A note on persistency of excitation

Jan C. Willems[a], Paolo Rapisarda[b], Ivan Markovsky[a,*], Bart L.M. De Moor[a]

[a]*ESAT, SCD/SISTA, K.U. Leuven, Kasteelpark Arenberg 10, B 3001 Leuven, Heverlee, Belgium*
[b]*Department of Mathematics, University of Maastricht, 6200 MD Maastricht, The Netherlands*

Abstract

We prove that if a component of the response signal of a controllable linear time-invariant system is persistently exciting of sufficiently high order, then the windows of the signal span the full system behavior. This is then applied to obtain conditions under which the state trajectory of a state representation spans the whole state space. The related question of when the matrix formed from a state sequence has linearly independent rows from the matrix formed from an input sequence and a finite number of its shifts is of central importance in subspace system identification.
© 2004 Elsevier B.V. All rights reserved.

Keywords: Behavioral systems; Persistency of excitation; Lags; Annihilators; System identification

1. Introduction

Persistency of excitation of an input or a noise signal is of importance in system identification and adaptive control, see, for example, [1,3–6]. In this paper, we examine consequences of persistency of excitation using the behavioral language.

The problem studied may be posed as follows. Assume that a response

$$\tilde{w}(1), \tilde{w}(2), \ldots, \tilde{w}(T)$$

of a linear time-invariant system is observed. Now consider, for some L, $1 \leqslant L \leqslant T$, the 'windows' of length L:

$$[\tilde{w}(1), \tilde{w}(2), \ldots, \tilde{w}(L)]$$
$$[\tilde{w}(2), \tilde{w}(3), \ldots, \tilde{w}(L+1)]$$
$$\cdots$$
$$[\tilde{w}(T-L+1), \tilde{w}(T-L+2), \ldots, \tilde{w}(T)]. \qquad (1)$$

Under which conditions do these windows span the whole space of all possible windows of length L which the system can produce?

We will show that a sufficient condition for this is that a component (typically the input component) of the observed signal is persistently exciting of order $L+n$, where n equals the dimension of the state space of the system.

* Corresponding author. Tel.: +32 16 321710;
fax: +32 16 321970.

E-mail addresses: jan.willems@esat.kuleuven.ac.be
(J.C. Willems), p.rapisarda@math.unimaas.nl (P. Rapisarda),
ivan.markovsky@esat.kuleuven.ac.be (I. Markovsky),
bart.demoor@esat.kuleuven.ac.be (B.L.M. De Moor).

2. Linear time-invariant systems

We use the behavioral language [7,11]. A dynamical system is defined as $\Sigma = (\mathbb{T}, \mathbb{W}, \mathfrak{B})$, with $\mathbb{T} \subseteq \mathbb{R}$ the time axis, $\mathbb{W}$ the signal space, and $\mathfrak{B} \subseteq \mathbb{W}^{\mathbb{T}}$ the *behavior*. In the present paper, we deal exclusively with discrete-time systems with time axis $\mathbb{T} = \mathbb{N}$, $\mathbb{W}$ a finite-dimensional real vector space (generic notation $\mathbb{W} = \mathbb{R}^w$), and a behavior $\mathfrak{B}$ that is (i) linear, (ii) shift-invariant ($\mathfrak{B} \subseteq \sigma\mathfrak{B}$, where σ denotes the shift: $(\sigma f)(t) := f(t+1)$), and (iii) complete, i.e., $\mathfrak{B}$ is closed in the topology of pointwise convergence. We denote the class of systems $\Sigma = (\mathbb{N}, \mathbb{R}^w, \mathfrak{B})$ satisfying (i)–(iii) by $\mathfrak{L}^w$. Since throughout the time axis equals $\mathbb{N}$, we use both notations $\Sigma \in \mathfrak{L}^w$ and $\mathfrak{B} \in \mathfrak{L}^w$. If w is not specified, we use $\Sigma \in \mathfrak{L}^{\bullet}$ or $\mathfrak{B} \in \mathfrak{L}^{\bullet}$ whence $\mathfrak{L}^{\bullet} = \bigcup_{w \in \mathbb{Z}_+} \mathfrak{L}^w$.

It is well known (see [10, Theorem 5]) that $\mathfrak{B} \in \mathfrak{L}^w$ if and only if there exists a real polynomial matrix $R \in \mathbb{R}^{\bullet \times w}[\xi]$ (this notation means: R is a matrix of polynomials in the indeterminate ξ, with real coefficients, w columns, and any finite number of rows) such that

$$\mathfrak{B} = \{w : \mathbb{N} \to \mathbb{R}^w \mid R(\sigma)w = 0\}.$$

Equivalently, $\mathfrak{B} = \ker(R(\sigma))$. We call $R(\sigma)w = 0$ a *kernel representation* of this $\mathfrak{B}$ or of $\Sigma = (\mathbb{R}, \mathbb{R}^w, \mathfrak{B})$. The kernel representation associated with a given $\mathfrak{B} \in \mathfrak{L}^w$ is not unique, and there exists always one in which the polynomial matrix R has full row rank. Such kernel representations are called *minimal*.

Let $\mathfrak{B} \in \mathfrak{L}^{\bullet}$. Denote, for $T \in \mathbb{N}$, by $\mathfrak{B}|_{[1,T]}$ the $w \in \mathfrak{B}$ restricted to $[1, T]$, i.e.

$$\mathfrak{B}|_{[1,T]} := \{w : [1, T] \to \mathbb{R}^w \mid \exists v \in \mathfrak{B} :$$
$$w(t) = v(t) \text{ for } 1 \leqslant t \leqslant T\}.$$

Closely associated with a kernel representation $\mathfrak{B} \in \mathfrak{L}^w$ is the module of *annihilators* of $\mathfrak{B}$, $\mathfrak{N}_{\mathfrak{B}} \subseteq \mathbb{R}^w[\xi]$, defined by

$$\mathfrak{N}_{\mathfrak{B}} := \{n \in \mathbb{R}^w[\xi] \mid n^{\top}(\sigma)\mathfrak{B} = 0\}.$$

It is easy to see that $\mathfrak{N}_{\mathfrak{B}}$ is a submodule of $\mathbb{R}^w[\xi]$. $\mathfrak{N}_{\ker(R(\sigma))}$ is, in fact, the submodule generated by the rows of R. Consider also for nonnegative integers $\Delta \in \mathbb{Z}_+$ the annihilators of degree less than Δ,

$$\mathfrak{N}_{\mathfrak{B}}^{\Delta} := \{n \in \mathbb{R}^w[\xi] \mid \text{each element of } n \text{ is of degree} \leqslant \Delta\}.$$

Observe that there holds (with apologies for the slight abuse of notation), for $L \in \mathbb{Z}_+$,

$$\mathfrak{N}_{\mathfrak{B}}^{L-1} = \ker(\mathfrak{B}_{[1,L]}).$$

There are a number of important 'integer invariants' associated with $\mathfrak{L}^{\bullet}$. The following are of interest to us in this paper.

- $\mathbf{w} : \mathfrak{L}^{\bullet} \to \mathbb{Z}_+$, the *variable cardinality*. If $\mathfrak{B} \in \mathfrak{L}^w$, then $\mathbf{w}(\mathfrak{B}) := w$.
- $\mathbf{m} : \mathfrak{L}^{\bullet} \to \mathbb{Z}_+$, the *input cardinality*. This may be defined as follows. $\mathbf{m}(\mathfrak{B}) = m$ if there exists an m-dimensional subvector that is free in $\mathfrak{B}$. An m-dimensional subvector w_1 is free in $\mathfrak{B}$ if after permutation $w = (w_1, w_2)$ with $w_1 : \mathbb{N} \to \mathbb{R}^{w_1}$, $w_2 : \mathbb{N} \to \mathbb{R}^{w_2}$, and $w_1 + w_2 = w$, then for all $w_1 : \mathbb{N} \to \mathbb{R}^{w_1}$ there exists a $w_2 : \mathbb{N} \to \mathbb{R}^{w_2}$, such that $(w_1, w_2) \in \mathfrak{B}$.
- $\mathbf{p} : \mathfrak{L}^{\bullet} \to \mathbb{Z}_+$, the *output cardinality*, defined as $\mathbf{p}(\mathfrak{B}) := \mathbf{w}(\mathfrak{B}) - \mathbf{m}(\mathfrak{B})$.
- $\mathbf{n} : \mathfrak{L}^{\bullet} \to \mathbb{Z}_+$, the *state cardinality*. This may be defined as follows. Every $\mathfrak{B} \in \mathfrak{L}^w$ admits a state representation, i.e., there exists $n \in \mathbb{Z}_+$ (called the state dimension) and $\mathfrak{B}' \in \mathfrak{L}^{w+n}$, such that

$$\mathfrak{B} = \{w \mid \exists x : (w, x) \in \mathfrak{B}'\}$$

and such that $\mathfrak{B}'$ satisfies the *state axiom*. This means that if $(w_1, x_1), (w_2, x_2) \in \mathfrak{B}$ and $t_0 \in \mathbb{N}$ satisfy $x_1(t_0) = x_2(t_0)$, then $(w, x) = (w_1, x_1) \wedge_{t_0} (w_2, x_2) \in \mathfrak{B}$. $\wedge_{t_0}$ denotes concatenation at t_0, defined by

$$(w_1, x_1) \wedge_{t_0} (w_2, x_2) := \begin{cases} (w_1, x_1)(t) & \text{for } t < t_0, \\ (w_2, x_2)(t) & \text{for } t \geqslant t_0. \end{cases}$$

The smallest state-space dimension among all state representations of $\mathfrak{B}$ is the *state cardinality* $\mathbf{n}(\mathfrak{B})$ of $\mathfrak{B}$.

- $\mathbf{L} : \mathfrak{L}^{\bullet} \to \mathbb{Z}_+$, the *lag*. This may be defined as follows. Let $R(\sigma)w = 0$ be a kernel representation of $\mathfrak{B}$. The maximum of the degrees of the polynomial elements of R is called the lag associated with this particular kernel representation. $\mathbf{L}(\mathfrak{B})$ is the smallest possible lag over all kernel representations of $\mathfrak{B}$. In fact, $\mathbf{L}(\mathfrak{B})$ is also the smallest Δ such that $\mathfrak{N}_{\mathfrak{B}}^{\Delta}$ generates the module $\mathfrak{N}_{\mathfrak{B}}$. In particular, there exist a kernel representation of $\mathfrak{B}$ with equation lags less than or equal to $\mathbf{L}(\mathfrak{B})$.

- $\mathbf{l} : \mathfrak{L}^{\bullet} \to \mathbb{Z}_+$, the *shortest lag*. This may be defined as follows. Let $R(\sigma)w = 0$ be a kernel representation of $\mathfrak{B}$. Define the degree of a vector of polynomials to be equal to the largest of the degrees of the entries. The minimum of the degrees of the rows of R is called the minimal lag associated with this kernel representation: each equation in $R(\sigma)w = 0$ involves lags at least equal to the minimal lag. $\mathbf{l}(\mathfrak{B})$ is the smallest possible minimal lag over all kernel representations of $\mathfrak{B}$. Every equation in every kernel representation of $\mathfrak{B}$ has lag at least $\mathbf{l}(\mathfrak{B})$.

These integers are all readily computable from a kernel representation, and certainly from an input/state/output representation of $\mathfrak{B}$ (see [10, Section 7]). It is, for example, possible to prove that

$$\mathbf{n}(\mathfrak{B}) \geqslant \mathbf{L}(\mathfrak{B}).$$

Also,

$$L \geqslant \mathbf{L}(\mathfrak{B}) \Leftrightarrow \dim(\mathfrak{B}|_{[1,L]}) = \mathbf{m}(\mathfrak{B})L + \mathbf{n}(\mathfrak{B})$$
$$\Leftrightarrow \dim(\mathfrak{N}_{\mathfrak{B}}^{L-1}) = \mathbf{p}(\mathfrak{B})L - \mathbf{n}(\mathfrak{B}),$$

and

$$L \leqslant \mathbf{l}(\mathfrak{B}) \Leftrightarrow \dim(\mathfrak{B}|_{[1,L]}) = \mathbf{w}(\mathfrak{B})L$$
$$\Leftrightarrow \dim(\mathfrak{N}_{\mathfrak{B}}^{L-1}) = 0.$$

Recall that $\mathfrak{B} \in \mathfrak{L}^{\bullet}$ is said to be *controllable* $:\Leftrightarrow$ for all $T \in \mathbb{N}$, $w_1 \in \mathfrak{B}|_{[1,T]}$, and $w_2 \in \mathfrak{B}$, there exists $v \in \mathfrak{B}$ and $T' \in \mathbb{N}$, such that $v|_{[1,T]} = w_1$ and $w_2(t - T - T') = v(t)$ for $t > T + T'$. Denote by $\mathfrak{L}^{\bullet}_{\text{controllable}}$ and $\mathfrak{L}^{w}_{\text{controllable}}$ of the controllable elements $\mathfrak{L}^{\bullet}$ and $\mathfrak{L}^{w}$, respectively.

3. Sequences with spanning windows

Let $\mathfrak{B} \in \mathfrak{L}^{\bullet}$, and assume that a finite trajectory $\tilde{w} \in \mathfrak{B}|_{[1,T]}$ is 'observed'. *Under which conditions it is possible to recover from $\tilde{w}$ the laws of the system $\mathfrak{B}$ that generated $\tilde{w}$?* This question is closely related to the question asked in the introduction: *Under which conditions do the observed windows of length L span the space of all possible windows of length L which the system can produce?*

Define the Hankel matrix of *depth L* associated with the vector signal $f(1), f(2), \ldots, f(T)$ by

$$\mathfrak{H}_L(f)$$
$$:= \begin{bmatrix} f(1) & f(2) & \cdots & f(T - L + 1) \\ f(2) & f(3) & \cdots & f(T - L + 2) \\ \vdots & \vdots & & \vdots \\ f(L) & f(L+1) & \cdots & f(T) \end{bmatrix}.$$

Note that the columns of the Hankel matrix $\mathfrak{H}_L(\tilde{w})$ correspond to the windows of $\tilde{w}$ displayed in the introduction. Of course, since $\tilde{w} \in \mathfrak{B}|_{[1,T]}$, any $n \in \mathfrak{N}_{\mathfrak{B}}^{L-1}$, $n(\xi) = n_0 + n_1\xi + \cdots + n_{L-1}\xi^{L-1}$, is such that

$$[n_0^{\top} \ n_1^{\top} \ \cdots \ n_{L-1}^{\top}]\mathfrak{H}_L(\tilde{w}) = 0.$$

Therefore, the left kernel of $\mathfrak{H}_L(\tilde{w})$ contains the vectors generated by the elements of $\mathfrak{N}_{\mathfrak{B}}^{L-1}$. The question is: *When are there no other annihilators?* Equivalently (with a very slight abuse of notation): When is leftkernel($\mathfrak{H}_L(\tilde{w})$) $= \mathfrak{N}_{\mathfrak{B}}^{L-1}$, equivalently, rowspan($\mathfrak{H}_L(\tilde{w})$) $= \mathfrak{B}|_{[1,L]}$?

Crucial in our result is the persistency of excitation of a component (typically, the input component) of $\tilde{w}$. The signal $f = [1, T] \cap \mathbb{N} \to \mathbb{R}^f$ is said to be *persistently exciting of order L* $:\Leftrightarrow$ rank($\mathfrak{H}_L(f)$) $= Lf$, i.e., if there exist no non-trivial linear relations of order L among the $f(t)$'s. In other words, there are no $a_1, a_2, \ldots, a_L \in \mathbb{R}^f$, not all zero, such that $a_1^{\top} f(t) + a_2^{\top} f(t+1) + \cdots + a_L^{\top} f(t + L - 1) = 0$, for $t = 1, 2, \ldots, T - L$.

The following is the main result of the paper.

Theorem 1. *Consider* $\mathfrak{B} \in \mathfrak{L}^{w}_{\text{controllable}}$. *Let* $\tilde{u} : [1, T] \to \mathbb{R}^{\mathbf{m}(\mathfrak{B})}$, $\tilde{y} : [1, T] \to \mathbb{R}^{\mathbf{p}(\mathfrak{B})}$, *and* $\tilde{w} = (\tilde{u}, \tilde{y})$. *Assume that* $\tilde{w} \in \mathfrak{B}|_{[1,T]}$. *Then, if* $\tilde{u}$ *is persistently exciting of order* $L + \mathbf{n}(\mathfrak{B})$,

$$\text{leftkernel}(\mathfrak{H}_L(\tilde{w})) = \mathfrak{N}_{\mathfrak{B}}^{L}, \tag{K}$$

and

$$\text{rowspan}(\mathfrak{H}_L(\tilde{w})) = \mathfrak{B}|_{[1,L]}. \tag{I}$$

Proof. We only need to prove (K). The inclusion leftkernel($\mathfrak{H}_L(\tilde{w})$) $\supseteq \mathfrak{N}_{\mathfrak{B}}^{L}$ is obvious.

Consider the reverse inclusion: leftkernel($\mathfrak{H}_L(\tilde{w})$) $\subseteq \mathfrak{N}_{\mathfrak{B}}^{L}$. Assume, to the contrary, that

$$0 \neq r^{\top} = [r_0^{\top} \ r_1^{\top} \ \cdots \ r_{L-1}^{\top}] \in \text{leftkernel}(\mathfrak{H}_L(\tilde{w}))$$

but $r(\xi) = r_0 + r_1 \xi + \cdots + r_{L-1} \xi^{L-1} \notin \mathfrak{N}_{\mathfrak{B}}^{L-1}$. Consider $\mathfrak{H}_{L+\mathbf{n}(\mathfrak{B})}(\tilde{w})$. Obviously, leftkernel $(\mathfrak{H}_{L+\mathbf{n}(\mathfrak{B})}(\tilde{w}))$ contains $\mathfrak{N}_{\mathfrak{B}}^{L+\mathbf{n}(\mathfrak{B})-1} + \mathfrak{R}$, with $\mathfrak{R} \subset \mathbb{R}^w[\xi]$ the linear span of

$$\mathfrak{R} = \operatorname{span}\{r^\top(\xi), \xi r^\top(\xi), \ldots, \xi^{\mathbf{n}(\mathfrak{B})} r^\top(\xi)\}.$$

Recall that

$$\dim(\mathfrak{N}_{\mathfrak{B}}^{L+\mathbf{n}(\mathfrak{B})-1}) = (L + \mathbf{n}(\mathfrak{B}))\mathbf{p}(\mathfrak{B}) - \mathbf{n}(\mathfrak{B}).$$

Clearly, $\dim(\mathfrak{R}) = \mathbf{n}(\mathfrak{B}) + 1$. We now show that the persistency of excitation assumption implies $\mathfrak{R} \cap \mathfrak{N}_{\mathfrak{B}}^{L+\mathbf{n}(\mathfrak{B})} \neq \{0\}$. If $\mathfrak{R} \cap \mathfrak{N}_{\mathfrak{B}}^{L+\mathbf{n}(\mathfrak{B})} = \{0\}$, then

$$\dim(\mathfrak{N}_{\mathfrak{B}}^{L+\mathbf{n}(\mathfrak{B})-1} + \mathfrak{R}) = (L + \mathbf{n}(\mathfrak{B}))\mathbf{p}(\mathfrak{B}) + 1.$$

But the persistency of excitation implies

$$\operatorname{rank}(\mathfrak{H}_{L+\mathbf{n}(\mathfrak{B})}(\tilde{w})) \geqslant (L + \mathbf{n}(\mathfrak{B}))\mathbf{m}(\mathfrak{B}).$$

Hence

$$\dim(\mathfrak{N}_{\mathfrak{B}}^{L+\mathbf{n}(\mathfrak{B})-1} + \mathfrak{R})$$
$$= (L + \mathbf{n}(\mathfrak{B}))\mathbf{p}(\mathfrak{B}) + 1$$
$$\leqslant \dim(\operatorname{leftkernel}(\mathfrak{H}_{L+\mathbf{n}(\mathfrak{B})}(\tilde{w})))$$
$$\leqslant (L + \mathbf{n}(\mathfrak{B}))\mathbf{p}(\mathfrak{B}).$$

Therefore $\mathfrak{R} \cap \mathfrak{N}_{\mathfrak{B}}^{L+\mathbf{n}(\mathfrak{B})} \neq \{0\}$.

Consequently, there is a linear combination of

$$r^\top(\xi), \xi r^\top(\xi), \ldots, \xi^{\mathbf{n}(\mathfrak{B})} r^\top(\xi),$$

that is contained in $\mathfrak{N}_{\mathfrak{B}}^{L+\mathbf{n}(\mathfrak{B})}$. In terms of the minimal kernel representation $R(\frac{d}{dt})w = 0$ of $\mathfrak{B}$, this means that there is $0 \neq f \in \mathbb{R}[\xi]$, such that $fr = FR$, for some $0 \neq F \in \mathbb{R}^{1 \times \operatorname{rowdim}(R)}[\xi]$. If $\deg(f) \geqslant 1$, then there is $\lambda' \in \mathbb{C}$, such that $f(\lambda') = 0$, hence $F(\lambda')R(\lambda') = 0$. Now use the well-known fact [11] that $R(\sigma)w = 0$ of $\mathfrak{B}$ is a minimal kernel representation of a controllable behavior if and only if $R(\lambda)$ has full row rank for all $\lambda \in \mathbb{C}$. Hence controllability implies $F(\lambda') = 0$. This implies that f and each element of F have a common root λ'. Cancel this common factor. Proceed until $\deg(f) = 0$. But then $r = FR$. This contradicts the assumption $r^\top \notin \mathfrak{N}_{\mathfrak{B}}^{L-1}$. Hence leftkernel$(\mathfrak{H}_L(\tilde{w})) \subseteq \mathfrak{N}_{\mathfrak{B}}^{L-1}$, and (K) holds. $\square$

4. Comments and corollaries

1. The interesting, and somewhat surprising, part of Theorem 1 is that persistency of excitation of order $L + \mathbf{n}(\mathfrak{B})$ is needed in order to be able to deduce that the observed sequences (1) of length L have the 'correct' annihilators and the 'correct' span. In other words, we have to assume a 'deeper' persistency of excitation on $\tilde{u}$ than the width of the windows of $(\tilde{u}, \tilde{y})$ which are considered.
2. Note that Theorem 1 holds for all L (and not just for $L > \mathbf{L}(\mathfrak{B})$). So, in particular, if $L \leqslant \mathbf{l}(\mathfrak{B})$, and under persistency of excitation of order $L + \mathbf{n}(\mathfrak{B})$, $\mathfrak{H}_L(\tilde{w})$ has full row rank. Also, if $L > \mathbf{L}(\mathfrak{B})$, and under persistency of excitation of order $L + \mathbf{n}(\mathfrak{B})$, the left kernel of $\mathfrak{H}_L(\tilde{w})$ (identified in the obvious way with polynomial vectors) generates the full annihilator module $\mathfrak{N}_{\mathfrak{B}}$. The observed system signal then completely specifies the laws of the system.
3. An interesting special case is when $\mathfrak{B}$ is the usual state space system $\sigma x = Ax + Bu$. Note that for this system, $\mathbf{L}(\mathfrak{B}) = 1$. Theorem 1 yields the following corollary.

Corollary 2. *Assume that $\sigma x = Ax + Bu$ is controllable. Consider a trajectory $\tilde{u}(1), \tilde{u}(2), \ldots, \tilde{u}(T)$; $\tilde{x}(1), \tilde{x}(2), \ldots, \tilde{x}(T)$ of this system. Then*

(i) *If $\tilde{u}$ is persistently exciting of order $\dim(x)$, then* rank$[\tilde{x}(1) \ \tilde{x}(2) \ \cdots \ \tilde{x}(T)] = \dim(x) + 1$.

(ii) *If $\tilde{u}$ is persistently exciting of order $\dim(x) + 1$, then rank*

$$\begin{bmatrix} \tilde{u}(1) & \cdots & \tilde{u}(T) \\ \tilde{x}(1) & \cdots & \tilde{x}(T) \end{bmatrix} = \dim(x) + \dim(u).$$

(iii) *If $\tilde{u}$ is persistently exciting of order $\dim(x) + L$, then rank*

$$\begin{bmatrix} \mathfrak{H}_L(\tilde{u}) \\ \tilde{X} \end{bmatrix} = \dim(x) + L\dim(u),$$

where

$$\tilde{X} := [\tilde{x}(1) \ \cdots \ \tilde{x}(T - L + 1)].$$

In [9, Section 3.3] the conditions rank $\begin{bmatrix} \mathfrak{H}_L(\tilde{u}) \\ \tilde{X} \end{bmatrix} = \dim(x) + L\dim(u)$ is recognized to have a crucial

role in subspace system identification. To the best of our knowledge, however, a test to verify it from the given data $(\tilde{u}, \tilde{y})$ that is an arbitrary response of the system is not available in the literature. Special cases that were studied are: u white noise [2] and u periodic [9, Theorem 2]. Corollary 4 gives such a test for an arbitrary u.

4. The matrices A, B, C, D of the system

$$\sigma x = Ax + Bu, \quad y = Cx + Du$$

can be recovered from the input/state/output trajectory

$$\begin{bmatrix} \tilde{u}(1) \\ \tilde{x}(1) \\ \tilde{y}(1) \end{bmatrix}, \begin{bmatrix} \tilde{u}(2) \\ \tilde{x}(2) \\ \tilde{y}(2) \end{bmatrix}, \ldots, \begin{bmatrix} \tilde{u}(T) \\ \tilde{x}(T) \\ \tilde{y}(T) \end{bmatrix}$$

(think of the input/output as measured directly, and the state computed using a subspace algorithm) if $\tilde{u}$ is persistently exciting of order $\dim(x) + 2$.

5. Let $w = (u, y)$ with u the input and y the output of $\mathfrak{B} \in \mathfrak{L}^\bullet_{\text{controllable}}$. Assume that the system is driven by a 'random' input $\tilde{u}$, meaning an input that is persistently exciting of any order. How many (exact) data points

$$\begin{bmatrix} \tilde{u}(1) \\ \tilde{y}(1), \end{bmatrix}, \begin{bmatrix} \tilde{u}(2) \\ \tilde{y}(2) \end{bmatrix}, \ldots, \begin{bmatrix} \tilde{u}(T) \\ \tilde{y}(T) \end{bmatrix},$$

input/output measurements, do we need in order to be able to identify the system? The left kernel of $\mathfrak{H}_{\mathbf{L}(\mathfrak{B})+1}\left(\begin{bmatrix} \tilde{u} \\ \tilde{y} \end{bmatrix}\right)$ will give us the laws, provided $\tilde{u}$ is persistently exciting of order $\mathbf{L}(\mathfrak{B}) + \mathbf{n}(\mathfrak{B}) + 1$. This yields the inequality

$$T \geqslant (\mathbf{L}(\mathfrak{B}) + \mathbf{n}(\mathfrak{B}) + 1)\mathbf{m}(\mathfrak{B}) + \mathbf{L}(\mathfrak{B}) + \mathbf{n}(\mathfrak{B}).$$

Adapted for the case $D = 0$ and known zero initial conditions, our bound of T is the same as the one derived in [8].

Acknowledgements

References

[1] M. Cadic, J.W. Polderman, I.M.Y. Mareels, Set membership identification for adaptive control: input design, in: Proceedings of the 42nd IEEE Conference on Decision and Control, Maui, Hawaii, 2003, pp. 5011–5026.
[2] B. Gopinath, On the identification of linear time-invariant systems from input–output data, Bell System Tech. J. 48 (5) (1969) 1101–1113.
[3] I.M.Y. Mareels, Sufficiency of excitation, Systems Control Lett. 5 (1984) 159–163.
[4] I.M.Y. Mareels, R.R. Bitmead, M. Gevers, C.R. Johnson, R.L. Kosut, M.A. Poubelle, How exciting can a signal really be? Systems Control Lett. 8 (1987) 197–204.
[5] I.M.Y. Mareels, M. Gevers, Persistence of excitation criteria for linear, multivariable, time-varying systems, Math. Control Signals Systems 1 (1988) 203–226.
[6] I.M.Y. Mareels, J.W. Polderman, Adaptive Systems: An Introduction, Birkhäuser, Basel, 1996.
[7] J.W. Polderman, J.C. Willems, Introduction to Mathematical Systems Theory, Springer, New York, 1998.
[8] E. Sontag, On the length of inputs necessary in order to identify a deterministic linear system, IEEE Trans. Automat. Control 25 (1) (1980) 120–121.
[9] M. Verhaegen, P. Dewilde, Subspace model identification, part I: the output error state space model identification class of algorithms, Int. J. Control 56 (1992) 1187–1210.
[10] J.C. Willems, From time series to linear system—part I. Finite dimensional linear time invariant systems, Automatica 22 (5) (1986) 561–580.
[11] J.C. Willems, Paradigms and puzzles in the theory of dynamical systems, IEEE Trans. Automat. Control 36 (3) (1991) 259–294.