Proceedings of the 28th Conference
on Decision and Control
Tampa, Florida • December 1989

A DETERMINISTIC APPROACH TO APPROXIMATE MODELLING OF INPUT–OUTPUT DATA

Paul Van den Hof

Lab. Measurement and Control, Dept. Mechanical Engineering and Marine Technology
Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands.

Abstract

The problem of system identification is reconsidered as a problem of deterministic approximate modelling on the basis of input–output data. In the approach presented, system identification methods are required to yield models that are well-defined, in the sense that the models obtained proceed from the available data sequence and from specified users' choices, and not from implicit (statistical) assumptions on the data and the underlying process.
Based on the system theoretic concept of dynamical system behaviour, a framework is presented in which the identification problem as considered above can be formulated properly. In this framework the different components of an identification method: model set, parametrization, and identification criterion, are defined in a fundamental and natural way. A clear distinction is made between the problems of identification and parametrization. For the popular class of equation error identification methods, it is shown that the construction of parametrizations that are identifiable by a least squares identification criterion, requires specific users' choices that not have been recognized before and that influence the optimal models obtained.

Introduction

System identification can be defined as the problem of creating mathematical models of dynamical processes on the basis of measurement data of the processes concerned. Current methods of system identification have been mainly developed from a viewpoint of statistics. Assumptions on statistical properties of the available data sequences, play a crucial role in the methodology of constructing models from time series, see e.g. [1],[3],[5]. Inherent in this approach is the implicit assumption that one indeed is able to describe the process at hand exactly with a model of restricted complexity. However identification methods might be very sensitive with respect to the assumptions underlying the methods. From a methodological point of view this situation is not satisfactory. The presence of statistical assumptions in the fundamental formulation of the identification problem has brought a number of authors to express their opinion on the state of the art quite strongly, as illustrated in the following quotations: *"The subject is so underdeveloped at present that it is not possible to say very much about the identification of dynamical systems"*, (Kalman [4]), and *"Notwithstanding the fact that identification theory and time series analysis have produced some very useful algorithms and important applications, it can be stated that there is a need to put a clear and rational foundation under the problem of obtaining models from time series. It is very much of an area where some of the first principles still need to be sorted out. In particular one should start by formalizing what is meant by an optimal (approximate) model"* (Willems [10]).
In the course of years the conviction has been growing that processes to be modelled are in general far too complex to be modelled exactly by linear, time–invariant and finite dimensional models. On the other hand, the resulting models have to be relatively simple, in order to be applicable in a manageable way, e.g. in control system design. Consequently the modelling errors that we have to deal with, will not be caused mainly by random effects like measurement noise, but will rather be due to the fact that our models are not complex enough.
In this paper the opinion of Willems [11] will be supported, stating that although there may be many situations in which a statistical framework indeed is a suitable one, it has many fundamental drawbacks as a general philosophy. As an alternative approach the problem of system identification will be considered as a problem of deterministic approximation, in which the model(s) finally obtained, should be an optimal approximation of the process at hand, in a prespecified and well–defined sense. Explicit users' choices have to replace the implicit statistical assumptions discussed before.
In this paper we will direct our attention to the construction of (a) model(s) out of measurement data of input and output signals. We will not make any reference to a data generating process, since this process is principally unknown. The step from data to model is a well–defined problem that allows a formal treatment without having a priori knowledge available. In view of these starting points the main parts of an identification procedure are reflected by the following three choices:

1. the choice of a model set
2. the choice of a parametrization
3. the choice of an identification criterion

In order to discuss the construction of models in a fundamental way, there is a need of having a proper definition of the notion of model, and, even more general, a need of having a proper definition of a dynamical system. Linear time–invariant and finite dimensional dynamical systems are commonly defined in terms of transfer functions, state space representations or difference equations, all having their own specific properties. However, all these definitions have one thing in common: they impose restrictions on the external signals of the dynamical system. In a dynamical system with inputs and outputs, not all pairs of input and output signals are admissible, but only those signals that are related through the laws of the dynamical system. In the recent system theoretic work of Willems [10], [12] this notion has been formalized and a dynamical system correspondingly is defined in terms of its behaviour, i.e. the set of admissible signal trajectories. This fundamental concept has been adopted in this paper and has been used to construct a framework for the formulation of the identification problem. Within this framework the basic choices of model set ($\mathcal{M}$), parametrization ($\tilde{M}$) and identification criterion (J) will be defined, and their mutual relations and distinctions will be stated clearly. The purpose of the identification criterion is to select, given an available data sequence, that (those) model(s) out of the model set, that can be considered to be optimal for the available data sequence. Consequently the optimal models obtained are dependent on the specific model set and the identification criterion chosen. The parametrization takes care of a (unique) representation of the elements of the model set in terms of parameters; this is a problem of representation and consequently the parametrization itself should not influence the optimal models obtained.
First appropriate definitions will be given of the models that will be considered in this paper. The basic elements of an identification procedure: model set, parametrization and identification criterion will be defined subsequently. Next it is shown which kind of restrictions have to be imposed in order to arrive at a useful identification problem, and the relation with the identifiability of parametrizations is discussed. Results are further specified for a least squares identification criterion.

Some notational conventions: $\mathbb{R}^{p\times m}(z)$ is the field of [p×m] rational matrices; $\mathbb{R}^{p\times m}[z,z^{-1}]$ is the ring of [p×m] polynomial matrices in the indeterminates z and z^{-1}; $\mathbb{R}\backslash\{0\}$ is the set of

CH2642-7/89/0000-0659$1.00 © 1989 IEEE

real numbers excluding 0; $\det_{\mathbb{R}(z)}(\cdot)$, $\text{rank}_{\mathbb{R}(z)}(\cdot)$ denote the determinant and the rank of a rational matrix over the field $\mathbb{R}(z)$, and $\text{rank}_{\mathbb{C}}(\cdot)$ the rank of a matrix over $\mathbb{C}$.

Dynamical Systems and (i/o/pr)–Models

First we will define the notion of a dynamical system.

Definition 1. (Willems [10]). A *linear, time–invariant, finite dimensional dynamical system* S *on* $\mathbb{Z}$ is defined as a triple: $S = (T, W, \mathcal{B})$, with
- $T = \mathbb{Z}$, the *time set*;
- W is the *signal set*, being a vector space in which the variables that are related to the system take on their values;
- $\mathcal{B} \subset W^T$ is the *behaviour* of the system, denoted by $\mathcal{B}(S)$, being a linear closed subspace of W^T in the topology of pointwise convergence. It is the space of all signal trajectories w: $T \to W$ that are compatible with the system, and it satisfies the shift invariance property:
$$w \in \mathcal{B}(S) \Leftrightarrow \sigma w \in \mathcal{B}(S) \Leftrightarrow \sigma^{-1} w \in \mathcal{B}(S)$$
with the *shift operators* σ, σ^{-1} defined by: $(\sigma w)(t)=w(t+1)$, $t\in\mathbb{Z}$ and $(\sigma^{-1}w)(t)=w(t-1)$, $t\in\mathbb{Z}$. □

For a dynamical system having m inputs and p outputs the signal set W will generally be fixed to $W=\mathbb{R}^{p+m}$. Having specified T and W the essential characteristic of a dynamical system is reflected by its behaviour $\mathcal{B}$. This concept of a linear time-invariant finite dimensional system encompasses various definitions used in the literature such as the transfer function approach, representations with difference equations, and the well known state space approach. For the purpose of identification we will use models that are linear, time–invariant and finite dimensional dynamical systems, describing relations between three types of model variables: inputs (u), outputs (y) and residuals (e).

Definition 2. An *input–output–processing residual model* or (i/o/pr)–model $M_{p,m}$ is defined as a linear time–invariant finite dimensional dynamical system on $\mathbb{Z}$, with $W = Y\times U\times E$, $Y = \mathbb{R}^p$, $U = \mathbb{R}^m$, and $E = \mathbb{R}^p$, with the properties that
- u is free, i.e. $\mathcal{B}_u(M)=(\mathbb{R}^m)^{\mathbb{Z}}$ with
$\mathcal{B}_u(M):=\{u\in U^{\mathbb{Z}} \mid \exists (y,e)\in (Y\times E)^{\mathbb{Z}},\ (y,u,e)\in\mathcal{B}(M)\}$, e is free and y processes (u,e), where "w_1 processes w_2" means that once w_2 has been specified, the space of admissible signals w_1 such that $(w_1,w_2)\in\mathcal{B}(M)$ is finite dimensional; and
- (u,y) are free and e processes (y,u). □

The notions of "free variables" and "processing variables" have been introduced in [10]. The expression "w_2 is free" and "w_1 processes w_2" implies that the system with w_1 as input and w_2 as output can be considered as a "classical" input–output system. The signals $y\in Y^{\mathbb{Z}}$, $u\in U^{\mathbb{Z}}$, $e\in E^{\mathbb{Z}}$ will be denoted as output, input and *residual*. Signals $w\in\mathcal{B}(M)$ will be written as $w = (y,u,e)$ or as $w = (v,e)$ with $v\in V^{\mathbb{Z}}$, $V = Y\times U$, and $v=(y,u)$. The variables v will be the variables that can be measured, while residuals e in fact are artificially added to the model description to deal with modelling errors.
Restrictions of signal variables and behaviours to time sets $\mathbb{Z}_+$ or $T_N=\mathbb{Z}\cap[0,N-1]$ will be denoted by w^+, $\mathcal{B}^+$ or w^N, $\mathcal{B}^N$ respectively, with $\mathbb{Z}_+=\mathbb{Z}\cap[0,\infty)$.

An (i/o/pr)–model is a special form of a dynamical system with auxiliary variables, as defined in [10]. It can be interpreted in different ways:
1. In terms of *linear input–output systems with disturbances*, an (i/o/pr)–model can be viewed as a dynamical system with input signals u, disturbance inputs e and output signals y.
2. In terms of *dynamical models used for identification* it is much more appropriate to consider an (i/o/pr)–model as a model with the measured data consisting of u and y as inputs, and the residual e as output.

Residual signals will play an important role in the modelling procedures that will be treated. They reflect the error terms that are required for making a measured data sequence compatible with a given model. The residual signals determine the locations in the models where differences between models and measured data sequences are discounted for; consequently they can act as a basis for a measure of fit between a model and a data sequence. For an (i/o/pr)–model $M_{p,m}$ the time set and the signal set are fixed and the essential part of the model is reflected by its behaviour. This behaviour can be represented in different ways. We will mainly restrict attention to a representation in polynomial matrix form.

Proposition 3. For any (i/o/pr)–model $M_{p,m}$ there exists a polynomial matrix $T = [P|-Q|-R]$, $T\in\mathbb{R}^{p\times(p+m+p)}[z,z^{-1}]$ with $\det_{\mathbb{R}(z)}P[z,z^{-1}] \neq 0$ and $\det_{\mathbb{R}(z)}R[z,z^{-1}] \neq 0$, such that
$$\mathcal{B}(M_{p,m}) = \{w\in W^{\mathbb{Z}} \mid w=(y,u,e),\ P(\sigma,\sigma^{-1})y-Q(\sigma,\sigma^{-1})u-R(\sigma,\sigma^{-1})e=0\}. \quad \square$$

The (i/o/pr)–model $M_{p,m}$ will be said to be induced by the polynomial matrix $T(z,z^{-1})$. Note that polynomial matrices are considered in the two shift operators σ and σ^{-1}, generalizing the descriptions in either one of the two shift operators.
By definitions 1 and 2 two (i/o/pr)–models $M^{(1)}_{p,m}$ and $M^{(2)}_{p,m}$ are equal, i.e. $M^{(1)}_{p,m} = M^{(2)}_{p,m}$, if and only if their behaviours are equal. In terms of the corresponding polynomial matrices, this can be formulated as follows (see [10]):

Proposition 4. Two (i/o/pr)–models $M^{(1)}_{p,m}$ and $M^{(2)}_{p,m}$, induced by full row rank polynomial matrices $T^{(1)}(z,z^{-1})$ and $T^{(2)}(z,z^{-1})$ respectively, have the same behaviour if and only if $T^{(1)}(z,z^{-1})=U(z,z^{-1})T^{(2)}(z,z^{-1})$ with $U(z,z^{-1})$ unimodular over the ring $\mathbb{R}[z,z^{-1}]$, i.e. $\det_{\mathbb{R}(z)}U(z,z^{-1})=cz^d$ with $c\in\mathbb{R}\setminus\{0\}$ and $d\in\mathbb{Z}$.

Note that the unimodularity of U over the ring $\mathbb{R}[z,z^{-1}]$ allows that its determinant is a function of z.
In this signal based framework the notion of controllability becomes a property of the behaviour of the system. It is defined as the property that any two admissible trajectories w_1, $w_2 \in \mathcal{B}(S)$ can be concatenated into a new admissible trajectory, by insertion of a finite time trajectory [12]. In terms of a polynomial matrix representation of an (i/o/pr)–model, the property of controllability can be formulated as follows.

Proposition 5. An (i/o/pr)–model $M_{p,m}$ is controllable if and only if it can be induced by a full row rank polynomial matrix $T(z,z^{-1})$, that satisfies $\text{rank}_{\mathbb{C}}T(\lambda,\lambda^{-1})=p$ for all $\lambda\in\mathbb{C}\setminus\{0\}$. □

The condition as formulated in this proposition is equivalent with the condition of left coprimeness of the polynomial matrices $\{P(z,z^{-1}),Q(z,z^{-1}),R(z,z^{-1})\}$.
An (i/o/pr)–model $M_{p,m}$ generates two transfer functions: one transfer function $H_y(z)$ that considers y as the processing variable, and one transfer function $H_e(z)$ considering e as the processing variable. If $M_{p,m}$ is induced by a full row rank polynomial matrix $T=[P|-Q|-R]\in\mathbb{R}^{p\times(p+m+p)}[z,z^{-1}]$, these two transfer functions can be denoted by:
$$H_y(z) = [H_{yu}(z)\,|\,H_{ye}(z)] := [\ P^{-1}Q \mid P^{-1}R\ ], \text{ and} \quad (1)$$
$$H_e(z) = [H_{ey}(z)\,|\,H_{eu}(z)] := [\ R^{-1}P \mid -R^{-1}Q\ ] \text{ with} \quad (2)$$
$H_y, H_e\in\mathbb{R}^{p\times m}(z)$.
The behaviour of $M_{p,m}$ is completely characterized by either one of the two transfer functions if and only if $M_{p,m}$ is

controllable, i.e. T does not contain nontrivial left factors.
In order to be able to separate the input–output part of an (i/o/pr)–model, we will denote the i/o part of its behaviour by $\mathcal{B}^{io}(M) := \{v \in V^{\mathbb{Z}} \mid (v,e) \in \mathcal{B}(M), e=0\}$. If an (i/o/pr)–model M equals $(\mathbb{Z}, W, \mathcal{B}(M))$, then its i/o–part M^{io} is defined by $M^{io} = (\mathbb{Z}, V, \mathcal{B}^{io}(M))$.

Residual–Based System Identification

In the problem of modelling input–output data of a dynamical system, three central aspects will be distinguished:

1. The set of models that is considered; i.e. the set of all models among which (a) best model(s) is (are) searched for, given the measured data;
2. a parametrization, representing the models in the model set with (real valued) parameters, and
3. an identification criterion that selects "best" or "optimal" models from the set of models.

Given the measurement data, the models that are finally obtained as a result of the modelling procedure should be determined by the set of models taken into account and by the identification criterion, and should not be dependent on other choices, like e.g. the parametrization of the model set. We will now present a formal introduction of the three notions mentioned above.

Definition 6. Denote with $\mathcal{M}_{p,m}$ any collection of controllable (i/o/pr)–models $\{M_{p,m}^{(\alpha)}, \alpha \in I_\alpha\}$ with I_α an index set. □

If the subscripts are clear from the context, $\mathcal{M}_{p,m}$ will be written as $\mathcal{M}$. The restriction that is made to controllable (i/o/pr)–models in $\mathcal{M}$ will be motivated later on. Note that, because of this controllability, the models can be characterized by their transfer function $H_y(z)$ or $H_e(z)$ as discussed in the previous section.

Definition 7. A *parametrization* $\tilde{M}$ for a model set $\mathcal{M}_{p,m}$, is a surjective mapping $\tilde{M}: \Theta \to \mathcal{M}_{p,m}$, with $\Theta \subset \mathbb{R}^d$ the parameter set, and $\mathcal{M}_{p,m}$ the parametrized set of models. □

The image of a parametrization is a model set. The representation of models in terms of polynomial matrices can also act as a parametrization, which formally can be denoted by:

$\tilde{M} = \tilde{M}_P \circ \tilde{G}_P$ with mappings:

$\tilde{G}_P: \Theta \to \Theta_P$ with $\Theta \subset \mathbb{R}^d$, and $\Theta_P \subset \mathbb{R}^{p\times(p+m+p)}[z,z^{-1}]$, and

$\tilde{M}_P: \Theta_P \to \mathcal{M}_{p,m}$, with $\tilde{G}_P$ bijective, $\tilde{M}_P$ surjective, and

$M = \tilde{M}_P(T)$ specified by $\mathcal{B}(M) := \{w \in (\mathbb{R}^{p+m+p})^{\mathbb{Z}} \mid T(\sigma,\sigma^{-1})w = 0\}$.
With slight abuse of the definitions we will also speak about the polynomial matrix parametrization $\tilde{M}_P$ (in stead of $\tilde{M}$). Without loss of generality it will be implicitly assumed that the polynomial matrix parametrization is restricted to polynomial matrices having full row rank, $\operatorname{rank}_{\mathbb{R}(z)} T(z,z^{-1}) = p$. Since any (i/o/pr)–model $M_{p,m}$ admits such a representation, this assumption can be made without loss of generality.
We will now focus on the identification criterion (see also [6],[7]).

Definition 8. An *identification criterion* J^N on a model set $\mathcal{M}$ over a time set $T_N = \mathbb{Z} \cap [0,N-1]$ is defined as an operator

$J^N: V^{T_N} \times 2^{\mathcal{M}} \to 2^{\mathcal{M}}$ with $2^{\mathcal{M}}$ the set of all subsets of $\mathcal{M}$, such that for all $\mathcal{M}^* \subset \mathcal{M}$ and $v^N \in V^{T_N}$ it holds that $J^N(v^N, \mathcal{M}^*) \subset \mathcal{M}^*$. □

For a given data sequence $v^N \in V^{T_N}$, a model $\hat{M} \in \mathcal{M}$ is called an *optimal model* within $\mathcal{M}$ *with respect to* J^N if $\hat{M} \in J^N(v^N, \mathcal{M})$. Actually the identification criterion is defined as a selection rule. Given a time series and a model set, the criterion J selects one or more optimal models from the model set. If J is defined on model set $\mathcal{M}$, then by definition it also can be applied to any subset of $\mathcal{M}$. This definition is a very general one. It allows a unifying approach to a number of system identification methods, and moreover it offers the possibility to incorporate extensions to them.
In this paper a restriction will be made to identification criteria that are based on a *residual function* $\ell^N: E^{T_N} \to \mathbb{R} \cup \{\infty\}$. Such criteria take residual signals that are compatible with the measured data v^N for a model M, as "measures of fit" between the data and the model; The residual function is examined on the residual signals that, together with the measured data sequence v^N, constitute an admissible trajectory $w^N \in \mathcal{B}(M)$ for $M \in \mathcal{M}$; these residuals are contained in $\mathcal{E}^N(v^N, M)$. In formal terms this can be denoted as follows:

$$J^N(v^N, \mathcal{M}) = \arg \min_{M \in \mathcal{M}} \min_{e^N \in \mathcal{E}^N(v^N, M)} \ell^N(e^N) \tag{3}$$

$$:= \{\hat{M} \in \mathcal{M} \mid \exists \hat{e}^N \in \mathcal{E}^N(v^N, \hat{M}),\ \ell^N(\hat{e}^N) \le \ell^N(e^N) \text{ for all } e^N \in \mathcal{E}^N(v^N, M) \text{ and } M \in \mathcal{M}\} \tag{4}$$

with $\mathcal{E}^N(v^N, M) := \{e^N \in E^{T_N} \mid (v^N, e^N) \in \mathcal{B}^N(M)\}$. (5)
The identification criterion has been formulated as a minimization problem. It has to be stressed that this is a specific choice, and that other formulations are possible [8]. In this paper a restriction will be made to the most straightforward choice for the identification criterion, i.e. the minimization of a sum of squared residual terms:

$$\ell^N(e^N) = \frac{1}{N} \sum_{t=0}^{N-1} e^T(t) e(t) \tag{6}$$

For $N \to \infty$ this expression will be denoted by $\ell^+(e^+)$ and the corresponding identification criterion by J^+_{LS}.

Returning to the set of models $\mathcal{M}$, one of the important properties of a model $M \in \mathcal{M}$ is the way in which the residual signal is connected to the measured input and output signals of the model. The connection between i/o signals and residual signals determines the "location" in the model where disturbance terms between data and the i/o–part of the model are represented. In residual–based identification methods the residual signals form the basis for the "measure of fit" between measured i/o–data and (i/o/pr)–models. Consequently the residual signals have a great influence on the properties of the optimal models finally obtained.
We will characterize three different types of (i/o/pr)–models with respect to their residual types: prediction error (PE) models, output error (OE) or simulation error models, and equation error (EE) models. These expressions are generally known in the literature, see e.g. [5]. In this paper the different types of residuals will be presented as general properties of controllable (i/o/pr)–models, formulated in terms of their polynomial matrix representations. For the basic definitions of these types of models in terms of their behaviour, one is referred to [8], [9].

Proposition 9. Let $M_{p,m}$ be a controllable (i/o/pr)–model, that is induced by a full row rank polynomial matrix $T(z,z^{-1})$ with $T = [P \mid -Q \mid -R]$. Then:

a. $M_{p,m}$ is a k–step ahead prediction error (PE) model if and only if
   1. $P(z,z^{-1})^{-1}Q(z,z^{-1})$ is a proper rational matrix, and
   2. there exists a nonsingular matrix $L \in \mathbb{R}^{p\times p}$ such that $\{P(z,z^{-1})^{-1}R(z,z^{-1}) - L\}z^k$ is proper.
b. $M_{p,m}$ is an output error (OE) or simulation error model if

and only if $P(z,z^{-1})^{-1}R(z,z^{-1}) = L \in \mathbb{R}^{p\times p}$, nonsingular.

c. $M_{p,m}$ is an equation error (EE) model if and only if $R(z,z^{-1})$ is unimodular with respect to $\mathbb{R}^{p\times p}[z,z^{-1}]$. □

Note that the different model properties are not conflicting. Any combination of (PE), (OE) and (EE) properties is also possible. These different residual type models will be distinguished in the sequel of this paper.
In this section the three basic components of an identification procedure have been introduced. In the following section it will be discussed which interrelations between these three notions have to be taken into account.

Discriminable model sets and identifiable parametrizations

Let us consider the following example in which a least squares identification criterion is applied to a set of equation error models.

Example 1.
Consider a set $\mathcal{M}$ of (i/o/pr)–models with one input and one output, and consequently one residual signal (SISO situation):

$$\mathcal{M} = \{M_{1,1} \mid M_{1,1} \text{ is induced by } T(z,z^{-1})=[a_0+a_1z \mid -b_0-b_1z \mid -c],\ a_0,a_1,b_0,b_1,c\in\mathbb{R},((a_0+a_1z),(b_0+b_1z)) \text{ coprime}\}. \quad (7)$$

Because of the fact that we are dealing with (i/o/pr)–models the parameter c in $T(z,z^{-1})$ is not allowed to be equal to 0, and at least one of the parameters $\{a_0,a_1\}$ has to be unequal to 0.
As an identification criterion we will consider the well known least squares criterion J^+_{LS} as formulated in (3)–(6).

When given a data sequence v^+ the following situations can be distinguished.

a. Data sequences v^+ for which holds that $f^+(v^+,M)>0$ for all $M\in\mathcal{M}$; for these data sequences it follows that $J^+_{LS}(v^+,\mathcal{M})=\emptyset$ since $\ell^+(e^+)\to 0$ for $a_0,a_1,b_0,b_1\to 0$, or similarly $c\to\infty$, and the limit point $\ell^+(e^+)=0$ can not be reached within $\mathcal{M}$.

b. Data sequences v^+ for which holds that $f^+(v^+,M)=0$ for some $M\in\mathcal{M}$; any such data sequence can be written as $v^+\in\mathcal{B}^+(M^{io})$ with $M\in\mathcal{M}$ and for such a data sequence the following expression holds true:
$M_1\in J^+_{LS}(v^+,\mathcal{M}) \Leftrightarrow \{\exists d\in\mathbb{R}\setminus\{0\}$ such that $(v,e)\in\mathcal{B}(M) \Leftrightarrow (v,e')\in\mathcal{B}(M_1)$ with $e=de'\}$.
Consequently all models that lead to residual signals that are related by scaling factors, are always selected together in the identification criterion. □

In the generic situation (a) of the example, there will not be selected any optimal model, i.e. the identification criterion $J^+_{LS}(v^+,\mathcal{M})$ will be empty. Apparently the chosen $\mathcal{M}$ and J^+_{LS} do not lead to a useful identification problem. The classical way to circumvent this problem is to restrict the considered model set by requiring that the polynomial matrix $T(z,z^{-1})$ in eq. (7) satisfies the additional constraint $a_0/c=1$. However it should be stressed that this is a specific choice out of many different alternatives and, moreover, that each different choice of restricting the model set will generally lead to different selected optimal models. Apparently restriction of the model set is required in order to come up with a sensible identification problem, but it is also necessary in order to guarantee that all models in the model set can be distinguished by the identification criterion. The ability to distinguish the different models in a model set during identification experiments will now be formalized.

Definition 10. A model set $\mathcal{M}$ is called *discriminable by an identification criterion* J^N, defined on $\mathcal{M}$, if for all $M\in\mathcal{M}$ there exists a data sequence $v^N\in V^{T_N}$ such that $J^N(v^N,\mathcal{M})=\{M\}$. □

If $\mathcal{M}$ is discriminable by J^N, then J^N is called *discriminating* on $\mathcal{M}$. An identification criterion that is discriminating on a model set can distinguish between the different models in this set. This aspect of discriminability can also be approached from a different point of view, using a slightly weaker formulation in terms of an identification criterion–based equivalence relation on $\mathcal{M}$, see [6] and [7].

Dealing with model sets that are not discriminable by the identification criterion is an undesirable situation; this is exactly the situation as illustrated in the example shown.
Apart from the "usefulness" of the identification problem, the concept of discriminability is also important for the construction of identifiable parametrizations. Identifiability of a parametrization will be considered as the ability in identification experiments to distinguish between different parameter values $\theta_1, \theta_2 \in \Theta$ in a parametrization $\tilde{M}: \Theta \to \mathcal{M}$. In this line of thought it is very natural to relate the identifiability of a parametrization to the identification criterion that is applied. This is formulated in the following definition.

Definition 11. A parametrization $\tilde{M}: \Theta \to \mathcal{M}$ with $\Theta\subset\mathbb{R}^d$ is *strictly identifiable by an identification criterion* J^N defined on $\mathcal{M}$ if:
* $\tilde{M}$ is a bijective mapping, and
* J^N is discriminating on $\mathcal{M}$. □

The strict identifiability of a parametrization deals with the question whether it is possible at all to find a unique parameter $\hat{\theta}$ as the solution to an identification problem, irrespective of the data. In line with this philosophy, identifiability is defined as a property of a parametrization and an identification criterion, in contrast with the current approach in the literature where it is only directed towards the bijective mapping $\tilde{M}$, see e.g. [2], [5]. In line with this approach, a parametrization will be called *identifiable* if, in stead of $\tilde{M}$ being bijective, it holds that $\tilde{M}(\theta_1) = \tilde{M}(\theta_2) \Rightarrow \theta_1 = \theta_2$ for almost all $\theta_1,\theta_2\in\Theta$, reflected by the expression that $\tilde{M}$ is *almost bijective*.

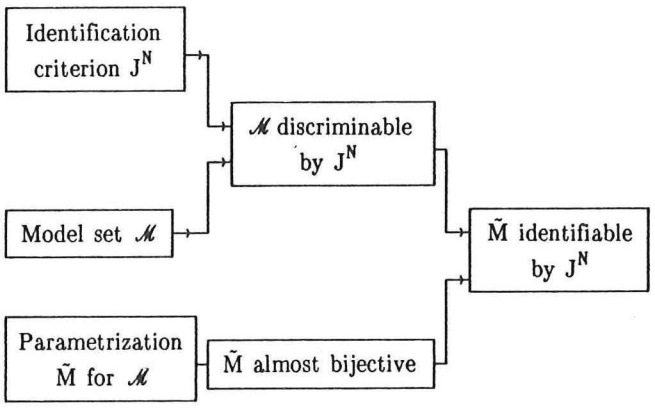

Figure 1 Schematic representation of the requirements for a parametrization $\tilde{M}$ of $\mathcal{M}$ to be identifiable by J^N.

The situation as formulated in this section is represented in Figure 1. In the following section the discriminability of specific types of model sets will be discussed in view of the least squares identification criterion presented before.

Discriminability of model sets with respect to J^+_{LS}

In this section it will be discussed which specific choices can be made in order to guarantee that model sets are discriminable by the asymptotic least squares identification criterion J^+_{LS}.

We will consider model sets of controllable (i/o/pr)–models. Any such model set $\mathcal{M}_{p,m}$ can be uniquely represented by a set of rational functions $\Theta_t \subset \mathbb{R}^{px(p+m)}(z)$ with $H_e(z) \in \Theta_t$ the transfer function of a model $M \in \mathcal{M}$ considering the residual e as the output (processing) variable. According to (2) this transfer function is written as $H_e(z) = [H_{ey}(z)|H_{eu}(z)]$.

We consider two elements $M_1, M_2 \in \mathcal{M}$ being induced by rational matrices $H_e^{(1)}(z)$, $H_e^{(2)}(z) \in \Theta_t$. It can be shown [8] that a sufficient condition for discriminability of $\mathcal{M}$ with respect to J^+_{LS} is given by the requirement that for any such M_1, M_2 it holds that the rational function $H_{ey}^{(2)}(H_{ey}^{(1)})^{-1}$, if it is stable, can be written as:

$$H_{ey}^{(2)}(H_{ey}^{(1)})^{-1}(z) = I + \sum_{\substack{k=t \\ k\neq 0}}^{\infty} M(k)z^{-k} \quad \text{for some } t \in \mathbb{Z}. \qquad (8)$$

It seems to be a straightforward choice to satisfy this condition by imposing restrictions on the transfer function $H_{ey}(z)$, or its inverse $H_{ye}(z)$, within Θ_t.

Considering the classification of models in terms of their residual–type properties, as presented in proposition 9, the following observations can be made. K–step ahead prediction error models and output error models intrinsically contain a nonsingular matrix $L \in \mathbb{R}^{pxp}$ in their representation. This matrix L is directed towards the relation between outputs and residuals in the corresponding models, reflected by the expression $L = \lim_{z\to\infty} \{z^{k-1} H_{ye}(z)\}$. In view of the statements made above, this property advocates a special restriction of model sets in order to arrive at discriminability by J^+_{LS}, by fixing this matrix L for all models in a model set on beforehand. This is formulated in the following corollary [8].

Corollary 12.

Denote with $\mathcal{M}^{pe}_{p,m}(L)$, $\mathcal{M}^{oe}_{p,m}(L)$ any collection of controllable k–step ahead prediction error models, c.q. output error models, with a prespecified –and for all models equal– nonsingular matrix $L \in \mathbb{R}^{pxp}$ as defined in proposition 9. Then $\mathcal{M}^{pe}_{p,m}(L)$, $\mathcal{M}^{oe}_{p,m}(L)$, is discriminable by J^+_{LS}. □

Considering model sets with a fixed nonsingular matrix L=I is the most common situation in prediction error and output error modelling. For prediction error models this situation can actually be compared with the application of a (stochastic) innovations representation in state space form. In contrast with these prediction error and output error models, the definition of equation error models does not point to a specific restriction of the corresponding model sets. However also for equation error models discriminability can be achieved. One of the options is to choose prediction error or output error form restrictions, i.e. by requiring the models in the EE model set to satisfy $\lim_{z\to\infty} \{z^{k-1} H_{ye}(z)\} = L$ with the nonsingular $L \in \mathbb{R}^{pxp}$ fixed over the model set (PE–form), or even $H_{ye}(z) = L$ (OE–form).

Apart from these possibilities, equation error models give extra opportunities to restrict corresponding model sets to be discriminable, by exploiting the fact that for an equation error model the transfer function $H_e(z)=[H_{ey}(z)|H_{eu}(z)]$ is a polynomial matrix. This special property points to alternative ways for obtaining discriminability, through restrictions on the polynomial degrees in this transfer function. In this line of thought we will present one result for obtaining discriminability of equation error model sets. Without loss of generality we will restrict attention to models having a transfer function $H_e(z)$ that is polynomial in one indeterminate, i.e. $H_e(z) \in \mathbb{R}^{px(p+m)}[z]$.

Theorem 13.

Let L be a nonsingular matrix $L \in \mathbb{R}^{pxp}$ and let $\Theta_t(L)$ be a set of polynomial transfer function matrices: $\Theta_t(L) \subset \mathbb{R}^{px(p+m)}[z]$ with $H_e \in \Theta_t(L)$ written as $H_e=[H_{ey}|H_{eu}]$, satisfying $\text{rank}_{\mathbb{R}(z)} H_{ey}=p$ and $\Gamma_{hc}(H_{ey})=L$ for all $H_e \in \Theta_t(L)$, with $\Gamma_{hc}(H_{ey})$ the leading column coefficient matrix of H_{ey}.

Let $\Theta_t(L)$ induce an equation error model set $\mathcal{M}^{ee}_{\Theta_t(L)}$.

Denote with $\mu_i(H_{ey})$ the i^{th} column degree of H_{ey} and consider the following conditions.

(i) $H_e^{(1)}, H_e^{(2)} \in \Theta_t \Rightarrow \mu_i(H_{ey}^{(1)}) = \mu_i(H_{ey}^{(2)})$ for i=1,.,p;

(ii) $H_e^{(1)}, H_e^{(2)} \in \Theta_t \Rightarrow \sum_{i=1}^{p} \mu_i(H_{ey}^{(1)}) = \sum_{i=1}^{p} \mu_i(H_{ey}^{(2)})$; and $H_e \in \Theta_t \Rightarrow \text{rank}[H_{ey}(0)]=p$.

If Θ_t satisfies (i) or (ii) then $\mathcal{M}^{ee}_{\Theta_t(L)}$ is discriminable by J^+_{LS}. □

Proof. See [8].

The restriction that is proposed for obtaining discriminability of equation error model sets, is not presented in the form of a fixed Markov parameter of the transfer function $H_{ye}(z)$, but in the form of a fixed leading column coefficient matrix of the polynomial transfer function $H_{ey}(z)$. Some additional conditions have to be fulfilled, such as a fixed set of column degrees of $H_{ey}(z)$, or a fixed sum of column degrees of $H_{ey}(z)$ and absence of poles in z=0 of the transfer function $H_y(z)$. It has to be stressed that the result as presented in this theorem reflects only one choice out of a great number of possibilities for obtaining discriminability of equation error model sets in general. A dual approach directed towards row degrees in stead of column degrees can be found in [8]. It has to be stressed that every other choice of arriving at discriminability will lead to different model sets, and generally also to different results for the identification procedure, i.e. to different sets of selected optimal models.

The specific restrictions as formulated in the theorem are chosen because of their very close relationship with representations of model sets of all models having a specified McMillan degree.

The two different approaches for obtaining discriminability with respect to J^+_{LS} for PE and OE model sets on the one hand, and EE model sets on the other hand, also imply a separate treatment of the problem of constructing identifiable parametrizations (see fig. 1). These consequences will be briefly glanced at in the final section.

Some consequences for identifiable parametrizations

When constructing identifiable parametrizations for model sets that satisfy the conditions of discriminability as presented in the previous section, we have to distinguish between the different residual type of models.

Any set of prediction error or output error models, satisfying the conditions of corollary 12, can be parametrized in a polynomial matrix parametrization by means of a set of polynomial matrices $\Theta \subset \mathbb{R}^{px(p+m+p)}[z,z^{-1}]$ with $T \in \Theta$ written as $T=[P|-Q|-R]$, satisfying the properties as formulated in proposition 9a or 9b.

In order to obtain an identifiable parametrization it is required

that the parametrization is (almost) bijective, i.e. that Θ does not contain distinct elements that generate the same model. With proposition 4 it follows that obtaining an identifiable parametrization comes down to restricting Θ in such a way that there do not exist two distinct elements that are related through premultiplication with a unimodular matrix. Note that the restrictions for obtaining discriminability with respect to J^+_{LS}, do not solve the parametrization problem. A parametrization problem remains in terms of removing unimodular premultiplication within Θ.

For equation error model sets that are discriminable by J^+_{LS} the situation is different. Any set of equation error models, satisfying the conditions of theorem 13 can be parametrized in a polynomial matrix parametrization by means of a set of polynomial matrices $\Theta_{ee} \subset \mathbb{R}^{px(p+m+p)}[z,z^{-1}]$ with $T \in \Theta_{ee}$ written as $T=[P|-Q|-R]$. Because of the unimodularity of matrix R (by definition), it follows directly that any such set of models can also be represented by a set of polynomial matrices $\Theta^*_{ee} \subset \mathbb{R}^{px(p+m+p)}[z,z^{-1}]$ with $T^* \in \Theta^*_{ee}$ written as $T^*=[P^*|-Q^*|-I]$. The parametrization induced by Θ^*_{ee} is bijective since the identity matrix on the position of matrix R^* does not allow a unimodular premultiplication. The identifiability problem for this kind of EE model sets has now been solved straightforwardly.

For a summary of the results of the previous sections, it is illustrative to reconsider figure 1. The upper part of the diagram refers to the discriminability of the model set, being an essential choice that influences the optimal models finally obtained by J^+_{LS}. The lower part refers to the uniqueness of the parametrization which is just a matter of representation. For output error and prediction error model sets it has been shown that there exists a straightforward choice for obtaining discriminable model sets (corollary 12) and that a problem of parametrization remains when identifiability with respect to J^+_{LS} is required. For equation error model sets this situation is different. There are many different ways for obtaining discriminability with respect to J^+_{LS}, and each different choice generally yields different identified models. As shown in this section, there actually does not exist a parametrization problem for the equation error model sets discussed. Where the essential problem of PE and OE model sets is a problem of parametrization, the main problem for EE model sets is a problem of obtaining discriminability.
For equation error identification methods the aspect of discriminability is generally not recognized as an explicit users' choice; this leads to optimal models that are partly determined by the (arbitrary) way in which discriminability was achieved. Moreover it means that the experimenter – without notice – can have an essential influence on the optimal models finally obtained. In view of the fact that the identification methods discussed are required to be based on explicit users' choices, the statement made above makes the equation error approach rather questionable. This is especially remarkable if we take into account that equation error methods are by far the most popular methods in the application of system identification techniques.

Conclusions

In this paper the problem of system identification is considered as a problem of deterministic approximate modelling of input–output data. In order to deal with this problem a framework has been constructed in which the basic components of a modelling procedure are reformulated. Residual signals are incorporated in the model descriptions in order to deal with modelling errors and to determine a measure of fit between a measured data sequence and a model. A clear distinction has been made between the problems of identification and parametrization. The identifiability of parametrizations has been reconsidered and has been shown to consist of two different aspects: the discriminability of the model set with respect to the identification criterion, and the uniqueness of the parametrization of the models in the set. Whereas for output error and prediction error model sets the construction of unique parametrizations is the main problem, it is shown that for equation error model sets the essential problem is to obtain model sets that are discriminable by a (least squares) identification criterion.

References

[1] Eykhoff, P. (1974), *System Identification: Parameter and State Estimation.* Wiley & Sons, London.

[2] Gevers, M. and V. Wertz (1987), "Techniques for the selection of identifiable parametrizations for multivariable linear systems". In: C.T. Leondes (Ed.), *Control and Dynamic Systems*, Vol.24, pp.35–86. Ac. Press, New York.

[3] Goodwin, G.C. and R.L. Payne (1977), *Dynamic System Identification: Experiment Design and Data Analysis.* Academic Press, New York.

[4] Kalman, R.E. (1982), "System identification from noisy data." In: A.R. Bednarek and L. Cesari (Eds.), *Dynamical systems II.* Academic Press, New York. pp. 135–164.

[5] Ljung, L. (1987), *System Identification: Theory for the User.* Prentice–Hall, Englewood Cliffs, N.J.

[6] Van den Hof, P.M.J. (1989a), "A criterion based approach to parametrization and identification of multivariable systems". In: Han–Fu Chen (Ed.), *Identification and System Parameter Estimation 1988*, IFAC Proceedings Series, 1989, No. 8, pp. 715–720.

[7] Van den Hof, P.M.J. (1989b), "Criterion based equivalence for equation error models", *IEEE Trans. Autom. Control*, Vol. AC–34, No. 2, pp. 191–193.

[8] Van den Hof, P.M.J. (1989c), *On Residual–Based Parametrization and Identification of Multivariable Systems.* Dr. Dissertation, Dept. Electrical Engineering, Eindhoven University of Technology, The Netherlands, 1989.

[9] Van den Hof, P.M.J. (1990), "On discriminable model sets in residual–based system identification", Proc. Int. Symposium on the Mathem. Theory on Networks and Systems, (MTNS–89), Amsterdam, June 19–23, 1989. In: *Progress in Systems and Control Theory*, to be published by Birkhauser Boston Inc.

[10] Willems, J.C. (1986a), "From time series to linear system – Part I: Finite dimensional linear time–invariant systems", *Automatica*, Vol. 22, pp. 561–580.

[11] Willems, J.C. (1986b), "From time series to linear system – Part II: Exact modelling." *Automatica*, Vol. 22, 675–694.

[12] Willems, J.C. (1988), "Models for dynamics", In: Kirchgraber, U. and H.O. Walther (Eds.), *Dynamics Reported*", Vol. 2, pp. 171–269. Wiley and Teubner, 1988.