EXPLORING THE LINEAR RELATIONS IN THE ESTIMATION OF MATRICES B AND D IN SUBSPACE IDENTIFICATION METHODS

Catarina J. M. Delgado
Faculdade de Economia do Porto
Rua Dr. Roberto Frias, 4200-464 Porto - Portugal

P. Lopes dos Santos
Faculdade de Engenharia da Universidade do Porto
Rua Dr. Roberto Frias, 4200-465 Porto - Portugal

Keywords: System Identification, Subspace methods, State-space models, Linear Systems, Parameter estimation.

Abstract: In this paper we provide a different way to estimate matrices B and D, in subspace identification algorithms. The starting point was the method proposed by Van Overschee and De Moor (1996) – the only one applying subspace ideas to the estimation of those matrices. We have derived new (and simpler) expressions and we found that the method proposed by Van Overschee and De Moor (1996) can be rewritten as a weighted least squares problem, involving the future outputs and inputs.

1 INTRODUCTION

In subspace identification methods, there are two main steps: in the first step, a basis for the column space of a certain matrix, the extended observability matrix, is determined from the input-output data. The dimension of this subspace is equal to n, the order of the system to be identified. If we know the extended observability matrix, then we can estimate (explicitly or implicitly) the state sequence.

In the second main step of these algorithms, the system matrices are estimated. Several strategies exist, in order to estimate A and C and B and D, but we will focus our attention in the one proposed by Van Overschee and De Moor (van Overschee and de Moor, 1996), for the algorithm R-MOESP (Robust MOESP). We show in this paper that, for the estimation of B and D matrices, the R_MOESP method can be simplified, thus allowing a significant improvement on the numerical efficiency of the estimation procedure, without any loss of accuracy.

On the other hand, we manage to relate the R-MOESP algorithm to a different (geometric) approach (**?**), thus proving that these two different approaches are not that different – which can be seen as an extension of .the unifying theorem, for the estimation of B and D matrices step, in Subspace Identification Algorithms. This kind of relation has already been suggested for the matrices A and C (Chiuso and Picci, 2001) but has never been proposed for the estimation of matrices B and D, since the two approaches appear to be very different.

In this paper, we will focus our attention to the problem of estimating matrices B and D, knowing the extended observability matrix and matrices A and C. Therefore, the paper is organized as follows: in section 2, we introduce the subspace identification problem, notation, main concepts behind subspace methods and we describe the technique proposed by Van Overschee and De Moor (van Overschee and de Moor, 1996) for the estimation of B and D. In section 3, we provide new expressions for the estimation of the input matrices and in section 4 we show that the technique presented by Van Overschee and De Moor (van Overschee and de Moor, 1996) is merely a projection on the null space a certain matrix. Finally, in section 5, some simulation results are introduced and, in secton 6, the conclusions are presented.

2 BACKGROUND

2.1 Subspace Identification Problem

Subspace Identification algorithms aim to estimate, from measured input / output data sequences ($\{u_k\}$ and $\{y_k\}$, respectively), the system described by:

$$\left\{ \begin{array}{l} x_{k+1} = Ax_k + Bu_k + Ke_k \\ y_k = Cx_k + Du_k + e_k \end{array} \right. \tag{1}$$

$$E\left[e_p e_q^T\right] = R_e \delta_{pq} \geqslant 0 \tag{2}$$

Delgado C. and dos Santos P. (2004).
EXPLORING THE LINEAR RELATIONS IN THE ESTIMATION OF MATRICES B AND D IN SUBSPACE IDENTIFICATION METHODS.
In *Proceedings of the First International Conference on Informatics in Control, Automation and Robotics*, pages 168-175
DOI: 10.5220/0001146701680175

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{l\times n}$, $D \in \mathbb{R}^{l\times m}$, $K \in \mathbb{R}^{n\times l}$ and $x_k \in \mathbb{R}^n$. The sequence $\{e_k\} \in \mathbb{R}^l$ is a white noise stochastic process and the input data sequence is assumed to be a persistently exciting quasi-stationary deterministic sequence(Ljung, 1987)

2.2 Block Hankel Matrices

The notation used will be based on the notation of Van Overschee and De Moor(van Overschee and de Moor, 1996): U and Y are two block Hankel matrices built with $2i$ row-blocks and j column-blocks (for N, the number of measurements, greater or equal than $2i + j - 1$:

$$U = \begin{bmatrix} U_{(1)} \\ ... \\ U_{(2i)} \end{bmatrix} \in \mathbb{R}^{2mi\times j}, Y = \begin{bmatrix} Y_{(1)} \\ ... \\ Y_{(2i)} \end{bmatrix} \in \mathbb{R}^{2li\times j}$$

where $U(k)$ and $Y(k)$ are the $k-th$ row-blocks of, respectively, U and Y. Matrix U can be partitioned as:

$$\begin{aligned} U_p &= \begin{bmatrix} U_{(1)} \\ ... \\ U_{(i)} \end{bmatrix}, U_f = \begin{bmatrix} U_{(i+1)} \\ ... \\ U_{(2i)} \end{bmatrix} \\ U_{p+} &= \begin{bmatrix} U_{(1)} \\ ... \\ U_{(i+1)} \end{bmatrix}, U_{f-} = \begin{bmatrix} U_{(i+2)} \\ ... \\ U_{(2i)} \end{bmatrix} \end{aligned}$$

where the subscripts p and f denote "past" and "future", respectively.

The same happens to matrix Y and to the input/output data matrices: $H = \begin{bmatrix} U \\ Y \end{bmatrix}$, $H_p = \begin{bmatrix} U_p \\ Y_p \end{bmatrix}$, $H_{p+} = \begin{bmatrix} U_{p+} \\ Y_{p+} \end{bmatrix}$, $HU = \begin{bmatrix} H_p \\ U_f \end{bmatrix}$ and $H^+U^- = \begin{bmatrix} H_{p+} \\ U_{f-} \end{bmatrix}$.

When the input-output data is organized into matrices with this special block Hankel structure, then (1) can be written as

$$Y_f = \Gamma_i X_i + H_i^d U_f + E_f \quad (3)$$

$$Y_{f-} = \Gamma_{i-1} X_{i+1} + H_{i-1}^d U_{f-} + E_{f-} \quad (4)$$

and also as

$$\begin{bmatrix} \widehat{X}_{i+1} \\ Y_{(i)} \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} \widehat{X}_i \\ U_{(i)} \end{bmatrix} + \begin{bmatrix} \rho_w \\ \rho_v \end{bmatrix} \quad (5)$$

where:

1. Matrix $\widehat{X}_i$ is the state sequence generated by a bank of Kalman filters, working in parallel on each of the columns of the block Hankel matrix of past inputs and outputs:
$$\begin{aligned} \widehat{X}_i &= \begin{bmatrix} \widehat{x}_{i+1|1} & \dots & \widehat{x}_{N-i+1|N-2i+1} \end{bmatrix} \\ \widehat{X}_{i+1} &= \begin{bmatrix} \widehat{x}_{i+2|2} & \dots & \widehat{x}_{N-i+2|N-2i+2} \end{bmatrix} \end{aligned}$$
2. $\Gamma_i \in \mathbb{R}^{li\times n}$ is the extended observability matrix (since i> n), where the subscript i denotes the number of row-blocks
3. $H_i^d \in \mathbb{R}^{li\times mi}$ is a block Toeplitz matrix, built with Markov parameters
4. $\rho_{wk} = \widehat{x}_{k+1} - \widehat{A}\widehat{x}_k - \widehat{B}u_k$, $\rho_{vk} = y_k - \widehat{C}\widehat{x}_k - \widehat{D}u_k$ and
$$\begin{bmatrix} \rho_w \\ \rho_v \end{bmatrix} = \begin{bmatrix} \rho_{wi} & \cdots & \rho_{wi+j-1} \\ \rho_{vi} & \cdots & \rho_{vi+j-1} \end{bmatrix}$$

2.3 The projection theorem

The main idea behind the subspace theory is stated in the "projection theorem" (Van Overschee and de Moor, 1996): given (3) and (4) then, under certain conditions, there is a connection between an estimated kalman filter state sequence and the orthogonal projection of the row space of Y_f (future outputs) into the row space of the past inputs, past outputs and future inputs row space U_f:

$$Z_i = Y_f /_{HU} = \Gamma_i \widehat{X}_i + H_i^d U_f \quad (6)$$

$$Z_{i+1} = Y_{f-} /_{H^+U^-} = \Gamma_{i-1} \widehat{X}_{i+1} + H_{i-1}^d U_{f-} \quad (7)$$

where A/B denotes an orthogonal projection of the row space of A into the row space of B and the state sequences are given by

$$\widehat{X}_i = \begin{bmatrix} \theta_1 & \theta_2 & \theta_3 \end{bmatrix} \begin{bmatrix} \widehat{X}_0 \\ U_p \\ Y_p \end{bmatrix} \quad (8)$$

$$\widehat{X}_{i+1} = \begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 \end{bmatrix} \begin{bmatrix} \widehat{X}_0 \\ U_{p+} \\ Y_{p+} \end{bmatrix} \quad (9)$$

where $\theta_1, \theta_2, \theta_3, \alpha_1, \alpha_2, \alpha_3$ are functions of the system matrices, and $\widehat{X}_0$ a function of U. In fact, as the inputs are possibly correlated, one can obtain information about $\widehat{X}_0$ from the inputs U, by projecting the initial state sequence X_0 (exact but unknown) into U: $\widehat{X}_0 = X_0/U$. A similar relation has been establish between a second estimated state sequence $\widetilde{X}_i$ and the oblique projection of the row space of Y_f (future outputs), along the future inputs row space U_f, into the row space of the past inputs and outputs H_p:

$$O_i = Y_f /_{U_f} H_p = \Gamma_i \widetilde{X}_i \quad (10)$$

$$O_{i+1} = Y_{f-} /_{U_{f-}} H_{p+} = \Gamma_{i-1} \widetilde{X}_{i+1} \quad (11)$$

There is a slight difference between Z_i and O_i. In fact, O_i can be computed from Z_i by just ignoring the information given by U_f. The consequences are clear: part of the information required to estimate $\widehat{X}_0$

is no longer available so, the estimated state sequence $\left(\widetilde{X}_i\right)$ is different from $\widehat{X}_i$. Although $\widetilde{X}_i$ and $\widehat{X}_i$ are not the same estimates, they are still very similar and, actually, under some special conditions ($i \to \infty$ or $\{u_k\}$ is white noise or the system is purely deterministic) they are the same. This approximation of the state sequences is used to obtain the more elegant and simple algorithm presented in next section. Unlike the algorithm that considers the "exact" Kalman state estimates by implementing some orthogonal projections (unbiased for $j \to \infty$), this approximate algorithm is biased for finite i, except under certain special cases (Van Overschee and De Moor, 1996)

2.4 Subspace Identification Algorithms

There are two main steps in subspace identification algorithms:

1. determine the model order n and estimate the extended observability matrix through the singular value decomposition of a weighted oblique projection, $W_L O_i W_R$, and
2. solve a least squares problem, in order to obtain the state space matrices:

$$min \left\| \begin{bmatrix} \widehat{X}_{i+1} \\ Y_{(i)} \end{bmatrix} - \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} \widehat{X}_i \\ U_{(i)} \end{bmatrix} \right\|_F^2 \qquad (12)$$

where $\|\Xi\|_F$ denotes the Frobenius norm of matrix Ξ.

In the first step, since Van Overschee and De Moor (van Overschee and de Moor, 1996) proved that (10) is a valid approximation, then one can estimate the observability matrix from the singular value decomposition of O_i. Since $rank(\Gamma_i) = n$ (we assume $\{A, C\}$ observable), then $rank(O_i)$ should also be n, and

$$\begin{aligned} O_i &= USV^T = \\ &= \begin{bmatrix} U_n & U_{n^\perp} \end{bmatrix} \begin{bmatrix} S_n & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_n^T \\ V_{n^\perp}^T \end{bmatrix} = \\ &= U_n S_n V_n^T \end{aligned} \qquad (13)$$

where $O_i \in \mathbb{R}^{li \times j}$, $S \in \mathbb{R}^{li \times j}$ (diagonal matrix with the singular values of O_i) , $U \in \mathbb{R}^{li \times li}$, $V \in \mathbb{R}^{j \times j}$ (U and V are orthonormal matrices), $U_n \in \mathbb{R}^{li \times n}$, $S_n \in \mathbb{R}^{n \times n}$, $V_n^T \in \mathbb{R}^{n \times j}$, $rank(O_i) = n$. The order n of the system should, therefore, be determine by the number of the nonzero singular values of O_i , $dim(S_n)$ (van Overschee and de Moor, 1996). However, in many practical situations, when the measurements are noise corrupted, it can not be straightforward to distinguish the "nonzero" from the "zero" singular values – one must then make a decision by comparing the values or by assuming different orders and comparing simulation errors.

As the column spaces of Γ_i and $U_n S_n^{1/2}$ are the same, we compute

$$\Gamma_i = U_n S_n^{1/2} \qquad (14)$$

and then $\underline{\Gamma}_i$=Γ_{i-1}, by removing the last l rows from Γ_i.

In CVA and $MOESP$ approaches, W_L and W_R are given by

	W_L	W_R
CVA	$\left(Y_f \Pi_{U_f^\perp} Y_f^T\right)^{-1/2}$	$\Pi_{U_f^\perp}$
$MOESP$	I_{li}	$\Pi_{U_f^\perp}$

where $\Pi_{U_f^\perp} = \left(I_j - U_f^+ U_f\right)$.

Knowing an estimate of Γ_i $\left(\Gamma_i = U_n S_n^{1/2}\right)$, matrices A and C are obtained by solving a linear equation, in a least squares sense:

$$\begin{bmatrix} \Gamma_{i-1}^+ Z_{i+1} \\ Y_{(i)} \end{bmatrix} = \begin{bmatrix} A & \\ C & K_{BD} \end{bmatrix} \begin{bmatrix} \Gamma_i^+ Z_i \\ U_f \end{bmatrix} + \rho \qquad (15)$$

Lopes dos Santos and Martins de Carvalho (dos Santos and de Carvalho, 2004) have shown that these estimates and the estimates of A and C produced by the shift-invariant property of Γ_i are the same. The matrix K_{BD} is then used to estimate B and D, since

$$\begin{aligned} K_{BD} &= \begin{bmatrix} K_A \\ K_C \end{bmatrix} = \begin{bmatrix} K_1 & \dots & K_i \end{bmatrix} = \\ &= \begin{bmatrix} B & \Gamma_{i-1}^+ H_{i-1}^d \\ D & 0 \end{bmatrix} - \begin{bmatrix} A \\ C \end{bmatrix} \Gamma_i^+ H_i^d = \\ &= \begin{bmatrix} N_1 \begin{bmatrix} D \\ B \end{bmatrix} & \dots & N_i \begin{bmatrix} D \\ B \end{bmatrix} \end{bmatrix} \end{aligned} \qquad (16)$$

Equation (16) can be rewritten as:

$$\begin{bmatrix} K_1 \\ \dots \\ K_i \end{bmatrix} = \begin{bmatrix} N_1 \\ \dots \\ N_i \end{bmatrix} \begin{bmatrix} D \\ B \end{bmatrix} \qquad (17)$$

and B and D estimated in the least squares sense. Lopes dos Santos and Martins de Carvalho (dos Santos and de Carvalho, 2003) have shown that K_A can be written as $K_A = K_p K_C = \left(-A\left(\overline{\Gamma}_i^T \overline{\Gamma}_i\right)^{-1} C^T\right) K_C$. Therefore, we can work only with

$$\begin{aligned} K_{C(B,D)} &= \begin{bmatrix} K_p \\ I_l \end{bmatrix}^+ K_{BD} = \\ &= \begin{bmatrix} N_{C1} \begin{bmatrix} D \\ B \end{bmatrix} & \dots & N_{Ci} \begin{bmatrix} D \\ B \end{bmatrix} \end{bmatrix} \end{aligned} \qquad (18)$$

where

$$\begin{aligned} N_{C1} &= \left[\begin{array}{cccc} I_l - L_{C1} & -L_{C2} & \dots & -L_{Ci} \end{array}\right] G_i \\ N_{Ck} &= \left[\begin{array}{cccc} -L_{Ck} & \dots & -L_{Ci} & 0 \end{array}\right] G_i \ (k > 1) \\ L_A &= \left[\begin{array}{cccc} L_{A1} & L_{A2} & \dots & LA \end{array}\right] = A\Gamma_i^+ \\ L_C &= \left[\begin{array}{cccc} L_{C1} & L_{C2} & \dots & L_{Ci} \end{array}\right] = C\Gamma_i^+ \\ G_i &= \left[\begin{array}{cc} I_l & 0 \\ 0 & \Gamma_{i-1} \end{array}\right] \end{aligned}$$

Equation (18) can be rewritten as:

$$\left[\begin{array}{c} K_{C1} \\ \dots \\ K_{Ci} \end{array}\right] = \left[\begin{array}{c} N_{C1} \\ \dots \\ N_{Ci} \end{array}\right] \left[\begin{array}{c} D \\ B \end{array}\right] \quad (19)$$

and B and D estimated in the least squares sense.

When matrix $\left(U_f U_f^T\right)$ is almost a singular matrix, (17) provides bad results. It is better then to avoid the explicit estimation of K_{BD} and obtain B and D directly from:

$$\begin{aligned} P &= \left(\left[\begin{array}{c} \Gamma_{i-1}^+ Z_{i+1} \\ Y_{(i)} \end{array}\right] - \left[\begin{array}{c} A \\ C \end{array}\right] \Gamma_i^+ Z_i\right) \quad (20) \\ &= \left[\begin{array}{cccc} K_1 & K_2 & \dots & K_i \end{array}\right] U_f = \\ &= \sum_{k=1}^{i} K_k U_k = \sum_{K=1}^{i} N_k \left[\begin{array}{c} D \\ B \end{array}\right] U_k \end{aligned}$$

where $U_k \in \mathbb{R}^{m\times j}$ is the k-th block row of U_f.

In order to determine D and B, one has to apply the vector operation, $vec(.)$, and the Kronecker product, $\otimes$, to (20). The new equation can now be solve in the least squares sense:

$$vec(P) = \left(\sum_{k=1}^{i} \left(U_k^T \otimes N_k\right)\right) vec\left(\left[\begin{array}{c} D \\ B \end{array}\right]\right) \quad (21)$$

3 ESTIMATION OF MATRICES B AND D

In this section, we will analyze both sides of (20). First, we will consider matrix P and the relation between the orthogonal projections Z_i and Z_{i+1}, and then provide a new expression for P. Then, we will consider matrix K_{BD}, also providing a new expression for this matrix. Finally, we will relate P and K_{BD}.

3.1 The orthogonal projections

Theorem 1 *Given matrices* $A \in \mathbb{R}^{n\times m}$, $B \in \mathbb{R}^{b\times m}, C \in \mathbb{R}^{c\times m}$ *then*

$$\begin{aligned} A/_{\left[\begin{array}{c} B \\ C \end{array}\right]} - A/_C &= A/_{\left(B/_{C^\perp}\right)} = \quad (22) \\ &= A/_C B \Pi_{C^\perp} \quad (23) \end{aligned}$$

Proof. We can define the first orthogonal projection as

$$\begin{aligned} A/_{BC} &= A\left[B^T \ C^T\right] \left[\begin{array}{cc} BB^T & BC^T \\ CB^T & CC^T \end{array}\right]^{-1} \left[\begin{array}{c} B \\ C \end{array}\right] = \\ &= A\left[B^T \ C^T\right] \left[\begin{array}{cc} \Delta^{-1} & \gamma_1 \\ \gamma_2 & \varphi^{-1} \end{array}\right] \left[\begin{array}{c} B \\ C \end{array}\right] = \\ &= \left[\theta_B \ \theta_C\right] \left[\begin{array}{c} B \\ C \end{array}\right] \end{aligned}$$

where, by the Matrix Inversion Lemma (Kailath, 1980),

$$\begin{aligned} \Delta^{-1} &= \left(B\Pi_{C^\perp} B^T\right)^{-1} \\ \varphi^{-1} &= \left(CC^T\right)^{-1} + \left(CC^T\right)^{-1} CB^T \\ &\quad \times \left(B\Pi_{C^\perp} B^T\right)^{-1} BC^T \left(CC^T\right)^{-1} \\ \gamma_1 &= -\left(B\Pi_{C^\perp} B^T\right)^{-1} BC^T \left(CC^T\right)^{-1} \\ \gamma_2 &= -\left(CC^T\right)^{-1} CB^T \left(B\Pi_{C^\perp} B^T\right)^{-1} \end{aligned}$$

On the other hand,

$$A/_{BC} = A/_B C + A/_C B \quad (24)$$

where

$$A/_B C = \theta_C C = A\left(B^T \gamma_1 + C^T \varphi^{-1}\right) C \quad (25)$$

and, after some manipulation,

$$A/_B C = \left(A - A/_C B\right) \Pi_C \quad (26)$$

Therefore,

$$A/_{BC} = A/_C + A/_C B \left(I - \Pi_C\right) \quad (27)$$

■

Corollary 2 *Given* $Z_i = Y_f/_{HU}$ *and* $\overline{Z}_{i+1} = Y_f/_{H^+U^-}$*, then*

$$\overline{Z}_{i+1} = Z_i + \triangle Z_i = Z_i + Y_f/_{\left(Y_{(i)}/_{HU^\perp}\right)} \quad (28)$$

Proof. Since the rowspace of H^+U^- is spanned by the rows of $\left[\begin{array}{c} HU \\ Y(i) \end{array}\right]$, we can explore (27), with $A = Y_f$, $B = HU$, $C = Y_{(i)}$.

Another way to prove this is through the LQ decomposition of H^+U^- :

$$\left[\begin{array}{c} UH \\ Y_{(i)} \end{array}\right] = \left[\begin{array}{ccc} L_{UH} & 0 & 0 \\ L_{Y1} & L_{Y2} & 0 \end{array}\right] \left[\begin{array}{c} Q_{LUH} \\ Q_{LY} \\ Q_{L^\perp} \end{array}\right]$$

where

$$\begin{aligned} Y_f.Q_L^T &= \begin{bmatrix} Y_{(i)} \\ Y_{f-} \end{bmatrix} Q_L^T = \\ &= [B_{UH} \;\; B_Y \;\; B_{UHYp}] \\ &= \begin{bmatrix} b_{UH} & b_Y & b_{UHYp} \\ B_{mUH} & B_{mY} & B_{mUHYp} \end{bmatrix} \end{aligned}$$

In fact,

$$Z_i = \begin{bmatrix} b_{UH}Q_{LUH} \\ B_{mUH}Q_{LUH} \end{bmatrix} \quad (29)$$

$$\overline{Z}_{i+1} = \begin{bmatrix} b_{UH} & b_Y \\ B_{mUH} & B_{mY} \end{bmatrix} \begin{bmatrix} Q_{LUH} \\ Q_{LY} \end{bmatrix} \quad (30)$$

$$\begin{aligned} \triangle Z_i &= \begin{bmatrix} b_Y Q_{LY} \\ B_{mY} Q_{LY} \end{bmatrix} = \quad (31) \\ &= Y_f/_{\left(Y_{(i)}/_{UH^\perp}\right)} = \overline{Z}_{i+1} - Z_i \end{aligned}$$

■

3.2 Simplifying matrix P

Theorem 3 *Given (1), where* $\{A,C\}$ *is observable and A is non-singular, then*

$$P = \begin{bmatrix} \Gamma_{i-1}^+ Z_{i+1} \\ Y_{(i)} \end{bmatrix} - \begin{bmatrix} A \\ C \end{bmatrix} \Gamma_i^+ Z_i \quad (32)$$

can be rewritten as:

$$P = M\, Z_i + N \triangle Z_i \quad (33)$$

where

$$Zi = Y_f/_{UH} = Y_f/_{\begin{bmatrix} H_P \\ U_f \end{bmatrix}} \quad (34)$$

$$\triangle Z_i = Y_f/_{\widehat{Y}_{(i)}} = Y_f/_{\left(Y_{(i)}/_{UH^\perp}\right)} \quad (35)$$

and

$$N = \begin{bmatrix} 0 & \Gamma_{i-1}^+ \\ I_l & 0 \end{bmatrix} \quad (36)$$

$$M = \begin{bmatrix} -AP_iC^T & (P_{i-1} - AP_iA^T)\Gamma_{i-1}^T \\ I_l - CP_iC^T & -CP_iA^T\Gamma_{i-1}^T \end{bmatrix} \quad (37)$$

Proof. Since we assume $\{A, C\}$ to be observable, Γ_i and Γ_{i-1} are full column rank matrices and we can replace their pseudo-inverse expressions with

$$\Gamma_{i-i}^+ = \left(\Gamma_{i-1}^T\Gamma_{i-1}\right)^{-1}\Gamma_{i-1}^T = P_{i-1}\Gamma_{i-1}^T \quad (38)$$

$$\Gamma_i^+ = \left(\Gamma_i^T\Gamma_i\right)^{-1}\Gamma_i^T = P_i\Gamma_i^T \quad (39)$$

On the other hand, if A is a non-singular matrix, then, by the shift-invariance property of Γ_i,

$$A = \Gamma_{i-1}^+\overline{\Gamma}_i \quad (40)$$

with $\overline{\Gamma}_i = [\,0 \quad I_l\,]\,\Gamma_i$, and

$$P_{i-1} = A\overline{P}_iA^T \quad (41)$$

Therefore,

$$\begin{aligned} P &= \begin{bmatrix} P_A \\ P_C \end{bmatrix} = \quad (42) \\ &= \begin{bmatrix} P_{i-1}\Gamma_{i-1}^T Z_{i+1} - AP_i\Gamma_i^T Z_i \\ Y_{(i)} - CP_i\Gamma_i^T Z_i \end{bmatrix} = \\ &= \begin{bmatrix} 0 & P_{i-1}\Gamma_{i-1}^T \\ I_l & 0 \end{bmatrix} \overline{Z}_{i+1} - \begin{bmatrix} A \\ C \end{bmatrix} P_i\Gamma_i^T Z_i \end{aligned}$$

$$\begin{aligned} P_C &= [I_l \;\; 0]\,\overline{Z}_{i+1} - CP_i\Gamma_i^T Z_i = \quad (43) \\ &= [I_l \;\; 0]\,\overline{Z}_{i+1} - \left[CP_iC^T \;\; CP_iA^T\Gamma_{i-1}^T\right] Z_i \\ &= \left[I_l - CP_iC^T \quad -CP_iA^T\Gamma_{i-1}^T\right] Z_i + \\ &\quad + [I_l \;\; 0] \triangle Z_i \end{aligned}$$

where $[I_l \;\; 0] \triangle Z_i = \triangle ZY_i = Y_{(i)}/_{UH^\perp}$, and

$$\begin{aligned} P_A &= \left[0 \;\; P_{i-1}\Gamma_{i-1}^T\right] \overline{Z}_{i+1} - AP_i\Gamma_i^T Z_i = \quad (44) \\ &= \left[-AP_iC^T \;\; (P_{i-1} - AP_iA^T)\Gamma_{i-1}\right] Z + \\ &\quad + \left[0 \;\; \Gamma_{i-1}^+\right] \triangle Z_i \end{aligned}$$

Then,

$$\begin{aligned} P &= \begin{bmatrix} -AP_iC^T & (P_{i-1} - AP_iA^T)\Gamma_{i-1}^T \\ I_l - CP_iC^T & -CP_iA^T\Gamma_{i-1}^T \end{bmatrix} Z_i \\ &\quad + \begin{bmatrix} 0 & \Gamma_{i-1}^+ \\ I_l & 0 \end{bmatrix} \triangle Z_i \end{aligned}$$

■

3.3 Simplifying K_{BD}

Theorem 4 *Given matrix* K_{BD}*, defined in (16), then*

$$K_{BD} = MH_i^d \quad (45)$$

where M was introduced in (37).

Proof. As mentioned before,

$$K_{BD} = \begin{bmatrix} B & \Gamma_{i-1}^+ H_{i-1}^d \\ D & 0 \end{bmatrix} - \begin{bmatrix} A \\ C \end{bmatrix} \Gamma_i^+ H_i^d \quad (46)$$

Since H_i^d can be given by

$$H_i^d = \begin{bmatrix} D & 0 \\ \Gamma_{i-1}B & H_{i-1}^d \end{bmatrix} = \begin{bmatrix} \varphi_i & \begin{matrix} 0 \\ H_{i-1}^d \end{matrix} \end{bmatrix}$$

the expression (46) can be expressed as a function of

φ_i and H^d_{i-1}:

$$K_{BD} = \begin{bmatrix} \mathcal{K}_{11} & \mathcal{K}_{12} \\ \mathcal{K}_{21} & \mathcal{K}_{22} \end{bmatrix} = \\ = \begin{bmatrix} Ma & Mb \\ Mc & Md \end{bmatrix} \begin{bmatrix} D & 0 \\ \Gamma_{i-1}B & H^d_{i-1} \end{bmatrix}$$

where

$$\mathcal{K}_{11} = B - AP_i\Gamma_i^T\varphi_i$$
$$\mathcal{K}_{12} = \left(P_{i-1}\Gamma_{i-1}^T - AP_i\overline{\Gamma}_i^T\right)H^d_{i-1}$$
$$\mathcal{K}_{21} = D - CP_i\Gamma_i^T\varphi_i$$
$$\mathcal{K}_{22} = CP_i\overline{\Gamma}_i^T H^d_{i-1}$$

$$M_a = -AP_iC^T$$
$$M_b = P_{i-1}\Gamma_{i-1}^T - AP_i\overline{\Gamma}_i^T = (P_{i-1} - AP_iA^T)\Gamma_{i-1}^T$$
$$M_c = I_l - CP_iC^T$$
$$M_d = -CP_i\overline{\Gamma}_i^T = -CP_iA^T\Gamma_{i-1}^T$$

A different way to prove this result can be found in (Delgado et al., 2004). ■

3.4 The estimation of B and D

Theorem 5 *The equation (20) can be written as*

$$MY_f = MH_i^dU_f \tag{47}$$

where

$$M = \begin{bmatrix} A(\overline{P}_i - P_i) \\ -CP_i \end{bmatrix}(\Gamma_i^T + \\ [-C^T(CP_iC^T)^{-1} \;\; 0]) \tag{48}$$
$$= \begin{bmatrix} -AP_iC^T & (P_{i-1} - AP_iA^T)\Gamma_{i-1}^T \\ I_l - CP_iC^T & -CP_iA^T\Gamma_{i-1}^T \end{bmatrix}$$

Proof. We start by assuming that M can be written as:

$$M = \begin{bmatrix} A\left(\overline{P}_i - P_i\right) \\ -CP_i \end{bmatrix}\begin{bmatrix} N_1 & N_2 \end{bmatrix} \tag{49}$$

and then will prove that

$$N_1 = C^T\left(I_l - (CP_iC^T)^{-1}\right) \tag{50}$$
$$N_2 = A^T\Gamma_{i-1}^T$$

In fact, when A is a non-singular matrix,

$$P_{i-1} - AP_iA^T = A\left(\overline{P}_i - P_i\right)A^T, \tag{51}$$

and, therefore,

$$\begin{bmatrix} A\left(\overline{P}_i - P_i\right) \\ -CP_i \end{bmatrix} N_2 = \begin{bmatrix} (P_{i-1} - AP_iA^T)\Gamma_{i-1}^T \\ -CP_iA^T\Gamma_{i-1}^T \end{bmatrix}$$

On the other hand, since

$$-CP_iN_1 = -CP_iC^T\left(I_l - (CP_iC^T)^{-1}\right) = \\ = -CP_iC^T + I_l$$

and

$$(-AP_iC^T) = (-A\overline{P}_iC^T)(I_l - CP_iC^T)$$
$$A\left(\overline{P}_i - P_i\right) = (-A\overline{P}_iC^T)(-CP_i)$$

we obtain

$$A\left(\overline{P}_i - P_i\right)N_1 = (-A\overline{P}_iC^T)(-CP_i)N_1 = \\ = (-A\overline{P}_iC^T)(I_l - CP_iC^T) = \\ = -AP_iC^T$$

■

If we consider all the previous results, we can see that knowing estimates of matrices A, C and Γ_i^T allows the estimation of H_i^d, in the least squares sense, and therefore, the estimation of B and D.

4 ORTHOGONAL PROJECTION INTO THE NULLSPACE OF Γ_i^T

An analysis of matrix M shows us that,

$$M = \Upsilon CP_i\left(\Omega + \Gamma_i^T\right) \tag{52}$$

where

$$\Upsilon = \begin{bmatrix} A\overline{P}_iC^T \\ -I_l \end{bmatrix} \tag{53}$$
$$\Omega = \begin{bmatrix} -C^T(CP_iC^T)^{-1} & 0 \end{bmatrix} \tag{54}$$

is such that

$$MY_f = M(\Gamma_iX_f + H_i^dU_f + E_f) = \tag{55}$$
$$= MH_i^dU_f + ME_f \tag{56}$$

which means that

$$rowspace(M) \perp colspace(\Gamma_i) \tag{57}$$

In fact,

$$M.\Gamma_i = \Upsilon CP_i\left(\Omega + \Gamma_i^T\right)\Gamma_i = \tag{58}$$
$$= \Upsilon\left(CP_i\Omega\Gamma i + CP_i\Gamma_i^T\Gamma_i\right) =$$
$$= \Upsilon\left(-CP_iC^T(CP_iC^T)^{-1}C + C.I_n\right) =$$
$$= \Upsilon(-C + C) = 0$$

This means that the rows of

$$CP_i\left(\begin{bmatrix} -C^T(CP_iC^T)^{-1} & 0 \end{bmatrix} + \Gamma_i^T\right)$$

are orthogonal to the columns of Γ_i ■

5 SIMULATION RESULTS

It was considered the following system with two inputs and two outputs, represented in the forward innovation model:

$$\begin{cases} x_{k+1} = Ax_k + Bu_k + Ke_k \\ y_k = Cx_k + Du_k + e_k \end{cases} \quad (59)$$

$$E\left[e_p e_q^T\right] = R_e \delta_{pq} \geqslant 0 \quad (60)$$

where:

$$A = \begin{bmatrix} 0.603 & 0.603 & 0 & 0 \\ -0.603 & 0.603 & 0 & 0 \\ 0 & 0 & -0.603 & -0.603 \\ 0 & 0 & 0.603 & -0.603 \end{bmatrix}$$

$$B = \begin{bmatrix} 1.1650 & -0.6965 \\ 0.6268 & 1.6961 \\ 0.0751 & 0.0591 \\ 0.3516 & 1.7971 \end{bmatrix}$$

$$C = \begin{bmatrix} 0.2641 & -1.4462 & 1.2460 & 0.5774 \\ 0.8717 & -0.7012 & -0.6390 & -0.3600 \end{bmatrix}$$

$$D = \begin{bmatrix} -0.1356 & -1.2704 \\ -1.3493 & 0.9846 \end{bmatrix}$$

$$K = \begin{bmatrix} 0.2820 & -0.3041 \\ -0.7557 & 0.0296 \\ 0.1919 & 0.1317 \\ -0.3797 & 0.6538 \end{bmatrix}$$

$$R_e = \begin{bmatrix} 0.1253 & 0.1166 \\ 0.1166 & 0.2170 \end{bmatrix}$$

As inputs, two white noise sequences with 1000 samples were generated.

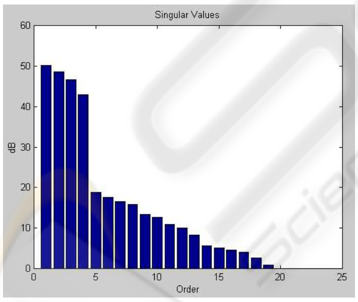

Figure 1: Singular values (MOESP approach).

Comparing the results obtained with the proposed and the original algorithms, we have similar values, for the matrices, frequency responses (figures 2 and 3) and for the simulation errors (van Overschee and de Moor, 1996):

$$\epsilon_{old} = \begin{bmatrix} 15.0104 \\ 14.6752 \end{bmatrix} (\%)$$

$$\epsilon_{new} = \begin{bmatrix} 15.0027 \\ 14.6731 \end{bmatrix} (\%)$$

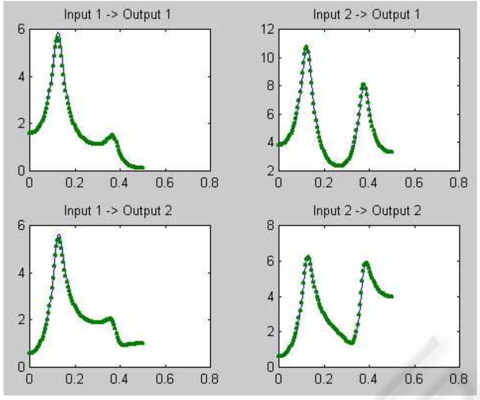

Figure 2: The frequency response of the original and the estimated system (MOESP approach).

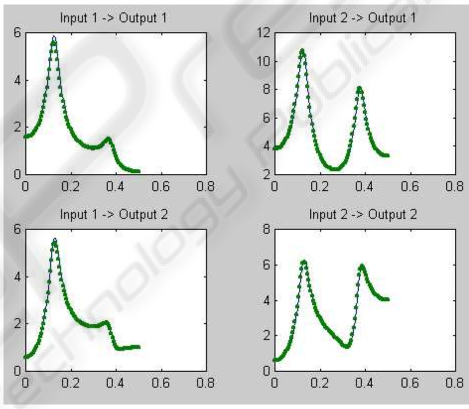

Figure 3: The frequency response of the original and the estimated system (proposed approach, Van Overschee e De Moor-based).

6 CONCLUSIONS

In this paper we describe an alternative approach for the estimation of matrices B and D in subspace identification. If we consider the methods used nowadays, both "simulation error method" and ""prediction error method" do not apply the subspace ideas, since they "go back to the data" (van Overschee and de Moor, 1996). As to the robust method proposed by Van Overschee and De Moor (van Overschee and de Moor, 1996), it is the slowest of the existing methods, due to its numerical complexity. We have shown that this robust subspace method can be just expressed as an orthogonal projection of the future outputs on the orthogonal complement of the column space of the extended observability matrix – thus providing a new sort of simpler (but equally accurate) subspace algorithms.

REFERENCES

Chiuso, A. and Picci, G. (2001). Asymptotic variance of subspace estimates. In *Proc. CDC'01 - 40th Conference on Decision and Control, Orlando, Florida, USA*, pages 3910–3915.

Delgado, C., dos Santos, P. L., and de Carvalho, J. L. M. (2004). Simplified estimation of matrices b and d in subspace identification algorithm. In *Proc. Controlo'04 – 6th Portuguese Conference on Automatic Control, Faro, Portugal*.

dos Santos, P. L. and de Carvalho, J. L. M. (2003). Improving the numerical efficiency of the b and d estimates produced by the combined deterministic-stochastic subspace identification algorithms. In *Proc. CDC'03 – 42nd Conference on Decision and Control, Maui, Hawaii, USA*, pages 3473–3478.

dos Santos, P. L. and de Carvalho, J. L. M. (2004). A new insight to the a and c matrices extraction in a MOESP type subspace identification algorithm. In *Proc. Controlo'04 – 6th Portuguese Conference on Automatic Control, Faro, Portugal*.

Kailath, T. (1980). *Linear Systems*. Prentice-Hall.

Ljung, L. (1987). *System Identification: Theory for the User*. Prentice-Hall.

van Overschee, P. and de Moor, B. (1996). *Subspace Identification for linear systems*. Kluwer Academic Publishers.