Systems & Control Letters 17 (1991) 185–189
North-Holland

Local parametrizations of ARMAX systems with nonlinear restrictions

Liqun Wang
Institut für Statistik und Oekonometrie, Universität Basel, Petersgraben 51, CH-4051 Basel, Switzerland

Received 18 August 1990
Revised 15 March 1991

Abstract: A necessary and sufficient condition of local identifiability is derived by using the well-known Rank Theorem. Using this condition a suitable local parametrization is defined. The topological and geometrical properties of this parametrization are investigated. The results obtained generalize the corresponding results of Deistler and Wang (1989), which deal with essentially the systems with linear restrictions. The treatment covers also the systems without a priori restrictions.

Keywords: Linear dynamic systems; structural identifiability; parametrizations; topological and geometrical properties.

1. Introduction

The problems of identifiability and parametrizations of linear dynamic systems have been intensively studied in past twenty years and rather complete structure theory for systems without a priori restrictions in system parameters has been obtained. See e.g. [7], Chapter 2. It turns out that most usual parametrizations share many common topological and geometrical properties which are important for system identification. These properties have been derived by different authors for different parametrizations. Recently Deistler and Wang [5] proposed a general framework to treat the problem and generalized these results to the ARMAX systems with general linear affine restrictions. The treatment covers most parametrizations discussed before.

However, there are some important cases, e.g. the case of overlapping parametrizations of the manifold of all transfer functions of a given order, in which the system parameters are subject to nonlinear restrictions due to the causality requirement of the transfer functions. See e.g. [3;7, Chapter 2; 5]. Although for this case the transformed form of systems (the MFD's $(\tilde{a}, \tilde{b})$ of the transfer functions $\tilde{k}(z) = k(z^{-1})$) are dealt with already in [5], a direct treatment of the original parameter space of systems is not yet available. Therefore nonlinear restrictions must be considered, in order to develop a more general framework of parametrization. In the case of nonlinear restrictions usually only the local behaviors of the model are considered, because the global identifiability cannot be guaranteed in general. Deistler [1] derived a condition of local identifiability for an ARMAX model at a minimal system, whereas a similar condition without minimality assumption was given in [4]. A condition of local identifiability for the minimal state-space model was derived in [6].

In this paper we restrict ourselves to the ARMAX systems. First we derive a necessary and sufficient condition of local identifiability at an arbitary system, which is proved to be equivalent to the condition of [4]. Then by using this condition a suitable local parametrization is defined. The topological properties of this parametrization, which are important for estimation and numerical calculation, are investigated. Our treatment includes models with linear restrictions or without a priori restrictions. Thus the results obtained generalize the corresponding results previously obtained in the literature, e.g. in [5].

Consider the ARMAX system

$$A(z)y(t) = B(z)u(t) \tag{1.1}$$

where $y(t) \in \mathbb{R}^n$ are the outputs, $u(t) = (\varepsilon(t)', x(t)')'$ are the inputs containing an unobserved white noise component $\varepsilon(t) \in \mathbb{R}^n$ and possibly an observed input $x(t) \in \mathbb{R}^{m-n}$ $(n \le m)$; $t \in \{\ldots, -1, 0, 1, \ldots\}$ is the time index; z denotes the backward shift operator as well as a complex variable and finally $A(z)$, $B(z)$ are the matrix polynomials

$$A(z)=\sum_{j=0}^{p} A_j z^j, \qquad B(z)=\sum_{j=0}^{p} B_j z^j$$

with $A_j \in \mathbb{R}^{n\times n}$ and $B_j \in \mathbb{R}^{n\times m}$. Throughout this paper n and m are fixed. However the distinction between the observed and unobserved inputs is not important in this paper and hence they are treated as a whole.

Given the statistical structure of the white noise input $\varepsilon(t)$, any system (1.1) is then described by the corresponding polynomial matrices $(A(z), B(z))$. In this paper we consider only the systems satisfying

$$\det A(0) \neq 0 \tag{1.2}$$

and

$$B(0) = (A(0), 0). \tag{1.3}$$

Then the corresponding transfer function of $(A(z), B(z))$,

$$K(z) = A^{-1}(z)B(z),$$

is causal in the sense that $K(z)$ has a power series expansion in some neighborhood of zero and has the form

$$K(z) = (I_n, 0) + \sum_{j=1}^{\infty} K_j z^j. \tag{1.4}$$

Let $\mathbb{M}$ be the set of all $(A(z), B(z))$ (with fixed n, m but arbitrary p) satisfying (1.2) and (1.3), let $\mathbb{U}$ be the set of all $n\times m$ rational matrices which are causal and have the form (1.4) and define the mapping $\pi : \mathbb{M} \to \mathbb{U}$ by

$$\pi(A(z,), B(z)) = A^{-1}(z)B(z).$$

Then clearly $\pi(\mathbb{M}) = \mathbb{U}$.

The identifiability considered in this paper means the (unique) determination of $(A(z), B(z))$ from its transfer function $K(z) = \pi(A(z), B(z))$. Thus we assume generally that the transfer function $K(z)$ of the system (1.1) and the second moment of $\varepsilon(t)$ are uniquely determined by the processes $\{y_t\}$ and $\{x_t\}$.

If $K(z) \in \mathbb{U}$, then the set $\pi^{-1}(K(z)) \subset \mathbb{M}$ is called the observational equivalence class of $K(z)$ in $\mathbb{M}$. A subset $\mathbb{M}_0 \subset \mathbb{M}$ is said to be (globally) identifiable, if π restricted to $\mathbb{M}_0$ is injective. The model $\mathbb{M}$ is said to be locally identfiable at $(A(z), B(z))$, if $(A(z), B(z)) \in \mathbb{M}$ and $(A(z), B(z))$ has a neighborhood in $\mathbb{M}$ (with respect to the Euclidean topology of the parameter space), which is identifiable. If $\mathbb{M}_0 \subset \mathbb{M}$ is identifiable, then there exists a bijective mapping $\psi : \pi(\mathbb{M}_0) \to \mathbb{M}_0$. ψ is called a parametrization of $\pi(\mathbb{M}_0)$.

2. Local identifiability

The order of each system (1.1) is defined as the maximum degree of the polynomials in $(A(z), B(z))$ and is denoted by $\delta(A(z), B(z))$. In the following we restrict ourselves to the systems of order less than or equal to p, for an arbitrarily given p. Let $\mathbb{M}_p$ be the set of all $(A(z), B(z)) \in \mathbb{M}$ satisfying $\delta(A(z), B(z)) \leq p$. Then any $(A(z), B(z)) \in \mathbb{M}_p$ is uniquely described by the vector of its coefficients,

$$a = \operatorname{vec}(A_0, \ldots, A_p, B_1, \ldots, B_p) \in \mathbb{R}^N$$

where $N = n(n+m)p + n^2$ and $\operatorname{vec}(A_0, \ldots, B_p)$ denotes the column vector consisting of the stacked rows of $(A_0, \ldots, B_p)$. We will identify each $(A(z), B(z)) \in \mathbb{M}_p$ with the corresponding $a = \operatorname{vec}(A_0, \ldots, A_p, B_1, \ldots, B_p)$. Thus we will write

$$\mathbb{M}_p = \{ a \in \mathbb{R}^N \mid (1.2) \}.$$

Now consider any given $K(z) \in \pi(\mathbb{M}_p)$. The observational equivalence class of $K(z)$ in $\mathbb{M}_p$ is the set of all solutions $(A(z), B(z))$ of

$$A(z)K(z) = B(z) \tag{2.1}$$

which satisfy (1.2) and $\delta(A(z), B(z)) \leq p$. In terms of coefficient matrices (2.1) is

$$\begin{aligned}
&A_0 K_0 = B_0,\\
&A_0 K_1 + A_1 K_0 = B_1,\\
&\vdots\\
&A_0 K_p + \cdots + A_p K_0 = B_p,\\
&A_0 K_{p+j} + \cdots + A_p K_j = 0, \quad j = 1, 2, \ldots.
\end{aligned} \tag{2.2}$$

It is shown in [2] that (2.2) has the same solution set with the system of the first $1 + p + np$ matrix equations of (2.2). Thus the observational equivalence class of $K(z)$ in $\mathbb{M}_p$ is just the set of all solutions a of

$$(I_n \otimes G_K)a = 0 \tag{2.3}$$

which satisfy (1.2), where $a = \mathrm{vec}(A_0, \ldots, B_p)$,

$$G'_K = \begin{pmatrix} K_1 & K_2 & \cdots & K_p & K_{p+1} & \cdots & K_{p+np} \\ K_0 & K_1 & \cdots & K_{p-1} & K_p & \cdots & K_{p+np-1} \\ 0 & K_0 & \cdots & K_{p-2} & K_{p-1} & \cdots & K_{p+np-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & K_0 & K_1 & \cdots & K_{np} \\ & & -I_{mp} & & & 0 & \end{pmatrix}$$

$\in \mathbb{R}^{[n(1+p)+mp]\times[(1+n)mp]}$

and $I_n \otimes G_K$ denotes the Kronecker product

$$I_n \otimes G_K = \begin{pmatrix} G_K & & & \\ & G_K & & \\ & & \ddots & \\ & & & G_K \end{pmatrix}.$$

Clearly $\pi^{-1}(K(z)) \cap \mathbf{M}_p$ is an open and dense subset of $\mathbb{R}^{N-nr}$ with $r = \mathrm{rank}(G_K)$.

Now assume that the (essentially nonlinear) restrictions in the system parameters be given by

$$f(a) = 0 \tag{2.4}$$

where the mapping $f \in C^1(\mathbb{R}^N, \mathbb{R}^M)$ (i.e., f is continuously differentiable in $\mathbb{R}^N$) and the kernel of f, $\mathrm{Ker}(f)$, is a C^1-manifold of dimension $N - M$. It is also assumed that the restrictions (2.4) do not contradict condition (1.2) in the sense that $\mathrm{Ker}(f)$ contains at least one element satisfying (1.2). Then the elements in $\mathrm{Ker}(f)$ satisfying (1.2) form an open subset because the determinant $\det A_0$ is a nonzero polynomial function of its entries.

In the following we will write (A, B) and K for $(A(z), B(z))$ and $K(z)$ respectively, when this does not cause confusion. Let $\mathbf{M}_p^f = \mathbf{M}_p \cap \mathrm{Ker}(f)$, and suppose $K \in \pi(\mathbf{M}_p^f)$. Consider the mapping $g_K : \mathbb{R}^N \to \mathbb{R}^{M+L}$ with $L = n(n+1)mp$ defined by

$$g_K(a) = \begin{pmatrix} f(a) \\ (I_n \otimes G_K)a \end{pmatrix}.$$

Clearly $g_K \in C^1(\mathbb{R}^N, \mathbb{R}^{M+L})$ and

$$\frac{\partial g_K}{\partial a} \begin{pmatrix} \dfrac{\partial f}{\partial a} \\ I_n \otimes G_K \end{pmatrix} := F_K(a).$$

Then the observational equivalence class of K in $\mathbf{M}_p^f$, $\pi^{-1}(K) \cap \mathbf{M}_p^f$, is just the set of all $a \in \mathrm{Ker}(g_K)$ satisfying (1.2) and hence is open and dense in $\mathrm{Ker}(g_K)$. Applying the (Global) Rank Theorem [8, pp. 178] we obtain the following result.

Theorem 1. *If $(A, B) \in \mathbf{M}_p^f$, $K = \pi(A, B)$ and the matrix $F_K(a)$ has constant rank in some neighborhood of $a = \mathrm{vec}(A_0, \ldots, B_p)$, then the model $\mathbf{M}_p^f$ is locally identifiable at (A, B) if and only if*

$$\mathrm{rank}\, F_K(a) = N. \tag{2.5}$$

Remark 1. Condition (2.5) is equivalent to the condition (14) in [4], which is

$$\hat{F}(I_n \otimes \hat{A}) \text{ has full column rank} \tag{2.6}$$

where $\hat{A}$ is the $[n(1+p+np)] \times [n + (n+m)(1+$

Table 1

$$\hat{A} = \begin{pmatrix} A_0 & \cdots & A_p & 0 & \cdots & 0 & B_1 & \cdots & B_p & 0 & \cdots & 0 \\ 0 & A_0 & \cdots & A_p & \cdots & 0 & B_0 & \cdots & B_{p-1} & B_p & \cdots & 0 \\ \vdots & & \ddots & & \ddots & \vdots & \vdots & \ddots & & & \ddots & \vdots \\ \vdots & & & \ddots & & A_p & \vdots & & \ddots & & & B_p \\ \vdots & & & & \ddots & \vdots & \vdots & & & \ddots & & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & A_0 & 0 & \cdots & \cdots & \cdots & \cdots & B_0 \end{pmatrix}$$

$$\hat{F} = \begin{pmatrix} F_{11} & 0 & F_{12} & 0 & \cdots & F_{n1} & 0 & F_{n2} & 0 \\ 0 & I_{n^2p} & 0 & 0 & \cdots & & 0 & & \\ 0 & 0 & 0 & I_{nmp} & \cdots & & 0 & & \\ & \vdots & & & \ddots & & \vdots & & \\ & & & & & 0 & I_{n^2p} & 0 & 0 \\ & 0 & & & \cdots & 0 & 0 & 0 & I_{nmp} \end{pmatrix}$$

$n)p]$ matrix given in Table 1 and $\hat{F}$ is the other matrix given there, with $F_{j1} \in \mathbb{R}^{M\times[n(p+1)]}$, $F_{j2} \in \mathbb{R}^{M\times[mp]}$ the submatrices in

$$\frac{\partial f}{\partial a} = (F_{11}, F_{12}, \dots, F_{n1}, F_{n2}).$$

The equivalence between (2.5) and (2.6) may be shown analogously to Remark 2 in [5]. However, the advantage of presentation (2.5) is that the conditions of identifiability are imposed on the transfer functions rather than the system parameters and hence is more convenient in defining and handling the parametrizations, which will be seen in the next section.

Remark 2. When the restrictions in (2.4) are linear, for instance, $f(a) = Ra - r$ with $R \in \mathbb{R}^{M\times N}$ and $r \in \mathbb{R}^M$. Then as

$$F_K(a) = \begin{pmatrix} R \\ I_n \otimes G_K \end{pmatrix},$$

condition (2.5) is exactly condition (ii) of Theorem 1 in [5]. In this case the local identifiability is equivalent to global identifiability. Thus the results derived in this paper generalize the corresponding ones in [5].

3. Local parametrizations

In this section we define a local parametrization by using condition (2.5). First for any $K(z) \in \pi(\mathbb{M}_p)$, it is easily seen from (2.2) that, under (1.2), the coefficient matrices

$$(K_{p+np+1}, K_{p+np+2}, \dots)$$

are uniquely determined by

$$(K_1, K_2, \dots, K_{p+np}) \text{ and } (A_0, \dots, A_p).$$

Thus $K(z)$ may be identified, as we do, with

$$k = \mathrm{vec}(K_1, K_2, \dots, K_{p+np}) \in \mathbb{R}^L,$$

where $L = n(n+1)mp$. Consider the mapping defined by

$$g: (k, a) \to \begin{pmatrix} f(a) \\ (I_n \otimes G_K)a \end{pmatrix}.$$

Cleary $g \in C^1(\mathbb{R}^L \times \mathbb{R}^N, \mathbb{R}^{L+M})$ and

$$\frac{\partial g}{\partial a} = F_K(a) := F(k, a).$$

Now we consider an arbitrarily fixed $a_0 \in \mathbb{M}_p^f$, such that the corresponding $k_0 = \pi(a_0)$ satisfies (2.5). Clearly there exists an (open) neighborhood of (k_0, a_0) in $\mathbb{R}^L \times \mathbb{R}^N$, over which the matrix $F(k, a)$ has constant rank N. By the (generalized) Implicit Function Theorem [8, pp. 199, Prob. 4.4d], there exists a neighborhood $U_0 \subset \mathbb{R}^L$ of k_0 and a neighborhood $\Theta_0 \subset \mathrm{Ker}(f)$ of a_0, such that for every $k \in U_0$ the equation

$$g(k, a) = 0$$

has unique solution $a := \psi(k) \in \Theta_0$ and the mapping $\psi \in C^1(U_0, \Theta_0)$. We assume that U_0 and Θ_0 are the largest neighborhoods with these properties. Note that not every $k \in U_0$ corresponds to a causal transfer function. Let U_f be the set of all $k \in U_0$ such that $\psi(k)$ satisfy (1.2) and let $\Theta_f = \psi(U_f)$. Then $\Theta_f \subset \mathbb{M}_p^f$ and ψ is a local parametrization of U_f. We will call U_f and Θ_f the local neighborhood of k_0 and a_0 respectively. The topological and geometrical properties of the local neighborhoods U_f, Θ_f and of the local parametrization $\psi: U_f \to \Theta_f$ are demonstrated in the following theorem. We will denote by $\bar{A}$ the closure of a subset A of a topological space.

Theorem 2. (1) *$\psi: U_f \to \Theta_f$ is a diffeomorphism.*

(2) *U_f is open in $\bar{U}_f$.*

(3) *Θ_f is open in* Ker(f), *which is a C^1-manifold of dimension $N - M$.*

(4) $\pi(\bar{\Theta}_f \cap \mathbb{M}_p^f) \subset \bar{U}_f$.

Proof. (1) By the Implicit Function Theorem and the definition of Θ_f, $\psi \in C^1(U_f, \Theta_f)$ and is bijective. As $\psi^{-1} = \pi$ when restricted to Θ_f, it is clearly a C^1-mapping on Θ_f.

(2) If $k \in U_f$, then k has an (open) neighborhood $O \subset U_0$. As ψ is continuous, O may be chosen such that all elements in $\psi(O)$ satisfy (1.2). Thus $O \subset U_f$. As U_0 is open in $\mathbb{R}^L$, so is U_f. It follows that U_f is open in $\bar{U}_f$.

(3) It has been shown in (2) that U_f is open in $\mathbb{R}^L$. Since $\pi: \mathbb{M}_p^f \to \mathbb{R}^L$ is continuous, $\pi^{-1}(U_f)$ is open in $\mathbb{M}_p^f$. As both $\mathbb{M}_p^f$ and Θ_0 are open in Ker(f), so is $\Theta_f = \pi^{-1}(U_f) \cap \Theta_0$.

(4) follows from the continuity of π. □

Next we consider the boundary points of Θ_f and U_f. We will call the mapping $h = (f', f_1')' \in C^1(\mathbb{R}^N, \mathbb{R}^{M+M_1})$ an extension of f, if h satisfies all conditions which are similar to the conditions for f below (2.4). For the extension h of f, the set Θ_h and U_h are similarly defined.

Theorem 3. *For every* $k \in \pi(\mathbf{M}_p^f)$, *if* rank $F_K(a) = r$ *for all* $a \in \operatorname{Ker}(g_K)$, *then*:

(1) *The observational equivalence class* $\pi^{-1}(k) \cap \mathbf{M}_p^f$ *is a* C^1*-manifold of dimension* $N - r$.

(2) *There is a set* $\mathscr{E}$ *consisting of finite number of extensions of* f *and there exists an* $h \in \mathscr{E}$, *such that* $k \in U_h$ *and* $U_h \subset \pi(\mathbf{M}_p^f)$.

(3) *every* $k \in U_0 - U_f$ *uniquely determines the corresponding (non-causal) transfer function* $K(z)$.

Proof. (1) follows from the fact that $\pi^{-1}(k) \cap \mathbf{M}_p^f$ is an open subset of $\operatorname{Ker}(f)$, which by the Rank Theorem is a C^1-manifold of dimension $N - r$.

(2) The proof is analogous to that of Theorem 2 (3) in [5]. In fact $\mathscr{E}$ may be chosen, such that it contains only 'linear extensions' of f.

(3) By definition k determines an unique a and hence an unique $K(z) = A^{-1}(z)B(z)$. Since a does not satisfy (1.2), $K(z)$ is essentially non-causal. □

Acknowledgement

This paper was completed at Fachbereich Statistik, Universität Dortmund, Germany. I would like to thank Professor Walter Krämer and the referee for helpful comments and suggestions. Financial support of Deutsche Forschungsgemeinschaft (DFG) is gratefully acknowledged.

References

[1] M. Deistler, The structural identifiability of linear models with autocorrelated errors in the case of cross-equation restrictions, *J. Econometrics* **8** (1978) 23–31.

[2] M. Deistler, The properties of the parametrization of ARMAX systems and their relevance for structural estimation and dynamic specification, *Econometrica* **51** (1983) 1187–1207.

[3] M. Deistler and E.J. Hannan, Some properties of the parametrization of ARMA systems with unknown order, *J. Multivariate Anal.* **11** (1981) 474–484.

[4] M. Deistler and J. Schrader, Linear models with autocorrelated errors: Structural identifiability in the absence of minimality assumptions, *Econometrica* **47** (1979) 495–504.

[5] M. Deistler and Liqun Wang, The common structure of parametrizations for linear systems, *Linear Algebra Appl.* **122/123/124** (1989) 921–941.

[6] K. Glover and J.C. Willems, Parametrizations of linear dynamical systems: Canonical forms and identifiability, *IEEE Trans. Automat. Control* **19** (1974) 640–645.

[7] E.J. Hannan and M. Deistler, *The Statistical Theory of Linear Systems* (Wiley, New York, 1988).

[8] E. Zeidler, *Nonlinear Functional Analysis and its Applications I* (Springer-Verlag, New York, 1986).