Interpretation of Subspace Methods: Consistency Analysis.

Lennart Ljung and Tomas McKelvey

Department of Electrical Engineering
Linköping University, S-581 83 Linköping, Sweden

WWW: <http://www.control.isy.liu.se>

Email: ljung,tomas@isy.liu.se

3 March 1999

Report no.: LiTH-ISY-R-2103

For the IFAC Symposium on System Identification, Fukuoka, Japan, 1997

Technical reports from the Automatic Control group in Linköping are available by anonymous ftp at the address [ftp.control.isy.liu.se](ftp://ftp.control.isy.liu.se). This report is contained in the compressed postscript file 2103.ps.Z.

INTERPRETATION OF SUBSPACE METHODS: CONSISTENCY ANALYSIS.

Lennart Ljung and Tomas McKelvey

*Dept. of Electrical Engineering, Linköping University
S-581 83 Linköping, Sweden
E-mail: ljung@isy.liu.se, tomas@isy.liu.se.*

Abstract: So called subspace methods for direct identification of linear state space models form a very useful alternative to maximum-likelihood type approaches, in that they are non-iterative and offer efficient numerical implementations. The algorithms consist of series of quite complex projections, and it is not so easy to intuitively understand how they work. The asymptotic analysis of them is also complicated. This contribution describes an interpretation of how they work in terms of k -step ahead predictors of carefully chosen orders. It specifically deals how consistent estimates of the dynamics can be achieved, even though correct predictors are not used. This analysis gives some new angles of attack to the problem of asymptotic behavior of the subspace algorithms.

Keywords: Identification algorithms, Consistency, Least squares, Predictors

1. INTRODUCTION

A linear system can always be represented in state space form as

$$\begin{aligned}x(t+1) &= Ax(t) + Bu(t) + w(t) \\ y(t) &= Cx(t) + Du(t) + \nu(t)\end{aligned}\quad (1)$$

We shall generally let n denote the dimension of x and let p be the number of outputs.

The so called subspace approach to identify the matrices A, B, C and D has been developed by a number of researchers, see, e.g. (Van Overschee and De Moor, 1993), (Van Overschee and De Moor, 1994), (E., 1983), (Verhaegen, 1994), (Viberg *et al.*, 1993), (Viberg, 1995) and (Jansson and Wahlberg, 1996). The idea behind these methods can be explained as first estimating

the state vector $x(t)$, and then finding the state space matrices by a linear least squares procedure. These methods are most often described in a geometric framework, which gives nice projection interpretations and often elegant formulas. It is however not so obvious how the approaches relate to the "classical" input-output methods, like least squares etc.

We shall in this contribution describe the subspace approach in a conventional least squares estimation framework. This gives some complementary insight, which could be useful for development of alternative algorithms and for the asymptotic analysis. Our focus on k -step ahead predictors is similar to the analysis in (Jansson and Wahlberg, 1996).

2. K -STEP AHEAD PREDICTORS

Let us, at time t , define the following vector of past inputs and outputs:

¹ This work was supported in part by the Swedish Research Council for Engineering

$$\begin{aligned} \varphi_s(t) = & [y(t-1), u(t-1), \\ & \dots, y(t-s), u(t-s)]^T \end{aligned} \quad (2)$$

and the following vector of future inputs

$$\tilde{\varphi}_\ell(t) = [u(t), \dots, u(t+\ell-1)]^T \quad (3)$$

The algorithm and analysis focus on k -step ahead predictors, i.e., “intelligent guesses”, made at time t about the value of $y(t+k)$, given various type of information. To define these approximate predictors, set up the model

$$\begin{aligned} y(t+k) = & (\theta^{k,\ell,s})^T \varphi_s(t+1) \\ & + (\gamma^{k,\ell,s})^T \tilde{\varphi}_\ell(t+1) + \epsilon(t+k) \end{aligned} \quad (4)$$

Estimate θ and γ in (4) using least squares over the available measurements of y and u . Denote the estimates by $\hat{\theta}_N^{k,\ell,s}$ and $\hat{\gamma}_N^{k,\ell,s}$. The k -step ahead predictors we will discuss are then given by

$$\hat{y}_{s,\ell}(t+k|t) = (\hat{\theta}_N^{k,\ell,s})^T \varphi_s(t+1) \quad (5)$$

and

$$\begin{aligned} \bar{y}_{s,\ell}(t+k|t) = & (\hat{\theta}_N^{k,\ell,s})^T \varphi_s(t+1) \\ & + (\hat{\gamma}_N^{k,\ell,s})^T \tilde{\varphi}_\ell(t+1) \end{aligned} \quad (6)$$

Notice the difference! $\bar{y}_{s,\ell}(t+k|t)$ assumes that we know the future relevant inputs at time t . We are thus then only predicting the noise contributions over the indicated time span. $\hat{y}_{s,\ell}(t+k|t)$, on the other hand, just ignores the contributions from the input. No attempt is made to predict these values from the past. Non-trivial predictions of these future outputs could have been made if the input is not white noise. Note also that such predictions of future inputs would automatically have been included if the term with $\tilde{\varphi}$ had not been present in (4).

The predictions are approximate, since they only involve a certain, fixed, amount of past data. Only if the true system can be described as an ARX-model of order s , the predictions are correct. Otherwise, s has to be chosen as a sufficiently large number in order to achieve “good” predictions. Nevertheless, the approximate predictors obey certain exact relationships, that are defined in terms of the true system. These will be dealt with in the next section.

3. SOME RELATIONSHIPS FOR THE APPROXIMATE K -STEP AHEAD PREDICTORS

We start by establishing two lemmas for $\hat{y}$ and $\bar{y}$, which show properties analogous to Levinson type recursions.

Suppose that the true system can be written as a difference equation

$$A_0(q)y(t) = B_0(q)u(t) + C_0(q)e(t) \quad (7)$$

where the polynomials in the shift operator are all of degree at most $\bar{n}$:

$$A_0(q) = I + A_1 q^{-1} + \dots + A_{\bar{n}} q^{-\bar{n}} \quad (8)$$

and $e(t)$ is a white noise sequence. Clearly, if the system is given by the state space description (1), it can always be represented in the input output ARMAX form (7) for some order $\bar{n}$.

We then have the following results, (Ljung and McKelvey, 1996):

Lemma 1 Suppose that the true system can be described by (7) with $\bar{n}$ as the maximal order of the polynomials, and that the system operates in open loop, so that e and u are independent. Let $\hat{y}_{s,\ell}^*(t+k|t)$ and $\bar{y}_{s,\ell}^*(t+k|t)$ be the limits of (5) and (6) as $N \rightarrow \infty$. Then for any s , any $r > \bar{n}$ and any $\ell \geq r$

$$\begin{aligned} \hat{y}_{s,\ell}^*(t+r|t) + A_1 \hat{y}_{s,\ell}^*(t+r-1|t) + \\ \dots + A_{\bar{n}} \hat{y}_{s,\ell}^*(t+r-\bar{n}|t) = 0 \end{aligned} \quad (9)$$

and

$$\begin{aligned} \bar{y}_{s,\ell}^*(t+r|t) + A_1 \bar{y}_{s,\ell}^*(t+r-1|t) + \\ \dots + A_{\bar{n}} \bar{y}_{s,\ell}^*(t+r-\bar{n}|t) \\ = B_0 u(t+r) + B_1 u(t+r-1) + \\ \dots + B_{\bar{n}} u(t+r-\bar{n}) \end{aligned} \quad (10)$$

Proof: Consider the equation (4). Suppress the indices ℓ and s and let

$$\phi(t+1) = \begin{bmatrix} \tilde{\varphi}_\ell(t+1) \\ \varphi_s(t+1) \end{bmatrix} \quad (11)$$

Let $\zeta(t+1)$ be any vector of the same dimension as $\phi(t+1)$ such that

$$\bar{E}\zeta(t+1)C_0(q)e(t+r) = 0 \quad (12)$$

Suppose that θ^k and γ^k are estimated from (4) using the IV-method with instruments $\zeta(t+1)$. Then the limiting estimate are given by

$$\begin{bmatrix} \gamma^k \\ \theta^k \end{bmatrix}^T = \bar{E}y(t+k)\zeta^T(t+1)[\bar{E}\phi(t+1)\zeta^T(t+1)]^{-1}$$

Note also that we can write, for some β_0

$$B_0(q)u(t+r) = \beta_0^T \tilde{\varphi}_\ell(t+1) \quad (13)$$

if $r > n$ and $\ell \geq r$, Hence

$$\begin{aligned} & \begin{bmatrix} \gamma^r \\ \theta^r \end{bmatrix}^T + A_1 \begin{bmatrix} \gamma^{r-1} \\ \theta^{r-1} \end{bmatrix}^T + \dots + A_n \begin{bmatrix} \gamma^{r-n} \\ \theta^{r-n} \end{bmatrix}^T = \\ & = \bar{E}[(A_0(q)y(t+r))\zeta^T(t+1)][\bar{E}\phi(t+1)\zeta^T(t+1)]^{-1} \\ & = \bar{E}[(B_0(q)u(t+r) + C_0(q)e(t+r))\zeta^T(t+1)] \times \\ & \quad \times [\bar{E}\phi(t+1)\zeta^T(t+1)]^{-1} \\ & = \beta_0^T [\bar{E}\tilde{\varphi}_\ell(t+1)\zeta^T(t+1)] \begin{bmatrix} \bar{E}\tilde{\varphi}_\ell(t+1)\zeta^T(t+1) \\ \bar{E}\varphi_s(t+1)\zeta^T(t+1) \end{bmatrix}^{-1} \\ & = \beta_0^T [I \ 0] = [\beta_0^T \ 0] \end{aligned}$$

Here we used (13) and (12) in the third last step, and the definition of a matrix inverse in the second last step. Since

$$\begin{aligned}\bar{y}_{s,\ell}^*(t+k|t) &= \begin{bmatrix} \gamma^k \\ \theta^k \end{bmatrix}^T \phi(t+1) \\ \hat{y}_{s,\ell}^*(t+k|t) &= \begin{bmatrix} \gamma^k \\ \theta^k \end{bmatrix}^T \begin{bmatrix} 0 \\ \varphi(t+1) \end{bmatrix}\end{aligned}$$

we just need to multiply the above expression with $\phi(t+1)$ to obtain the stated result.

It now only remains to show that $\zeta(t+1) = \phi(t+1)$ obeys (12), so that the result holds for the least squares estimates. The vector $\phi(t+1)$ contains a number of inputs, which are uncorrelated with the noise, under open loop operation. It also contains $y(t)$ and older values of y , which are uncorrelated with $C_0(q)e(t+r)$ if $r > n$, since the order of C_0 is at most n . This concludes the proof.

Corollary: Suppose that the true system is given by

$$A_0(q)y(t) = B_0(q)u(t) + v(t) \quad (14)$$

and that the parameters of the predictors are estimated from (4) using an instrumental variable method with instruments $\zeta(t+1)$ that are uncorrelated with $v(t+r)$, which may be any noise sequence. Then the result of the lemma still holds.

Notice that the Lemma holds for any s , which could be smaller than n .

Lemma 2 Let $\bar{y}$ and $\hat{y}$ be defined as above. (These thus depend on N , but this subscript is suppressed.) Then for any N, s, k and any ℓ

$$y(t) = \bar{y}_{s,\ell+1}(t|t-1) + \epsilon(t) \quad (15)$$

$$\begin{aligned}\bar{y}_{s+1,\ell}(t+k|t) &= \bar{y}_{s,\ell+1}(t+k|t-1) \\ &+ \tilde{h}_{k,N}^{s+1,\ell} \epsilon(t)\end{aligned} \quad (16)$$

$$\begin{aligned}\hat{y}_{s+1,\ell}(t+k|t) &= \hat{y}_{s,\ell+1}(t+k|t-1) + \hat{b}_{k,N}^{s+1,\ell} u(t) \\ &+ \beta_{k,N,s,\ell}^T \tilde{\varphi}_\ell(t+1) + \tilde{h}_{k,N}^{s+1,\ell} \epsilon(t)\end{aligned} \quad (17)$$

where $\epsilon(t)$ (same in the three expressions) is uncorrelated with $\varphi_s(t), u(t)$ and $\tilde{\varphi}_\ell(t+1), t = 1, \dots, N$. If the input sequence $\{u(t)\}$ is white, then

$$\hat{b}_{k,N}^{s+1,\ell} \rightarrow h_u(k) \text{ and } \beta_{k,N,s,\ell} \rightarrow 0 \text{ as } N \rightarrow \infty \quad (18)$$

where $h_u(k)$ is the true impulse response coefficient.

Proof: Let

$$\phi_1(t+1) = \begin{bmatrix} \tilde{\varphi}_\ell(t+1) \\ \varphi_{s+1}(t+1) \end{bmatrix}$$

and

$$\phi_2(t+1) = \begin{bmatrix} \tilde{\varphi}_{\ell+1}(t) \\ \varphi_s(t) \end{bmatrix}$$

The vector $\phi_2(t+1)$ contains the values $u(i), i = t+\ell, \dots, t-s$ and $y(i), i = t-1, \dots, t-s$. The

vector $\phi_1(t+1)$ contains the same values, and in addition $y(t)$. Define ϵ as the residuals from the least squares fit

$$y(t) = L_2 \phi_2(t+1) + \epsilon(t)$$

so that $\epsilon(t) \perp \phi_2(t+1)$. With this we mean that

$$\sum_{t=1}^N \epsilon(t) \phi_2(t+1) = 0$$

Note that $\epsilon(t)$ will depend on ℓ and s , but not on k . Moreover, by definition we find that

$$L_2 \phi_2(t+1) = \bar{y}_{s,\ell+1}(t|t-1) \quad (19)$$

so the first result of the lemma has been proved.

Let

$$\phi_3(t+1) = \begin{bmatrix} \phi_2(t+1) \\ \epsilon(t) \end{bmatrix}$$

It is clear that ϕ_1 and ϕ_3 span the same space, so that for some matrix R (built up using L_2) we can write

$$\phi_1(t+1) = R \phi_3(t+1)$$

Now write

$$y(t+k) = \hat{K}_1 \phi_1(t+1) + \varepsilon(t+k) \quad (20)$$

where $\hat{K}_1$ is the LS-estimate, so that

$$\varepsilon(t+k) \perp \phi_1(t+1)$$

Let

$$\hat{K}_1 R = \begin{bmatrix} K_2 & K_3 \end{bmatrix} \quad (21)$$

Clearly, by definition

$$\begin{aligned}\bar{y}_{s+1,\ell}(t+k|t) &= \hat{K}_1 \phi_1(t+1) = \hat{K}_1 R \phi_3(t+1) \\ &= K_2 \phi_2(t+1) + K_3 \epsilon(t)\end{aligned} \quad (22)$$

Now rewrite (20) as

$$\begin{aligned}y(t+k) &= \hat{K}_1 \phi_1(t+1) + \varepsilon(t+k) \\ &= \hat{K}_1 R \phi_3(t+1) + \varepsilon(t+k) \\ &= K_2 \phi_2(t+1) + K_3 \epsilon(t) + \varepsilon(t+k)\end{aligned}$$

Both $\epsilon(t)$ and $\varepsilon(t+k)$ are orthogonal to $\phi_2(t+1)$, so K_2 must be the least squares fit of $y(t+k)$ to $\phi_2(t+1)$, which means that

$$\bar{y}_{s,\ell+1}(t+k|t-1) = K_2 \phi_2(t+1) \quad (23)$$

Comparing (22) with (23) we have shown (16), (with $\tilde{h}_{k,N}^{s+1,\ell} = K_3$). Moreover,

$$\begin{aligned}
\bar{y}_{s+1,\ell}(t+k|t) &= \hat{y}_{s+1,\ell}(t+k|t) + \beta_1 \tilde{\varphi}_\ell(t+1) \\
\bar{y}_{s,\ell+1}(t+k|t-1) &= \hat{y}_{s,\ell+1}(t+k|t-1) + \beta_2 \tilde{\varphi}_{\ell+1}(t) \\
&= \hat{y}_{s,\ell+1}(t+k|t-1) + b_k u(t) \\
&\quad + \beta_3 \tilde{\varphi}_\ell(t+1)
\end{aligned}$$

(with b_k being the first column of β_2). Applying (16) to the two left hand sides of these expressions, we have also proved (17) (with $\beta_{k,N,s,\ell}^T = \beta_2 - \beta_3$ and $b_{k,N}^{s+1,\ell} = b_k$. The proof of (18) is straightforward and omitted here, since we will not need this result for the ensuing discussion. This concludes the proof of Lemma 2.

We note that using Lemma 1 we can obtain consistent estimates of the dynamic part of the system (1)=(7) even though the underlying ARX-models (5) may be too simple. (Note, though, that consistency from (9) and (10) also requires that $\hat{y}^*$ and $\bar{y}^*$ be persistently exciting of sufficient order.) This is in analogy with the IV-method for the same problem.

4. SOME VECTOR NOTATION

Let us first define

$$\hat{Y}_r^{s,\ell}(t+1) = \begin{bmatrix} \hat{y}_{s,\ell}(t+1|t) \\ \vdots \\ \hat{y}_{s,\ell}(t+r|t) \end{bmatrix} \quad (24)$$

and

$$\bar{Y}_r^{s,\ell}(t+1) = \begin{bmatrix} \bar{y}_{s,\ell}(t+1|t) \\ \vdots \\ \bar{y}_{s,\ell}(t+r|t) \end{bmatrix} \quad (25)$$

Clearly, we can treat all predictors simultaneously: Let

$$Y_r(t+1) = \begin{bmatrix} y(t+1) \\ \vdots \\ y(t+r) \end{bmatrix} \quad (26)$$

and stack r equations like (4) on top of each other:

$$\begin{aligned}
Y_r(t+1) &= \Theta^{r,\ell,s} \varphi_s(t+1) \\
&\quad + \Gamma^{r,\ell,s} \tilde{\varphi}_\ell(t+1) + E(t+1)
\end{aligned} \quad (27)$$

Estimate the $pr \times s(p+m)$ matrix Θ (together with Γ) by least squares and then form

$$\hat{Y}_r^{s,\ell}(t+1) = \hat{\Theta}_N^{r,\ell,s} \varphi_s(t+1) \quad (28)$$

$$\begin{aligned}
\bar{Y}_r^{s,\ell}(t+1) &= \hat{\Theta}_N^{r,\ell,s} \varphi_s(t+1) \\
&\quad + \hat{\Gamma}_N^{r,\ell,s} \tilde{\varphi}_\ell(t+1)
\end{aligned} \quad (29)$$

In fact, these quantities can be efficiently calculated by projections using the data vectors, without explicitly forming the matrices $\hat{\Theta}$ and $\hat{\Gamma}$.

5. RELATIONSHIP TO STATE-SPACE MODELING

For notational convenience we now focus on the SISO case. To arrive at the subspace approach to state-space modeling from an uncommon direction, let us first introduce

$$\bar{x}(t+1) = \bar{Y}_n^{s+1,\ell}(t+1) \quad (30)$$

$$x(t) = \bar{Y}_n^{s,\ell+1}(t) \quad (31)$$

Suppose we estimate A, C and β in

$$\begin{bmatrix} \bar{x}(t+1) \\ y(t) \end{bmatrix} = \begin{bmatrix} A \\ C \end{bmatrix} x(t) + \beta \tilde{\varphi}_{\ell+1}(t) \quad (32)$$

using the least squares method. What would then the result be?

In view of the exact relationships (for any finite N) in Lemma 2 we will get:

$$C = [1 \ 0 \ \dots \ 0] \quad (33)$$

exactly, and

$$A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -\hat{a}_n^N & -\hat{a}_{n-1}^N & -\hat{a}_{n-2}^N & \dots & -\hat{a}_1^N \end{bmatrix} \quad (34)$$

Moreover, all rows of β will be exactly zero, except for the second last one which will be

$$[\hat{b}_n^N \ \dots \ \hat{b}_0^N] \quad (35)$$

As $N \rightarrow \infty$ it follows from Lemma 1 that the estimates $\hat{a}$ and $\hat{b}$ will converge to their true values. We have thus constructed a method that consistently estimates A and C in the standard observability form (see, e.g. (Kailath, 1980)). Moreover the β estimate contains correct information and could by standard procedures be transformed to a consistent B -matrix estimate.

Suppose now that instead of (30) and (31) we define x and $\bar{x}$ as

$$\bar{x}(t+1) = \hat{Y}_n^{s+1,\ell}(t+1) \quad (36)$$

$$x(t) = \hat{Y}_n^{s,\ell+1}(t) \quad (37)$$

This would imply no change at all for the estimates of C and A , since $\hat{y}$ and $\bar{y}$ differ only by something proportional to $\tilde{\varphi}_{\ell+1}(t)$.

Notice that the estimates $\hat{a}$ describe the linear relationship between $\bar{y}_{s,\ell+1}$ (which in the limit $N \rightarrow \infty$ is given by Lemma 1).

Assume now that we instead use

$$\bar{x}(t+1) = L \bar{Y}_n^{s+1,\ell}(t+1) \quad (38)$$

$$x(t) = L \bar{Y}_n^{s,\ell+1}(t) \quad (39)$$

for some invertible matrix L . Again, since this is just a change of variables, we would obtain exactly the same linear system as before, although in another coordinate representation.

6. THE SUBSPACE METHODS

A possible description of the subspace methods in (Van Overschee and De Moor, 1993), (Van Overschee and De Moor, 1994), (E., 1983), (Verhaegen, 1994) (see, e.g. (Ljung and McKelvey, 1996)) is to estimate $\hat{Y}_r^{s+1,\ell}(t)$, $\hat{Y}_r^{s,\ell+1}(t)$ and $\hat{Y}_r^{s,\ell+1}(t)$ using (28) and (29) with $\ell \geq r > \bar{n}$ where $\bar{n}$ is an upper bound of the system order.

Then estimate the rank of $\hat{Y}_r^{s,\ell+1}$ and determine the $n|r$ matrix L so that

$$\bar{x}(t+1) = L\bar{Y}_r^{s+1,\ell}(t+1) \quad (40)$$

$$x(t) = L\bar{Y}_r^{s,\ell+1}(t) \quad (41)$$

become well-conditioned bases spanning the “dominating” n -dimensional subspace of the range space of $\hat{Y}_r$. Then proceed as in the previous section. (There are several different variants of the subspace methods, essentially differing in the choice of L and how B and D are determined, once A and C are estimated.)

If the true order of the system is n , then, in the limit as $N \rightarrow \infty$, the rank of $\hat{Y}_r$ will be n , according to Lemma 1. This means, in the limit, that all choices of x in (40) are contained in the subset of choices (38). This also proves the consistency of the general method.

7. CONCLUSIONS

We have in this contribution explicitly shown and exploited how the subspace identification algorithms are primarily based on estimates of the k -step ahead output predictors. We have also pointed to the main technical relationships (Lemmas 1 and 2) that are responsible for the consistency of the method, also in the case where too simple k -step ahead predictors are used. These relationships are also present in the previous literature and correspond to shift properties of a suitably estimated extended observability matrix. Our direct proof may be technically more extensive, but the basic value lies in connecting the asymptotic theory to well known methods like the ordinary least squares. We thus offer additional insights and the results might also help in developing an asymptotic second order analysis of these methods’ statistical properties.

8. REFERENCES

- E., Larimore W. (1983). System identification, reduced order filtering and modeling via canonical variate analysis. In: *Proc. of the American Control Conference*. San Francisco, CA.
- Jansson, M. and B. Wahlberg (1996). “A Linear Regression Approach to State-Space Subspace System Identification”. *Signal Processing (EURASIP), Special Issue on Subspace Methods, Part II: System Identification* **52**(2), 103–129.
- Kailath, T. (1980). *Linear Systems*. Prentice-Hall. Englewood Cliffs, New Jersey.
- Ljung, L. and T. McKelvey (1996). A least squares interpretation of sub-space methods for system identification. In: *Proc. IEEE Conference on Decision and Control, CDC.*. Kobe, Japan.
- Van Overschee, P. and B. De Moor (1993). Subspace algorithms for the stochastic identification problem. *Automatica* **29**(3), 649–660.
- Van Overschee, P. and B. De Moor (1994). N4SID: Subspace algorithms for the identification of combined deterministic-stochastic systems. *Automatica* **30**(1), 75–93.
- 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.
- Viberg, M., B. Ottersten, B. Wahlberg and L. Ljung (1993). Performance of subspace based state-space system identification methods. In: *Proc. 12th World Congress International Federation of Automatic Control, Sydney, Australia*. Vol. 7. pp. 369–372.