PARAMETRIZATION ISSUES IN SYSTEM IDENTIFICATION

M. Gevers and V. Wertz (Chargé de recherches FNRS)
Laboratoire d'Automatique, Dynamique et Analyse des Systèmes
Université Catholique de Louvain, Bâtiment Maxwell
Place du Levant, 3, B-1348 LOUVAIN-LA-NEUVE (BELGIUM)

Abstract. We present a brief introduction to the problems of parametrization
and identifiability. We introduce a distinction, which we believe useful,
between identifiability of specific parameters and identifiability of a model
structure. For the latter we present the structure, and some possible
parametrizations, for the set of all linear multivariable systems of given
McMillan degree.

Keywords. System identification; multivariable systems; identifiability.

1. INTRODUCTION

At the earliest stage of any identification
experiment the would be identifier has to
consider the question of identifiability. This
question contains two aspects:

- an input signal aspect: is the input signal
  rich enough (or informative enough) to
distinguish between different models? For
  example, if the input signal is constant, it is
  impossible to distinguish between the models

\[ Y(s) = K U(s) \] \hspace{1cm} \[ Y(s) = \frac{K}{1 + sT} U(s) \]

from input-output measurements.

- a parametrisation aspect: loosely speaking,
  the problem is whether the model set is
  parametrized in such a way that every
  input-output map is represented by a unique value of
  the parameter vector.

In this paper we shall not discuss the problem of
sufficiently rich or persistently exciting input
signals. We shall concentrate on the
parametrisation aspects of identifiability. We
shall distinguish between two concepts:
parameter identifiability and identifiability of
the model structure (or structural
identifiability). The two concepts arise from two
different problems.

Parameter identifiability deals with the question
of whether a set of specific and physically
meaningful parameters that appear in the
description of a physical system can be
identified from input-output experiments. The
outcome of the investigation might be that some
parameters can be identified but others not.

The question of identifiability of a model
structure (or structural identifiability) arises
when a parametrized black box model set is chosen
to identify or approximate a system. The
parametrisation must then be chosen so that every
input-output map (or transfer function)
corresponds to a unique value of the parameter
vector. This notion is a property of the
parametrisation without any reference to a true
system.

The main thrust of this paper will be concerned
with the identifiability of model sets, and in
particular with the choice of identifiable
parametrisations for multivariable linear
systems. This is because parameter
identifiability can only be checked on a case by
case basis, while a comprehensive theory is now
available for the parametrisation of multivariable black box models.

To introduce these two concepts of
identifiability in a rigorous way, we first
present some definitions on models, model sets
and model structures in section 2. In section 3
we shall very briefly deal with the concept of
identifiability, the motivation being mainly to
distinguish parameter identifiability from the
identifiability of a model structure. In section
4 we describe the structure of all linear
multivariable systems of given McMillan degree.
This then leads directly to a presentation of
identifiable parametrisations for such systems,
which is the subject of section 5. A much more
detailed presentation of these concepts, together
with a brief survey of results on structural
estimation, can be found in Gevers and Wertz

2. MODELS, MODEL SETS AND MODEL STRUCTURES

Because the problems of parametrisation are
essentially algebraic, we shall for simplicity
assume that we deal with deterministic input-
output (I/O) models. We shall also limit
ourselves to linear discrete time models. Our
development is inspired by but not identical to

Definition 2.1: A model \( M \) is a stable algebraic
operator that transforms a given input process
\( u(t) \) into a unique output process \( y(t) \).

Therefore the following are all models by our
definition, provided specific values (rational
functions or real-valued parameters, or
polynomials) are put into the operators \( G(z) \), \( A \),
\( B \), \( C \), etc.

Transfer function model (TF):

\[ y(t) = G(z) u(t) \] \hspace{1cm} (2.1)
where $y(t) \in \mathbb{R}^p$, $u(t) \in \mathbb{R}^n$ and $G(z)$ is a strictly causal rational stable transfer function matrix,

$$ G(z) = \frac{z^{-1}}{1 - z^{-1}}$$

State space model (SS):

$$ \begin{cases} x(t+1) = Ax(t) + Bu(t) \\ y(t) = Cx(t) \end{cases} \quad (2.2) $$

where $x(t) \in \mathbb{R}^m$, $u(t)$ and $y(t)$ are as before, $A$, $B$, $C$ are real matrices of appropriate dimensions and $dim x(t)$ is minimal.

Matrix fraction description (MFD):

$$ P(z)y(t) = Q(z)u(t) \quad (2.3) $$

where $u(t)$ and $y(t)$ are as before, and $P(z)$, $Q(z)$ are left coprime polynomial matrices in $z$, with $z^{-1}u(t) = (u(t-1))$.

One can think of other model descriptions such as an impulse response representation or an ARMA model. Whatever the description chosen for our model (e.g., SS, MFD, ARMA), one can always compute the TF $G(z)$ from it, e.g.,

$$ G(z) = C(zI - A)^{-1}B = P^{-1}(zIQ(z)) $$

One can therefore introduce the following definition.

**Definition 2.2:** Two models $N^1$ and $N^2$ are equivalent (we write $N^1 \equiv N^2$) if and only if their transfer functions are identical: $G(z) = G(z)$ for all $z$.

In identification a search will typically be conducted over a model set, which is obtained as the range of a smoothly parameterized algebraic operator where the parameter vector $\theta$, of dimension $d_\theta$, is allowed to cover a subset $D_\theta$ of $\mathbb{R}^n$. The operator is called a model structure.

**Definition 2.3:** A model set $\mathbb{N}$ is a set of models:

$$ \mathbb{N} = \{N(\theta) \mid \theta \in D_\theta \} $$

**Definition 2.4:** A model structure $\tilde{N}$ is a differentiable mapping from a subset $D_\tilde{N}$ of $\mathbb{R}^n$ to a model set:

$$ \tilde{N} : \theta \in D_\tilde{N} \rightarrow N(\theta) \in \mathbb{N} $$

**Example:**

1) $\mathbb{N}^1 : y(t) = 0.85 y(t-1) + 1.5 u(t-1)$

2) $\mathbb{N}^2 : y(t) = ay(t-1) + bu(t-1) \quad (a, b) \in D_\mathbb{N} \subset \mathbb{R}^2$

with $D_\mathbb{N} = \{(a, b) \mid |a|, |b| \leq 100\}$

3) $\tilde{N} : (a, b) \in D_\tilde{N} \rightarrow \tilde{N} = \{y(t) = ay(t-1) + bu(t-1) \mid (a, b) \in D_\tilde{N}\}$

**3. IDENTIFIABILITY**

Roughly speaking, a model structure is called identifiable if the mapping of Definition 2.4 is invertible. We now give a precise definition of structural identifiability.

**Definition 3.1:** A model structure $\tilde{N}$ is globally identifiable at $\theta$ if

$$ \tilde{N}(\theta) = \tilde{N}(\theta') \Rightarrow \theta = \theta' \quad (\theta, \theta' \in D_\theta) $$

It is strictly globally identifiable if it is globally identifiable for all $\theta \in D_\theta$.

**Comment 3.1:** We stress that this definition covers only one aspect of identifiability: it is a property of the model structure (i.e., the parametrization) only and is totally independent of a possible 'true system'. In addition it does not say anything about the richness of the data set.

To stress the relevance of these definitions to our problem of constructing identifiable model structures, we point out the following facts:

1) The set of all SISO models of order up to $n$ cannot be described as the range of a strictly globally identifiable model structure because of pole-zero cancellations. It can, however, be described as the range of a globally identifiable model structure.

2) The set of all multivariable systems of order up to $n$ (or even exactly $n$) cannot be described as the range of a unique globally identifiable model structure, but it can be described as a union of ranges of model structures.

$$ \mathbb{N} = \bigcup_{i=1}^{M} \mathbb{N}^i $$

The problem of structural identifiability arises when, in a blackbox identification context, it is desired to select a model structure whose range covers a desired model set, e.g., the set of all linear multivariable systems of order $n$ with $p$ outputs and $m$ inputs. If a structure is globally identifiable and if the data are informative enough, the parameter estimation problem will be well posed. Notice that this does not guarantee convergence of the estimated model to the true system since that would require, in addition, that the true system be contained in the model set.

The problem of parameter identifiability arises in a different context, namely when a model of a true system has been obtained from physical principles. The question of parameter identifiability is then whether some specific physically meaningful parameters within that model can be identified from input-output experiments. Parameter identifiability is also dependent on the invertibility of the model structure, but with two essential differences:

1) if a set of parameters is called identifiable, it means that these parameters converge to their true values, while structural identifiability is independent of any notion of true values.

2) some parameters of interest may be identifiable whilst the structure itself may not be globally identifiable. The following example, from Godfrey and Distefero (1985), illustrates this.

**Example 1:** Suppose that a physical system gives rise to the following model:

$$ \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ y \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -a_12 & -a_13 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u $$

$$ y = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} $$

283
From I/O experiments, the following parameters can be identified: \( a_{ij}, b \) and the product \( a_{ij}b \). The structure parameters \( a_{ij}, a_{jk} \), and \( b \) are identifiable, but the structure is not identifiable, unless \( a_{ij}b \) or \( a_{ij} \) is known from the physics of the problem.

The problem of checking the parameter identifiability of a physical model is essentially a case by case study. A number of methods have been developed. See Walter (1982) and Godfrey and Distefano (1985) for recent surveys.

4. THE STRUCTURE OF MULTIVARIABLE LINEAR SYSTEMS

Definition 4.1: We call \( S(n) \) the set of all strictly proper stable rational transfer function matrices \( G(z) \) of order \( n \) (i.e. McMillan degree) \( n \), with \( m \) inputs and \( p \) outputs.

From now on we look at the problem of parametrizing \( S(n) \). The reason for looking at this problem is twofold:
- The set \( S(n) \) cannot be described as the range of a single model structure; hence the parametrization problem is not trivial;
- The structure of \( S(n) \) is by now well understood.

We consider the system (2.1) with a strictly proper transfer function matrix (TPM) \( G(z) \) and we introduce the following Hankel matrix:

\[
\begin{bmatrix}
G_1 & G_2 & \ldots & G_N \\
G_2 & G_3 & \ldots & G_{N+1} \\
\vdots & \vdots & \ddots & \vdots \\
G_N & G_{N+1} & \ldots & G_{2N-1}
\end{bmatrix}
\] (4.1)

The Hankel matrix \( H_n[G] \) is made up of blocks of \( p \) rows. The rank of \( H_n[G] \) is called order or McMillan degree of \( G(z) \). We assume that the rows of the first block of \( H_n[G] \) are linearly independent. (This only eliminates degenerate transfer functions whose rows would be linearly dependent). We call \( r_{iz} \) the \( i \)-th row of the \( j \)-th block of \( H_n[G] \).

Our objective now is to find a parametrization of the elements of \( S(n) \), i.e. we would like to be able to represent an arbitrary element of \( S(n) \) by a model (SS or MF or ARMA, ... ) that depends on a finite parameter vector \( \theta \), as opposed to the infinite set of matrices \( G_1 \). To solve this problem we need to state a few results about the structure of \( S(n) \).

Result 4.1: If \( G(z) \in S(n) \), then there exists at least one partition \( n = n_1 + n_2 + \ldots + n_p \) of \( n \) such that the set of rows \( \{ r_{i1}; i = 1, \ldots, p \} \) *\( (4.2) \) form a basis of \( H_n[G] \).

This follows immediately from the Hankel structure. Note that the rows in (4.2) are chosen in such a way that if \( i \neq j \), then \( r_{ij} \in R(n_1, \ldots, n_p) \). The set of the rows selected in \( R(n_1, \ldots, n_p) \) is completely determined by the partition \( \mu = [n_1, \ldots, n_p] \).

The indices \( n_1 \) are called structure indices. We also denote \( n \) as \( [n] = [n_1, \ldots, n_p] = n \). It is easy to see that there are \( \binom{n+p-1}{p} \) such partitions.

Definition 4.2: We call \( U \) the set of all elements in \( S(n) \) for which the corresponding set of rows \( (4.2) \) forms a basis for the rows of \( H_n[G] \).

We now show that \( U \) can be completely coordinatized by \( n(p+m) \) coordinates. Row \( r_{i1} \), \( j = 1, \ldots, p \), is a unique combination of the basis rows.

\[
r_{i1} = \sum_{j=1}^{p} \sum_{k=1}^{n} a_{ik} r_{kj}
\] (4.3)

It follows again from the Hankel structure that knowing the first block of elements (i.e. the first \( s \) elements) of each basis row \( (4.2) \) and the coefficients \( a_{ij} \), \( i, j = 1, \ldots, p \), \( i = 1, \ldots, p \), \( j = 1, \ldots, n \), allows one to compute any other row of \( H_n[G] \) and therefore to specify \( G(z) \) completely. Therefore any element \( G(z) \in U \) can be mapped into a vector \( \phi \in \mathbb{R}^d \) by the following \( d = n(p+m) \) coordinates:

\[
\phi = \left[ \phi_1(1, j); i = 1, \ldots, p; j = 1, \ldots, n \right]
\] (4.4)

where \( \phi_k(1, j) \) is the \( (i, j) \)-th element of the matrix \( G_k \).

With this mapping, which we denote by \( \phi \), every \( G(z) \in U \) can be mapped into a point in a subspace \( \phi \mu \) of \( \phi \mu \):

\[
\phi : (z) \in U \rightarrow \phi = \phi(G(z)) \in \phi \mu \subset \mathbb{R}^d
\] (4.5)

We now have the following important result.

Result 4.2: see Clark (1976)

1) \( \phi(\mathbb{R}) \) is a real analytic manifold of dimension \( n(p+m) \).

2) \( S(n) \) is the union of all \( U \) for which \( |\phi| = n \). Each \( U \) is open and dense in \( S(n) \); \( |\phi| \) is open and dense in \( \mathbb{R}^d \).

Comment 4.1: Since \( U \), with \( |\phi| = n \) is open and dense in \( S(n) \), it follows that almost all points of \( S(n) \) in \( U \) for any \( \mu \). The choice of \( \mu \) specifies a local coordinate system and, as we shall see in Section 5, a parametrization (in SS or MF form, say) called pseudoeuclidean form. The message therefore is that a given system of order \( \epsilon \) can be represented almost surely by one of the \( (p+m) \) pseudoeuclidean forms

\[
\phi = [\phi(1, j); i = 1, \ldots, p; j = 1, \ldots, n]
\] (4.6)

of \( n \), i.e. the \( U \) overlap. However, no unique set \( U \) can cover all \( S(n) \). It means that no unique identifiable parametrization can represent all systems in \( S(n) \).

Having described the structure of \( S(n) \), one can now think of the identification problem in the following terms. Estimate the order \( n \) and then take any partition \( \mu \) of \( n \) such that \( |\mu| = n \) and compute the maximum likelihood estimate \( \phi \) of the corresponding vector \( \phi \) that completely specifies the system. Almost surely the system can be represented in that coordinate system. The computation of the parameter vector by, say, a
maximum likelihood method necessitates an I/O description of the point $r$. In section 5 we shall present pseudo-canonical SS or MDF models which can be written directly as functions of $r$.

An alternative is to cover $S(n)$ by disjoint subsets $V$ which can be coordinatized by non-overlapping coordinate systems. These will give rise to the somewhat simpler canonical SS or MDF forms.

Definition 4.3: We call $V$ the subset of $U$ for which the rows (4.2) specified by $\mu$ are the first $n$ linearly independent rows of $x$. Since the row $r_{i, n+1}$ is now a linear combination of the basis rows above it, (4.3) is replaced by

$$r_{i, n+1} = \sum_{j=1}^{p} \sum_{k=1}^{n} a_{ijk} r_{jk}^i \quad i = 1, \ldots, p$$

(4.6)

where

$$a_{ijk} = \min(n_1, n_j) \quad i < j$$

(4.7)

$$a_{ijk} = \min(n_{i+1}, n_j) \quad i > j$$

It follows, by the same argument as before, that any element $G(z) \in G(V)$ can be mapped into a vector $\rho_\mu \in R^p$ defined by the following coordinates:

$$\rho_\mu = \left[ a_{11}, \ldots, a_{1n_1}, \ldots, a_{i1}, \ldots, a_{in_i}, \ldots, a_{p1}, \ldots, a_{pnp} \right]$$

(4.8)

where

$$d(\mu) = n(n+1) + \sum_{i=1}^{n} \left( \min(n_i, n_{i+1}) + \min(n_i, n_{i+1}) \right)$$

(4.9)

We call this mapping $\psi_\mu$. It maps every $G(z) \in G(V)$ into a point in a subspace $X_\mu \subset R^p$:

$$\psi_\mu : G(z) \in G(V) \rightarrow \rho_\mu = \psi_\mu (G(z)) \in X_\mu \subset R^p$$

(4.10)

Result 4.3: See Harwin and Kalman (1976)

1) The $V$ are disjoint, $V \subset U$ and

$$U \cup V = S(n) \quad d(\mu)$$

2) $X_\mu$ is an open and dense subset of $R^p$

Comment 4.2: For $p > 1$, $S(n)$ is partitioned into the $n+p-1$ disjoint sets $V_1, \ldots, V_{n+p-1}$, which are of different dimensions $d(\mu)$. Hence every system $\sigma \in S(n)$ belongs to one and only one of the $V$ and has therefore a set of structure indices $\mu = (n_1, \ldots, n_p)$ attached to it. Those structure indices are usually called observability indices or left Kronecker indices; they determine the local coordinate system in which that system is described by the vector $\rho$. They in turn define canonical SS or MDF forms as we shall see in the next section. Note that in all cases $d(\mu) = d = n(p+1)$. The canonical forms therefore will generally have fewer parameters than the corresponding pseudo-canonical ones. This is one advantage of canonical forms.

Recall now that a model structure was defined (see Definition 2.4) as a mapping from a parameter vector $\eta$ to a particular model $M(\eta)$, and that a structure was called globally identifiable if that mapping was injective for almost all $\theta$ in a subset $D_\mu$. Now we have shown that

(i) given a system $G(z) \in S(n)$, there exists a uniquely defined set of structure indices $\mu = (n_1, \ldots, n_p)$ and a uniquely defined mapping $\psi_\mu$ from $G(z)$ to a parameter vector $\rho_\mu$. This in turn yields a unique canonical form as we shall see next.

(ii) given a system $G(z) \in S(n)$, then for almost any arbitrary set of structure indices $\mu = (n_1, \ldots, n_p)$ such that $d(\mu) = d = n(p+1)$ there exists a uniquely defined mapping $\Phi_\mu$ from $G(z)$ to a parameter vector $\eta_\mu$. This parameter vector in turn defines a pseudo-canonical form.

In the next section we define model structures (in SS or MDF form) that are entirely specified by the integer-valued structure indices and the real-valued parameter vectors ($\rho$ or $\eta$). Hence these model structures will be identifiable since the sought after inverse mappings are precisely $\psi_\mu$ and $\Phi_\mu$.

5. CANONICAL AND PSEUDO-CANONICAL FORMS

We now describe canonical and pseudo-canonical SS and MDF forms for a strictly proper $p \times m$ TFM $G(z) \in S(n)$.

5.1. Canonical forms

It follows from section 4 that if $G(z) \in S(n)$, then $G(z) \in V_\mu$ for a unique $\mu$ and that it is represented by $\mu$ coordinate vector $\rho_\mu$. See (4.8). From $\rho_\mu$ the following canonical SS model $A, B, C$ can be obtained (see (2.2)):

$$A = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \\ 0 & \cdots & 0 & \cdots & 0 \\ a_{11} & \cdots & a_{1n_1} & 0 & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ a_{p1} & \cdots & a_{pnp} & 0 & 0 \end{bmatrix}$$

(5.1b)
\[ C = \begin{pmatrix} \mathbf{1} & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \mathbf{1} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{1} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \mathbf{1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & \mathbf{1} \end{pmatrix} \] (5.1a)

\[ B = \begin{pmatrix} g_1(1,1) & \cdots & g_1(1,m) \\ g_2(1,1) & \cdots & g_2(1,m) \\ \vdots & \ddots & \vdots \\ g_p(1,1) & \cdots & g_p(1,m) \\ \vdots & \ddots & \vdots \\ g_1(p,1) & \cdots & g_1(p,m) \\ \vdots & \ddots & \vdots \\ g_p(p,1) & \cdots & g_p(p,m) \end{pmatrix} \] (5.1c)

\[ K_{ij} = \begin{pmatrix} -\alpha_{i1} & -\alpha_{i2} & \cdots & -\alpha_{im} & 0 & 0 \\ -\alpha_{i1} & \mathbf{1} & \cdots & 0 & 0 & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \\ -\alpha_{i1} & 0 & \cdots & \mathbf{1} & 0 & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \\ 0 & 0 & \cdots & 0 & \mathbf{1} & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \\ 0 & 0 & \cdots & 0 & 0 & \cdots \end{pmatrix} \] (5.1b)

Then:
\[ G = NB \] (5.1d)

Note that the parameters appearing in \( A, B, C \) are precisely the elements of \( \mu \); \( A \) and \( C \) have a specific structure, specified by the structure indices \( n_k \), while \( B \) is fully parametrized. Any state space representation of a \( n \)th order system with Kronecker indices \( (n_k) \), \( n_p \) can be transformed to this canonical form by a similarity transformation: see e.g. Guidorzi (1981).

A canonical MFID form \( P(z), Q(z) \) such that \( K(z) = P(z)Q(z) \) is obtained as follows from \( \mu \).

Let \( P(z) = \left[ p_{ij}(z) \right] \) and \( Q(z) = \left[ q_{ij}(z) \right] \). Then
\[ p_{ij}(z) = z^{n_i} - \alpha_{i1}z^{n_i-1} - \cdots - \alpha_{ii} \] (5.2a)
\[ q_{ij}(z) = \begin{cases} -\alpha_{ij}z^{n_j-1} - \alpha_{ij} & \text{for } i \neq j \\ \beta_{ij} & \text{for } i = j \end{cases} \] (5.2b)
with the coefficients \( \beta_{ij} \) are bilinear functions of the coefficients \( \alpha_{i,j} \) and \( g_k(1,j) \) obtained as follows. Let
\[ G = \begin{pmatrix} G_0 & G_1 & \cdots & G_p \end{pmatrix} \]
\[ N = \begin{pmatrix} N_{1,n-p} \end{pmatrix} \]
\[ M = \begin{pmatrix} M_{1,n} \end{pmatrix} \]
(5.3)
(5.4)
(5.5)

where \( M_{1,n} \) with \( \text{dim}A_{11} = n \times n \) (5.7)
where the \( A_{11} \) are as in (5.1b) and where
\[ A_{11} = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \\ q_{11} & \cdots & q_{1mn} \end{pmatrix} \] (5.8)

See e.g. Gavers and Wertz (1984) for a derivation of this form. The free parameters are exactly the coordinates of \( \mu \), which uniquely describe the system in the coordinate system defined by \( \mu \). Hence this form is identifiable.

A pseudo-canonical MFID form for \( P(z), Q(z) \) is obtained from the coordinates of \( \mu \) (compare with (5.2)-(5.6)).

\[ p_{ij}(z) = z^{n_i} - \alpha_{i1}z^{n_i-1} - \cdots - \alpha_{ii} \] (5.9a)
\[ q_{ij}(z) = -\alpha_{ij}z^{n_j-1} - \cdots - \alpha_{ij} \text{ for } i \neq j \] (5.9b)

See Guidorzi (1981) for a proof and for further properties of this form. In an identification context, once the Kronecker indices have been estimated, the structure of \( P(z) \) and \( Q(z) \) is completely specified by (5.2): the \( \alpha_{ij} \) and \( \beta_{ij} \) are free parameters, whose number is again \( \text{dim}(\mu) \).

The main disadvantage of canonical forms is that one has to estimate the Kronecker indices. Hannan and Kavalieris (1984) have proposed a consistent procedure for the estimation of the \( n_k \). An alternative is to use pseudo-canonical forms: this requires the estimation of only one integer-valued parameter, the order \( n \).

5.2. Pseudo-canonical forms

If \( G(z) \in \mathbb{S}(n) \), then \( G(z) \in \mathbb{U} \), for almost any \( \mu \) such that \( |\mu| = n \) and for each such \( \mu \) it is represented uniquely by a coordinate vector \( \tau \); see (4.4). A pseudo-canonical SS form for \( G(z) \) is obtained from \( \tau \) by taking \( C \) as in (5.1a), \( B \) as in (5.1c) and

\[ A = \begin{pmatrix} A_{11} \end{pmatrix} \text{ with dim}A_{11} = n_1 \times n_1 \]
(5.7)

where the \( A_{11} \) are as in (5.1b) and where

\[ A_{11} = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \\ q_{11} & \cdots & q_{1mn} \end{pmatrix} \] (5.8)

See e.g. Gavers and Wertz (1984) for a derivation of this form. The free parameters are exactly the coordinates of \( \mu \), which uniquely describe the system in the coordinate system defined by \( \mu \). Hence this form is identifiable.

A pseudo-canonical MFID form for \( P(z), Q(z) \) is obtained from the coordinates of \( \mu \) (compare with (5.2)-(5.6)).
identifiable parametrized black box model must be selected to represent or approximate a system (structural identifiability).

Most efforts so far have been spent on solving the latter problem, which is by no means trivial because the structure of $S(n)$, the set of systems of order $n$, is fairly complicated: the set of systems in $S(n)$ cannot be described by a single parametrization. In this brief tutorial introduction we have tried to explain the structure of $S(n)$ in simple terms. We have illustrated the canonical and pseudo-canonical forms, and their connections with the structure of $S(n)$, by two models: SS and MFD models. Other parametrizations are possible: in particular, much recent work has been done on the degree properties (Gevens, 1986) and the parametrization (Deistler and Gevers, 1987) of monic ARMA models. Many other problems have not been touched on in this brief tutorial paper, the most important being structure and order estimation procedures. We refer to Gevers and Wertz (1987) for a comprehensive survey and to Hannan and Kavalieris (1984) and Deistler (1986) for technical details.

**REFERENCES**

Clark J.M.C. (1976), "The consistent selection of parametrizations in system identification", JACC, Purdue University.


