Conditioning analysis of a continuous time subspace-based model identification algorithm

G. H. Salázar-Silva△, J. C. Martínez-García◊[1] and R. Garrido△
△Departamento de Control Automático
CINVESTAV-IPN
A.P. 14-740
07300 México D.F.
MÉXICO
◊Programa de Investigación en Matemáticas Aplicadas y Computación
Instituto Mexicano del Petróleo
Eje Central Lázaro Cárdenas 152
Col. San Bartolo Atepehuacan
07730 México D.F.
MÉXICO

Abstract

We present in this paper a study concerning the conditioning analysis of a continuous-time deterministic subspace-based model identification algorithm. We show that the conditioning number of the associated extended observability matrices depends on an exponential way from both: the estimated order of the system and the dimension of the system output vector.

1 Introduction

As far as linear time-invariant multivariable systems are concerned, subspace-based model identification algorithms constitute a broad family of identification methods mainly characterized by the use of geometric information (see for instance [3], [4], and [5]). The subspace-based methods compute the estimated parameters, i.e., a state space realization of the system, from an approximation of the observability subspace of the concerned system. This approximation is obtained from a discrete-time set of input-output measurements. Continuous-time data is filtered and sampled in order to obtain discrete-time information (see [1]). We present in this paper an algorithm to identify a continuous linear time-invariant model from a given set of discrete-time input-output measurements. Our algorithm is based on the method proposed in [3], and depends essentially on the pseudo-inverse of the so-called extended observability matrices. We present then the analysis of the conditioning number of these observability matrices, which is the main purpose of this paper.

In Section 2 we present our proposed algorithm, while in Section 3 we discuss its numerical properties.

[1] Corresponding author, on leave from Departamento de Control Automático, CINVESTAV-IPN, México. E-mail address: martinez@ctrl.cinvestav.mx.

2 The CN4SID algorithm

Consider the continuous linear time-invariant system $(A,\, B,\, C,\, D)$ described by:

$$\left\{ \begin{array}{l} \mathbf{p}x(t) = Ax(t) + Bu(t)\,, \\ y(t) = Cx(t) + Du(t)\,, \end{array} \right. \tag{2.1}$$

where: $\mathbf{p}$ denotes the differential operator, i.e. $\mathbf{p} = d/dt$; $x(\cdot) \in \mathbb{R}^n$ denotes the state; $u(\cdot) \in \mathbb{R}^m$ denotes the input, and $y(\cdot) \in \mathbb{R}^p$ denotes the output. A, B, C and D are linear maps represented by real constant matrices. It is assumed that $(C,\, A)$ is observable.

Consider the scalar causal stable operator:

$$\lambda = \frac{1}{1 + \mathbf{p}\tau}, \tag{2.2}$$

where τ is a scalar such that $\tau > 0$. Given a $\mathbb{R}^p$ valued signal $y(t)$, we define the i^{th}-filtered signal $[\lambda^i y](t)$ as follows:

$$\left[\lambda^i y\right](t) = \left\{ \begin{array}{ll} \lambda\left[\lambda^{i\text{-}1} y\right](t)\,, & \text{for } i > 1, \\ y(t)\,, & \text{when } i = 0. \end{array} \right.$$

Applying this filtering action to the system $(A,\, B,\, C,\, D)$ we obtain the modified system:

$$\left\{ \begin{array}{l} x(t) = A_\lambda [\lambda x](t) + B_\lambda [\lambda u](t)\,, \\ y(t) = Cx(t) + Du(t)\,, \end{array} \right. \tag{2.3}$$

where $A_\lambda = I + \tau A$ and $B_\lambda = \tau B$. Since $(A,\, B)$ and $(A_\lambda,\, B_\lambda)$ are linked through a biyective map, the obtention of estimates for A and B results in the obtention of estimates for A and B.

Suppose constant sampling and a sampling time sequence $\{t_k\}_{k=0}^{N\text{-}1}$ be given, where $t_k = t_0 + kh$ (with h denoting the sampling period). Let input-output measurements $\{u_k\}_{k=0}^{N\text{-}1}$ and $\{y_k\}_{k=0}^{N\text{-}1}$ on system (2.3) be given. The deterministic **C**ontinuous **S**ubspace **Id**entification (**CSId**) problem is then defined as follows:

Definition 2.1. ***CSId*** *problem: Consider the filtered model (2.3) and the input-ouput measurements* $\{u_k\}_{k=0}^{N\text{-}1}$ *and* $\{y_k\}_{k=0}^{N\text{-}1}$*, estimate then the unknown matrices* A*,* B*,* C*, and* D*.*

We can then build the following matrices:

$$Y := \begin{bmatrix} [\lambda^{i\text{-}1}y]_0 & [\lambda^{i\text{-}1}y]_1 & \cdots & [\lambda^{i\text{-}1}y]_{N\text{-}1} \\ [\lambda^{i\text{-}2}y]_0 & [\lambda^{i\text{-}2}y]_1 & \cdots & [\lambda^{i\text{-}2}y]_{N\text{-}1} \\ \vdots & \vdots & \ddots & \vdots \\ y_0 & y_1 & \cdots & y_{N\text{-}1} \end{bmatrix} \tag{2.4}$$

and:

$$U := \begin{bmatrix} [\lambda^{i\text{-}1}u]_0 & [\lambda^{i\text{-}1}u]_1 & \cdots & [\lambda^{i\text{-}1}u]_{N\text{-}1} \\ [\lambda^{i\text{-}2}u]_0 & [\lambda^{i\text{-}2}u]_1 & \cdots & [\lambda^{i\text{-}2}u]_{N\text{-}1} \\ \vdots & \vdots & \ddots & \vdots \\ u_0 & u_1 & \cdots & u_{N\text{-}1} \end{bmatrix}. \tag{2.5}$$

In what follows we present the **CN4SID** algoritm, which is based on: the **N4SID** algorithm [3], and the space trasformation discussed in [1].

CN4SID Algorithm: Consider a continous linear-time invariant system described by (2.1) and the corresponding filtered model (2.3). Let the input-output sequences $\{u_k\}_{k=0}^{N\text{-}1}$ and $\{y_k\}_{k=0}^{N\text{-}1}$ be given. Let i be an the estimated order of the system such that $N >> i > n$, where n denotes the real system. It is then possible to find matrices $\widehat{A}$, $\widehat{B}$, $\widehat{C}$, and $\widehat{D}$ (i.e., the estimates of A, B, C, and D, respectively), according to the procedure:

1. Build matrices Y and U as defined in (2.4) and (2.5), respectively.
2. Build matrices $[\lambda^i Y]$ and $[\lambda^i U]$.
3. Build matrices $[\lambda^i W]$ and $[\lambda^i W]$ defined as follows:

$$[\lambda^i W] := \begin{bmatrix} [\lambda^i U] \\ [\lambda^i Y] \end{bmatrix}, \text{ and } [\lambda^{i+1} W] := \begin{bmatrix} [\lambda^{i+1} U] \\ [\lambda^{i+1} Y] \end{bmatrix}.$$

4. Compute the proyections:

$$Y / \begin{bmatrix} [\lambda^i W] \\ [\lambda^i U] \end{bmatrix} = \begin{bmatrix} L_w & L_u \end{bmatrix} \begin{bmatrix} [\lambda^i W] \\ [\lambda^i U] \end{bmatrix}$$

and:

$$[\lambda Y] / \begin{bmatrix} [\lambda^{i+1} W] \\ [\lambda^{i+1} U] \end{bmatrix} = \begin{bmatrix} L_{\overline{w}} & L_{\overline{u}} \end{bmatrix} \begin{bmatrix} [\lambda^{i+1} W] \\ [\lambda^{i+1} U] \end{bmatrix}.$$

5. From the singular value decomposition of L_w:

$$\Lambda_i X_i = L_w \left[\lambda^i W\right] = \begin{bmatrix} U_N & U_0 \end{bmatrix} \begin{bmatrix} S_n & 0 \\ 0 & S_0 \end{bmatrix} \begin{bmatrix} V_n' \\ V_0' \end{bmatrix} [\lambda^i W]$$

compute the i^{th}-Λ-extended observabilily matrices:

$$\Lambda_i = U_n S_n^{1/2}, \tag{2.6}$$

where $S_n^{1/2}$ stands for the Cholesky factor of S_n.

6. From extended observabilily matrices Λ_i compute the state sequences:

$$X_i = \Lambda_i^{\dagger} L_w \left[\lambda^i W\right]$$

and:

$$[\lambda X_i] = \Lambda_i^{\dagger} L_{\overline{w}} \left[\lambda^{i+1} W\right].$$

$\Lambda_i^{\dagger}$ stands for the pseudo-inverse of Λ_i.

7. Solve the following optimization problem:

$$\min_k \left\| \begin{bmatrix} X_i \\ [\lambda^i y] \end{bmatrix} \text{-} K \begin{bmatrix} [\lambda X_i] \\ [\lambda^i u] \end{bmatrix} \right\|_F^2 ,$$

where:

$$K := \begin{bmatrix} A_\lambda & B_\lambda \\ \widehat{C} & \widehat{D} \end{bmatrix}.$$

and $\|\alpha\|_F$ stands for the Frobenius norm of α.

8. Finally:

$$\widehat{A} = \frac{A_\lambda \text{-} I}{\tau}$$

and:

$$\widehat{B} = \frac{B_\lambda}{\tau}.$$

In what follows we discuss the numerical properties of the **CN4SID** algoritm.

3 Numerical properties

Since the **CN4SID** algoritm depends essentially on the pseudo-inverse of the Λ-extended observability matrices Λ_i given by (2.6), in this section we discuss the conditioning of these matrices.

We first present some preliminary results concerning the nature of Λ_i. Let us define the i^{th}-extended observability matrix Γ_i as follows:

$$\Gamma_i := \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{i\text{-}1} \end{bmatrix}. \tag{3.7}$$

Then:

Lemma 3.1. *Let Γ_i be as defined in (3.7), then:*

$$\Lambda_i = T\Gamma_i,$$

where:

$$T = \begin{bmatrix} I & 0 & 0 & \cdots & 0 \\ I & \tau I & 0 & \cdots & 0 \\ I & 2\tau I & \tau^2 I & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ I & \binom{i\text{-}1}{1}\tau I & \binom{i\text{-}1}{2}\tau^2 I & \cdots & \tau^{i\text{-}1} I \end{bmatrix} \in \mathbb{R}^{ip\times ip}, \tag{3.8}$$

with:

$$\binom{v}{z} := \frac{v!}{(v\text{-}z)!z!}.$$

Proof. Because of the definition of Λ_i we have that:

$$\Lambda_i = \begin{bmatrix} C \\ CA_\lambda \\ \vdots \\ CA_\lambda^{i\text{-}1} \end{bmatrix}, \tag{3.9}$$

where:

$$A_\lambda = I + \tau A. \tag{3.10}$$

Thus:

$$\Lambda_i = \begin{bmatrix} C \\ C + \tau CA \\ C + 2\tau CA + \tau^2 CA^2 \\ \vdots \\ C + \binom{i\text{-}1}{1}\tau CA + \binom{i\text{-}1}{2}\tau^2 CA^2 + \cdots + \tau^{i\text{-}1} CA^{i\text{-}1} \end{bmatrix} \tag{3.11}$$

which implies:

$$\Lambda_i = T\Gamma_i. \tag{3.12}$$

□

Thus, the numerical conditioning of Λ_i, in terms of the Frobenius norm, is given by:

$$k_F(\Lambda_i) = \|\Lambda_i\|_F \left\|\Lambda_i^{\text{-}1}\right\|_F,$$

and from Lemma 3.1:

$$k_F(\Lambda_i) = \|T\Gamma_i\|_F \left\|(T\Gamma_i)^{-1}\right\|_F .$$

Consequently:

$$k_F(\Lambda_i) \leq \|T\|_F \left\|T^{-1}\right\|_F \|\Gamma_i\|_F \left\|\Gamma_i^{-1}\right\|_F ,$$

i.e.:

$$k_F(\Lambda_i) \leq k_F(T)\, k_F(\Gamma_i) . \tag{3.13}$$

Remark 3.1. *Because of its dependence on the input-output measurements, it is not possible to have* a priori *knowledge of $k_F(\Gamma_i)$. However, it is suitable to have a small value for this conditioning number. As far as $k_F(T)$ is concerned, it is possible to have* a priori *knowledge of its value. For the above reason, we focus our analysis on the study of the conditioning number of T.*

Lemma 3.2. *Let T be as defined in (3.8), then:*

$$\|T\|_F^2 = p \sum_{k=1}^{i} \sum_{j=1}^{k} \binom{k\text{-}1}{j\text{-}1}^2 \tau^{2(j\text{-}1)}.$$

Proof. Let matrix T be as defined in (3.8). Then, the block matrix $T_{kj} \in \mathbb{R}^{p \times p}$ can be written as follows:

$$T_{kj} = \binom{k\text{-}1}{j\text{-}1} \tau^{j\text{-}1} I_{p\times p} \tag{3.14}$$

and its corresponding transpose block matrix is given by:

$$T'_{jl} = \binom{l\text{-}1}{j\text{-}1} \tau^{j\text{-}1} I_{p\times p} \tag{3.15}$$

and consequently the kl-block matrix of TT', i.e. $(TT')_{kl} \in \mathbb{R}^{p\times p}$, is given by:

$$\begin{aligned} (TT')_{kl} &= \sum_{j=1}^{i} \binom{k\text{-}1}{j\text{-}1} \tau^{j\text{-}1} I_{p\times p} \binom{l\text{-}1}{j\text{-}1} \tau^{j\text{-}1} I_{p\times p} \\ &= \sum_{j=1}^{i} \binom{k\text{-}1}{j\text{-}1} \binom{l\text{-}1}{j\text{-}1} \tau^{2(j\text{-}1)} I_{p\times p}. \end{aligned} \tag{3.16}$$

Moreover:

$$\begin{aligned} tr\,(TT') &= \sum_{k=l=1}^{i} \sum_{j=1}^{i} \binom{k\text{-}1}{j\text{-}1} \binom{l\text{-}1}{j\text{-}1} \tau^{2(j\text{-}1)} p \\ &= \sum_{k=1}^{i} \sum_{j=1}^{k} \binom{k\text{-}1}{j\text{-}1}^2 \tau^{2(j\text{-}1)} p, \end{aligned} \tag{3.17}$$

and because of the definition of the Frobenius norm:

$$\begin{aligned} \|T\|_F^2 &= tr\,(TT') \qquad (3.18)\\ &= p\sum_{k=1}^{i}\sum_{j=1}^{k}\binom{k\text{-}1}{j\text{-}1}^2\tau^{2(j\text{-}1)} \qquad (3.19)\end{aligned}$$

□

Remark 3.2. *From the previous lemma we have that $\|T\|_F^2$ is a $2\,(j-1)$ polynomial function of τ. It is usual that $0<\tau<1$. Thus, $\|T\|_F^2$ usually has a small value.*

As far as the Frobenius norm of T^{-1} is concerned, we have the following result, which gives the Frobenius norm of *adj* (T), i.e., the adjunct of T:

Lemma 3.3. *Let T be as defined in (3.8), then:*

$$\|adj\,(T)\|_F^2 = p\tau^{n(n\text{-}1)}\sum_{k=1}^{i}\sum_{j=1}^{k}\binom{k\text{-}1}{j\text{-}1}^2\tau^{\text{-}2(k\text{-}1)},$$

where $p>0$ denotes the dimension of the output vector.

Proof. Let matrix T be as defined in (3.8). Then, the adjunct of T can be written in terms of its kj-block matrix, i.e. $(adjT)_{kj}\in\mathbb{R}^{p\times p}$, as follows:

$$(adjT)_{kj} = (-1)^{k+j}\binom{k\text{-}1}{j\text{-}1}\tau^{\frac{i(i\text{-}1)}{2}\text{-}(k\text{-}1)}I_{p\times p} \qquad (3.20)$$

and the corresponding block matrices of the transpose of *adj* (T) are given by:

$$(adjT)'_{jl} = (-1)^{l+j}\binom{l\text{-}1}{j\text{-}1}\tau^{\frac{i(i\text{-}1)}{2}\text{-}(l\text{-}1)}I_{p\times p}. \qquad (3.21)$$

Then, the block matrices of the matrix product *adj* (T) *adj* $(T)'$ can be written as:

$$\begin{aligned} (adjT\,(adjT'))_{kl} &= \sum_{j=1}^{i}(-1)^{k+j}\binom{k\text{-}1}{j\text{-}1}\tau^{\frac{i(i\text{-}1)}{2}\text{-}(k\text{-}1)}I_{p\times p}\\ &\quad\cdot(-1)^{l+j}\binom{l\text{-}1}{j\text{-}1}\tau^{\frac{i(i\text{-}1)}{2}\text{-}(l\text{-}1)}I_{p\times p}\\ &= \sum_{j=1}^{i}(-1)^{k+l+2j}\binom{k\text{-}1}{j\text{-}1}\binom{l\text{-}1}{j\text{-}1}\tau^{i(i\text{-}1)\text{-}2(k\text{-}1)}I_{p\times p}. \qquad (3.22)\end{aligned}$$

As far as the trace of the matrix product *adj* (T) *adj* $(T)'$ is concerned we have:

$$\begin{aligned} tr\left(adj\,(T)\,adj\,(T)'\right) &= \sum_{k=l=1}^{i}\sum_{j=1}^{i}(-1)^{k+l+2j}\binom{k\text{-}1}{j\text{-}1}\\ &\quad\cdot\binom{l\text{-}1}{j\text{-}1}\tau^{\frac{i(i\text{-}1)}{2}\text{-}(k\text{-}1)}\tau^{\frac{i(i\text{-}1)}{2}\text{-}(l\text{-}1)}p\\ &= \sum_{k=1}^{i}\sum_{j=1}^{i}\binom{k\text{-}1}{j\text{-}1}^2\tau^{i(i\text{-}1)\text{-}2(k\text{-}1)}p. \qquad (3.23)\end{aligned}$$

Consequently:

$$\begin{aligned}\|adj\,(T)\|_F^2 \quad : & = tr\left(adj\,(T)\, adj\,(T)'\right) \\ & = p\tau^{i(i\text{-}1)} \sum_{k=1}^{i}\sum_{j=1}^{k} \binom{k\text{-}1}{j\text{-}1}^2 \tau^{\text{-}2(k\text{-}1)}. \end{aligned} \tag{3.24}$$

□

We can at this level present our main result. Combining Lemma 3.2 and Lemma 3.3 we have:

Theorem 3.1. *Let T be as defined in (3.8), then:*

$$k_F\,(T) = p\tau^{\frac{i(i\text{-}1)(1\text{-}p)}{2}} \sqrt{\left(\sum_{k=1}^{i}\sum_{j=1}^{k}\binom{k\text{-}1}{j\text{-}1}^2 \tau^{2(j\text{-}1)}\right)\left(\sum_{k=1}^{i}\sum_{j=1}^{k}\binom{k\text{-}1}{j\text{-}1}^2 \tau^{\text{-}2(j\text{-}1)}\right)}.$$

Proof. By definition:

$$k_F\,(T) := \|T\|_F \left\|T^{-1}\right\|_F, \tag{3.25}$$

i.e.:

$$\begin{aligned} k_F\,(T) & = \sqrt{tr\,(TT')\, tr\left(T^{-1}\,(T^{-1})'\right)} \\ & = \sqrt{tr\,(TT')\,\frac{tr\left(adj\,(T)\, adj\,(T)'\right)}{\det\,(TT')}}. \end{aligned} \tag{3.26}$$

Now, taking into account Lemma 3.2, Lemma 3.3, and since $\det\,(T) = \tau^{\frac{ip(i\text{-}1)}{2}}$, we have:

$$\begin{aligned} & k_F\,(T) \\ & = \sqrt{\left(p\sum_{k=1}^{i}\sum_{j=1}^{k}\binom{k\text{-}1}{j\text{-}1}^2 \tau^{2(j\text{-}1)}\right)\left(p\tau^{i(i\text{-}1)}\sum_{k=1}^{i}\sum_{j=1}^{k}\binom{k\text{-}1}{j\text{-}1}^2 \tau^{\text{-}2(k\text{-}1)}\right)\tau^{ip(i\text{-}1)}} \\ & = p\tau^{\frac{i(i\text{-}1)(1\text{-}p)}{2}} \sqrt{\left(\sum_{k=1}^{i}\sum_{j=1}^{k}\binom{k\text{-}1}{j\text{-}1}^2 \tau^{2(j\text{-}1)}\right)\left(\sum_{k=1}^{i}\sum_{j=1}^{k}\binom{k\text{-}1}{j\text{-}1}^2 \tau^{\text{-}2(k\text{-}1)}\right)}. \end{aligned} \tag{3.27}$$

□

Remark 3.3. *As is established in Theorem 3.1, the conditioning number of T increases (and so the conditioning number of Λ_i) in an exponential way with both the estimated order of the system, i, and the number of outputs, p.*

References

[1] R. Johansson, M. Verhaegen, and C. T. Chou: "Stochastic Theory of continuous-time state-space identification", *IEEE Transactions on Automatic Control*, Vol. 47, No. 1, pp. 41-51, 1999.

[2] M. Moonen, B. De Moor, L. Vandenberghe, and J. Vandewalle: "On- and off-line identification of linear state-space models", *International Journal of Control*, Vol. 49, No. 1, pp. 219-232, 1989.

[3] P. V. Overschee and B. de Moor: "N4SID: Subspace algorithms for the identification of combined deterministic-stochastic systems", *Automatica*, Vol. 30, No. 1, pp. 75-93, 1994.

[4] M. Verhaegen and P. Dewilde: "Subspace model identification. Part I: The output-error state-space model identification class of algorithms", *International Journal of Control*, Vol. 56, No. 5, pp. 1187-1210, 1992.

[5] M. Verhaegen and P. Dewilde: "Subspace model identification. Part II: The analysis of the elementary output-error state-space model identification algorithm", *International Journal of Control*, Vol. 56, No. 5, pp. 1187-1210, 1992.