A method for the identification of state space models from input and output measurements†

DAVID DI RUSCIO‡

In this paper we present a simple and general algorithm for the combined deterministic stochastic realization problem directly from known input and output time series. The solution to the pure deterministic as well as the pure stochastic realization problem are special cases of the method presented.

1. Introduction

A method for identification of state space models from arbitrary input output time series is presented in this paper, which is an extended version of a paper by Di Ruscio (1994).

The method is based on the observation that the underlying model (assumed to be a combined deterministic stochastic, discrete time, linear innovations model), can be written as an extended state space model where the states are known. The known input and output time series are stacked in vectors which yield the inputs and the states for the extended model.

The stacked vectors are collected into two data matrices which define a set of linear equations where the extended system matrices are unknowns. The extended system matrices (which share the underlying system matrices) can be determined from a least squares solution. The system order and the underlying system matrices are then determined by simple singular value analysis and factorization on the extended system matrices. An 'output normal' or 'balanced' minimal model realization results naturally from this method because it is based on the Singular Value Decomposition (SVD).

In our opinion, the method that is presented here for the realization of combined deterministic stochastic systems from known time series is simple, compared to existing methods. It should be noted that the purely stochastic realization problem, as well as the purely deterministic, stands out as a special case in this context.

Other results exist and are presented in the literature. Some methods are based on the factorization of Hankel matrices constructed from correlation matrices, and others are based on constructing a state sequence for the underlying model, either from a projection of the time series or from instrumental variables (Aoki 1991). The state

space matrices are then determined from a set of linear equations defined by the estimated state sequence and the known input output sequences.

A method of this last type is the 'N4SID subspace algorithm' by Van Overschee and De Moor (1994). A simple example has been used to compare the method in this paper with the N4SID algorithm.

The rest of this paper is organized as follows. Section 2 presents the problem definitions and some preliminaries. A complete solution to the deterministic realization problem is presented in § 3. The solution to the combined deterministic stochastic realization problem is presented in § 4. Monte Carlo simulations are used to illustrate and compare the method with existing algorithms, N4SID and ARMAX, in § 5. Some concluding remarks are given in § 6.

2. Problem definition and preliminaries

2.1. *A state space model on innovations form*

Define the discrete-time, time invariant, linear state space model of the form

$$x_{x+1} = Ax_k + Bu_k + Ce_k \tag{1}$$

$$y_k = Dx_k + Eu_k + e_k \tag{2}$$

where $k \geqslant 0$ is discrete time, $x_k \in \Re^n$ is the state vector, $y_k \in \Re^m$ is the system output, $u_k \in \Re^r$ is the system input and $e_k \in \Re^m$ is an unknown innovations process, which is assumed to be standard white Gaussian noise. We will in this paper assume that the time series y_k and u_k are known for $k = 0, 1, \ldots, N-1$, i.e. a number N samples are known. The following theorem is important for the results in this paper.

Theorem 2.1. (*observability*)

If rank $(D) = d$, then the system represented by (1) and (2) is completely observable if the $(n - d + 1)m \times n$ 'observability' matrix

$$O = \begin{bmatrix} D \\ DA \\ DA^2 \\ \vdots \\ DA^{n-d} \end{bmatrix} \quad \text{and rank } (O) = n \tag{3}$$

$$\Rightarrow OO^T \text{ is nonsingular.}$$

The proof is the same as that presented for controllability by Kalman, Falb and Arbib (1969), Ch. 2, pp. 37.   △

2.2. *A matrix valued state space model*

Given (the possibility matrix) dynamic, discrete time, linear state space model

$$\mathcal{X}_{k+1} = A\mathcal{X}_k + B\mathcal{U}_k \quad \forall k \geqslant 0 \tag{4}$$

$$\mathcal{Y}_k = D\mathcal{X}_k + E\mathcal{U}_k \tag{5}$$

This model can be determined from the state space innovations model as follows: shift time index with $k = :i$ and $i = :i + k$ in Eqns (1) and (2), post multiply with u_i^T (or y_i^T) and take expectations. $\mathcal{Y}_k$, $\mathcal{U}_k$ and $\mathcal{X}_k$ are in this case correlation and cross-correlation

matrices given by

$$\mathscr{Y}_k = \mathbf{E}(y_{i+k}u_i^T) \in \mathfrak{R}^{m \times r} \tag{6}$$

$$\mathscr{U}_k = \mathbf{E}(u_{i+k}u_i^T) \in \mathfrak{R}^{r \times r} \tag{7}$$

$$\mathscr{X}_k = \mathbf{E}(x_{i+k}u_i^T) \in \mathfrak{R}^{n \times r} \tag{8}$$

where $k \geqslant 0$. Note that if the underlying system is assumed to be deterministic, then we can take $\mathscr{Y}_k =: y_i$, $\mathscr{X}_k =: x_i$ and $\mathscr{U}_k =: u_i$, and the two model representations (1), (2) and (4), (5) are identical.

We will in this paper assume that the sequences $\mathscr{Y}_k$ *and* $\mathscr{U}_k$ (or y_k and u_k, deterministic case) are known for $k = 0, 1, \dots N_k - 1$, i.e. a number N_k samples are known.

2.3. *Problem statement*

The main problem of this paper can be stated as follows: From known input and output time series, u_i and y_i, respectively, determine a state space model realization for the deterministic part of the underlying system, i.e. the quadruple (A, B, D, E). It is assumed that the output sequences, y_i, are generated by an unknown underlying combined deterministic stochastic state space model of the form (1) and (2) excited by the input sequences u_i. Note that the deterministic realization problem as well as the stochastic is a special case of the problem addressed.

The results in this paper are presented in two parts, §§ 3 and 4, respectively. A description is given in the following.

Part I. A solution to the identification (or realization) problem based on the matrix valued state space model given in (4) and (5) will be presented in § 3. The estimation problem of the correlation and cross-correlation matrices $\mathscr{U}_k$ and $\mathscr{Y}_k$ from the measured sequences, u_i and y_i, can in some circumstances be hard to obtain. However, our experience is that the unbiased estimates, viz.

$$\mathscr{Y}_k = \mathbf{E}(y_{i+k}u_i^T) =: \frac{1}{N-1-k} \sum_{i=0}^{N-k-1} y_{i+k}u_i^T$$

$$\mathscr{U}_k = \mathbf{E}(u_{i+k}u_i^T) =: \frac{1}{N-1-k} \sum_{i=0}^{N-k-1} u_{i+k}u_i^T$$

(where N is the number of samples in the time series of y and u and $k \geqslant 0$), work satisfactorily for many systems.

However, the inclusion of the correlations matrices is only a trick to transform the innovations model into a deterministic model. The main purpose in § (3) is to solve the deterministic realization problem.

Part II. A solution to the combined stochastic and deterministic identification (or realization) problem can be obtained directly from the state space innovations model formulation given by (1) and (2), directly from the output input time series, y_k and u_k, without the need of correlation and cross-correlation estimation. This result is presented in § 4. A solution to the stochastic realization problem which results as a special case will also be presented.

3. **Part I**

We will in this section as a starting point develop the solution of the realization

problem with the matrix valued state space model, Eqns (4) and (5). The solution coincides with the solution to the deterministic realization problem.

3.1. *The general extended model formulation*

Introduce L as the number of block rows in the 'extended' observability matrix. This definition will be clarified later.

The state space model, Eqns (4) and (5), can in general be written as the following extended state space model.

$$\tilde{\mathscr{Y}}_{k+1} = \tilde{A}\tilde{\mathscr{Y}}_k + \tilde{B}\tilde{\mathscr{U}}_k \tag{9}$$

where $k \geq 0$ and $\tilde{A}$ and $\tilde{B}$ are constant matrices given by

$$\tilde{A} = OA(O^{\mathrm{T}}O)^{-1}O^{\mathrm{T}} \tag{10}$$

$$\tilde{B} = [OB - \tilde{A}E_0 \; E_0 - \tilde{A}E_1 \; E_1 - \tilde{A}E_2 \ldots E_{L-1}] \tag{11}$$

The extended output sequences, $\tilde{\mathscr{Y}}_k \in \mathfrak{R}^{mL \times r}$ and $\tilde{\mathscr{U}} \in \mathfrak{R}^{r(L+1) \times r}$ respectively, are given by

$$\tilde{\mathscr{Y}}_k = \begin{bmatrix} \mathscr{Y}_k \\ \mathscr{Y}_{k+1} \\ \mathscr{Y}_{k+2} \\ \vdots \\ \mathscr{Y}_{k+L-1} \end{bmatrix} \quad \tilde{\mathscr{U}}_k = \begin{bmatrix} \mathscr{U}_k \\ \mathscr{U}_{k+1} \\ \mathscr{U}_{k+2} \\ \vdots \\ \mathscr{U}_{k+L} \end{bmatrix} \tag{12}$$

Note that $\tilde{A}$ contains the system matrix, at least as a submatrix. This can be clarified as follows. Assume that (A, B, D, E) is an output normal realization and let C be the controllability matrix for the pair (B, A). Let the product of the observability (O) and the controllability (C) matrices be factored into $OC = USV^{\mathrm{T}}$ by the singular value decomposition. In the case of an output normal realization, $O = U$ and $O^{\mathrm{T}}O = I$. this gives

$$\tilde{A} = UAU^{\mathrm{T}} \tag{13}$$

which shows, through similarity, that $\tilde{A}$ contains the eigenvalues of A.

Proof of Equation (9)
The system model, (4) and (5), can be written as

$$\mathscr{X}_{k+1} = A\mathscr{X}_k + B\mathscr{U}_k \tag{14}$$

$$\tilde{\mathscr{Y}}_k = O\mathscr{X}_k + E_0\mathscr{U}_k + E_1\mathscr{U}_{k+1} + E_2\mathscr{U}_{k+2} + \ldots + E_{L-1}\mathscr{U}_{k+L-1} \tag{15}$$

where $\tilde{\mathscr{Y}}_k$ is defined above and

$$O = \begin{bmatrix} D \\ DA \\ DA^2 \\ \vdots \\ DA^{L-1} \end{bmatrix} \quad E_0 = \begin{bmatrix} E \\ DB \\ DAB \\ \vdots \\ DA^{L-2}B \end{bmatrix} \quad E_1 = \begin{bmatrix} 0 \\ E \\ DB \\ \vdots \\ DA^{L-3}B \end{bmatrix} \quad E_{L-1} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ E \end{bmatrix} \tag{16}$$

Express Eqn (15) for time $k + 1$ and substitute for $\mathscr{X}_{k+1}$ from Eqn (14). We have

$$\tilde{\mathscr{Y}}_{k+1} = OA(O^{\mathrm{T}}O)^{-1}O^{\mathrm{T}} \cdot O\mathscr{X}_k + OB\mathscr{U}_k$$

$$+ E_0\mathscr{U}_{k+1} + E_1\mathscr{U}_{k+2} + E_2\mathscr{U}_{k+3} + \ldots + E_{L-1}\mathscr{U}_{k+L} \tag{17}$$

because $O^{\mathrm{T}}O$ is nonsingular. Substitute $O\mathscr{X}_k$, defined from Eqn (15), into Eqn (17). We have

$$\tilde{\mathscr{Y}}_{k+1} = \tilde{A}\tilde{\mathscr{Y}}_k + (OB - \tilde{A}E_0)\mathscr{U}_k$$

$$+ (E_0 - \tilde{A}E_1)\mathscr{U}_{k+1} + (E_1 - AE_2)\mathscr{U}_{k+2} + \ldots + E_{L-1}\mathscr{U}_{k+L} \tag{18}$$

which is identical with Eqns (9), (10) and (11).   QED.

The system dynamics can now be isolated from Eqn (9). Assume that the sequences $\tilde{\mathscr{Y}}_k$ and $\tilde{\mathscr{U}}_k$ are known for $k = 0, 1, 2, \ldots, K$, i.e. the sequences are of length $K + 1$. This requires that the output sequences $\mathscr{Y}_k$ and $\mathscr{U}_k$ are known for $k = 0, 1, 2, \ldots, K + L - 1$, i.e. the sequences must be of length $K + L$. The sequences $\mathscr{Y}_k$ and $\mathscr{U}_k$ are by definition known for N_k samples, this gives $K = N_k - L$. We will also define L such that $L \geqslant 1$. Note that L coincides with the number of $(m \times n)$ block rows in the extended observability matrix (O).

The known sequences $\mathscr{Y}_k$ and $\mathscr{U}_k$ can now be collected into data matrices Y_1, Y_0 and U_0 which satisfy the following linear matrix equation

$$Y_1 = \Gamma \begin{bmatrix} Y_0 \\ U_0 \end{bmatrix} \tag{19}$$

where

$$\Gamma = [\tilde{A} \quad \tilde{B}] \tag{20}$$

$$Y_1 = [\tilde{\mathscr{Y}}_1 \quad \tilde{\mathscr{Y}}_2 \quad \ldots \quad \tilde{\mathscr{Y}}_K] \tag{21}$$

$$Y_0 = [\tilde{\mathscr{Y}}_0 \quad \tilde{\mathscr{Y}}_1 \quad \ldots \quad \tilde{\mathscr{Y}}_{K-1}] \tag{22}$$

$$U_0 = [\tilde{\mathscr{U}}_0 \quad \tilde{\mathscr{U}}_1 \quad \ldots \quad \tilde{\mathscr{U}}_{K-1}] \tag{23}$$

The system of linear equations, Eqn (19), can be solved for Γ in a least squares sense. The system quadruple (A, B, D, E) can then be extracted from Γ. This will be shown later.

The matrices Y_1, Y_0 and U_0 which are defined from the known sequences can be more precisely defined as follows.

$$Y_1 = \begin{bmatrix} \mathscr{Y}_1 & \mathscr{Y}_2 & \ldots & \mathscr{Y}_K \\ \mathscr{Y}_2 & \mathscr{Y}_3 & & \mathscr{Y}_{K+1} \\ \vdots & \vdots & & \vdots \\ \mathscr{Y}_L & \mathscr{Y}_{L+1} & \ldots & \mathscr{Y}_{K+L-1} \end{bmatrix} \tag{24}$$

$$Y_0 = \begin{bmatrix} \mathscr{Y}_0 & \mathscr{Y}_1 & \ldots & \mathscr{Y}_{K-1} \\ \mathscr{Y}_1 & \mathscr{Y}_2 & & \mathscr{Y}_K \\ \vdots & \vdots & & \vdots \\ \mathscr{Y}_{L-1} & \mathscr{Y}_L & \ldots & \mathscr{Y}_{K+L-2} \end{bmatrix} \tag{25}$$

D. Di Ruscio

$$
U_0 = \begin{bmatrix} \mathcal{U}_0 & \mathcal{U}_1 & \dots & \mathcal{U}_{K-1} \\ \mathcal{U}_1 & \mathcal{U}_2 & & \mathcal{U}_K \\ \mathcal{U}_2 & \mathcal{U}_3 & & \mathcal{U}_{K+1} \\ \vdots & \vdots & & \vdots \\ \mathcal{U}_L & \mathcal{U}_{L+1} & \dots & \mathcal{U}_{K+L-1} \end{bmatrix} \tag{26}
$$

where $Y_1, Y_0 \in \mathfrak{R}^{Lm \times Kr}$, $U_0 \in \mathfrak{R}^{(L+1)r \times Kr}$ and $\Gamma \in \mathfrak{R}^{Lm \times (Lm + (L+1)r)}$. Note that this result also is valid if $\mathcal{Y}_k =: y_i$ and $\mathcal{U}_k =: u_i$. In this case $Y_1,\ Y_0 \in \mathfrak{R}^{Lm \times K}$ and $U_0 \in \mathfrak{R}^{(Lm + (L+1)r) \times K}$.

Note that the complexity of the problem can be reduced if the system matrix E is assumed to be zero, i.e. when the system is *strictly proper*. The last column in $\tilde{B}$, (i.e. E_{L-1}), and the corresponding last row in U_0 can be removed in this case (i.e. when $E = 0$).

3.2. *Determination of the system matrices*

Define, for simplicity

$$
Q_0 = \begin{bmatrix} Y_0 \\ U_0 \end{bmatrix} \tag{27}
$$

Post multiply (19) with Q_0^T and divide on both sides with K for consistency with the expectation operator. The linear matrix equation can then be written as

$$
\frac{1}{K} Y_1 Q_0^T = \Gamma \frac{1}{K} Q_0 Q_0^T \tag{28}
$$

The known matrices on the left and right hand sides of the least squares problem, Eqn (28), can be expressed as follows

$$
\frac{1}{K} Y_1 Q_0^T \overset{\text{def}}{=} [\Lambda_1 \quad \Sigma_1] = \begin{bmatrix} \dfrac{1}{K} \displaystyle\sum_{k=0}^{K-1} \tilde{\mathcal{Y}}_{k+1} \tilde{\mathcal{Y}}_k^T & \dfrac{1}{K} \displaystyle\sum_{k=0}^{K-1} \tilde{\mathcal{Y}}_{k+1} \tilde{\mathcal{U}}_k^T \end{bmatrix} \tag{29}
$$

$$
\frac{1}{K} Q_0 Q_0^T \overset{\text{def}}{=} \begin{bmatrix} \Lambda_0 & \Sigma_0 \\ S_0 & U_0 \end{bmatrix} = \begin{bmatrix} \dfrac{1}{K} \displaystyle\sum_{k=0}^{K-1} \mathcal{\tilde{Y}}_k \tilde{\mathcal{Y}}_k^T & \dfrac{1}{K} \displaystyle\sum_{k=0}^{K-1} \tilde{\mathcal{Y}}_k \tilde{\mathcal{U}}_k^T \\ \dfrac{1}{K} \displaystyle\sum_{k=0}^{K-1} \tilde{\mathcal{U}}_k \tilde{\mathcal{Y}}_k^T & \dfrac{1}{K} \displaystyle\sum_{k=0}^{K-1} \tilde{\mathcal{U}}_k \tilde{\mathcal{U}}_k^T \end{bmatrix} \tag{30}
$$

We will now clarify the link between the above matrices and the known data $\mathcal{Y}_k$ and $\mathcal{U}_k$. It is necessary to point out that the above matrices can be formulated in terms of symmetrical and nonsymmetrical block Toeplitz matrices where the block matrices in Eqns (29) and (30) are given from following submatrices

$$
\Lambda_{jl} = \frac{1}{K} \sum_{k=0}^{K-1} \mathcal{Y}_{k+j} \mathcal{Y}_{k+l}^T
$$

$$
U_{jl} = \frac{1}{K} \sum_{k=0}^{K-1} \mathcal{U}_{k+j} \mathcal{U}_{k+l}^T
$$

$$
\Sigma_{jl} = \frac{1}{K} \sum_{k=0}^{K-1} \mathcal{Y}_{k+j} \mathcal{U}_{k+l}^T
$$

$$S_{jl} = \frac{1}{K} \sum_{k=0}^{K-1} \mathcal{U}_{k+j} \mathcal{Y}_{k+l}^{\mathrm{T}}$$

From the partitioning defined above we have the following two equations which define $\tilde{A}$ and $\tilde{B}$.

$$\Lambda_1 = \tilde{A}\Lambda_0 + \tilde{B}S_0 \tag{31}$$

$$\Sigma_1 = \tilde{A}\Sigma_0 + \tilde{B}\mathbf{U}_0 \tag{32}$$

Note that $\Sigma_0 = S_0^{\mathrm{T}}$. From this we get

$$\Lambda_1 - \Sigma_1\mathbf{U}_0^{-1}S_0 = \tilde{A}(\Lambda_0 - \Sigma_0\mathbf{U}_0^{-1}S_0) \tag{33}$$

Note that the pseudo inverse should be used instead of $\mathbf{U}_0^{-1}$ in the case when $\mathbf{U}_0$ is not of full rank.

We will now go for the system order n and the system matrices. For the sake of simplicity, define the matrices Z_1 and Z_0 such that Eqn (33) is identical with

$$Z_1 = \tilde{A}Z_0 \tag{34}$$

The system order n

The system order is determined by a singular value decomposition of the matrix on the right hand side of (33), defined as Z_0. Obtain the singular value decomposition for Z_0, i.e.

$$Z_0 = USV^{\mathrm{T}} \tag{35}$$

where

$$U = [U_1 \quad U_2] \quad S = \begin{bmatrix} S_n & 0 \\ 0 & 0 \end{bmatrix} \quad V = [V_1 \quad V_2] \tag{36}$$

where $S_n \in \mathfrak{R}^{n \times n}$ and n is the number of non-zero singular values of Z_0, which is equal to the system order. n is determined by inspection of the non-zero diagonal elements of S or SS^{T}.

The extended observability matrix for the pair (D, A)

The (extended) observability matrix can be taken directly as the first left part in U, i.e. U_1. We have

$$O = U(1:Lm, 1:n) \tag{37}$$

The system matrix A

The system matrix A is determined as follows

$$A = U_1^{\mathrm{T}} Z_1 V_1 S_n^{-1} \tag{38}$$

The system output matrix D

The system output matrix D can be taken as the $m \times n$ upper submatrix in the observability matrix U. This can be formulated as follows

$$D = U(1:m, 1:n) \tag{39}$$

The extended system matrices $\tilde{A}$ *and* $\tilde{B}$

$\tilde{A}$ can now be determined from (34) and (35). We have

$$\tilde{A} = Z_1 V_1 S_n^{-1} U_1^{\mathrm{T}} \tag{40}$$

The matrix $\tilde{B}$ is defined from (32) because $\tilde{A}$ is known from (40).

The system matrices B *and* E

The system matrix E can be extracted directly as the lower right ($m \times r$) corner of the matrix $\tilde{B}$. The matrix OB is extracted recursively from block columns of the matrix $\tilde{B}$ and $\tilde{A}$. The system input matrix B can then be taken as

$$B = (O^{\mathrm{T}}O)^{-1}O^{\mathrm{T}} \cdot OB = U_1^{\mathrm{T}} \cdot OB \tag{41}$$

We have now determined a complete realization for the underlying system (A, B, D, E). We have for simplicity, chosen $O = U_1$ in Step 2, because $O^{\mathrm{T}}O = I$. This gives an output normal realization when $L \to \infty$. The algorithm can also be formulated with $O = US^{1/2}$, i.e. a balanced realization when $L \to \infty$. $O^{\mathrm{T}}O$ is equal to the observability grammian as L tends to infinity because, in this case, DA^{L-1} tends to zero.

4. Part II

We will in this section solve the combined stochastic and deterministic identification problem directly from y_k and u_k, without the need to form the correlation and cross-correlation matrices.

4.1. *The combined deterministic stochastic problem*

The innovations model, Eqns (1) and (2), can generally be written as the following 'extended' state space model. Note also the connection with a MIMO ARMAX model formulation with constant polynomial matrices. However, the main point of writing the model in this form is that the 'state' description is known.

$$\tilde{y}_{k+1} = \tilde{A}\tilde{y}_k + \tilde{B}\tilde{u}_k + \tilde{C}\tilde{e}_k \ \forall \ k \geq 0 \tag{42}$$

where $\tilde{y}_k \in \mathfrak{R}^{mL \times 1}$, $\tilde{u}_k \in \mathfrak{R}^{r(L+1) \times 1}$ and $\tilde{e}_k \in \mathfrak{R}^{m(L+1) \times 1}$, are given by

$$
\tilde{y}_k = \begin{bmatrix} y_k \\ y_{k+1} \\ y_{k+2} \\ \vdots \\ y_{k+L-1} \end{bmatrix} \quad
\tilde{u}_k = \begin{bmatrix} u_k \\ u_{k+1} \\ u_{k+2} \\ \vdots \\ u_{k+L} \end{bmatrix} \quad
\tilde{e}_k = \begin{bmatrix} e_k \\ e_{k+1} \\ e_{k+2} \\ \vdots \\ e_{k+L} \end{bmatrix} \tag{43}
$$

The matrices $\tilde{A} \in \mathfrak{R}^{mL \times mL}$, $\tilde{B} \in \mathfrak{R}^{mL \times r(L+1)}$ and $\tilde{C} \in \mathfrak{R}^{mL \times m(L+1)}$ are constant and defined as follows.

$$\tilde{A} = OA(O^{\mathrm{T}}O)^{-1}O^{\mathrm{T}} \tag{44}$$

$$\tilde{B} = [OB - \tilde{A}E_0 \ E_0 - \tilde{A}E_1 \ E_1 - \tilde{A}E_2 \dots E_{L-1}] \tag{45}$$

$\tilde{C}$ have the same structure as $\tilde{B}$ but with B substituted with C and E with the ($m \times m$) dimensional identity matrix.

Note that it is expected that the stacked noise process, $\tilde{e}_k$, satisfy

$$\mathbf{E}(\tilde{e}_{k+j}\tilde{e}_k^{\mathrm{T}}) = 0 \ \forall \ j \geq L + 1 \tag{46}$$

Define the data matrices from the stacked vectors, (43), as follows

$$Y_k = [\tilde{y}_k \ \tilde{y}_{k+1} \dots \tilde{y}_{k+K-1}] \tag{47}$$

$$U_k = [\tilde{u}_k \ \tilde{u}_{k+1} \dots \tilde{u}_{k+K-1}] \tag{48}$$

$$E_k = [\tilde{e}_k \tilde{e}_{k+1} \ldots \tilde{e}_{k+K-1}] \tag{49}$$

Let the matrix Γ be defined as

$$\Gamma = [\tilde{A} \; \tilde{B}] \tag{50}$$

We then have

$$Y_k = \Gamma \begin{bmatrix} Y_{k-1} \\ U_{k-1} \end{bmatrix} + \tilde{C}E_{k-1} \quad \forall k \geq 1 \tag{51}$$

From the definition that the white noise e_i is uncorrelated with e_{i+k}, $(k \neq i)$, then, we have:

$$\lim_{K \to \infty} \frac{1}{K} E_{k+L} E_{k-1}^{\mathrm{T}} = \lim_{K \to \infty} \frac{1}{K} \sum_{j=0}^{K-1} \tilde{e}_{k+L+j} \tilde{e}_{k-1+j}^{\mathrm{T}} = 0 \tag{52}$$

where $k \geq 1$. Note that (52) coincides with (46) if ergodicity is assumed. We then have the definition $\mathrm{E}(\cdot) = \lim_{K \to \infty} 1/K \Sigma_{j=0}^{K-1}(\cdot)$.

Several sets of linear equations which define Γ, can now be defined from Eqn (51). Define the following two linear matrix equations from (51), with $k = 1$ and $k = L + 1$.

$$Y_1 = \Gamma \begin{bmatrix} Y_0 \\ U_0 \end{bmatrix} + \tilde{C}E_0 \tag{53}$$

$$Y_{L+1} = \Gamma \begin{bmatrix} Y_L \\ U_L \end{bmatrix} + \tilde{C}E_L \tag{54}$$

Define

$$Q_k = \begin{bmatrix} Y_k \\ U_k \end{bmatrix} \tag{55}$$

and the matrix of 'instrumental variables'

$$W_i = \begin{bmatrix} Y_0 \\ U_i \end{bmatrix} \quad \forall i \geq 0 \tag{56}$$

Post multiply Eqn. (54) with W_i and divide with the scalar parameter K for consistency with the expectation operator. We have

$$\frac{1}{K} Y_{L+1} W_i^{\mathrm{T}} = \Gamma \frac{1}{K} Q_L W_i^{\mathrm{T}} \tag{57}$$

because

$$\frac{1}{K} E_L W_i^{\mathrm{T}} = \frac{1}{K} E_L \begin{bmatrix} Y_0 \\ U_i \end{bmatrix}^{\mathrm{T}} = 0 \tag{58}$$

when K approaches infinity.

Equation (57) can now be solved for Γ in a least squares sense and the quadruple (A, B, D, E) extracted from the extended system matrices A and B. The procedure is the same as that presented in § 3. The realization problem, details and special cases are discussed further in the next section.

Assume that the time series y_k and u_k are of length N. The parameter K, which defines the number of columns in the data matrices Y_0, Y_L, Y_{L+1}, U_0 and U_L, is of size $K = N - 2L - 1$ in this case.

4.2. Details and special cases

The system matrices (A, B, D, E) can be determined exactly in the same way as in § 3. However, it is necessary to clarify some details and special cases. Note that the pure deterministic as well as the pure stochastic realization problem are special cases of the problem addressed.

The known matrices on the left and right hand sides of the least squares problem, Eqn (57), can be expressed as follows. However, note that we have introduced the lag parameter M. The justification for this will be clarified later.

$$\frac{1}{K} Y_{M+1} W_i^T \overset{\text{def}}{=} [\Lambda_{M+1} \quad \Sigma_{M+1}] = \left[\frac{1}{K} \sum_{k=0}^{K-1} \tilde{y}_{k+M+1} \tilde{y}_k^T \quad \frac{1}{K} \sum_{k=0}^{K-1} \tilde{y}_{k+M+1} \tilde{u}_{k+i}^T \right] \tag{59}$$

$$\frac{1}{K} Q_M W_i^T \overset{\text{def}}{=} \begin{bmatrix} \Lambda_M & \Sigma_M \\ \mathbf{S}_M & \mathbf{U}_M \end{bmatrix} = \begin{bmatrix} \frac{1}{K} \sum_{k=0}^{K-1} \tilde{y}_{k+M} \tilde{y}_k^T & \frac{1}{K} \sum_{k=0}^{K-1} \tilde{y}_{k+M} \tilde{u}_{k+i}^T \\ \frac{1}{K} \sum_{k=0}^{K-1} \tilde{u}_{k+M} \tilde{y}_k^T & \frac{1}{K} \sum_{k=0}^{K-1} \tilde{u}_{k+M} \tilde{u}_{k+i}^T \end{bmatrix} \tag{60}$$

From the partitioning defined above we have the following two linear equations which define $\tilde{A}$ and $\tilde{B}$.

$$\Lambda_{M+1} = \tilde{A}\Lambda_M + \tilde{B}S_M \tag{61}$$

$$\Sigma_{M+1} = \tilde{A}\Sigma_M + \tilde{B}U_M \tag{62}$$

The justification for the lag parameter M will now be clarified. For variations of M, special cases concerning the type of the system result.

$M = L$: Combined deterministic stochastic systems

$$\Lambda_{L+1} = \tilde{A}\Lambda_L + \tilde{B}S_L \tag{63}$$

$$\Sigma_{L+1} = \tilde{A}\Sigma_L + \tilde{B}U_L \tag{64}$$

We will first discuss the case with $i = 0$ in the matrix of instrumental variables, Eqn (56), i.e. W_0. Consider the rare case where the system input is standard white Gaussian noise. In this case U_L is singular as the number of samples approaches infinity. The system matrices $(\tilde{A})$ and $(\tilde{B})$ defined from (63) and (64) in any case. We mention it because it is obscure to invert U_L in this case. However, this problem is avoided by using $i = L$ in the matrix of instrumental variables, i.e. W_L.

$M = 0$: Purely deterministic systems

$$\Lambda_1 = \tilde{A}\Lambda_0 + \tilde{B}S_0 \tag{65}$$

$$\Sigma_1 = \tilde{A}S_0^T + \tilde{B}U_0 \tag{66}$$

Note that $\Sigma_0 = S_0^T$ in this case. Note also that $U_0 \neq 0$ in the rare case where the system input is white noise.

4.3. *Purely stochastic systems*

From (63) and (64) with $u_k = 0$ we have

$$\Lambda_{L+1} = \tilde{A}\Lambda_L \tag{67}$$

It is important to point out that the (D, A) matrices, in the stochastic realization problem, are defined from two finite matrices (Eqn (67)), instead from one Hankel matrix as with the traditional approach. The matrices Λ_{L+1} and Λ_L will have a Toepliz structure when the number of samples approaches infinity.

5. Monte Carlo simulation

5.1. *Example 1*

A scalar example is chosen in order to compare the algorithm presented in this paper (DSR) with two other algorithms, N4SID (Van Overschee and De Moor (1994)) and ARMAX (MATLAB: System identification toolbox, Ljung (1991)).

$$x_{k+1} = 0.9x_k + 0.5u_k + 0.6e_k \tag{68}$$

$$y_k = 1.0x_k - 1.0u_k + e_k \tag{69}$$

Three types of input signals were used. One input equal to a sum of four sinusoid signals, u^1, one input equal to a white noise signal with unit covariance, u^2, and one equal to a sine, u^3.

$$u^1 \quad u_k = 0.2\left(\sin\left(\frac{k}{25}\right) + \sin\left(\frac{k}{10}\right) + \sin\left(\frac{k}{5}\right) + \sin(k)\right) \tag{70}$$

$$u^2 \quad \text{White noise, unit covariance} \tag{71}$$

$$u^3 \quad u_k = \sin(k) \tag{72}$$

For each input the time series (y_k, u_k) was generated by simulating the model with 100 different white noise sequences e_k also with unit covariance.

The DSR algorithm parameter L was changed from 1 to 5 and the N4SID parameter I from 2 to 6. For each L and I, the mean and standard deviation of the parameters of the 100 different estimated models are presented in Tables 1 to 6. The results obtained by the ARMAX algorithm are also shown in the tables. See Ljung (1991) for the description of the parameters $nn = [1, 2, 1, 0]$ which is used as arguments to ARMAX.

The true deterministic system quadruple is denoted $(a, b, d, e) := (0.9, 0.5, 1, -1)$ and the deterministic steady state gain and deterministic zero are denoted $H^d(1) = 4.0$ and $p_d(1) = 1.4$, respectively. The parameters in the stochastic part of the model are $(c, \Delta) := (0.6, 1)$. The stochastic steady state gain and stochastic zero are denoted $H^s(1) = 7.0$ and $p_s(1) = 0.3$, respectively. The signal to noise ratio is as low as 0.4, hence, the identification problem is not simple.

The N4SID algorithm sometimes estimated systems with negative (b, d) parameters, i.e. sometimes an estimated quadruple (a, b, d, e) and some times (a, $-$b, $-$d, e). This happened with the algorithm parameter $I = 3$ and with an random input signal.

The results are very good both for the DSR method in this paper and for the N4SID method, Tables 1 and 2. There are small differences in the estimated models for both methods when N is large, see Tables 1 and 2.

D. Di Ruscio

Alg.	Par.	a	b	d	e	$p_d(1)$	$H^d(1)$	c	Δ	$p_s(1)$	$H^s(1)$
DSR	L=1	0.8998	0.4965	1.0000	-1.0021	1.3975	3.9598	0.6088	1.0045	0.2910	7.0910
DSR	L=2	0.8995	0.4971	1.0000	-1.0021	1.3979	3.9516	0.6003	0.9989	0.2992	6.9867
DSR	L=3	0.8998	0.4965	1.0000	-1.0021	1.3975	3.9604	0.5992	0.9985	0.3006	6.9940
DSR	L=4	0.8996	0.4967	1.0000	-1.0019	1.3977	3.9572	0.5992	0.9981	0.3004	6.9875
DSR	L=5	0.8998	0.4963	1.0000	-1.0020	1.3975	3.9644	0.5989	0.9977	0.3009	6.9976
N4SID	I=2	0.8996	0.4967	1.0000	-1.0021	1.3977	3.9545	0.6008	1.0046	0.2988	7.0015
N4SID	I=3	0.8996	0.4972	1.0000	-1.0026	1.3980	3.9572	0.5995	1.0000	0.3001	6.9849
N4SID	I=4	0.8998	0.4952	1.0000	-1.0034	1.3961	3.9471	0.5990	0.9994	0.3008	6.9930
N4SID	I=5	0.8996	0.4984	1.0000	-1.0012	1.4003	3.9714	0.5992	0.9993	0.3003	6.9830
N4SID	I=6	0.8996	0.4975	1.0000	-1.0016	1.3992	3.9659	0.5993	0.9993	0.3003	6.9872
ARMAX	nn=[1,2,1,0]	0.8995	0.4982	1	-1.0024	1.3987	3.9604	0.5995		0.3000	6.9769

Table 1. Mean of the parameters of 100 different estimated models: input type u^1, $N = 10\,000$.

Alg.	Par.	a	b	d	e	$p^d(1)$	$H^d(1)$	c	Δ	$p_s(1)$	$H^s(1)$
DSR	L=1	0.0054	0.0269	0	0.0677	0.0436	0.3020	0.0101	0.0153	0.0130	0.3170
DSR	L=2	0.0052	0.0261	0	0.0678	0.0434	0.3063	0.0099	0.0152	0.0121	0.3021
DSR	L=3	0.0054	0.0257	0	0.0678	0.0434	0.3190	0.0102	0.0152	0.0128	0.3211
DSR	L=4	0.0056	0.0259	0	0.0678	0.0435	0.3236	0.0105	0.0151	0.0135	0.3309
DSR	L=5	0.0059	0.0261	0	0.0677	0.0433	0.3328	0.0111	0.0151	0.0145	0.3447
N4SID	I=2	0.0052	0.0266	0	0.0678	0.0437	0.2959	0.0101	0.0153	0.0124	0.2994
N4SID	I=3	0.0051	0.0294	0	0.0691	0.0471	0.3423	0.0097	0.0152	0.0118	0.3009
N4SID	I=4	0.0053	0.0333	0	0.0702	0.0517	0.3920	0.0096	0.0152	0.0118	0.3155
N4SID	I=5	0.0055	0.0381	0	0.0704	0.0572	0.4456	0.0096	0.0151	0.0119	0.3236
N4SID	I=6	0.0058	0.0376	0	0.0699	0.0563	0.4554	0.0096	0.0152	0.0122	0.3391
ARMAX	nn=[1,2,1,0]	0.0051	0.0269	0	0.0662	0.0420	0.2925	0.0095		0.0115	0.2983

Table 2. Standard deviation of the parameters of 100 different estimated models: input type u^1, $N = 10\,000$.

Alg.	Par.	a	b	d	e	$p_d(1)$	$H^d(1)$	c	Δ	$p_s(1)$	$H^s(1)$
DSR	L=1	0.8861	0.5074	1.0000	-0.9922	1.4047	4.0874	0.6093	0.9961	0.2768	7.0385
DSR	L=2	0.8884	0.5023	1.0000	-0.9945	1.3991	3.9833	0.6045	0.9699	0.2839	6.9509
DSR	L=3	0.8886	0.5016	1.0000	-0.9971	1.3980	3.9781	0.6094	0.9476	0.2792	7.0266
DSR	L=4	0.8891	0.4998	1.0000	-0.9932	1.3985	4.0110	0.6128	0.9218	0.2763	7.1207
DSR	L=5	0.8902	0.4999	1.0000	-0.9940	1.3992	4.0741	0.6139	0.8991	0.2763	7.2495
N4SID	I=2	0.8865	0.5077	1.0000	-0.9920	1.4056	3.9629	0.6027	1.0102	0.2838	6.8462
N4SID	I=3	0.8865	0.5031	1.0000	-0.9954	1.3996	3.8955	0.6072	0.9953	0.2793	6.8840
N4SID	I=4	0.8859	0.5004	1.0000	-0.9974	1.3957	3.8534	0.6111	0.9916	0.2749	6.9022
N4SID	I=5	0.8852	0.4994	1.0000	-0.9973	1.3938	3.8399	0.6122	0.9844	0.2730	6.9141
N4SID	I=6	0.8840	0.5007	1.0000	-0.9307	1.3960	3.7969	0.6145	0.9801	0.2695	6.8643
ARMAX	nn=[1,2,1,0]	0.8864	0.5036	1.0000	-0.9935	1.3993	3.8848	0.5987	0.9971	0.2877	6.7799

Table 3. Mean of the parameters of 100 different estimated models: input type u^2, $N = 200$.

Alg.	Par.	a	b	d	e	$p_d(1)$	$H^d(1)$	c	Δ	$p_s(1)$	$H^s(1)$
DSR	L=1	0.0379	0.0957	0	0.0725	0.1320	2.3261	0.0101	0.0907	0.0884	2.4984
DSR	L=2	0.0351	0.0907	0	0.0716	0.1222	1.8766	0.0131	0.0689	0.0915	1.9967
DSR	L=3	0.0357	0.0931	0	0.0793	0.1234	1.8585	0.0158	0.0887	0.0964	2.0583
DSR	L=4	0.0363	0.0945	0	0.0844	0.1227	1.9096	0.0183	0.0872	0.1046	2.1884
DSR	L=5	0.0361	0.0977	0	0.0895	0.1212	1.9666	0.0210	0.0865	0.1095	2.4102
N4SID	I=2	0.0354	0.0956	0	0.0726	0.1325	1.9081	0.0104	0.0936	0.0844	1.9693
N4SID	I=3	0.0351	0.0987	0	0.0737	0.1374	1.8680	0.0106	0.0959	0.0839	2.0161
N4SID	I=4	0.0357	0.0992	0	0.0768	0.1375	1.8757	0.0115	0.0931	0.0871	2.0303
N4SID	I=5	0.0367	0.0994	0	0.0778	0.1347	1.9511	0.0119	0.0951	0.0904	2.1405
N4SID	I=6	0.0376	0.0983	0	0.0785	0.1331	1.8513	0.0123	0.0955	0.0974	2.0864
ARMAX	nn=[1,2,1,0]	0.0348	0.0921	0	0.0675	0.1225	1.7760	0.0723	0.0933	0.0867	1.8969

Table 4. Standard deviation of the parameters of 100 different estimated models: input type u^2, $N = 200$.

Alg.	Par.	a	b	d	e	$p_d(1)$	$H^d(1)$	c	Δ	$p_s(1)$	$H^s(1)$
DSR	L=1	0.8952	0.4927	1.0000	-1.0011	1.3923	4.0591	0.6013	1.0006	0.2939	7.0975
DSR	L=2	0.8938	0.4925	1.0000	-1.0021	1.3900	3.9881	0.5961	0.9879	0.2977	6.9587
DSR	L=3	0.8934	0.4923	1.0000	-1.0026	1.3893	3.9382	0.6000	0.9784	0.2934	6.9510
DSR	L=4	0.8931	0.4924	1.0000	-1.0022	1.3893	3.9376	0.6018	0.9695	0.2914	6.9572
DSR	L=5	0.8936	0.4930	1.0000	-1.0024	1.3903	3.9738	0.6030	0.9612	0.2906	7.0060
N4SID	I=2	0.8944	0.4923	1.0000	-1.0017	1.3906	3.9908	0.5944	1.0047	0.2999	6.9579
N4SID	I=3	0.8933	0.4921	1.0000	-1.0013	1.3894	3.9558	0.5959	0.9998	0.2974	6.9273
N4SID	I=4	0.8931	0.4915	1.0000	-1.0000	1.3893	3.9105	0.5966	0.9992	0.2965	6.8970
N4SID	I=5	0.8930	0.4916	1.0000	-1.0003	1.3892	3.9064	0.5967	0.9986	0.2963	6.8926
N4SID	I=6	0.8931	0.4905	1.0000	-0.9997	1.3886	3.8941	0.5972	0.9977	0.2960	6.8930
ARMAX	nn=[1,2,1,0]	0.8936	0.4927	1.0000	-1.0024	1.3898	3.9354	0.5952	0.9980	0.2983	6.8954

Table 5. Mean of the parameters of 100 different estimated models: input type u^3, $N = 500$.

Alg.	Par.	a	b	d	e	$p_d(1)$	$H^d(1)$	c	Δ	$p_s(1)$	$H^s(1)$
DSR	L=1	0.0277	0.0658	0	0.0768	0.1019	1.6168	0.0409	0.0644	0.0542	1.4790
DSR	L=2	0.0274	0.0660	0	0.0752	0.0996	1.6257	0.0442	0.0637	0.0580	1.4551
DSR	L=3	0.0268	0.0660	0	0.0765	0.0993	1.5491	0.0456	0.0638	0.0583	1.4129
DSR	L=4	0.0272	0.0661	0	0.0757	0.1000	1.5699	0.0470	0.0626	0.0600	1.4067
DSR	L=5	0.0275	0.0665	0	0.0760	0.1000	1.6054	0.0479	0.0640	0.0609	1.4341
N4SID	I=2	0.0269	0.0659	0	0.0759	0.1002	1.5547	0.0417	0.0653	0.0542	1.4149
N4SID	I=3	0.0275	0.0658	0	0.0755	0.0991	1.5903	0.0431	0.0642	0.0560	1.4462
N4SID	I=4	0.0268	0.0666	0	0.0748	0.0995	1.5164	0.0415	0.0645	0.0532	1.3767
N4SID	I=5	0.0267	0.0664	0	0.0752	0.0990	1.5091	0.0420	0.0645	0.0539	1.3780
N4SID	I=6	0.0262	0.0667	0	0.0760	0.1007	1.5003	0.0433	0.0660	0.0539	1.3668
ARMAX	nn=[1,2,1,0]	0.0260	0.0663	0	0.0747	0.0995	1.5207	0.0421	0.0643	0.0535	1.3581

Table 6. Standard deviation of the parameters of 100 different estimated models: input type u^3, $N = 500$.

Alg.	Par.	a	$p_d(1)$	$H^d(1)$	Δ	$p_s(1)$	$H^s(1)$
DSR	L=1	-	+	-	+	-	+
DSR	L=2	+	-	+	-	+	+
DSR	L=3	+	+	+	-	+	+
DSR	L=4	+	+	+	-	+	-
DSR	L=5	+	+	+	-	+	-

Table 7. Comparison of DSR and N4SID. + indicate better result with DSR: input type u^2, $N = 200$.

The DSR method is at least as good as N4SID when the number of samples are small (for this example and with $N = 200$ and $N = 500$ samples) see Tables 3 to 6.

It is also interesting to observe that the results from DSR are as good as the results from ARMAX, even for a simple SISO system. Note that the ARMAX method is based on iterative optimization but that the DSR method only is based on SVD and QR decompositions. Note also that the ARMAX method is rather complicated for MIMO systems while the DSR method is very simple.

5.2. *Example 2*

A two input two output system with the following model matrices is considered.

$$A = \begin{bmatrix} 1{\cdot}5 & 1 & 0{\cdot}1 \\ -0{\cdot}7 & 0 & 0{\cdot}1 \\ 0 & 0 & 0{\cdot}85 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{73}$$
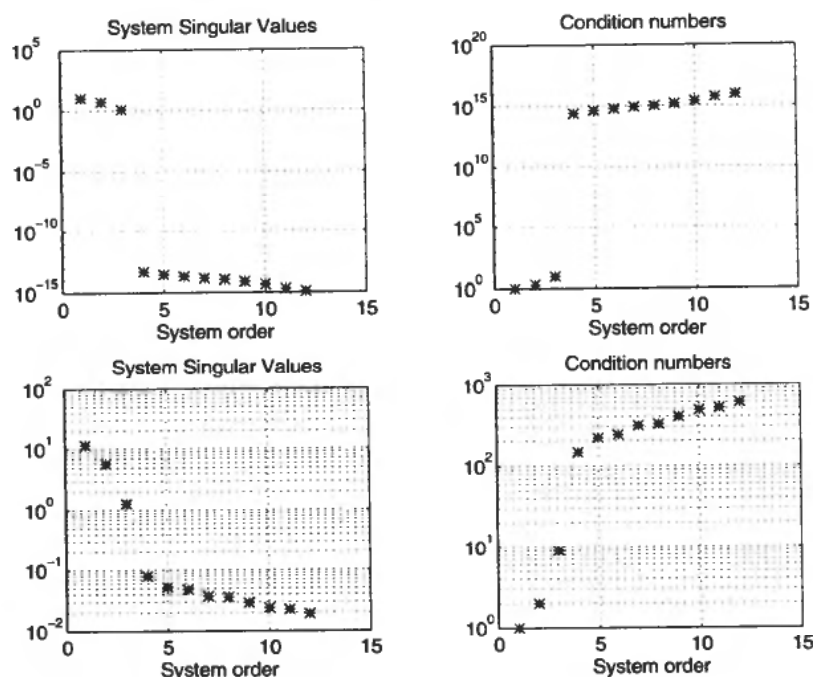
Figure 1.   Singular values and condition numbers for system order estimation of the system in
5.2. The noise free (deterministic) case with $\Delta = 0$, upper. The case with $\Delta = I$, lower.

$$D = \begin{bmatrix} 3 & 0 & -0.6 \\ 0 & 1 & 1 \end{bmatrix} \quad E = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \tag{74}$$

$$C = \begin{bmatrix} 0 & 0.1 \\ 0.1 & \\ 0 & 0.2 \end{bmatrix} \quad \Delta = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{75}$$

The system have a deterministic zero at -3, i.e. the system is non-minimum phase. The eigenvalues of A are 085, $0.75 \pm 0.3708i$.

The deterministic and stochastic gain matrices are given by

$$H^d(1) = \begin{bmatrix} 16 & 15 \\ 2.6667 & -2.5 \end{bmatrix} \quad H^s(1) = \begin{bmatrix} 2.5 & 4.7 \\ -0.25 & 1.1833 \end{bmatrix} \tag{76}$$

The algorithm gives 'exact' results when $\Lambda = 0$ (deterministic case). Hence, this result is not presented. The time series y_k, u_k was generated by simulating the model with one particular random noise process e_k with covariance Δ. The input was $u = [u^2 u^1]^T$. The DSR parameter was fixed to $L = 6$ and the number of samples was $N = 10\,000$. The system order can be estimated by inspection of the singular values or the condition numbers illustrated in Figure 1. The following estimates are obtained by DSR.

$$\hat{H}^d(1) = \begin{bmatrix} 15.6575 & 14.6168 \\ 2.5576 & -2.4188 \end{bmatrix} \quad \hat{H}^s(1) = \begin{bmatrix} 2.3625 & 4.5583 \\ -0.2436 & 1.1275 \end{bmatrix}$$

$$\hat{\Lambda} = \begin{bmatrix} 1.0531 & -0.0244 \\ -0.0244 & 0.9859 \end{bmatrix} \tag{77}$$

6. Remarks

We have in this paper presented a general method for the combined deterministic stochastic realization problem. The method needs only the input output time series for the construction of the state space model matrices. The pure deterministic as well as the pure stochastic realization problem pulls out as special cases of the method presented.

A general method for realization of deterministic linear state space models is presented in Section 3. The method is derived from a matrix valued state space model. Note that the correlation approach is a trick to transform the innovations model, (1) and (2), to a deterministic model. We can then present some important results in a deterministic environment. These results are also used to solve the combined deterministic stochastic realization problem in Section 4.

The method gives 'exact' results in the deterministic case provided the input signal is sufficiently 'rich' with respect to excitations, and does not require time series from several experiments. However, one can in some circumstances have sets of input output time series which are the result of process experiments. The time series can be collected into matrices and assumed to satisfy a matrix values state space model.

The system matrices (D, A) will usually be properly excited from the innovations process which is assumed to be standard white Gaussian noise (systems of opposite nature do of course exists). Assume that the input signal is poor with respect to excitations, this will only result in poor estimates of the (B, E) matrices.

REFERENCES

AOKI, M. (1991). Two complementary representations of multiple time series in state space innovation forms. Invited paper presented at seminar on *Recent Advances on Time Series Analysis and their Impact on Economic Forecasting*, Madrid, Spain, December 16–18.

DI RUSCIO, D. (1994). Methods for the identification of state space models from input and output measurements. Preprints *SYSID 94, The 10th IFAC Symposium on System Identification*, Copenhagen, July 4–6, pp. 2.377–2.84.

KALMAN, R. E., FALB, P. L. and ARBIB, M. A. (1969). *Topics in mathematical system theory*. (McGraw-Hill Book Company).

LJUNG, L. (1991). *System Identification Toolbox* (The Mathworks, Inc).

VAN OVERSCHEE, P. and DE MOOR, B. (1994). N4SID: subspace algorithms for the identification of combined deterministic stochastic systems. *Automatica*, Special Issue on Statistical Processing and Control. **30**, 75–94.