AN INTERPRETATION OF SUBSPACE IDENTIFICATION METHODS

Pouliquen Mathieu * **M'Saad Mohammed** *

* *Control Group, GREYC CNRS UMR 6072*
ENSICAEN, 06 Bd du Maréchal Juin - 14050 Caen Cedex, France

Abstract: Subspace-based system identification methods have been developed over the last two decades. The available methods are algebraic in nature as they are not based on a suitable identification criterion minimization. The aim of this paper is to investigate the subspace identification methods in the spirit of the well known prediction error methods and output error methods. *Copyright ©2005 IFAC*

Keywords: Identification algorithm, subspace methods, identification criterion, prediction error methods, output error methods, iterative algorithm

1. INTRODUCTION

There has been an increasing interest in subspace identification algorithm over the last decade as pointed out by the available contributions. The early contributions have been proposed in (Ho and Kalman, 1966), (Zeiger and Mc Ewen, 1974) and (Kung, 1978). The concept has been further developed in (Willems, 1986), (Moonen *et al.*, 1989), (Verhaegen, 1994), (Viberg, 1995), (Van Overschee and De Moor, 1995), (Van Overschee and De Moor, 1996), (Jansson, 1997), (Chui, 1997), (Bauer, 1998) and (Bauer, 2003).

Subspace methods have been mainly motivated by a set of interesting properties: the simplicity, the intrinsic numerical robustness and their straightforward application to multivariable systems. Of practical interest, they provide an estimation of the model structure.

However the subspace identification methods of the open literature lack of an appropriate identification criterion as the widespread classical prediction error methods (Ljung, 1999). Furthermore, the subspace system identification consists in a set of a well established methods which are not closely related in spite of unifying theorems established in ((Van Overschee and De Moor, 1995), (Van Overschee and De Moor, 1996), (Jansson, 1997) and (Bauer, 1998).

The motivation of this paper is twofold. Firstly, one proposes two classes of subspace methods based on the well known prediction error identification criterion and the output error identification criterion. Secondly, one shows that some well known subspace identification methods can be interpreted as particular cases of the two previous classes of methods.

Section 2 is devoted to the usual notations. The proposed identification subspace approaches are presented and appropriately commented in section 3 and section 4. Simulation results are given in section 5 and some concluding remarks end the paper.

2. STATE SPACE MODELS

2.1 *System and model*

Consider a discrete-time linear time-invariant system $\mathcal{S}_o$ which dynamical behavior can be described by the following equations

$$\mathcal{S}_o \begin{cases} x_o(t+1) = A_o x_o(t) + B_o u(t) + K_o e(t) \\ y(t) \quad\quad = C_o x_o(t) + D_o u(t) + e(t) \end{cases}$$

where $(A_o, B_o, C_o, D_o, K_o)$ is a system state realization of dimension n. $\{u(t)\} \in \mathbb{R}^m$, $\{y(t)\} \in \mathbb{R}^p$ and $\{x_o(t)\} \in \mathbb{R}^n$ respectively denote the known system input sequence, the measurable system output sequence and the system state sequence. $\{e(t)\} \in \mathbb{R}^p$ represents the noise sequence and is an unobserved sequence of independent random variables with zero

mean values and finite variances. The input and noise sequences are assumed to be uncorrelated.

The aim of the subspace identification methods is to estimate a model $\mathcal{M}$ with the following structure

$$\mathcal{M}\begin{cases} x(t+1) &= Ax(t)+Bu(t)+K\widetilde{y}(t) \\ y(t) &= Cx(t)+Du(t)+\widetilde{y}(t) \end{cases} \quad (1)$$

where (A,B,C,D,K) is a state realization of the model corresponding to the state sequence $\{x(t)\} \in \mathbb{R}^n$. $\widetilde{y}(t) \in \mathbb{R}^p$ is the prediction error defined as

$$\widetilde{y}(t) = y(t) - \widehat{y}(t)$$

with $\widehat{y}(t) = Cx(t) + Du(t)$.

Using the forward shift operator formulation, one can easily write (1) as follows

$$y(t) = G(q)u(t) + H(q)\widetilde{y}(t) \quad (2)$$

with $G(q) = C(qI_n - A)^{-1}B + D$ and $H(q) = C(qI_n - A)^{-1}K + I_p$.

The model $\mathcal{M}$ is such that its matrix pair $\{A, C\}$ is assumed to be observable, and its matrix pair $\{A, (B\ K)\}$ is assumed to be controllable. The model is supposed to be asymptotically stable, i.e. the eigenvalues of A are assumed to lie inside the unit circle, and $H(q)$ is supposed to be inversely asymptotically stable, i.e. the eigenvalues of $(A - KC)$ are assumed to lie inside the unit circle.

2.2 *Notations*

Most subspace methods use notations introduced by (Moonen *et al.*, 1989). The output block Hankel matrices are defined as

$$\begin{pmatrix} Y_p \\ \hline Y_f \end{pmatrix} = \left(\begin{array}{cccc} y(t-\beta) & y(t-\beta+1) & \dots & y(t-\beta+j-1) \\ y(t-\beta+1) & y(t-\beta+2) & \dots & y(t-\beta+j) \\ \dots & \dots & \dots & \dots \\ y(t-1) & y(t) & \dots & y(t+j-2) \\ \hline y(t) & y(t+1) & \dots & y(t+j-1) \\ y(t+1) & y(t+2) & \dots & y(t+j) \\ \dots & \dots & \dots & \dots \\ y(t+\alpha-1) & y(t+\alpha) & \dots & y(t+\alpha+j-2) \end{array}\right)$$

Y_p is called the past output Hankel matrix while Y_f is called the future output Hankel matrix. U_p, U_f , $\widetilde{Y}_p$ and $\widetilde{Y}_f$ are defined in the same way. α is the observability horizon and β is the controllability horizon while j is the number of column on Hankel matrices and it will be often assumed to be infinite. Notice that α, β and j are user-defined index.

Let introduce Γ_α the extended ($\alpha > n$) observability matrix of the model and H_α^d a lower triangular Toeplitz matrix containing the first α Markov parameters of $G(q)$.

$$\Gamma_\alpha = \begin{pmatrix} C \\ CA \\ \dots \\ CA^{\alpha-1} \end{pmatrix} \quad H_\alpha^d = \begin{pmatrix} D & 0 & \cdots & 0 \\ CB & D & \cdots & 0 \\ \cdots & \cdots & \cdots & 0 \\ CA^{\alpha-2}B & \cdots & CB & D \end{pmatrix}$$

H_α^s is defined in a similar way, substituting K for B and I_p for D.

3. SUBSPACE PREDICTION ERROR METHODS

In this section one proposes a first class of subspace methods: the subspace prediction error methods. Identification criterion and solution of the underlying identification problem are first presented. An iterative algorithm is then proposed.

3.1 *Identification criterion*

The widespread prediction error methods consist in considering a model structure of the form (2). The key idea in these methods is to compute a model such that its prediction error sequence is as small as possible (Ljung, 1999). There are several manners to qualify what "small" should mean. Basically the size of the prediction error is measured via the Frobenius norm $|.|_F$. The underlying identification problem is

$$(G(q),H(q)) = \underset{\underline{(G(q),H(q))}}{\text{ARG MIN}} \left\{ \lim_{j\to\infty} \frac{1}{j} \sum_{t=1}^{j} |\widetilde{y}(t)|_F^2 \right\} \quad (3)$$

It is worth noticing that the optimization problem (3) is non-linear in the state realization (A, B, C, D, K). Moreover the order n of the system is unknown and the solution is by no means unique due to the infinite number of basis in the state space. It turns out that this problem can't be straightforwardly solved.

However, it's possible to reformulate the identification problem (3) using Hankel matrices as follows

$$(A,B,C,D,K) = \underset{\underline{(A,B,C,D,K)}}{\text{ARG MIN}} \left\{ \lim_{j\to\infty} \frac{1}{\alpha j} \left|\widetilde{Y}_f\right|_F^2 \right\} \quad (4)$$

with

$$\begin{aligned} \widetilde{Y}_f &= (H_\alpha^s)^{-1}(Y_f - \Gamma_\alpha X_t - H_\alpha^d U_f) \\ X_t &= \begin{pmatrix} x(t) & x(t+1) & \dots & x(t+j-1) \end{pmatrix} \end{aligned}$$

Such an optimization problem is the corner stone of the subspace prediction error approach we are concerned by. Table 1 summarizes the similarity between the proposed approach and the usual prediction error approach. As in most subspace identification methods, the considered problem (4) is break down into two main steps:

(1) Estimation of the order n and of the extended observability matrix Γ_α.
(2) Extraction of a state realization from the knowledge of Γ_α or the recovered state matrix X_t.

In the next subsection one concentrates on the first step. This step has given its name to the identification method because it consists in analyzing the subspace spanned by rows of Γ_α.

	usual prediction error methods: based on transfer function	subspace prediction error methods: based on state space realization
model	$y(t) = G(q)u(t) + H(q)\widetilde{y}(t)$	$\left\{ \begin{array}{rcl} x(t+1) & = & Ax(t) + Bu(t) + K\widetilde{y}(t) \\ y(t) & = & Cx(t) + Du(t) + \widetilde{y}(t) \end{array} \right.$
criterion	$(G(q),H(q))= \underset{\underline{(G(q),H(q))}}{\text{ARG MIN}} \left\{ \lim_{j\to\infty} \frac{1}{j} \sum_{t=1}^{j} \left\| \widetilde{y}(t) \right\|_F^2 \right\}$	$(A,B,C,D,K)= \underset{\underline{(A,B,C,D,K)}}{\text{ARG MIN}} \left\{ \lim_{j\to\infty} \frac{1}{\alpha j} \left\| \widetilde{Y}_f \right\|_F^2 \right\}$

Table 1. Usual and subspace prediction error methods

3.2 *Subspace estimation*

Using the pseudolinear regression principe, the identification criterion (4) can be rewritten as follows

$$(A,B,C,D,K)= \underset{\underline{(A,B,C,D,K)}}{\text{ARG SOL}} \left\{ \lim_{j\to\infty} \frac{1}{j} H_\alpha^s \widetilde{Y}_f Z^T = 0 \right\} \tag{5}$$

with $Z = \begin{pmatrix} Z_p \\ U_f \end{pmatrix}$ and $Z_p = \begin{pmatrix} U_p \\ Y_p \end{pmatrix}$.

Due to hypothesis on input and noise sequences, problem (4) implies problem (5). Provided that the input sequence is persistently exciting of order $(\alpha + \beta)$ and that $j \to \infty$ then application of (5) gives

$$\lim_{j\to\infty} \frac{1}{j} Y_f /^Z Z = \Gamma_\alpha X_t /^Z Z + H_\alpha^d U_f$$

where $./^Z = Z^T(ZZ^T)^{-1}$ is the orthogonal projection operator. Notice that this projection could be interpreted as an asymptotic rejection of the noise Hankel matrix E_f on the future output Hankel matrix Y_f.

The following optimization problem allows to determine the model order n and the observability matrix Γ_α

$$\Gamma_\alpha= \underset{\underline{\Gamma_\alpha}}{\text{ARG MIN}} \left\{ \left| Y_f /^Z Z - \underline{\Gamma_\alpha} X_t /^Z Z - H_\alpha^d U_f \right|_F \right\} \tag{6}$$

The next theorem aims at giving solution on Γ_α.

Theorem 1.

Consider the class of systems and models respectively described by $\mathcal{S}_o$ and $\mathcal{M}$. If the following assumptions holds

(1) the matrix Ω, defined below, has rank $(n + \alpha m)$

$$\Omega = \lim_{j\to\infty} \frac{1}{j} \begin{pmatrix} X_t \\ U_f \end{pmatrix} \begin{pmatrix} Z_p^T & U_f^T \end{pmatrix}$$

(2) the input sequence is persistently exciting of order $(\alpha + \beta)$;
(3) $\alpha \geq n$;
(4) $\beta \geq n$.

Then the model $\mathcal{M}$ which is solution of the problem (4) is such that

(1) $\mathcal{M} = \mathcal{S}_o$;
(2) The SVD of $\lim_{j\to\infty} Y_f /_{U_f}^{Z_p}$ is of the form

$$\lim_{j\to\infty} Y_f /_{U_f}^{Z_p} = \begin{pmatrix} U_1 & U_2 \end{pmatrix} \begin{pmatrix} S_1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V_1^T \\ V_2^T \end{pmatrix}$$

with $S_1 \in \mathbb{R}^{n\times n}$ and where $./_{U_f}^{Z_p}$ is the oblique projection operator defined as

$$./_{U_f}^{Z_p} = Z^T(ZZ^T)^{-1} \begin{pmatrix} I_{\beta(p+m)} \\ 0_{\alpha m \times \beta(p+m)} \end{pmatrix}$$

(3) n is equal to the number of singular values different from zero ;
(4) Γ_α is given by

$$\Gamma_\alpha = U_1 S_1^{1/2} T$$

where T is a similarity transformation. □

Proof: See (Pouliquen and M'saad, 2004).

The rank condition (1) has been considered in (Jansson, 1997) where it is viewed as the critical relation for consistency of subspace methods. It involves some conditions on the choice of the index α and β, the degree of persistence excitation and the the system complexity. A wide discussion on how to practically satisfy this condition can be found in (Jansson, 1997) and (Chui, 1997). This rank condition is "generically" satisfied if the input is persistently exciting enough according to (Gustafsson, 2002).

If all assumptions of theorem 1 are satisfied, then $rank(Y_f /_{U_f}^{Z_p}) = n$. However, when j isn't infinite, the Hankel matrix E_f isn't asymptotically rejected and hence $Y_f /_{U_f}^{Z_p}$ is a full rank matrix. This means that the solution on Γ_α proposed by theorem 1 is nothing than an approximation. To improve this estimation a rank condition on Γ_α can be incorporated into the optimization problem (6) as follows:

$$\Gamma_\alpha= \underset{\substack{\underline{\Gamma_\alpha} \\ rank(\underline{\Gamma_\alpha})=n}}{\text{ARG MIN}} \left\{ \left| Y_f /^Z Z - \underline{\Gamma_\alpha} X_t /^Z Z - H_\alpha^d U_f \right|_F \right\} \tag{7}$$

A suboptimal solution consists in breaking down problem (7) in the following manner

$$\Gamma_\alpha= \underset{\substack{\underline{\Gamma_\alpha} \\ rank(\underline{\Gamma_\alpha})=n}}{\text{ARG MIN}} \left\{ \left| Y_f /_{U_f}^{Z_p} Z_p - \underline{\Gamma_\alpha} X_t /_{U_f}^{Z_p} Z_p \right|_F \right\} \tag{8}$$

The optimal solution on Γ_α of problem (8) is given by

$$\Gamma_\alpha = U_1 S_1^{1/2} T$$

with

$$Y_f /_{U_f}^{Z_p} Z_p = \begin{pmatrix} U_1 & U_2 \end{pmatrix} \begin{pmatrix} S_1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V_1^T \\ V_2^T \end{pmatrix}$$

This corresponds to N4SID algorithm described in (Van Overschee and De Moor, 1993) and (Van Overschee and De Moor, 1995).

method	W_1	W_2
standard	$I_{\alpha p}$	$I_{\beta(p+m)}$
N4SID	$I_{\alpha p}$	Z_p
PO-MOESP	$I_{\alpha p}$	$Z_p\Pi_{U_f}^{\perp}$
CVA	$\left(Y_f\Pi_{U_f}^{\perp}Y_f^T\right)^{-1/2}$	$Z_p\Pi_{U_f}^{\perp}$
IVM	$\left(Y_f\Pi_{U_f}^{\perp}Y_f^T\right)^{-1/2}$	$\left(Z_p\Pi_{U_f}^{\perp}Z_p^T\right)\left(Z_pZ_p^T\right)^{-1}$

Table 2. Some common subspace methods as special cases of theorem 2

An optimal solution of problem (7) is obtained with the PO-MOESP subspace method proposed in (Verhaegen, 1994). It's given by

$$\Gamma_\alpha = U_1 S_1^{1/2} T$$

with

$$Y_f/_{U_f}^{Z_p} Z_p \Pi_{U_f}^{\perp} = \begin{pmatrix} U_1 & U_2 \end{pmatrix} \begin{pmatrix} S_1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V_1^T \\ V_2^T \end{pmatrix}$$

where $\Pi_{U_f}^{\perp}$ is defined from the orthogonal projection operator as follows

$$\Pi_{U_f}^{\perp} = I_j - /^{U_f} U_f$$

In the following, one proposes an unifying theorem for the estimation of the order and the observability matrix according to the prediction error identification problem (4).

Theorem 2.

Consider the context of theorem 1. If W_1 and W_2 are two matrices such that W_1 is no singular and $rank\left(X_t/_{U_f}^{Z_p} W_2\right) = rank\left(X_t/_{U_f}^{Z_p}\right)$, then one has

(1) The SVD of $\lim_{j\to\infty} W_1 Y_f/_{U_f}^{Z_p} W_2$ is of the form

$$\lim_{j\to\infty} W_1 Y_f/_{U_f}^{Z_p} W_2 = \begin{pmatrix} U_1 & U_2 \end{pmatrix} \begin{pmatrix} S_1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V_1^T \\ V_2^T \end{pmatrix}$$

with $S_1 \in \mathbb{R}^{n\times n}$;

(2) n is equal to the number of singular values different from zero ;

(3) Γ_α is given by

$$\Gamma_\alpha = (W_1)^{-1} U_1 S_1^{1/2} T$$

where T is a similarity transformation. □

This theorem follows from theorem 1 as it has been proven in (Pouliquen and M'saad, 2004). Similar results have been presented in (Van Overschee and De Moor, 1995), (Jansson, 1997) and (Bauer, 1998). There are however several features that could be pointed out.

- It's based on an identification criterion similar to the prediction error criterion used in classical identification (Ljung, 1999).
- It includes unifying theorems presented in (Van Overschee and De Moor, 1995), (Jansson, 1997) and (Bauer, 1998).

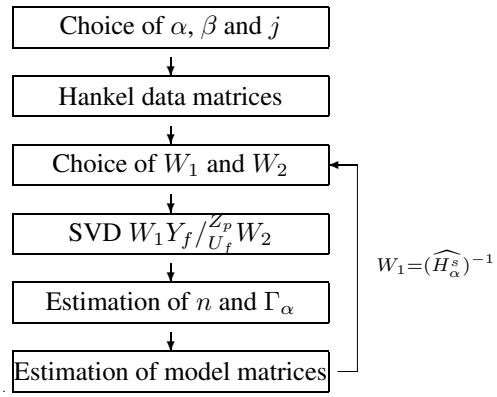

Fig. 1. Iterative subspace algorithm

- It includes some well known subspace methods up to the choice of two weighting matrices W_1 and W_2. Some special cases are summarized in table 2. CVA means Canonical Variate Analysis (Larimore, 1990) and IVM means Instrumental Variable Method (Viberg, 1995).

3.3 *An iterative subspace algorithm*

The purpose of this subsection is to outline a subspace algorithm based on the following choice of weighting matrix

$$W_1 = \left(H_\alpha^s\right)^{-1}$$

Condition on W_1 in theorem 2 shows consistency of that choice. Such a choice depends on unknown quantities, i.e. the first α Markov parameters of the noise model $H(q)$. As a consequence one proposes the iterative algorithm summarized in figure 1. $\left(\widehat{H_\alpha^s}\right)^{-1}$ represents an estimation of $\left(H_\alpha^s\right)^{-1}$ obtained from model estimated at previous iteration.

It seems difficult to verify analytically if the above algorithm converge and if the underlying solution is a better estimation than the one obtained at the first iteration. In section 5 simulation results will show the interest of this iterative subspace algorithm.

4. SUBSPACE OUTPUT ERROR METHODS

In this section one proposes a second class of subspace methods: the subspace output error methods. As is the previous section, the identification criterion is first introduced, then the underlying identification problem and its solution are presented.

4.1 *Identification criterion*

The widespread output error methods consist in considering a model structure of the form (2) with $K = 0$.

The rational behind these methods is a judicious correlation approach that is comprehensively presented in (Ljung, 1999). The system model is validated provided that the prediction error is independent of the input sequence. This identification framework assumes that there is no modelling errors in the prediction error and that the output disturbances are zero mean. The underlying identification problem is

$$\underset{\underline{G(q)}}{G(q)= \text{ ARG SOL}} \left\{ \lim_{j\to\infty} \frac{1}{j} \sum_{t=1}^{j} \widetilde{y}(t) u^T(t+\tau) = 0 \ ; \ \forall \tau \right\} \tag{9}$$

As in the previous section, it is possible to reformulate the correlation based identification problem (9) using Hankel matrices as follows

$$\underset{\underline{(A,B,C,D)}}{(A,B,C,D)= \text{ ARG SOL}} \left\{ \lim_{j\to\infty} \frac{1}{\alpha j} \widetilde{Y}_f Z^T = 0 \right\} \tag{10}$$

with $Z = \begin{pmatrix} U_p \\ U_f \end{pmatrix}$.

Such an optimization problem is the corner stone of the subspace output error approach we are concerned by. Table 3 summarizes the similarity between the proposed approach and the usual output error approach.

4.2 *Subspace estimation*

In the following, one proposes an unifying theorem for the estimation of the order and the observability matrix according to the output error identification problem (10).

Theorem 3.

Consider the class of systems and models respectively described by $\mathcal{S}_o$ and $\mathcal{M}$ with $K = 0$. If the following assumptions holds

(1) the matrix Ω, defined below, has rank $(n + \alpha m)$

$$\Omega = \lim_{j\to\infty} \frac{1}{j} \begin{pmatrix} X_t \\ U_f \end{pmatrix} \begin{pmatrix} U_p^T & U_f^T \end{pmatrix}$$

(2) the input sequence is persistently exciting of order $(\alpha + \beta)$;
(3) $\alpha \geq n$;
(4) $\beta \geq n$;
(5) W_1 and W_2 are two matrices such that W_1 is no singular and $rank\left(X_t/^{U_p}_{U_f} W_2\right) = rank\left(X_t/^{U_p}_{U_f}\right)$.

Then the model $\mathcal{M}$ which is solution of the problem (10) is such that

(1) $G(q) = C_o(qI_n - A_o)^{-1}B_o + D_o$;
(2) The SVD of $\lim_{j\to\infty} W_1 Y_f/^{U_p}_{U_f} W_2$ is of the form

$$\lim_{j\to\infty} W_1 Y_f/^{U_p}_{U_f} W_2 = \begin{pmatrix} U_1 & U_2 \end{pmatrix} \begin{pmatrix} S_1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V_1^T \\ V_2^T \end{pmatrix}$$

with $S_1 \in \mathbb{R}^{n\times n}$
(3) n is equal to the number of singular values different from zero
(4) Γ_α is given by

$$\Gamma_\alpha = (W_1)^{-1} U_1 S_1^{1/2} T$$

where T is a similarity transformation. □

This theorem has two main interest:

- It's based on the output error identification problem (10) similar to the usual output error problem studied in (Ljung, 1999).
- It includes the well known subspace method PI-MOESP (Verhaegen, 1994). To recover this method, all one has to do is to take $W_1 = I_{\alpha p}$ and $W_2 = U_p \Pi^{\perp}_{U_f}$.

5. EXAMPLE

In this section one illustrates the performance of the iterative algorithm proposed in subsection 3.3 with some simulation results. To this end, let consider the following example proposed in (Van Overschee and De Moor, 1995)

$$\mathcal{S}_o : \begin{cases} x_o(t+1) = \begin{pmatrix} 0.67 & 0.67 & 0 & 0 \\ -0.67 & 0.67 & 0 & 0 \\ 0 & 0 & -0.67 & -0.67 \\ 0 & 0 & 0.67 & -0.67 \end{pmatrix} x_o(t) \\ \qquad + \begin{pmatrix} 0.6598 \\ 1.9698 \\ 4.3171 \\ -2.6436 \end{pmatrix} u(t) + \begin{pmatrix} -0.1099 \\ 0.5667 \\ 0.3652 \\ -0.5288 \end{pmatrix} e(t) \\ y(t) = \begin{pmatrix} -0.5749 & 1.0751 & -0.5225 & 0.1830 \end{pmatrix} x_o(t) \\ \qquad -0.7139 u(t) + e(t) \end{cases}$$

$\{u(t)\}$ and $\{e(t)\}$ are white noise sequences with variance 1 and 8 respectively.

The iterative algorithm has been tested over 20 iterations: at the first iteration one chooses $W_1 = I_{\alpha p}$ and $W_2 = Z_p \Pi^{\perp}_{U_f}$, then $W_1 = (\widehat{H^s_\alpha})^{-1}$. The design parameters are $j = 100$, $\alpha = 8$ and $\beta = 16$.

Let define $G_o(q)$ and $G_i(q)$ as respectively the transfer function between sequences $\{u(t)\}$ and $\{y(t)\}$ and its estimation at i^{th} iteration.

To deal with the performances of the algorithm, one considers the behavior of the algorithm between two successive iterations. To feature this, one computes the following index for $2 \leq i \leq 20$:

$$\delta(i) = log_{10}\left(\frac{\|G_i(q) - G_{i-1}(q)\|_\infty}{min_i(\|G_i(q) - G_{i-1}(q)\|_\infty)} \right)$$

Results are shown in figure (2). It is clear that there's a convergence of the algorithm in a subset of models.

Figure (3) compares Bode magnitude of $G_o(q)$, $G_1(q)$, and $G_{20}(q)$. It shows that, in this example, the iterative

	usual output error methods based on transfer function	subspace output error methods based on state space realization
model	$y(t) = G(q)u(t) + \widetilde{y}(t)$	$\left\{ \begin{array}{rcl} x(t+1) & = & Ax(t) + Bu(t) \\ y(t) & = & Cx(t) + Du(t) + \widetilde{y}(t) \end{array} \right.$
criterion	$G(q) = \underset{G(q)}{\text{ARG SOL}} \left\{ \lim_{j\to\infty} \frac{1}{j} \sum_{t=1}^{j} \widetilde{y}(t) u^T(t+\tau) = 0 \ ; \ \forall \tau \right\}$	$(A,B,C,D) = \underset{(A,B,C,D)}{\text{ARG SOL}} \left\{ \lim_{j\to\infty} \frac{1}{\alpha j} \widetilde{Y}_f Z^T = 0 \right\}$

Table 3. Usual and subspace output error methods

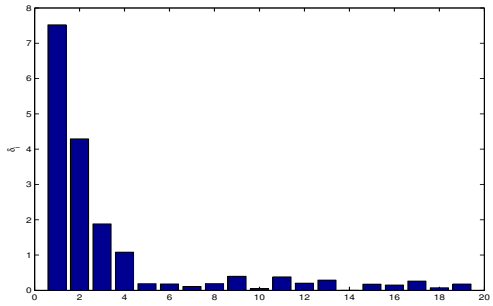

Fig. 2. The convergence of $\delta(i)$

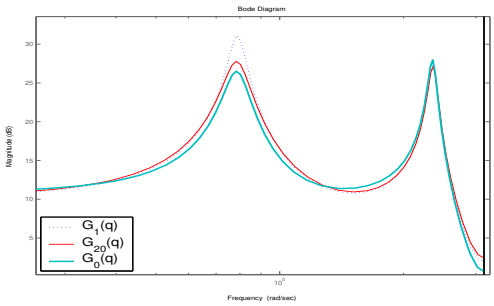

Fig. 3. Bode magnitude of $G_o(q)$, $G_1(q)$, and $G_{20}(q)$

algorithm increases estimation quality compared with the first estimated model.

6. CONCLUSION

This paper proposed an interpretation of subspace methods. Fundamental results have been established which allow to show that some well known subspace methods can be included into two main classes of subspace identification approaches. These classes are based on criteria in the spirit of well known identification criteria: the prediction error criterion and the output error criterion.

7. REFERENCES

Bauer, D. (1998). Some asymptotic theory for the estimation of linear systems using maximum likelihood methods or subspace algorithms. Phd thesis. TU Wien.

Bauer, D. (2003). Subspace algorithms. *Proceedings* 13^{th} *IFAC Symposium on System Identification-Rotterdam*.

Chui, N.L.C. (1997). Subspace method and informative experiments for system identification. Phd thesis. Pembroke College Cambridge.

Gustafsson, T. (2002). Subspace-based system identification: weighting and pre-filtering of instruments. *Automatica* **38**(3), 433–443.

Ho, B.L. and R.E. Kalman (1966). Effective construction of linear state-variable models from input/output function. *Regelungstechnick* **14**(12), 545–592.

Jansson, M. (1997). On subspace methods in system identification and sensor array signal processing. Phd thesis. Royal institute of technology, Stockholm.

Kung, S.Y. (1978). A new identification and model reduction algorithm via singular value decomposition. *Proceedings* 12^{th} *Asimolar conference on Circuits, Systems and Computers, Pacific Grove*.

Larimore, W.E. (1990). Canonical variate analysis in identification, filtering and adaptive control. *Proceedings* 29^{th} *Control and Decision Conference - Honolulu*.

Ljung, L. (1999). *System identification: theory for the user*. Prentice Hall.

Moonen, M., B. De Moor, L. Vandenberghe and J. Vandewalle (1989). On and off line identifications of linear state space models. *International Journal of Control* **49**(1), 219–232.

Pouliquen, M. and M. M'saad (2004). A class of subspace identification methods: Subspace prediction error methods. Technical report. Control Group, GREYC CNRS UMR 6072.

Van Overschee, P. and B. De Moor (1993). N4sid: subspace algorithms for the identification of combined deterministic-stochastic systems. *Automatica* **30**(1), 75–93.

Van Overschee, P. and B. De Moor (1995). A unifying theorem for three subspace system identification algorithms. *Automatica* **31**(12), 1853–1864.

Van Overschee, P. and B. De Moor (1996). *Subspace identification for linear systems. Theory, implementation applications*. Kluver Academic Publishers.

Verhaegen, M. (1994). Identification of the deterministic part of mimo state space models given in innovations form from input output data. *Automatica* **30**(1), 61–74.

Viberg, M. (1995). Subspace based methods for the identification of linear time invariant systems. *Automatica* **31**(12), 1835–1851.

Willems, J. C. (1986). From time series to linear system. part 2 - exact modelling. *Automatica*.

Zeiger, H.P. and A.J. Mc Ewen (1974). Approximate linear realizations of given dimension via ho's algorithm. *IEEE Transaction on Automatic Control* **19**, 153.