \bar{x}_0\|_{P_0}^{-1} + \sum_{i=0}^{k-1} \|z_{i+1} - \hat{H} \hat{x}_{i+1/k}\|_{V}^{-1} + \|\hat{E} \hat{x}_{i+1/k} - A \hat{x}_{i/k}\|_{W}^{-1} )) (4)$$

The following theorem gives the conditions of the estimability and the uniqueness of the solution of the problem (4).

**Theorem 1:**
System (1)-(2) is estimable, given the initial state $x_0$, and problem (4) has a unique solution if and only if matrices

$$\begin{bmatrix} sE - A \\ H \end{bmatrix} \text{ and } \begin{bmatrix} E \\ H \end{bmatrix}$$

are of full column rank.

**Proof:**
From equation (3), the estimability condition is given by

$$\text{rank}(\Phi_N) = (N+1) n$$  (5)

Now assume that

$$\begin{bmatrix} sE - A \\ H \end{bmatrix} \text{ or } \begin{bmatrix} E \\ H \end{bmatrix}$$

is not a full column rank matrix, this is equivalent to the existence of a column vector $q \neq 0$ such that

$$\begin{bmatrix} sE - A \\ H \end{bmatrix} q = 0$$  (7)

or

$$\text{rank} \begin{bmatrix} E \\ H \end{bmatrix} < n$$  (8)

where $s$ is a complex variable and $q$ is a finite polynomial vector in variable $s$ given by

$$q(s) = q_0 + s q_1 + s^2 x_2 + .... + s^k q_k$$  (9)

where $k$ is the minimal index such that $q_k \neq 0$ (Gantmacher (1959)).

Substituting (9) into (7) gives
\[ T_k \begin{bmatrix} q_k \\ q_{k-1} \\ \vdots \\ q_0 \end{bmatrix} = 0 \] (10)

where
\[
T_k = \begin{bmatrix}
E & 0 & \ldots & 0 & 0 \\
-A & E & \ldots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \ldots & -A & E \\
H & 0 & \ldots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \ldots & 0 & H
\end{bmatrix}
\] (11)
is a \( p \times (n+1) \) matrix. Equation (10) is equivalent to matrix \( T_k \) is not of full column rank. Now from (3) and (11), we can write
\[
\Phi_N = \begin{bmatrix}
-A & 0 \\
0 & \vdots \\
\vdots & \vdots \\
0 & \vdots \\
0 & E \\
H & T_{N-2} & 0 \\
\vdots & \vdots & \vdots \\
0 & \vdots \\
0 & H
\end{bmatrix}
\] (12)
and from (12) and the assumption (6), we can see that \( \Phi_N \) is not a full column rank matrix, thus system (1)-(2) is not estimable, which proves the theorem.

**Theorem 2 :**
System (1)-(2) is estimable, given the initial state \( x_0 \), and problem (4) has unique solution if and only if the matrix
\[
\begin{bmatrix}
E \\
H
\end{bmatrix}
\] (13)
is of full column rank.

**Proof :**
If \( x_0 \) is known, (3) becomes
\[
\Phi_N' = \begin{bmatrix}
    x_1 \\
    \vdots \\
    x_N
\end{bmatrix} = \begin{bmatrix}
    Ax_0 \\
    0 \\
    \vdots \\
    z_0 \\
    \vdots \\
    z_N
\end{bmatrix}
\]  

where

\[
\Phi_N' = \begin{bmatrix}
    E & 0 & \cdots & 0 \\
    -A & E & \cdots & 0 \\
    \vdots & \vdots & \ddots & \vdots \\
    0 & 0 & \cdots & -A \\
    H & 0 & \cdots & 0 \\
    \vdots & \vdots & \ddots & \vdots \\
    0 & 0 & \cdots & 0 \\
\end{bmatrix}
\]  

which can be written as

\[
\Phi_N' = U \begin{bmatrix}
    E & 0 & 0 & \cdots & 0 \\
    H & 0 & 0 & \cdots & 0 \\
    -A & E & 0 & \cdots & 0 \\
    0 & -A & E & \cdots & 0 \\
    \vdots & \vdots & \vdots & \ddots & \vdots \\
    0 & 0 & -A & E & \cdots \\
    0 & \cdots & 0 & H & \cdots \\
\end{bmatrix}
\]  

where \( U \) is a row permutation matrix.

From the echelon form (16) we can easily deduce that condition (13) is necessary and sufficient for the matrix \( \Phi_N' \) to be of full column rank.

\( \square \)

3. Problem solution

The solution to problem (4) is given by the following theorem.

**Theorem 3**:

If matrices \( \begin{bmatrix} sE - A \\ H \end{bmatrix} \) and \( \begin{bmatrix} E \\ H \end{bmatrix} \) are of full column rank, then estimate \( \hat{x}_{k/k} \) that minimizes criterion \( J_k \) is

\[
\hat{x}_{k/k} = P_{k/k} E^T (W + A P_{k-1/k-1} A^T)^{-1} A \hat{x}_{k-1/k-1} + P_{k/k} H^T V^{-1} z_k
\]

where

\[
P_{k/k} = (E^T (W + A P_{k-1/k-1} A^T)^{-1} E + H^T V^{-1} H)^{-1}
\]

with \( P_{0/0} = P_0 \) and \( \hat{x}_{0/0} = \bar{x}_0 \).

\( \square \)
Proof:

Differentiating the cost function (4) with respect to \( \hat{x}_{i/k} \) and equating it to zero, yields a two-point boundary-value problem:

\[
\begin{align*}
D_1 \hat{x}_{0/k} &= P_0 \bar{x}_0 \\
- C \hat{x}_{i-1/k} + D \hat{x}_{i/k} - C^T \hat{x}_{i+1/k} &= H^T V^{-1} z_i \\
- C \hat{x}_{k-1/k} + D_f \hat{x}_{k/k} &= H^T V^{-1} z_k 
\end{align*}
\]  

(19)

where

\[
\begin{align*}
D_1 &= P_0^{-1} + A^T W^{-1} A \\
D &= E^T W^{-1} E + H^T V^{-1} H + A^T W^{-1} A \\
D_f &= E^T W^{-1} E + H^T V^{-1} H \\
C &= E^T W^{-1} A 
\end{align*}
\]

and

\[
\begin{align*}
K_k &= (D - C K_{k-1} C^T)^{-1} \\
K_0^{-1} &= P_0^{-1} + A^T W^{-1} A 
\end{align*}
\]  

(20)

From equations (19), after a few manipulations, we obtain the following recursive equation for \( \hat{x}_{k/k} \):

\[
\hat{x}_{k/k} = (D_f C K_{k-1} C^T)^{-1} C K_{k-1} (D_f C K_{k-2} C^T) \hat{x}_{k-1/k-1} + (D_f C K_{k-1} C^T)^{-1} H^T V^{-1} z_k 
\]  

(21)

where

\[
K_k = (D - C K_{k-1} C^T)^{-1} 
\]

(22)

and

\[
K_0^{-1} = P_0^{-1} + A^T W^{-1} A 
\]

(23)

If we define the estimation error \( \varepsilon_k \) by:

\[
\varepsilon_k = x_k - \hat{x}_{k/k} 
\]

its covariance matrix \( P_{k/k} \) is:

\[
P_{k/k} = E(\varepsilon_k \varepsilon_k^T) 
\]

(24)

From equations (21), (23) and definition (24), we obtain after a few manipulations:

\[
P_{1/1} = (D_f - C K_0 C^T)^{-1} = (K_1^{-1} - A^T W^{-1} A)^{-1} 
\]

and

\[
K_1^{-1} = P_{1/1}^{-1} + A^T W^{-1} A 
\]

Now if we suppose that:

\[
K_k^{-1} = P_{k/k}^{-1} + A^T W^{-1} A 
\]

(25)
from relations (21), (24) and (25), it can be shown that:

\[ P_{k+1/k+1} = (D_f - C_k C_k^T)^{-1} \]  

(26)

and

\[ K_{k+1}^{-1} = P_{k+1/k+1}^{-1} + A^T W^{-1} A \]  

(27)

Substituting (22) and (26) into (21) gives (17). (18) can be obtained from (25) and (26).

Equation (18) represents a generalized Riccati difference equation.  

**Numerical example:**

As an example, we consider the singular discrete-time system described by the following equations:

\[ E x_{k+1} = A x_k + B u_k + w_k \]

\[ z_k = H x_k + v_k \]

where

\[
E = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 2 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix},
A = \begin{bmatrix} 1 & 1 & 0 & 0.59 \\ 0 & -1 & 0 & 0.50 \\ 1 & 0 & 1 & 0.09 \end{bmatrix},
B = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}
\text{and } H = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & -0.5 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}
\]

The vectors \( x_k, z_k, w_k \) and \( v_k \) have the same definition as in equations (1) and (2) \((n=4, m=3 \text{ and } p=3)\). \( u_k \) is the q-dimensional input vector \((q=2)\).

The variance matrices of \( x_0, w_k \) and \( v_k \) are:

\[ P_0 = \begin{bmatrix} 0.6 & 0 & 0 & 0 \\ 0 & 0.2 & 0 & 0 \\ 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0.7 \end{bmatrix},
W = \begin{bmatrix} 0.6 & 0 & 0 \\ 0 & 0.8 & 0 \\ 0 & 0 & 0.7 \end{bmatrix}
\text{and } V = \begin{bmatrix} 0.3 & 0 & 0 \\ 0 & 0.3 & 0 \\ 0 & 0 & 0.6 \end{bmatrix}
\]

The estimability conditions of theorem 1 are verified.

Inputs and outputs are plotted in figures 1 and 2. The true and estimated values of the state vector are shown in figures 3 to 6. The evolution of the norm \( ||P_{k/k}|| \) (the largest singular value) is plotted in figure 7.

**4. Conclusion**

By using the notion of estimability for the general discrete-time singular systems \((E x_{k+1} = A x_k)\), where matrices \( E \) and \( A \) are constant, not necessarily square, and by applying the least squares estimation method, we have established a simple algorithm for the state estimation of stochastic singular linear systems. A numerical example has been presented to illustrate the algorithm. The evolution of the norm \( ||P_{k/k}|| \) was plotted. The
sequence $P_k/k$ is the solution of a generalized Riccati equation. The convergence conditions of this sequence are under study.

References:
figure 1: inputs (— : $u_1$, ----- : $u_2$)

figure 2: outputs (— : $z_1$, ----- : $z_2$, ...... : $z_3$)
figure 3: true and estimated values of $x_1$ (solid: true values $x_1$, dashed: $\hat{x}_{1,k/k}$)

figure 4: true and estimated values of $x_2$ (solid: true values $x_2$, dashed: $\hat{x}_{2,k/k}$)
figure 5: true and estimated values of \( x_3 \) (——: true values \( x_3 \), -----: \( \hat{x}_{3,k/k} \))

figure 6: true and estimated values of \( x_4 \) (——: true values \( x_4 \), -----: \( \hat{x}_{4,k/k} \))
Figure 7: Evolution of the norm $\|\Sigma_{kk}^k\|$