SYS ID 1991 C52

Identifiability of scalar linearly parametrized polynomial systems *

Soura DASGUPTA⋆, Michel GEVERS, Georges BASTIN, Guy CAMPION, Libei CHEN
⋆Dept of Electrical Engineering, University of Iowa and
Laboratoire d'Automatique, de Dynamique et d'Analyse des Systèmes
Louvain University
Place du Levant, 3
B-1348 Louvain-la-Neuve, Belgium

Abstract - We present necessary and sufficient conditions for the identification of a vector θ of unknown parameters in scalar dynamical systems of the form $\dot{x} = f(x,\theta) + g(x,u)$, where $f(x,\theta)$ is a polynomial function of the state x that is affine in θ, and $g(x,u)$ is a polynomial in the state x and the input u. It is assumed that u and x are measured.
Keywords : Identifiability, polynomial systems.

1 Introduction

This paper is part of a continuing saga on identifiability and transfer of excitation in ever broader classes of systems that are parametrized in a linear or at least affine way. It is well known that the estimation (or, possibly, the tracking) of a parameter vector θ that appears linearly into a dynamical system hinges on two questions :

i) is the structure of the input-output dynamical system identifiable, i.e. are there no two different values of θ that produce the same input-output relationship? It can be shown that this is equivalent with the question of whether the regression vector (i.e. the vector made up of the coefficients of the components θ_i) can be made to span the whole space by a suitable choice of inputs. This is the identifiability question.

ii) the selection of "sufficiently rich inputs", i.e. the selection of inputs that will make the regression vector span the space. This is the input richness question.

In many cases, we will want the regression vector to be "persistently exciting" (i.e. to persistently span the whole space) rather than just spanning, either to guarantee exponential convergence of the estimate $\hat{\theta}$ in the case of a system with constant parameters, or to guarantee robustness of the estimation algorithm to parameter variations in the case of a time-varying system.

The identifiability of linearly parametrized linear time-invariant (LTI) systems is well established : it amounts to establishing that the dynamical system from the input to the regressor vector is completely reachable. For scalar input-output (or transfer function) models, this amounts to establishing that there are no common factors between the numerator and denominator polynomials. The sufficient richness conditions on the input for the identification of LTI systems are also classical : see e.g. Ljung (1987) or Narendra and Annaswamy (1989). Glad and Ljung (1990) have derived a procedure for detecting identifiability for a general class of non linear analytic systems, using concepts from differential algebra, and Ritt's algorithm. For the class of problems stated in the present paper, their procedure does not give simple explicit identifiability conditions like the ones we derive in this paper.

The tracking of a time-varying vector $\theta(t)$ in the case of linearly parametrized linear time-varying systems is an important problem in adaptive estimation and adaptive control, where various tricks have been proposed to cope with a transient loss of identifiability due to lack of excitation of the regressor. The transfer of excitation from input to regressor for time-varying systems has been studied in Mareels and Gevers (1988), Dasgupta, Anderson and Kaye(1988) and others.

Somewhat surprisingly perhaps, for linearly parametrized non linear systems the identifiability of the parameter vector θ is not equivalent to the reachability of the regression vector from the input. For example, a regression vector $\varphi^T(x) = (x\ x^2)$ is definitely not reachable whatever its connection to the input in a dynamical system, and yet a vector $\theta^T = (\theta_1 \theta_2)$ appearing in the form $\theta_1 x + \theta_2 x^2$ within a dynamical system may well be identifiable. This example indicates that the identifiability and input richness conditions may well be less stringent for some non linear systems than they are for linear systems.

An interesting class of non linear systems for which the problems of identifiability and input richness were completely solved are the set of discrete time systems described by difference equations which are polynomial in the input and linear in the output : see Dasgupta et al (1990). The bilinear systems are a subset of this class.

In this paper, we build on some of the ideas of Dasgupta et al (1990) to obtain necessary and sufficient identifiability conditions for a new class of systems, namely linearly

*The results presented in this paper have been obtained within the framework on Interuniversity Attraction Poles initiated by the Belgian State, Prime Minister's Office, Science Policy Programming. The Scientific responsibility rests with its authors. This work has been partially supported by NSF Grant-9001170.

In Definitions 1 and 2 we have required the regressor $\varphi(x(t))$ to span $\mathbf{R}^n$ for $t \in (0,\infty)$. Sometimes one will want to impose the stronger requirement that $\varphi(x(t))$ be *persistently exciting*: this will result in exponential convergence of $\hat{\theta}(t)$ to its true value with any reasonable recursive algorithm. We shall use the following definition for the persistency of excitation of $\varphi(.)$, which is now considered as a time function.

Definition 5. *A vector $\varphi(t) \in \mathbf{R}^n$ is called persistently exciting (PE) if $\exists \alpha > 0, \beta > 0, T > 0$ and $t_0 \geq 0$ such that*

$$\alpha I \leq \int_t^{t+T} \varphi(s)\varphi^T(s)ds \leq \beta I \quad \forall t \geq t_0 \tag{6}$$

We shall see later that, with model (1), if we can make $\varphi(t)$ span $\mathbf{R}^n$ by proper choice of the input function $u(t)$, we can also make it persistently exciting.

Having defined our model class and established a precise definition of identifiability, we can now move on to present our main results on identifiability of the model class (1).

3 Identifiability at a given θ.

In this section we examine the identifiability of (1) at a given θ (see Definition 1); we shall turn in the next section to structural identifiability (Definition 2).
The question of identifiability can be conceptually decomposed in two steps.

1. Can $\varphi(x)$ be made to span $\mathbf{R}^n$ if there are no constraints on x, i.e. independently of the fact that x is generated by the system (1)? This is a structural condition on $\varphi(x)$ only, independent of θ, of $x(0)$ or of $u(.)$. The answer to that question is easy and will be given by Lemma 1 below. Clearly, this will provide a necessary condition for identifiability: if $\varphi(x)$ cannot span $\mathbf{R}^n$ when x takes any values in $\mathbf{R}$, then (1) is not identifiable whatever the functions $m(x)$ and $g(x,u)$, and whatever the input function $u(.)$.

2. Subject to $\varphi(x)$ being able to generate $\mathbf{R}^n$ when x is free, the next question is whether the solution $x(t)$ generated by the particular model (1) can be made to produce a vector time function $\varphi(x(t))$ that spans $\mathbf{R}^n$. Here the value of θ, the choice of $u(t)$ and the initial condition $x(0)$ obviously play a role. Our way to handle this problem will be to use Definition 3, equivalent with Definition 1 as shown above. Following that definition, if (1) is not identifiable, then for every analytic input function $u(t)$, there exists a constant non-zero $\beta \in \mathbf{R}^n$ and a solution $x(t)$ of (1) such that

$$\beta^T\varphi(x(t)) \equiv 0 \quad \forall t \geq 0. \tag{7}$$

Since $x(t)$ is an analytic function of time, this is equivalent with

$$\beta^T\varphi(x) = \beta^T\dot{\varphi}(x) = \beta^T\ddot{\varphi}(x) = \ldots = 0 \tag{8}$$

at some particular value $x(t_0)$. Our (magic) tool will be to define $\beta^T\varphi(x(t))$, for any arbitrary t, as a function $h(x,\beta)$ and to examine whether for all u there exists a common solution x and β to all the equations (8) subject to the system equation (1). We shall write the equations (8) as functions of x and its derivatives and, through a process of successive substitutions using the system equation, transform these into a set of polynomial equations in x, β and u. Through the use of the analyticity Assumption A, the variable t will thus have disappeared from these equations, and the problem will have been reduced to one of searching for a common root to a set of polynomials in several variables.

Enough for the foreplay; let us now get into the real thing. As is often the case in such circumstances, we shall need some preliminary technical tools.

Lemma 1. *Let $\varphi(x) : \mathbf{R} \to \mathbf{R}^n$ be a polynomial vector and let q be the degree of $\varphi(x)$ (i.e. the degree of the highest degree polynomial in $\varphi(x)$). Then there exist $x_1, \ldots, x_n$ such that the $n \times n$ matrix*

$$L(x_1,\ldots,x_n) \triangleq [\varphi(x_1),\ldots,\varphi(x_n)] \tag{9}$$

is nonsingular if and only if

$$J_q(x) \triangleq [\varphi(x) \frac{\partial\varphi(x)}{\partial x} \frac{\partial^2\varphi(x)}{\partial x^2} \cdots \frac{\partial^q\varphi(x)}{\partial x^q}] \tag{10}$$

has full row rank $\forall x \in \mathbf{R}$.

Proof: The following are a sequence of equivalent statements.

(i) $L(x_1,\ldots,x_n)$ is nonsingular for some $x_1,\ldots,x_n$

(ii) $\forall \beta \in \mathbf{R}^n$ with $\beta \neq 0$, $\beta^T\varphi(x)$ is not the zero polynomial (it is a not-identically-zero polynomial of degree smaller than or equal to q)

(iii) $\forall \beta \in \mathbf{R}^n$ with $\beta \neq 0$, either $\beta^T\varphi(x)$ is a non-zero constant or one derivative of $\beta^T\varphi(x)$ w.r.t. x is a non-zero constant :

$$\forall x_0 \in \mathbf{R}, \beta^T\varphi(x_0) \neq 0 \text{ or } \beta^T\frac{\partial^i\varphi}{\partial x^i}\Big|_{x=x_0} \neq 0$$

(iv) $J_q(x)$ has full row rank $\forall x_0 \in \mathbf{R}$, noting that :

$$\frac{\partial^i\varphi(x)}{\partial x^i} \equiv 0 \quad \forall i > q$$

■

Lemma 2. *Consider the system (1). Let q be the degree of $\varphi(x)$ and let the matrix $J_q(x)$ be defined by (10). Then (1) is identifiable at θ (or structurally identifiable) only if $J_q(x)$ has rank n for all $x \in \mathbf{R}$. In particular this requires $q \geq n-1$.*

Proof: Follows immediately from Lemma 1 and Definitions 1 and 2.

Example 1. *Consider $\varphi_1^T = (x+1 \;\; 2x-1 \;\; 1)$, $\varphi_2^T = (1 \;\; x^3-1 \;\; x+1)$, $\varphi_3^T = (x^2 \;\; x+2 \;\; x-1)$, $\varphi_4^T = (x^2 \;\; -x^2+1 \;\; 1)$. Then a system of the form (1) will be unidentifiable whatever the value of θ if the regressor takes the form φ_1, or φ_4, but is possibly identifiable for the regressors φ_2 and φ_3.*

To examine the second part of the identifiability question we need the following technical lemma.

Lemma 3. *Consider a scalar polynomial $a(x)$ in $x \in \mathbf{R}$ and a set of scalar polynomials $\{b_i(x) + c_i^T(x)U, i = 1,\ldots,p\}$ where $b_i(x)$ are scalar, $c_i^T(x)$ are row vectors of polynomials in x, $a(x)$ and at least one of the $b_i(x)$ are not the zero polynomial, and U is a vector of monomials in some*

(i) of Theorem 1 will of course remain necessary, being independent of θ. As for condition (ii), we note that the coprimeness of $f(x,\theta)$ and $g_i(x), i = 1, ...l$, does of course depend on the value of θ. It follows immediately from Definition 2 and Theorem 1 that the system (1) will be structurally identifiable if these polynomials are coprime for all but a thin set of θ. We have the following necessary and sufficient conditions for structural identifiability.

Theorem 2
Suppose there exists $i \in 1,\dots,l$ such that $g_i(x)$ is not the zero polynomial. Then, (1) is structurally identifiable iff :

(i) $J_q(x)$ is full rank, for all $x \in \mathbf{R}$

(ii) $[\varphi^T(x), m(x), g_1(x), \dots, g_l(x)] \neq 0$ for all $x \in \mathbf{R}$

Proof: Follows from Theorem 1 and a direct application of Lemma 3, replacing $u_1,\dots,u_k$ by $\theta_1,\dots,\theta_n$, and $a(x)$ by one $g_i(x)$ which is non zero.

Corollary. *If one of the $g_i(x)$ is not the zero polynomial, then (1) is structurally identifiable iff there exists some $\theta \in \mathbf{R}^n$, at which (1) is identifiable.*

The importance of the assumption that at least one of the $g_i(x)$ is not the zero polynomial is illustrated by the following system with no input (i.e. $g(x) = 0$) :

$$\dot{x} = \theta_1(x^2 - 1) - \theta_2$$

Obvously, for any θ satisfying $\theta_1(\theta_1 + \theta_2) > 0$ the system is not identifiable, because the initial condition $x(0) = \sqrt{\frac{\theta_1+\theta_2}{\theta_1}}$ ensures $x(t) = x(0)$ $\forall t$. This condition on θ defines a thick set in the θ space, and the system is therefore not structurally identifiable. Nevertheless it is identifiable in a thick set which is complementary to the first one.

The corollary is a remarkable result in that it generalizes to polynomial systems a property of linear time invariant systems.

5 Conclusions

We have studied the identifiability of the vector θ in scalar polynomial systems of the form (1). We have shown that, since identifiability requires transfer of excitation from the input $u(.)$ to the polynomial regressor $\varphi(.)$, and since this transfer may depend on the particular value of the parameter vector θ, one has to distinguish between identifiability at a particular value θ^* and structural identifiability. We have produced necessary and sufficient conditions for both situations. These necessary and sufficient conditions are expressed in terms of remarkably simple structural conditions on $\varphi(x)$ and common root conditions on the polynomials $f(x,\theta)$, $m(x)$ and $g_i(x), i = 1, ..., l$. In addition to these identifiability results, we have also obtained sufficient conditions on the input richness for the excitation of $\varphi(x)$. These will be presented in a full version of this paper. Finally, we should mention that the extension of these results to polynomial systems with vector states is by no means trivial and is the subject of our present investigations.

References

Dasgupta S., B.D.O. Anderson , R.J. Kaye (1988) *"Identification of physical parameters in structured systems"*, Automatica, vol. 24, n 2, pp. 217-225 (March 1988).

Dasgupta S., Y. Shrisvastava ,G. Krenzer (1990), *"Persistent excitation in bilinear systems"*, to appear in IEEE Trans on Automatic Control.

Ljung L. (1987), *"System identification : theory for the user"*, Prentice Hall 1987.

Mareels I., M. Gevers (1988), *"Persistence of excitation criteria for linear multivariable time-varying systems"*, Mathematics of Control, Signals and Systems, vol.1, n 3, pp. 203-226 (1988).

Narendra K.S., M. Annaswamy. (1989) *"Stable Adaptive Systems"*, Prentice Hall, 1989.

Youla D.C., G. Gnavi.(1979), *"Notes on n-dimensional systems theory"*, IEEE Trans on Circuits and Systems, vol. CAS 26, n 2, pp. 105-111, Feb. 1979.

Glad S.T., Ljung L. (1990), *" Model structure identifiability and persistence of excitation"*, Proc. 29 th CDC Hawaii (1990) pp. 3236-3240.