Identification of Multivariable Linear Systems of Unknown Structure by the Prior Knowledge Fitting Method

Valis, Jaroslav

1970

Document Version:
Publisher's PDF, also known as Version of record

IDENTIFICATION OF MULTIVARIABLE LINEAR SYSTEMS OF UNKNOWN STRUCTURE BY THE PRIOR KNOWLEDGE FITTING METHOD

J.VALIŠ

REPORT 7005 SEPTEMBER 1970
LUND INSTITUTE OF TECHNOLOGY
DIVISION OF AUTOMATIC CONTROL

IDENTIFICATION OF MULTIVARIABLE LINEAR SYSTEMS OF UNKNOWN STRUCTURE BY THE PRIOR KNOWLEDGE FITTING METHOD †

J. Vališ ††

† This work has been supported by the Swedish Board for Technical Development under Contract No. 69-631/U489.

†† Now with L.M. Ericsson AB, MI-Divisionen, Stockholm, Sweden.

ABSTRACT

This report deals with identification of multivariable linear systems of finite order from measurements of input-output signals when the outputs are corrupted by additive random noises. The input-output relations of such systems can be described by a vector difference equation. A canonical form for the matrix coefficients of this equation is proposed and its existence and uniqueness are shown. This canonical form enables the identification of the structure of matrix coefficients by straightforward search, which drastically reduces the number of different structures, whose "fit" is otherwise to be tested.

An application of the prior knowledge fitting method for identification of the structure of the matrix coefficients and estimation of their elements, based on the (assumed) independence of noises and input signals, is described. No knowledge of statistical characteristics of the noises is needed, only their mean values are assumed to be constant in time. This method is finite (i.e. no iterations are used, no troubles with convergence etc.).

The required place in the computer memory does not grow with increasing length of observation of input-output data, as they need not to be stored; all information from the whole past history needed for identification of all examined structures is preserved in a relatively small matrix.

With increasing length of input-output data the obtained estimates converge to their true values with probability one.

A FORTRAN IV subroutine for this method, called MIMOID, was written and tested. The results confirm its usefulness.

TABLE OF CONTENTS

1. INTRODUCTION.

In recent years a considerable effort has been made in the field of experimental identification of linear discrete time (or sampled) single output-single (possibly multiple) input systems. As a result many useful algorithms have been described elsewhere and some of them have become very popular. Among them let us mention the maximum likelihood method (M.L.) by Åström and Bohlin [1] as an example of a statistically optimal, even if computationally expensive, method or its "counterpart", the extremely feasible computationally, even if not statistically optimal, prior knowledge fitting method (P.K.F.) by Peterka and Smuk [2], [3]. Both of these methods (and many others) approximate the examined system or process by a parametric model, whose input-output relation is described by a linear difference equation of finite order n, the parameters of which (i.e. the coefficients of the diff. eq.), denote them by θ_n, are chosen to minimize a loss function $V = V(\theta_n)$, which in some way penalizes the difference in the input-output behaviour of the model and the identified system or process. By these two methods (and all others known to the author) the order n is not identified directly as a parameter, but the models are determined for increasing orders i = 1, 2, ... with corresponding loss functions and as the "right" order n we accept the greatest i giving "significant" decrease of loss function. Rather exact criteria (as statistical tests, analysis of residuals) as well as more or less subjective judging was successfully used to determine whether the loss function is decreased significantly [3], [4], [5], [6].

The main disadvantage of this approach is that we have to evaluate (n+1) models to decide that the system is of n:th order which may sometimes be a time consuming task, depending mainly on the complexity of the used algorithm. (An example from [6]: Models up to 4th order were identified from 1000 pairs of input-output data from a paper machine. Total computing time for M.L. was 4 min., for P.K.F. 0.5 min.)

In spite of this, no better method for estimation of the order of a process is known at present time. In the next sections it shall be shown that a similar approach can be developed even for linear sampled systems with multiple outputs which can be described by a

vector difference equation (of higher order). The general properties of such an equation and its relation to the state space description shall be discussed in section 2.

In section 3 it shall be shown that there exists certain canonical form of the vector difference equation, which enables the determination of the order by similar simple algorithms as mentioned above.

In section 4 a multivariable version of the prior knowledge fitting method, which makes use of this canonical form, is described. This method has proved to be extremely suitable for this purpose and may even be used on-line.

Finally, in section 5 some practical results from identification of both simulated and industrial data are shown.

2. DESCRIPTION OF INPUT-OUTPUT RELATION BY A VECTOR DIFFERENCE EQUATION (V.D.E.).

In this section we shall derive the vector difference equation for a linear sampled system $\mathcal{S}$ of n:th order with r inputs and m outputs. Let us denote by $u_i(k)$, $y_i(k)$ the value of i:th input resp. output at time kT, T being the sampling period. Let us further define column vectors:

$$u(t) \triangleq \begin{bmatrix} u_1(t) \\ u_2(t) \\ \cdots\cdots \\ u_r(t) \end{bmatrix} ; \quad y(t) \triangleq \begin{bmatrix} y_1(t) \\ y_2(t) \\ \cdots\cdots \\ y_m(t) \end{bmatrix}$$

The state space description $\mathcal{S}(\phi, \Gamma, C, D)$ relates the output y(t) to the input u(t) by means of an n dimensional vector x(t) called the state of the system by the following equation:

$$\begin{aligned} x(t+1) &= \phi\, x(t) + \Gamma\, u(t) \\ y(t) &= C\, x(t) + D\, u(t) \end{aligned} \qquad (2.1)$$

where

ϕ is regular (n × n) matrix
Γ is (n × r)
C is (m × n) matrix of rank m with rows $c_1, c_2, \ldots, c_m$
D is (m × r)

In the following we assume that the system $\mathcal{S}$ is completely observable.

The vector difference equation relates the output y(t) to its past values and to present and past values of input, i.e.:

$$y(t) = L\{y(t-1), \ldots, y(t-p); u(t), u(t-1), \ldots, u(t-p)\} \quad (2.2)$$

where $L\{\cdot\}$ is a linear function.

In order to derive the V.D.E. (2.2) for a system $\mathcal{S}(\phi, \Gamma, C, D)$ we must exclude the state vector x(t) from (2.1). This can be done using the idea of observability: If the system $\mathcal{S}$ is observable, the state x(t) can be determined from the following values of output and input:

$$y(t+i), u(t+i); \qquad i = 0, 1, \ldots, p-1; \; p \leq n.$$

From (2.1) we get:

$$\begin{aligned}
y(t) &= C\,x(t) + D\,u(t) \\
y(t+1) &= C\phi x(t) + C\Gamma u(t) + D\,u(t+1) \\
y(t+2) &= C\phi^2 x(t) + C\phi\Gamma u(t) + C\Gamma u(t+1) + D\,u(t) \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
y(t+p-1) &= C\phi^{p-1}x(t) + C\phi^{p-2}\Gamma u(t) + C\phi^{p-3}\Gamma u(t+1) + \ldots \\
\hline
y(t+p) &= C\phi^{p}x(t) + C\phi^{p-1}\Gamma u(t) + C\phi^{p-2}\Gamma u(t+1) + \ldots \qquad (2.3)
\end{aligned}$$

Hence if the system $\mathcal{S}$ is observable we may from the upper part of (2.3) (above the line) calculate x(t), substitute it into the last equation (below the line) and so obtain the V.D.E. wanted [7], [8].

Let us carry this out in detail. First denote:

$$U_p \triangleq \begin{bmatrix} u(t) \\ u(t+1) \\ \hline\hline \\ u(t+p-1) \end{bmatrix};$$

(2.4)

$$\Psi_p \triangleq \begin{bmatrix} D & & & & & \\ C\Gamma & D & & & & \\ C\phi\Gamma & C\Gamma & & & 0 & \\ \cdots & \cdots & \cdots & \cdots & & \cdots \\ C\phi^{p-2}\Gamma & C\phi^{p-3}\Gamma & & C\Gamma & & D \end{bmatrix}$$

(2.4)

$$Y_p \triangleq \begin{bmatrix} y_1(t) \\ y_2(t) \\ \cdots\cdots\cdots \\ y_m(t) \\ \hline\hline \\ y_1(t+1) \\ y_2(t+1) \\ \cdots\cdots\cdots \\ y_m(t+1) \\ \hline\hline \\ \cdots\cdots\cdots \\ \hline\hline \\ y_1(t+p-1) \\ y_2(t+p-1) \\ \cdots\cdots\cdots \\ y_m(t+p-1) \end{bmatrix}; \quad \Omega_p \triangleq \begin{bmatrix} c_1 \\ c_2 \\ \cdots\cdots \\ c_m \\ \hline\hline \\ c_1\Phi \\ c_2\Phi \\ \cdots\cdots \\ c_m\Phi \\ \hline\hline \\ \cdots\cdots \\ \hline\hline \\ c_1\Phi^{p-1} \\ c_2\Phi^{p-1} \\ \cdots\cdots \\ c_m\Phi^{p-1} \end{bmatrix}$$

The smallest positive integer p such that Ω_p has rank n is called the observability index. The (2.3) may now be written as:

$$Y_p = \Omega_p x(t) + \Psi_p U_p \tag{2.5}$$

Because Ω_p has rank n we can always find n independent rows in it and form a set of n linear algebraic equations for x(t). The choice of independent rows in Ω_p need not be unique. Let our choice be:

$$T \triangleq \begin{bmatrix} c_{i_1}\phi^{j_1} \\ c_{i_2}\phi^{j_2} \\ \cdots\cdots \\ c_{i_n}\phi^{j_n} \end{bmatrix}; \qquad Y_p^* \triangleq \begin{bmatrix} y_{i_1}(t+j_1) \\ y_{i_2}(t+j_2) \\ \cdots\cdots\cdots \\ y_{i_n}(t+j_n) \end{bmatrix} \tag{2.6}$$

and similarly let Ψ_p^* be formed from the respective rows of Ψ_p. Thus from (2.5):

$$Y_p^* = Tx(t) + \Psi_p^* U_p$$

i.e.

$$x(t) = T^{-1}Y_p^* - T^{-1}\Psi_p^* U_p \tag{2.7}$$

which, after the substitution for x(t) into the last equation in (2.3), finally gives the wanted V.D.E.

$$y(t+p) = C\phi^p T^{-1}Y_p^* - C\phi T^{-1}\Psi_p^* U_p + \psi_p U_p + 1 \tag{2.8}$$

where

$$\psi_p \triangleq \left[C\phi^{p-1}\Gamma \;\middle|\; C\phi^{p-2}\Gamma \;\middle|\; \cdots \;\middle|\; C\Gamma \;\middle|\; D \right]$$

This form (2.8) is not very suitable. Let us rearrange it more clearly:

The last two terms on the right hand side may be united and written as:

$$G_{p+1}U_{p+1} \overset{\Delta}{=} \psi_p U_{p+1} - C\phi T^{-1}\psi^* U_{p+1}$$

where G_{p+1} has dimensions $(m \times (p+1)r)$. Using notation:

$$F_p \overset{\Delta}{=} C\phi p_T^{-1}$$

we get:

$$y(t+p) = F_p \cdot Y_p^* + G_{p+1}U_{p+1} \qquad (2.9)$$

For the sake of later convenience we prefer to write (2.9) using the whole vector of outputs Y_p instead of Y_p^*. To enable this, we must extend the $(m \times n)$ matrix F_p to an $(m \times (p \cdot m))$ dimensional matrix E_p by introducing appropriate dummy zero columns.

If for example $m = 2$, $p = 3$ and

$$Y_p^* = \left[y_1(t) \;\vdots\; y_2(t) \;\vdots\; y_2(t+1) \;\vdots\; y_1(t+2)\right]^T$$

then

$$E_p = \left[\; f_1 \;\vdots\; f_2 \;\vdots\; 0 \;\vdots\; f_3 \;\vdots\; f_4 \;\vdots\; 0 \;\right]$$

(f_i stands for the original columns of F).

The total number of possibly non-zero columns in E_p remains equal to n, i.e. to the order of the system. So we obtain:

$$y(t+p) = E_p \cdot Y_p + G_{p+1}U_{p+1} \qquad (2.10)$$

Further we split the matrix E_p into p square matrices $A_1, \ldots, A_p$ of dimensions $(m \times m)$ in the following manner:

$$E_p = \left[-A_p \;\vdots\; -A_{p-1} \;\vdots\; \cdots \;\vdots\; -A_2 \;\vdots\; -A_1 \right] \tag{2.11}$$

and similarly the matrix G_{p+1} into (p+1) rectangular matrices B_0, B_1, ..., B_p of dimensions (m × r), i.e.:

$$G_{p+1} = \left[B_p \;\vdots\; B_{p-1} \;\vdots\; \cdots \;\vdots\; B_1 \;\vdots\; B_0 \right] \tag{2.12}$$

Introducing this into (2.10) we obtain:

$$y(t+p) = - A_p y(t) - A_{p-1} y(t+1) - \ldots - A_1 y(t+p-1) +$$

$$+ B_p u(t) + B_{p-1} u(t+1) + \ldots + B_0 u(t+p) \tag{2.13}$$

and after shifting the time t+p => t and rearranging the equation we finally get:

$$y(t) + \sum_{i=1}^{p} A_i y(t-i) = \sum_{i=0}^{p} B_i u(t-i) \tag{2.14}$$

Notice that the total number of (possibly) non zero columns in all matrices A_1, ..., A_p together is still equal to n, i.e. to the order of the system $\mathcal{S}$. Their positions are uniquely given by the particular choice of rows from Ω_p to form the matrix T (2.6).

As it was already mentioned, this choice need not be unique (except for single output case [7]) and hence for the same system $\mathcal{S}(\phi, \Gamma, C, D)$ other equivalent V.D.E. with different structures of A matrices can be found. The number of possible structures depends on the properties of the particular pair of matrices ϕ,C. This implies that the knowledge of the order n is not sufficient to specify the form of V.D.E. (as it is in the single output case [7]). The information about the structure of A matrices must be given as well.

Formerly, this seemed to be a very crucial problem in connection with experimental identification of an unknown system where this information is of course not available and all possible alternative structures of A matrices have to be tested [8], [9].

3. IDENTIFIABLE CANONICAL FORM OF V.D.E.

In this section we show that there exists a canonical form for matrix coefficients in V.D.E. (2.14) which (even if not being unique) reduces the number of possible variants of A.

To do this we show that the following choice of transformation matrix T is always possible.

Let i_{obs} be the observability index of a linear sampled system $\mathcal{S}(\phi, \Gamma, C, D)$ and denote $q = i_{obs} - 1$.

Then among the vectors:

$$\begin{array}{ccccc} c_1\phi^q & c_1\phi^{q-1} & \cdot & c_1\phi & c_1 \\ c_2\phi^q & c_2\phi^{q-1} & \cdot & c_2\phi & c_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ c_s\phi^q & c_s\phi^{q-1} & \cdot & c_s\phi & c_s \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ c_m\phi^q & c_m\phi^{q-1} & \cdot & c_m\phi & c_m \end{array} \qquad (3.1)$$

according to the definition of the observability index there must be n independent vectors.

1. Let us examine them in the order indicated by the arrow, i.e. starting with $c_1\phi^q$, the ones found independent are chosen to create the matrix T. Hence $c_1\phi^q$ becomes the first row of T, $c_2\phi^q$ the second, etc. We continue in this way until a certain $c_s\phi^{q-k}$ is found to be linearly dependent on the rows $c_1\phi^q$, $c_2\phi^q$, ..., $c_{s-1}\phi^{q-k}$ previously chosen for matrix T. Assume this linear dependence in the form:

$$c_s\phi^{q-k} = \sum_{i=1}^{m} \sum_{j=0}^{k-1} \alpha_{i,j}\, c_i\phi^{q-j} + \sum_{i=0}^{s-1} \alpha_{i,k}\, c_i\phi^{q-k} \qquad (3.2$$

with not all $\alpha_{i,j}$ equal to zero.

The row $c_s\phi^{q-k}$ does not appear in T.

2. We proceed and examine the vectors $c_{s+1}\phi^{q-k}$, ..., $c_m\phi^{q-k}$, $c_1\phi^{q-(k+1)}$, ..., $c_{s-1}\phi^{q-(k+1)}$. Each of them may <u>either</u> be found independent of the rows chosen previously and appears in T <u>or</u> be found dependent, i.e. be a linear combination of rows previously chosen for T.

3. Having completed this we proceed to $c_s\phi^{q-(k+1)}$. The question is whether or not this may be independent of previously chosen rows and hence become a new row in T.

It follows from (3.2)

$$c_s\phi^{q-(k+1)} = \sum_{i=1}^{m}\sum_{j=0}^{k-1} \alpha_{i,j}\, c_i\phi^{q-(j+1)} + \sum_{i=1}^{s-1} \alpha_{i,k}\, c_i\phi^{q-(k+1)}$$

$$= \sum_{i=1}^{m}\sum_{j=1}^{k-1} \alpha_{i,j-1} c_i\phi^{q-j} + \sum_{i=1}^{m} \alpha_{i,k-1}\, c_i\phi^{q-k} +$$

$$+ \sum_{i=1}^{s-1} \alpha_{i,k}\, c_i\phi^{q-(k+1)}$$

$$= \sum_{i=1}^{m}\sum_{j=1}^{k-1} \alpha_{i,j-1}\, c_i\phi^{q-j} + \sum_{i=1}^{s-1} \alpha_{i,k-1}\, c_i\phi^{q-k} +$$

$$+ \alpha_{s,k-1}\, c_s\phi^{q-k} + \sum_{i=s+1}^{m} \alpha_{i,k-1}\, c_i\phi^{q-k} +$$

$$+ \sum_{i=1}^{s-1} \alpha_{i,k}\, c_i\phi^{q-(k+1)}$$

$$= \sum_{i=1}^{m}\sum_{j=1}^{k-1} \alpha_{i,j-1}\, c_i\phi^{q-k} + \sum_{i=1}^{s-1} \alpha_{i,k-1}\, c_i\phi^{q-k} +$$

$$+ \sum_{i=1}^{m}\sum_{j=0}^{k-1} \alpha_{i,j}\, c_i\phi^{q-j} + \sum_{i=0}^{s-1} \alpha_{i,k}\, c_i\phi^{q-k} +$$

$$+ \sum_{i=s+1}^{m} \alpha_{i,k-1}\, c_i \phi^{q-k} + \sum_{i=1}^{s-1} \alpha_{i,k}\, c_i \phi^{q-(k+1)}$$

$$= \sum_{i=1}^{m} \sum_{j=0}^{k-1} \beta_{i,j}\, c_i \phi^{q-j} + \sum_{i=1}^{s-1} \beta_{i,k}\, c_i \phi^{q-k} +$$

$$+ \sum_{i=s+1}^{m} \alpha_{i,k-1}\, c_i \phi^{q-k} + \sum_{i=1}^{s-1} \alpha_{i,k}\, c_i \phi^{q-(k+1)} \qquad (3.3)$$

where $\beta_{i,j} = \alpha_{i,j-1} + \alpha_{i,j}$; $\alpha_{i,-1} \triangleq 0$.

This means that even the vector $c_s \phi^{q-(k+1)}$ is a linear combination of the same rows as $c_s \phi^{q-k}$ was (and the other rows already investigated). Therefore it cannot appear in T.

Hence if $c_s \phi^{q-k}$ has not appeared in T then $c_s \phi^{q-(k+1)}$ cannot appear there either (and naturally none of $c_s \phi^{q-(k+i)}$; $i \geq 1$, as it may be shown by induction).

We can see that the n independent rows can always be found grouped on the left hand side of the array (3.1) and thus for a system $\mathcal{S}(\phi, \Gamma, C, D)$ there always exists a V.D.E. where the structure of the A matrices is such that if a certain, say j:th, column in the A_i matrix is equal to zero, then also the j:th columns in all matrices A_{i+1}, A_{i+2}, ... are equal to zero.

A V.D.E. with A matrices of this form will be called the identifiable canonical form of V.D.E. The particular structure of V.D.E. in this canonical form can be uniquely characterized by m numbers.

$$p_1, p_2, \cdots, p_m \qquad (3.4)$$

meaning that the non-zero column corresponding to the j:th output y_j is contained only in the first p_j matrices $A_1, A_2, \cdots, A_{pj}$.
The order of the system is then:

$$n = \sum_{j=1}^{m} p_j$$

and the maximal p_j is equal to the observability index:

$$i_{obs} = \max_j \{p_j\}$$

4. IDENTIFICATION OF MULTIVARIABLE SYSTEMS.

4.1. Introduction.

Being aware of the existence of the identifiable canonical form of V.D.E. we only need to consider models in this form and proceed quite similarly as in the single output case.

We start the identification of an unknown system with a model with structure characterized by $p_1 = p_2 = \ldots = p_m = 1$ (with order $n = m$), i.e. we suppose that the output $y(t)$ depends, apart from input, on the $y(t-1)$ only, and determine all unknown elements in the matrices A_1, B_0, B_1, (call them Θp_1, p_2, ..., p_m) and the loss function $V_1(\Theta p_1, p_2, \ldots, p_m)$.

Then we set $p_1 = 2$, i.e. we include $y_1(t-2)$ in the model, determine the new Θ and V_2 and test to see if the decrease of the loss function $\Delta V_2 = V_1 - V_2$ was significant.

If not, the $y_1(t-2)$ brought no new information about the state of the system, and neither $y_1(t-2)$ nor its more delayed values $y_1(t-3)$, $y_1(t-4)$, ... need to be considered further. Hence p_1 is set back to $p_1 = 1$ definitively.

If ΔV_2 was significant, $y_1(t-2)$ is included in the model and the same shall be repeated for $y_2(t-2)$, $y_3(t-2)$, ..., $y_m(t-2)$, $y_2(t-3)$, ... until in all elements of output such delayed values $y_i(t-p_i)$ are found so that $y_i(t-p_i-1)$ brings no significant reduction of the loss function. Such a model is accepted as final, its order being:

$$n = \sum_{i=1}^{m} p_i$$

and observability index

$$i_{obs} = \max_i \{p_i\}$$

To find it we had to evaluate and test only (n+1) models.

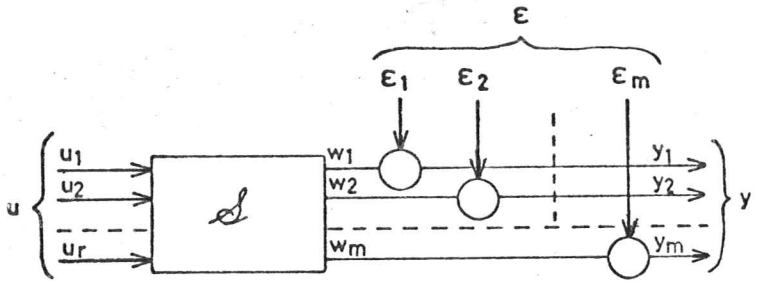
This algorithm is represented in Fig. 4.1.

Even though we reduced the number of different models to be examined to a quantity equal to the number of models for single output systems, the identification may still be rather time consuming when using the M.L. because of the greater number of unknown parameters and hence the necessity of the "hill climbing" in many dimensions [8], [9], [10]. (For example a system with 2 inputs, 2 outputs of 4th order could have 24 unknown elements in only the A,B matrices !!) The other disadvantage of M.L. is that the whole algorithm needs to be repeated for each model.

All this initiated the search for a more feasible method. The most suitable has proved to be the prior knowledge fitting [2], [3]. Because of the shortage of space only the principle ideas necessary for the extension of the P.K.F. to a multivariable case are given below. For the theoretical background we refer to the original papers [2], [3].

4.2. Application of P.K.F. to Multivariable Systems.

Consider a linear multivariable sampled system with r inputs and m outputs whose virtual outputs w are disturbed by random noises ε so that only the outputs affected by noise, y, are accessible for measurements.



Suppose that:

- the input signal {u(t); t = 1, 2, ..., L} is "persistently exciting" (see later),

- the virtual output w(t) can be related to the input u(t) by a V.D.E.

$$w(t) + \sum_{i=1}^{p} A_i\, w(t-i) = \sum_{i=0}^{p} B_i\, u(t-i) \qquad (4.1)$$

- the noise $\{\varepsilon(t);\ t = 1, 2, \ldots, L\}$ is some m dimensional ergodic random process with zero mean value (assumed for the sake of simplicity; it may be disregarded [3]) and finite covariance matrix,
- the actual output is given by:

$$y(t) = w(t) + \varepsilon(t) \qquad (4.2)$$

- finite length records of input-output signals are available

  $\{u(t), y(t);\ t = 1, 2, \ldots, L\}$

  or measurements of input and output may be done on-line.

All these assumptions are rather realistic for open loop systems.

Using (4.1) and (4.2) we get:

$$y(t) + \sum_{i=1}^{p} A_i\, y(t-i) = \sum_{i=0}^{p} B_i\, u(t-i) + \delta(t) \qquad (4.3)$$

where

$$\delta(t) = \varepsilon(t) + \sum_{i=1}^{p} A_i\, \varepsilon(t-i) \qquad (4.4)$$

is also an m dimensional ergodic process with zero mean value, and is independent of input signal u.

Thus if we express the conditional mean of $\delta(t)$ for given past values of input u(t), u(t-1), ..., u(t-N) using a linear regression model

$$E\{\delta(t)|u(t), u(t-1), \ldots, u(t-N)\} = \sum_{i=0}^{N} C_i\, u(t-i) \qquad (4.5)$$

then the $(m \times r)$ matrices of regression coefficients C_i must be zero for all i.

This fact is considered as the only "prior knowledge" available about the examined system.

In order to obtain estimates $\hat{A}_i$, $\hat{B}_i$ of matrices A_i, B_i in V.D.E. (4.3) according the principle of P.K.F. we must:

1. find unbiased estimates of regression matrices $\hat{C}_i$ for given input-output data as a function of $A_1, \ldots, A_p, B_0, \ldots, B_p$,
2. choose the estimates $\hat{A}_i$, $\hat{B}_i$ so that the estimates of regression coefficients are "as small as possible". More precisely so that:

$$V = \sum_{i=0}^{N} \|\hat{C}_i\|^2, \quad (\|C\| \text{ is suitable norm}) \qquad (4.6)$$

is minimal. So the V (4.6) is our loss function.

Let us carry this out in detail:

1. An unbiased estimate of C_i matrices can be obtained by the least squares method. Set

$$E\{\delta(t)|u(t), \ldots, u(t-N)\} = \delta(t) + e(t)$$

(where $e(t) \triangleq [e_1(t), \ldots, e_m(t)]$ is a vector of residuals) and substitute this and $\delta(t)$ from (4.3) into (4.5)

$$\sum_{i=0}^{N} C_i\, u(t-i) = \delta(t) + e(t)$$

$$= y(t) + \sum_{i=1}^{p} A_i\, y(t-i) - \sum_{i=0}^{p} B_i\, u(t-i) + e(t) \qquad (4.7)$$

Then the least squares estimates $\hat{C}_i$ are those minimizing the sum of squared residuals:

$$Q = \sum_{t=1}^{L} e^T(t)\, e(t) = \sum_{t=1}^{L} \sum_{s=1}^{m} e_s^2(t) = \sum_{s=1}^{m} \{ \sum_{t=1}^{L} e_s^2(t) \} \qquad (4.8)$$

Denoting by $c_{i,s}$, $a_{i,s}$, $b_{i,s}$ the s:th rows of matrices C_i, A_i, B_i respectively, then it follows for $e_s(t)$ from (4.7):

$$e_s(t) = \sum_{i=0}^{N} c_{i,s}\, u(t-i) - y_s(t) - \sum_{i=1}^{p} a_{i,s}\, y(t-i) +$$

$$+ \sum_{i=0}^{p} b_{i,s}\, u(t-i) \qquad (4.9)$$

We see that $e_s(t)$ depends only on those elements of matrices C_i, A_i, B_i, contained in their s:th rows. Hence the minimization of (4.8) can be done row by row in any arrangement.

Introducing:

$$\eta_s \triangleq \left[e_s(1) \mid e_s(2) \mid \ldots \mid e_s(L) \right]^T$$

$$\gamma_s \triangleq \left[c_{0,s} \mid c_{1,s} \mid \cdots \mid c_{N,s} \right]^T$$

$$\theta_s \triangleq \left[-a_{1,s} \mid -a_{2,s} \mid \cdots \mid -a_{p,s} \mid b_{0,s} \mid \cdot \mid b_{p,s} \right]^T$$

$$Z \triangleq \begin{bmatrix} u^T(1) & u^T(0) & & \\ u^T(2) & u^T(1) & & \\ \cdot\cdot & \cdot\cdot & & \\ u^T(L) & u^T(L-1) & \cdot\cdot & u^T(L-N) \end{bmatrix} \qquad (4.10)$$

$$X \triangleq \begin{bmatrix} y^T(0) & y^T(-1) & \cdot\cdot & u^T(1) & u^T(0) & \cdot\cdot \\ y^T(1) & y^T(0) & \cdot\cdot & u^T(2) & u^T(1) & \cdot\cdot \\ \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot \\ y^T(L-1) & y^T(L-2) & y^T(L-p-1) & u^T(L) & u^T(L-1) & u^T(L-p) \end{bmatrix}$$

then Q (4.8) can be written as:

$$Q = \sum_{s=1}^{m} \eta_s^T \eta_s \text{ where } \eta_s = Z \cdot \gamma_s + X\theta_s - Y_s \qquad (4.11)$$

i.e. the least squares solution

$$\hat{\gamma}_s = (Z^TZ)^{-1} Z^T(X\theta_s - Y_s) \qquad (4.11a)$$

(Note: The input signal must ensure that the matrix Z^TZ is regular. Such an input signal is called "persistently exciting".)

Notice that in (4.11) only the column vector Y_s depends on s. The other matrices Z, X are common for all s.

2. Having thusly expressed the $\hat{C}_i$ matrices as functions of A, B matrices, we must in step 2 choose the estimates $\hat{A}$, $\hat{B}$ so that the loss function V (4.6) is minimal. Using Euclidean norm we can write:

$$V = \sum_{i=0}^{N} \| \hat{C}_i \|^2 = \sum_{i=0}^{N} \sum_{s=1}^{m} \hat{c}_{i,s}^T \hat{c}_{i,s} = \sum_{s=1}^{m} \hat{\gamma}_s^T \hat{\gamma}_s \qquad (4.12)$$

Also here V can be minimized separately for each s = 1, 2, ..., m.

According to the assumed structure the matrices A_i may contain certain dummy columns so that in all rows θ_s the respective elements are equal to zero. Let us arrange the non-zero elements in θ_s (4.10) which are to be estimated into a new vector θ_s and arrange a matrix Ξ by letting out the respective columns in Z so that (4.11a) becomes:

$$\hat{\gamma}_s = (Z^TZ)^{-1} Z^T\Xi\, \theta_s - (Z^TZ)^{-1} Z^T Y_s \qquad (4.13)$$

The least squares solution is then:

$$\hat{\theta}_s = -\{\Xi^TZ(Z^TZ)^{-1}(Z^TZ)^{-1}Z^T\Xi\}^{-1} \cdot \Xi^TZ(Z^TZ)^{-1}(Z^TZ)^{-1}Z^T Y_s \qquad (4.14)$$

Notice that even here only the column vector Y_s depends on s, while the rest of the whole expression (4.14) is common for all s.

The impractical expressions (4.11a), (4.14) are not suitable for numerical computation - we use them just to express γ_s and θ_s in compact form. Actually both least squares solutions are obtained directly from (4.11) and (4.13) by their orthogonal transformation [12] to suitable form. Eq. (4.11) is transformed to the form:

$$\bar{\eta}_s = \bar{Z}\,\gamma_s + \bar{X}\,\theta_s - \bar{Y}_s \qquad (4.15)$$

where $\bar{Z}$ is an upper triangular matrix. All matrices in (4.15) are rather small, the greatest dimension is only (N+1) · r. The eq. (4.13) is treated similarly.

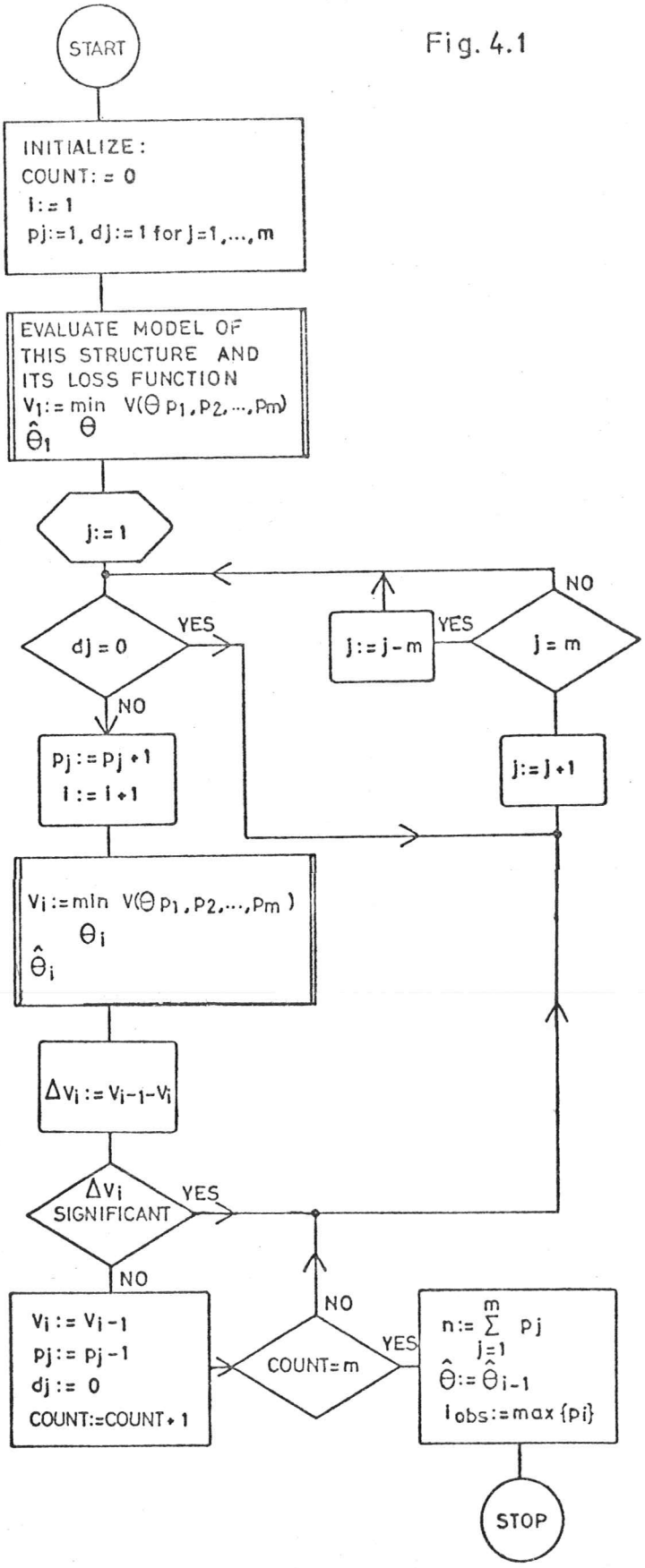
A very ingenious algorithm for this orthogonal transformation by means of elementary plane rotations is described in paper [3] and enables us to perform it even on-line.

So far, we can estimate $\hat{A}_1, \ldots, \hat{A}_p, \hat{B}_0, \ldots, \hat{B}_p$ and the corresponding minimal value of the loss function V for their assumed structures. Thus when identifying a system with an unknown structure the scheme (4.1) for determination of the order may be applied. Notice, that the laborious first stage in the estimation, i.e. the computation of the matrices $\bar{Z}$, $\bar{X}$, $\bar{Y}$ is common for all examined structures, and only the second stage needs to be repeated for each structure. The second stage is easily done because of the low dimensions of matrices occuring there.

The only difficulty is that the significance of the decrease of the loss function can be judged only subjectively (unless sufficient assumptions about the probability distribution of the noise ε are made).

Concerning the properties of the estimates, the proof of their strong consistence in the original paper [2] (with obvious trivial modifications) is valid here as well.

Fig. 4.1


START
INITIALIZE:
COUNT: = 0
i: = 1
pj:=1, dj:=1 for j=1,...,m
EVALUATE MODEL OF THIS STRUCTURE AND ITS LOSS FUNCTION
V1:= min V(θ p1, p2, ..., pm)
θ̂1 θ
j:= 1
dj = 0
YES
NO
pj:= pj + 1
i:= i + 1
j:= j-m
YES
j = m
NO
j:= j+1
Vi:= min V(θ p1, p2, ..., pm)
θi
θ̂i
ΔVi:= Vi-1-Vi
ΔVi SIGNIFICANT
YES
NO
Vi:= Vi-1
pj:= pj-1
dj:= 0
COUNT:=COUNT+1
COUNT=m
NO
YES
n:= Σ pj (j=1 to m)
θ̂:= θ̂i-1
iobs:=max{pi}
STOP

4.3. Description of the Subroutine MIMOID.

A FORTRAN IV subroutine called MIMOID (Multiple Input Multiple Output Identification) for identification of multivariable systems by P.K.F., based on ideas presented in previous sections, was written to get some insight into its properties from the practical point of view. The immediate aim was only the off-line use, but its modification for real-time applications is straight forward - only a few statements need to be changed. The algorithm described in section 4.2 was modified so that the mean value of the noise $\varepsilon(t)$ need not be zero nor assumed to be known in advance [3].

The least squares solutions of (4.11) and (4.13) are computed using the original Peterka and Šmuk's subroutine REDUCE [3](see Appendix B).

As it was mentioned previously, it is not possible to include an "objective" judging of the significance of the decrease of the loss function V into this subroutine. Hence in this version of MIMOID the structures of A matrices, which shall be estimated, must be specified by the numbers $p_1, \ldots, p_m$ (3.4) in the calling statement in the main program.

For each wanted structure, besides the estimates of A and B matrices, the respective minimal value of the loss function V (4.12) is calculated and printed out.

So the identification of a system with unknown structure can be done half-automatically, according to the flow chart (Fig. 4.1), in a man-machine loop.

The heading of the MIMOID and its dimensions part is as follows:

```
      SUBROUTINE MIMOID(U,IL,L,LU,Y,IM,M,MY,NN,LENGTH,NV,IP,NR,LP,EPS)
C
C     MULTIPLE INPUT - MULTIPLE OUTPUT LINEAR SYSTEM IDENTIFICATION
C     USING PETERKAS PRIOR KNOWLEDGE FITTING METHOD.  THE DESCRIPTION
C     OF THE SYSTEM IS ASSUMED IN THE FORM OF A VECTOR  DIFFERENCE
C     EQUATION WHICH MATRIX COEFFICIENTS ARE ESTIMATED FOR ALL THEIR
C     ASSUMED STRUCTURES
C     REFERENCE,VALIS, ON-LINE IDENTIFICATION OF MULTIVARIABLE LINEAR
C     SYSTEMS OF UNKNOWN STRUCTURE FROM INPUT/OUTPUT DATA,
C     IFAC SYMPOSIUM ON IDENTIFICATION PRAGUE 1970
C     AUTHOR,JAROSLAV VALIS 21/10-1969
C
C     U - ARRAY OF DIMENSION (IL,NN) CONTAINING ALL PHYSICAL INPUT
C         SIGNALS
C     IL - DIMENSION PARAMETER
C     L - NUMBER OF USED LOGICAL INPUT SIGNALS       (MAX 3)
C     LU - VECTOR OF DIMENSION ( 3).  (LU(I))-TH ROW OF ARRAY U SHALL BE
C          USED AS I-TH LOGICAL INPUT SIGNAL
C     Y - ARRAY OF DIMENSION (IM,NN) CONTAINING ALL PHYSICAL OUTPUT
C          SIGNALS
C     IM - DIMENSION PARAMETER
C     M - NUMBER OF USED LOGICAL OUTPUT SIGNALS      (MAX 3)
C     MY - VECTOR OF DIMENSION ( 3).  (MY(I))-TH ROW OF ARRAY Y SHALL BE
C          USED AS I-TH LOGICAL OUTPUT SIGNAL
C     MEAN VALUE OF I/O SIGNALS NEED NOT TO BE ZERO
C     NN - DIMENSION PARAMETER OF INPUT/OUTPUT ARRAYS U,RESP Y
C     LENGTH - LENGTH OF I/O DATA RECORD
C     NV - NUMBER OF INVESTIGATED VARIANTS OF STRUCTURES
C     IP - ARRAY OF DIMENSION (NV,IM) SPECIFYING  EXAMINED STRUCTURES
C          THE ABS VALUE OF IP(I,J) DETERMINES THE MAXIMAL TIME DELAY
C          OF J-TH OUTPUT IN I-TH STRUCTURE. IF IP(I,1) .LT. 0 THEN IN
C          THE I-TH STRUCTURE THE MATRIX B(0)=0,OTHERWISE B(0) .NE. 0
C     NR - NUMBER OF REGRESSION COEFFICIENTS. WHEN NR.LE.0 PROPER
C          VALUE SHALL BE DETERMINED AUTOMATICALLY BY PROGRAM
C          AND ITS VALUE WILL BE CHANGED DURING EXECUTION
C     LP - PRINT OUTS EVERY LP SAMPLES
C     EPS - ACCURACY CONST. - SEE REF
C
C     THE SUBROUTINE PRINTS OUT
C     1. NUMBER OF STRUCTURE
C     2. ASSUMED STATE SPACE DIMENSION (ORDER OF THE SYSTEM)
C        IF ASSUMED ORDER IS.GT. ACTUAL ORDER THE MESSAGE *THIS STRUCTU-
C        RE IS REDUNDANT* WILL BE PRINTED WHEN NOISE FREE I/O DATA WERE
C        USED
C     3. STRUCTURE OF THE TRANSFORMATION MATRIX (ROWS OF THE OBSERVABI-
C        LITY MATRIX CHOSEN TO OBTAIN THIS STRUCTURE FROM STATE SPACE
C        DESCRIPTION  S(A,B,C,D) )
C     4. ESTIMATES OF A-MATRICES
C     5. ESTIMATES OF B-MATRICES
C     6. SUMS OF SQUARES OF REGRESSIONS COEFFICIENTS FOR THE RESP. OUT-
C        PUTS AND THEIR TOTAL SUM
C        THIS MAY BE USED TO JUDGE THE ORDER OF THE SYSTEM
C
C     SUBROUTINE REQUIRED
C        REDUCE
C
```

For the correct use of MIMOID a more detailed explanation of its formal parameters must be added.

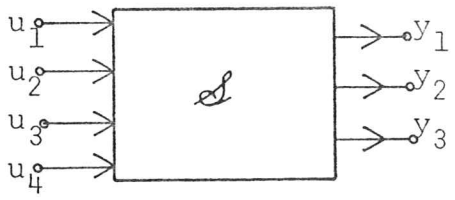


Fig. 4.2.

The records of all physical input and output signals of length LENGTH measured on "terminals" of the examined physical system (Fig. 4.2) must be stored in arrays U resp. Y in the following manner:

U =

	1	2	3		U_1(1) U_1(LENGTH) LENGTH		NN
1	$u_1(1)$	$u_1(2)$	$u_1(3)$	.....	u_1(LENGTH)	.X.	X
2	$u_2(1)$	$u_2(2)$	$u_2(3)$	.....	u_2(LENGTH)	.X.	X
3	$u_3(1)$	$u_3(2)$	$u_3(3)$	.....	u_3(LENGTH)	.X.	X
4	$u_4(1)$	$u_4(2)$	$u_4(3)$	.....	u_4(LENGTH)	.X.	X
.	X	X	X	X	X	X	X
IL	X	X	X	X	X	X	X

Y =

	1	2	3		LENGTH		NN
1	$y_1(1)$	$y_1(2)$	$y_1(3)$	.....	y_1(LENGTH)	.X.	X
2	$y_2(1)$	$y_2(2)$	$y_2(3)$	.....	y_2(LENGTH)	.X.	X
3	$y_3(1)$	$y_3(2)$	$y_3(3)$	.....	y_3(LENGTH)	.X.	X
.	X	X	X	X	X	X	X
IM	X	X	X	X	X	X	X

The subroutine enables us to determine models of dynamic responses between any L-tuple of physical inputs and any M-tuple of physical outputs. The inputs and outputs of such a model shall be referred to as "logical inputs" resp. "logical outputs". The logical inputs are numbered $u_1, u_2, \ldots, u_L$, the logical outputs are $y_1, y_2, \ldots, y_M$. Both L and M are limited to 3.

The assignment of a certain, say I-th, physical input to a certain, say J-th logical input is simply done by setting the J-th element of the array LU equal to I, i.e. LU(J) = I before the call of MIMOID.

The assignment of output is done similarly by the array MY.

One of the main advantages of the subroutine MIMOID is that more structures can be identified simultaneously during a single call of it. The number of wanted investigated structures is NV, the single structures then being numbered by $i_s = 1, 2, \ldots, NV$. The numbers $p_1, \ldots, p_m$ characterizing a certain, say i-th structure,must be stored in the i-th row of the array IP.

When we set the p_1 for a certain i-th structure, with a minus sign into IP (i,1), then the matrix B_0 in this structure is assumed to be zero and is not estimated.

The number of regression coefficients NR (= N in (4.5)) must be chosen so that the system of linear equation (4.13) is uniquely solvable in the least squares sense, i.e. the matrix in the { } brackets in (4.14) is regular.

If NR is assigned a negative value then it shall be automatically changed to a proper value during excention. The results of the estimation are computed (i.e. the second stage in the estimation is executed) and printed out every LP samples of input-output data. This may be useful for the check of the speed of the convergence or in on-line applications.

If the absolute value of a diagonal element of the transformed matrices in the least squares solutions (4.11) or (4.13) (for example $\bar{Z}$ in (4.15)) is less than EPS, the respective matrix is considered to be singular.

The following results are printed out:

1. Number of the structure
2. Assumed state space dimension
3. Structure of the transformation matrix T in (2.6)
4. Estimates of A matrices
5. Estimates of B matrices
6. The loss function V (4.6)

5. PRACTICAL RESULTS.

The subroutine MIMOID was used for identification of different test systems up to 6th order known from literature [8], [9], [13], from simulated input-output data. In all cases good results were obtained.

The following example may be considered as typical:

The continuous system [13]:

$$\dot{x} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -1 \end{bmatrix} x + \begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 0 & 1 \end{bmatrix} u$$

$$y = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} x \qquad (5.1)$$

driven by two 63 resp. 127 bits PRBS sequences with amplitudes ± 1, and sampled with T = 0.5s was simulated and both outputs were disturbed by pseudorandom N(0,0.1) sequences. 500 i/o data pairs were used. According to the flow chart (4.1) the following structures were examined:

i	p_1	p_2	n	$V_i \cdot 10^{-3}$	$\Delta V_i \cdot 10^{-3}$
1	1	1	2	18.55	-
2	2	1	3	1.22	17.33
3	2	2	4	1.13	0.09
4	3	1	4	0.63	0.59

Only the 2nd structure of 3rd order brought significant decrease of the loss function and hence this model should be accepted. The estimates are:

$$\hat{A}_1 = \begin{bmatrix} -0.832 & 0.002 \\ 0.036 & -0.597 \end{bmatrix}; \qquad \hat{A}_2 = \begin{bmatrix} 0.132 & 0 \\ -0.020 & 0 \end{bmatrix}$$

$$\hat{B}_1 = \begin{bmatrix} 0.398 & 0.515 \\ 0.396 & 0.394 \end{bmatrix}; \qquad \hat{B}_2 = \begin{bmatrix} -0.093 & -0.310 \\ -0.020 & 0.022 \end{bmatrix}$$

while the true values calculated directly from (5.1) are:

$$A_1 = \begin{bmatrix} -0.829 & 0.000 \\ 0.000 & -0.606 \end{bmatrix}; \qquad A_2 = \begin{bmatrix} 0.135 & 0 \\ 0.000 & 0 \end{bmatrix}$$

$$B_1 = \begin{bmatrix} 0.393 & 0.518 \\ 0.393 & 0.393 \end{bmatrix}; \qquad B_2 = \begin{bmatrix} -0.088 & -0.314 \\ 0.000 & 0.000 \end{bmatrix}$$

Total computing time 1.34 min on CDC 3600.

Results from another trial with this system (5.1) can be found in the Appendix A.

The later trials with industrial data showed that even if it was easy to estimate models for different structures, it was sometimes very difficult to judge whether the loss function decreased significantly. A typical example is the identification of dynamics of a distillation column from recorded input-output data, where the input was the reflux ratio perturbed by PRBS, the 1st output was the top product composition, the 2nd output the temperature on the top plate. This system was already investigated by different methods [14], [15]. According to [15] the dynamics of both systems can be described by models of 2nd order. Using P.K.F. we examined the following structures:

i	p_1	p_2	n	$V_i \cdot 10^{-3}$	$\Delta V_i \cdot 10^{-3}$
1	1	1	2	3.82	-
2	2	1	3	2.70	1.12
3	2	2	4	1.89	0.81
4	3	1	4	1.50	1.20

Total computing time 1.04 min on CDC 3600.

Here the significance of the decrease of the loss function cannot be judged directly from ΔV_i, and other facts (for example the roots, residuals) must be considered as well.

So we may choose the structure 1 of 2nd order (Table 5.1) or the structure 2 of 3rd order (Table 5.2). Both these models fit very close to the original measured data as it is shown on Figure 5.1 and Figure 5.2.

```
DATA FROM RAND 225 - FILES 13.14


TIME=  315


STRUCTURE  1

STATE SPACE DIMENSION  2

STRUCTURE OF TRANSFORMATION MATRIX

 C(1)*A**0
 C(2)*A**0


MATRIX A( 1)
    -0.08782817     -22.98165214
     0.00474181      -1.08914270


MATRIX B( 1)
     0.11905720
    -0.02940708


SUMS OF SQUARES OF REGRESSION COEFF.
SQR(1) =     3.811663-003
SQR(2) =     1.161621-005

TOTAL SUM     3.823280-003
```

Table 5.1 - Second order model of the distillation column.

```
STRUCTURE  2

STATE SPACE DIMENSION  3

STRUCTURE OF TRANSFORMATION MATRIX

 C(1)*A**0
 C(1)*A**1
 C(2)*A**1

MATRIX A( 1)
    -0.54768032     -13.21450403
     0.00315429      -1.09637449

MATRIX A( 2)
     0.08285040       0.00000000
     0.00208219       0.00000000

MATRIX B( 1)
     0.11186334
    -0.02937951

MATRIX B( 2)
    -0.34581735
     0.00045192

SUMS OF SQUARES OF REGRESSION COEFF.
SQR(1) =     2.692463-003
SQR(2) =     8.597067-006

TOTAL SUM     2.701060-003
```

Table 5.2 - Third order model of the distillation column.

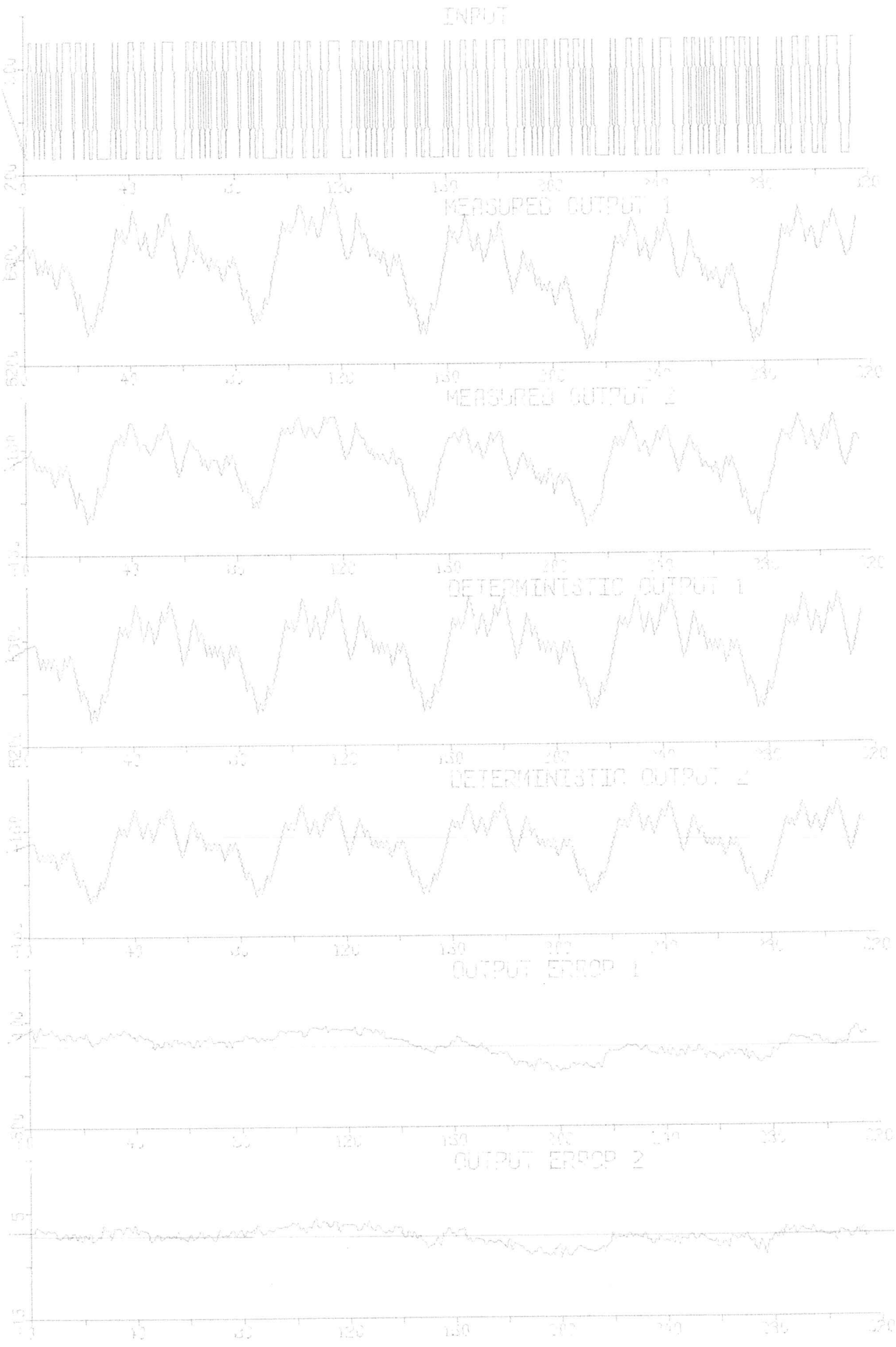


Figure 5.1 - Input-output behaviour of the 2nd order model of the distillation column.

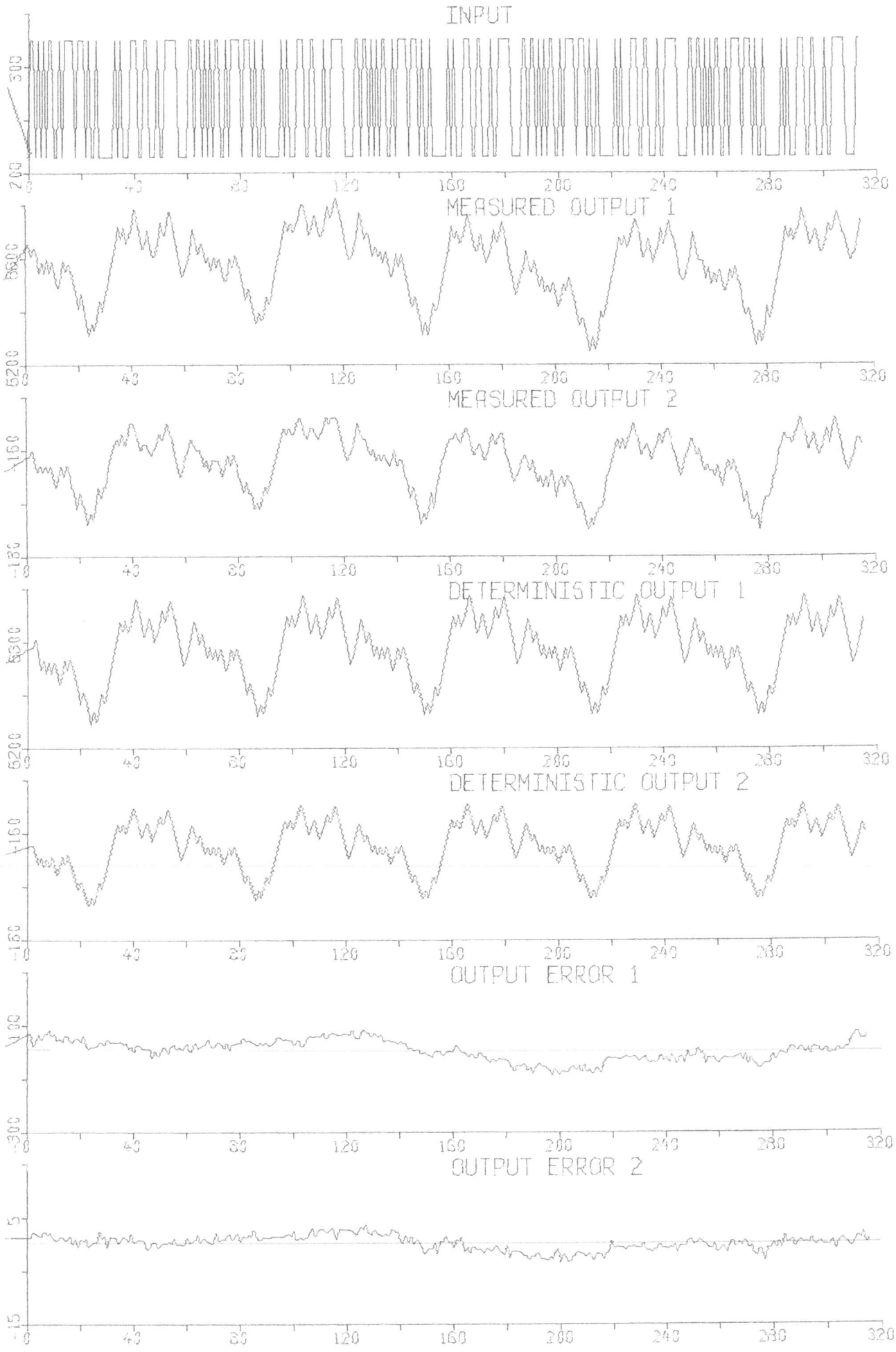


Figure 5.2 - Input-output behaviour of the 3rd order model of the distillation column.

6. ACKNOWLEDGEMENT.

This report was written during the author's stay at the Lund Institute of Technology. His work there in the Division of Automatic Control was made possible through the constant help of Professor Karl-Johan Åström and other members of the research staff.

The author is especially grateful to his friends:

- Ivar Gustavsson for continued cooperation and advice,
- Vaclav Peterka and Karel Šmuk for their generousity in granting permission to reproduce their original text in Appendix B,
- Patricia Coulburn for reading and correcting the manuscript.

7. REFERENCES.

[1] Åström, K.J., Bohlin, T.: Numerical Identification of Linear Dynamic Systems from Normal Operatic Records, IFAC Symposium - Theory of Self-Adaptive Systems, Teddington, 1965.

[2] Peterka, V.: New Approach to the Identification of Discrete Dynamic Systems (in Czech), Kybernetika No. 5, Vol. 4/ /1968, Academia Praha.

[3] Peterka, V., Smuk, K.: On-Line Estimation of Dynamic Model Parameters from Input Output Data, IV. IFAC Congress, Warsaw, 1969.

[4] Bohlin, T., Wensmark, S.: Process Modelling and Control Design Package, IBM Contributed Program Library 360D, August 1, 1967, IBM Nordic Laboratory, Lidingö 9, Sweden.

[5] Gustavsson, I.: Parametric Identification on Multiple Input, Single Output Linear Dynamic Systems, Report 6907, July, 1969, Lund Institute of Technology, Division of Automatic Control.

[6] Gustavsson, I.: Comparison of Different Methods for Identification of Linear Models for Industrial Processes, IFAC, 2nd Prague Symposium on Identification and Process Parameter Estimation, Prague, June, 1970.

[7] Jezek, J.: Prevody prenosové funkce lineární soustavy ze spojitého tvaru na diskretní a obrácene, Ph.D. thesis UTIA CSAV, Praha, 1969.

[8] Rowe, I.H.: Statistical Methods for the Identification and Control of Multivariate Stochastic Systems, Ph.D. thesis, Centre for Computing and Automation, Imperial College of Science and Technology, University of London, 1967.

[9] Tzafestas, S.G.: Estimation of Parameters of Multi-Input and Multi-Output Linear Dynamic Systems, Department of Electrical Engineering, Imperial College of Science and Technology, University of London, 1966.

[10] Eaton, J.H.: Identification for Control Purposes, Research paper, IBM Research Laboratory, San José, California, 1969.

[11] Vališ, J.: On-Line Identification of Multivariable Linear Systems of Unknown Structure from Input Output Data, IFAC Symposium on Identification and Process Parameter Estimation, Prague, 1970.

[12] Faddeev, D.K., Faddeeva, V.N.: Numerical Methods of Linear Algebra (in Russian), Moscow, 1960.

[13] Rosenbrock, H.H.: On the Design of Linear Multivariable Control Systems, Proceedings of the 3rd IFAC Congress, London, 1966.

[14] Hammond, P.M., Barber, D.L.A.: The Use of an On-Line Digital Computer to Evaluate the Dynamic Response of a Pilot Scale Distillation Column, Trans. of the Society of Instrument Technology, September, 1965.

[15] Gustavsson, I.: Identification of Dynamics of a Distillation Column, Report 6916, June, 1969, Lund Institute of Technology, Division of Automatic Control.

APPENDIX A

This appendix contains an example of the use of MIMOID for the identification of a 3rd order test system with two inputs and two outputs like in (5.1).

Noise free input output data are stored on magnetic tape (band 10, file 33).

During the first three calls of MIMOID the specified structured are identified. Then both output signals are disturbed by a pseudorandom gaussian noise with zero mean and stand. deviation 0.1 (RANSS) and the same calls of MIMOID as above are repeated.

The respective results are listed in the order they were printed by the computer.

```
      PROGRAM  MVPKF
C
C     THIS IS AN EXAMPLE HOW TO USE THE SUBROUTINE MIMOID
      DIMENSION U(2,500),Y(2,500),IP(4,2),LU(3),MY(3)
      PRINT 99
   99 FORMAT(///* SIMULATED IDENTIFICATION OF A 3RD ORDER TEST SYSTEM*)
      PRINT 100
  100 FORMAT(/* NOISE FREE DATA*)
      IREP=1
C
C     READ INPUT/OUTPUT DATA FROM BAND 10 FILE 33
      DO 1 I=1,32
    1 CALL SKIPFILE(10)
      READ TAPE 10,(U(1,I),U(2,I),Y(1,I),Y(2,I),I=1,500)
C
C     ASSIGNE PHYSICAL INPUTS
      LU(1)=1
      LU(2)=2
C
C     NOW THE RESPONSE OF THE OUTPUT Y(1,.) ON BOTH INPUTS IS IDENTIFIED
      MY(1)=1
      EPS=1.E-8
C
C     NR SHALL BE DETERMINED AUTOMATICALLY
    2 NR=-1
C
C     FOLLOWING STRUCTURES ARE EXAMINED
      IP(1,1)=1
      IP(2,1)=-1
      IP(3,1)=-2
      IP(4,1)=-3
      CALL MIMOID(U,2,2,LU,Y,2,1,MY,500,500,4,IP,NR,500,EPS)
C
C     NOW THE RESPONSE OF THE OUTPUT Y(2,.) ON BOTH INPUTS IS IDENTIFIED
      MY(1)=2
      NR=-1
      CALL MIMOID(U,2,2,LU,Y,2,1,MY,500,500,4,IP,NR,500,EPS)
C
C     NOW THE RESPONSES OF BOTH OUTPUTS ARE IDENTIFIED
      MY(1)=1
      MY(2)=2
      NR=-1
C
C     THE FOLLOWING STRUCTURES ARE EXAMINED
      IP(1,1)=-1
      IP(1,2)=1
      IP(2,1)=-2
      IP(2,2)=1
      IP(3,1)=-2
      IP(3,2)=2
      IP(4,1)=-3
      IP(4,2)=1
      CALL MIMOID(U,2,2,LU,Y,2,2,MY,500,500,4,IP,NR,500,EPS)
      IF(IREP) 5,5,4
    4 IT=3
      DO 3 I=1,2
```

```
      DO 3 K=1,500
      CALL RANSS(IT,XX)
    3 Y(I,K)=Y(I,K)+0.1*XX
      IREP =-1
      PRINT 101
  101 FORMAT(///* WITH ADDITIVE NOISE 0.1*)
      GO TO 2
    5 CALL EXIT
      END
```

```
SIMULATED IDENTIFICATION OF A 3RD ORDER TEST SYSTEM

NOISE FREE DATA


TIME=  500


STRUCTURE  1

STATE SPACE DIMENSION  1

STRUCTURE OF TRANSFORMATION MATRIX

 C(1)*A**0

MATRIX A( 1)
    -0.40107225

MATRIX B( 0)
    -0.00004307      -0.00000923

MATRIX B( 1)
     0.39342870       0.51791949

SUMS OF SQUARES OF REGRESSION COEFF.
SQR(1) =     1.925549-002

TOTAL SUM     1.925549-002
```

STRUCTURE 2

STATE SPACE DIMENSION 1

STRUCTURE OF TRANSFORMATION MATRIX

C(1)*A**0

MATRIX A(1)
-0.40107227

MATRIX B(1)
0.39342870 0.51791949

SUMS OF SQUARES OF REGRESSION COEFF.

SQR(1) = 1.925549-002

TOTAL SUM 1.925549-002

```
STRUCTURE  3

STATE SPACE DIMENSION  2

STRUCTURE OF TRANSFORMATION MATRIX

 C(1)*A**0
 C(1)*A**1


MATRIX A( 1)
    -0.82966082


MATRIX A( 2)
     0.13533528


MATRIX B( 1)
     0.39346934       0.51791323


MATRIX B( 2)
    -0.08779488      -0.31413025

SUMS OF SQUARES OF REGRESSION COEFF.
SQR(1) =     1.142618-021

TOTAL SUM     1.142618-021
```

```
STRUCTURE  4

STATE SPACE DIMENSION  3

STRUCTURE OF TRANSFORMATION MATRIX

 C(1)*A**0
 C(1)*A**1
 C(1)*A**2

THIS STRUCTURE IS REDUNDANT
```

```
STRUCTURE  1

STATE SPACE DIMENSION  1

STRUCTURE OF TRANSFORMATION MATRIX

 C(1)*A**0

MATRIX A( 1)
    -0.60653066

MATRIX B( 0)
     0.00000000       0.00000000

MATRIX B( 1)
     0.39346934       0.39346934

SUMS OF SQUARES OF REGRESSION COEFF.
SQR(1) =     1.140840-021

TOTAL SUM     1.140840-021
```

```
STRUCTURE  2

STATE SPACE DIMENSION  1

STRUCTURE OF TRANSFORMATION MATRIX

 C(1)*A**0

MATRIX A( 1)
    -0.60653066

MATRIX B( 1)
     0.39346934        0.39346934

SUMS OF SQUARES OF REGRESSION COEFF.
SQR(1) =     1.149607-021

TOTAL SUM     1.149607-021



STRUCTURE  3

STATE SPACE DIMENSION  2

STRUCTURE OF TRANSFORMATION MATRIX


 C(1)*A**0
 C(1)*A**1

THIS STRUCTURE IS REDUNDANT
```

STRUCTURE 4

STATE SPACE DIMENSION 3

STRUCTURE OF TRANSFORMATION MATRIX

C(1)*A**0
C(1)*A**1
C(1)*A**2

THIS STRUCTURE IS REDUNDANT

```
TIME=  500


STRUCTURE  1

STATE SPACE DIMENSION  2

STRUCTURE OF TRANSFORMATION MATRIX

 C(1)*A**0
 C(2)*A**0


MATRIX A( 1)
    -0.25472166      -0.16057777
    -0.00000000      -0.60653066


MATRIX B( 1)
     0.39345501       0.51792513
     0.39346934       0.39346934


SUMS OF SQUARES OF REGRESSION COEFF.
SQR(1) =     1.792346-002
SQR(2) =     8.223483-022

TOTAL SUM     1.792346-002
```

STRUCTURE 2

STATE SPACE DIMENSION 3

STRUCTURE OF TRANSFORMATION MATRIX

C(1)*A**0
C(1)*A**1
C(2)*A**1

MATRIX A(1)

-0.82966082	0.00000000
-0.00000000	-0.60653066

MATRIX A(2)

0.13533528	0.00000000
0.00000000	0.00000000

MATRIX B(1)

0.39346934	0.51791323
0.39346934	0.39346934

MATRIX B(2)

-0.08779488	-0.31413025
-0.00000000	-0.00000000

SUMS OF SQUARES OF REGRESSION COEFF.
SQR(1) = 1.016572-021
SQR(2) = 6.662314-022

TOTAL SUM 1.682804-021

STRUCTURE 3

STATE SPACE DIMENSION 4

STRUCTURE OF TRANSFORMATION MATRIX

C(1)*A**0
C(2)*A**0
C(1)*A**1
C(2)*A**1

THIS STRUCTURE IS REDUNDANT

STRUCTURE 4

STATE SPACE DIMENSION 4

STRUCTURE OF TRANSFORMATION MATRIX

C(1)*A**0
C(1)*A**1
C(1)*A**2
C(2)*A**2

THIS STRUCTURE IS REDUNDANT

WITH ADDITIVE NOISE 0.1

TIME= 500

STRUCTURE 1

STATE SPACE DIMENSION 1

STRUCTURE OF TRANSFORMATION MATRIX

C(1)*A**0

MATRIX A(1)
-0.40075945

MATRIX B(0)
0.00090549 -0.00040993

MATRIX B(1)
0.39425667 0.51768451

SUMS OF SQUARES OF REGRESSION COEFF.
SQR(1) = 1.839026-002

TOTAL SUM 1.839026-002

STRUCTURE 2

STATE SPACE DIMENSION 1

STRUCTURE OF TRANSFORMATION MATRIX

C(1)*A**0

MATRIX A(1)
-0.40075582

MATRIX B(1)
0.39425667 0.51768450

SUMS OF SQUARES OF REGRESSION COEFF.
SQR(1) = 1.839125-002

TOTAL SUM 1.839125-002

```
STRUCTURE  3

STATE SPACE DIMENSION  2

STRUCTURE OF TRANSFORMATION MATRIX

 C(1)*A**0
 C(1)*A**1

MATRIX A( 1)
    -0.86244777

MATRIX A( 2)
     0.15303719

MATRIX B( 1)
     0.39390641       0.51815336

MATRIX B( 2)
    -0.10500181      -0.32667125

SUMS OF SQUARES OF REGRESSION COEFF.
SQR(1) =     3.095060-004

TOTAL SUM     3.095060-004
```

```
STRUCTURE  4

STATE SPACE DIMENSION  3

STRUCTURE OF TRANSFORMATION MATRIX

 C(1)*A**0
 C(1)*A**1
 C(1)*A**2


MATRIX A( 1)
    -0.55366083


MATRIX A( 2)
    -0.06103810


MATRIX A( 3)
     0.01913163


MATRIX B( 1)
     0.39469234       0.51856049


MATRIX B( 2)
     0.01666739      -0.16683951


MATRIX B( 3)
    -0.00970799      -0.07662329


SUMS OF SQUARES OF REGRESSION COEFF.
SQR(1) =     1.843930-004

TOTAL SUM     1.843930-004
```

```
TIME=  500


STRUCTURE  1

STATE SPACE DIMENSION  1

STRUCTURE OF TRANSFORMATION MATRIX

 C(1)*A**0

MATRIX A( 1)
    -0.61106906

MATRIX B( 0)
     0.00207782       0.00166754

MATRIX B( 1)
     0.38718823       0.39355661

SUMS OF SQUARES OF REGRESSION COEFF.
SQR(1) =     3.240118-004

TOTAL SUM     3.240118-004
```

```
STRUCTURE  2

STATE SPACE DIMENSION  1

STRUCTURE OF TRANSFORMATION MATRIX

 C(1)*A**0

MATRIX A( 1)
    -0.61106586

MATRIX B( 1)
     0.38718824       0.39355661

SUMS OF SQUARES OF REGRESSION COEFF.
SQR(1) =     3.311099-004

TOTAL SUM     3.311099-004
```

```
STRUCTURE  3

STATE SPACE DIMENSION  2

STRUCTURE OF TRANSFORMATION MATRIX

 C(1)*A**0
 C(1)*A**1

MATRIX A( 1)
    -0.27112734

MATRIX A( 2)
    -0.20686383

MATRIX B( 1)
     0.38799558       0.39420158

MATRIX B( 2)
     0.13015831       0.13686085

SUMS OF SQUARES OF REGRESSION COEFF.
SQR(1) =    2.770300-004

TOTAL SUM    2.770300-004
```

STRUCTURE 4

STATE SPACE DIMENSION 3

STRUCTURE OF TRANSFORMATION MATRIX

C(1)*A**0
C(1)*A**1
C(1)*A**2

MATRIX A(1)
0.03563174

MATRIX A(2)
-0.02227477

MATRIX A(3)
-0.22748183

MATRIX B(1)
0.39126073 0.39670985

MATRIX B(2)
0.24987000 0.25828546

MATRIX B(3)
0.13842297 0.15307247

SUMS OF SQUARES OF REGRESSION COEFF.
SQR(1) = 1.734672-004

TOTAL SUM 1.734672-004

```
TIME=  500


STRUCTURE  1

STATE SPACE DIMENSION  2

STRUCTURE OF TRANSFORMATION MATRIX

 C(1)*A**0
 C(2)*A**0


MATRIX A( 1)
    -0.27598993      -0.13676893
    -0.00878592      -0.60250553



MATRIX B( 1)
     0.39379035       0.51720739
     0.38721821       0.39358961


SUMS OF SQUARES OF REGRESSION COEFF.
SQR(1) =     1.737165-002
SQR(2) =     3.266159-004

TOTAL SUM     1.769826-002
```

STRUCTURE 2

STATE SPACE DIMENSION 3

STRUCTURE OF TRANSFORMATION MATRIX

C(1)*A**0
C(1)*A**1
C(2)*A**1

MATRIX A(1)

-0.85034449	-0.05502109
0.01396729	-0.59003946

MATRIX A(2)

0.17897112	0.00000000
-0.01699193	0.00000000

MATRIX B(1)

0.39362723	0.51791187
0.38731336	0.39363610

MATRIX B(2)

-0.12163082	-0.34218151
0.01229603	0.01839621

SUMS OF SQUARES OF REGRESSION COEFF.
SQR(1) = 2.511449-004
SQR(2) = 3.061429-004

TOTAL SUM 5.572877-004

STRUCTURE 3

STATE SPACE DIMENSION 4

STRUCTURE OF TRANSFORMATION MATRIX

C(1)*A**0
C(2)*A**0
C(1)*A**1
C(2)*A**1

MATRIX A(1)

-0.85209869	-0.11597014
0.02440594	-0.22735123

MATRIX A(2)

0.17936203	0.03776766
-0.01931807	-0.22474324

MATRIX B(1)

0.39348238	0.51779779
0.38817530	0.39431500

MATRIX B(2)

-0.14590587	-0.36711471
0.15674903	0.16676570

SUMS OF SQUARES OF REGRESSION COEFF.
SQR(1) = 2.497945-004
SQR(2) = 2.583243-004

TOTAL SUM 5.081187-004

STRUCTURE 4

STATE SPACE DIMENSION 4

STRUCTURE OF TRANSFORMATION MATRIX

C(1)*A**0
C(1)*A**1
C(1)*A**2
C(2)*A**2

MATRIX A(1)

-0.50899114	-0.03074503
-0.36866782	-0.64763744

MATRIX A(2)

-0.09581023	0.00000000
0.27863583	0.00000000

MATRIX A(3)

0.03501126	0.00000000
-0.02322775	0.00000000

MATRIX B(1)

0.39475906	0.51861381
0.38604292	0.39289776

MATRIX B(2)

0.02224830	-0.15591131
-0.16087551	-0.20244776

MATRIX B(3)

-0.02022329	-0.09678035
0.00943659	0.09522862

SUMS OF SQUARES OF REGRESSION COEFF.
SQR(1) = 1.784939-004
SQR(2) = 1.545938-004

TOTAL SUM 3.330877-004

APPENDIX B

This appendix is the reprint of section 3 from the reference [3]:

V. Peterka and K. Šmuk:
ON-LINE ESTIMATION OF DYNAMIC MODEL PARAMETERS FROM INPUT OUTPUT DATA
IV. IFAC Congress, Warsaw, 1969.

Algorithm for the successive regression analysis with growing data and limited memory.

Let us consider the classical case of a least square linear regression, and the system of equations

$$\sum_{j=1}^{N} r_j \, y_{ij} = x_i + e_i \qquad (i = 1,2,\ldots L) \tag{6}$$

where $L > N$, x_i and y_{ij} are values obtained by observation, and e is an unknown random error. The latter can be interpreted as the deviation from the conditional expectation, i.e.

$$e_i = E\left[x/y_{i1}, y_{i2}, \ldots y_{iN}\right] - x_i \tag{7}$$

Now the task is to find the estimates $\hat{r}_j$ of regression coefficients r_j

minimising the sum of the squares of errors

$$Q = \sum_{i=1}^{L} e_i^2 \tag{8}$$

Relations (6) and (8) can be written in matrix form as follows

$$Y_{[L\times N]}\ r_{[N\times 1]} - x_{[L\times 1]} = e_{[L\times 1]} \tag{9}$$

$$Q = e^T_{[1\times L]}\ e_{[L\times 1]} \tag{10}$$

The subscript in the brackets indicates the dimensions of the respective matrix, whereas superscript T denotes transposition.

The classical solution of the given problem is (see e.g.[13])

$$\hat{r}_L = (Y^T Y)^{-1}\ Y^T x \tag{11}$$

provided that

$$\det\left[Y^T Y\right] \neq 0 \tag{12}$$

The numerical expression of formula (11) by ordinary matrix calculus (i.e. multiplication and inversion of matrices) is cumbersome when the number L of observations is large (e.g. hundreds or thousands); it requires a large capacity of memory and is numerically less stable[14]. Further on we shall derive an algorithm permitting the solution of the problem with substantially reduced demands on the memory of the computer used. This algorithm is based on orthogonal transformations of the system of linear equations[14,15,16] and is numerically very stable.

Let us arrange the system of equations (9) into the form

$$Z_{[L\times(N+1)]}\ \tilde{r}_{[(N+1)\times 1]} = e_{[L\times 1]} \tag{13}$$

where

$$Z = \left[Y_1 - x\right]_1 \qquad \tilde{r} = \begin{bmatrix} r \\ 1 \end{bmatrix} \tag{14}$$

and multiply from the left by square matrix T

By denoting

$$T_{[L\times L]}\ Z_{[L\times(N+1)]} = \tilde{Z}_{[L\times(N+1)]} \tag{15}$$

$$T_{[L\times L]}\ e_{[L\times 1]} = \tilde{e}_{[L\times 1]} \tag{16}$$

after this transformation we have (instead of (13))

$$\tilde{Z}_{[L\times(N+1)]}\ \tilde{r}_{[(N+1)\times 1]} = \tilde{e}_{[L\times 1]} \tag{17}$$

The sum of squares of the right sides of the transformed system of equations (17) will then be

$$\tilde{Q} = \tilde{e}^T \tilde{e} = e^T T^T T e \tag{18}$$

It is obvious that the sum of squares has not been changed by this transformation, i.e. $Q = Q$, provided that matrix T is orthogonal

$$T^T T = I \tag{19}$$

Let us further consider a special type of this orthogonal transformation where matrix T is a so called elementary matrix of rotation [15]. By representing only the non-zero elements this matrix can be written as follows:

$$T_{ij[L\times L]} = \begin{bmatrix} 1 & & & & & & & & \\ & 1 & & & & & & & \\ & & \ddots & & & & & & \\ & & & 1 & & & & & \\ & & & & C & & S & & i \\ & & & & & 1 & & & \\ & & & & & & \ddots & & \\ & & & & & & & 1 & \\ & & & & -S & & C & & j \\ & & & & & & & & 1 \end{bmatrix} \tag{20}$$

It will be easy to discover that matrix (20) will be orthogonal, and the sum of squares Q will not change, provided that condition is fulfilled that

$$C^2 + S^2 = 1 \tag{21}$$

It is obvious that the transformation of the system of equations (13) by multiplication by matrix (20) applies only to the i-th and j-th equations, in other words, only the i-th and j-th rows will change in matrices Z and e.

For $k \neq i$ $k \neq j$

$$\tilde{z}_{k\nu} = z_{k\nu}\ (\nu = 1, 2, \ldots N+1),\quad \tilde{e}_k = e_k$$

for k = i

$$\tilde{z}_{i\nu} = cz_{i\nu} + sz_{j\nu} \;, \quad \tilde{e}_i = ce_i + se_j \tag{22}$$

for k = j

$$\tilde{z}_{j\nu} = -sz_{i\nu} + cz_{j\nu} \;, \quad \tilde{e}_j = -se_i + ce_j \tag{23}$$

Coefficients c and s are bound by the condition of orthogonality (21), however, one of them can be selected. Let the selection make so that it holds that

$$\tilde{z}_{j\mu} = 0 \;.$$

i.e. according to (23) we obtain

$$-sz_{i\mu} + cz_{j\mu} = 0 \tag{24}$$

From (24) and (21) it follows that

$$c = \frac{z_{i\mu}}{\sqrt{z_{i\mu}^2 + z_{j\mu}^2}} \;, \quad s = \frac{z_{j\mu}}{\sqrt{z_{i\mu}^2 + z_{j\mu}^2}} \tag{25}$$

In this way we can annul any element in matrix Z without changing the sum of squares Q for any arbitrarily selected vector r. By the successive application of this transformation in a suitable sequence the original system of equations (13) can be arranged into the form

$$\begin{bmatrix} z_{11}^* & z_{12}^* & \cdots & z_{1N}^* & z_{1,N+1}^* \\ & z_{22}^* & \cdots & z_{2N}^* & z_{2,N+1}^* \\ & & & \vdots & \vdots \\ & & & z_{NN}^* & z_{N,N+1}^* \\ & & & & z_{N+1,N+1}^* \\ & & & 0 & \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_N \\ 1 \end{bmatrix} = \begin{bmatrix} e_1^* \\ e_2^* \\ \vdots \\ e_N^* \\ e_{N+1}^* \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{26}$$

while it still holds that

$$Q = \sum_{i=1}^{L} e_i^2 = \sum_{i=1}^{N+1} e_i^{*2} \tag{27}$$

for any arbitrarily selected r_j (j = 1,2,...N).

6

After this transformation the determination of estimates r ($j = 1, 2,...N$) minimising (27) becomes a simple matter. From equation (26) it is obvious that by selection of r ($j = 1,2,...N$) it is impossible to influence the last non-zero element of the right side $e^*_{N+1} = z^*_{N+1,N+1}$, however, it is possible to annul all the other elements e^*_j ($j = 1,2,...N$). The estimates of regression coefficients $\hat{r}$ ($j = 1,2,...N$) can thus be found by the solution of the system of linear equations

$$Z^*_{[N\times N]}\ \hat{r}_{[N\times 1]} + Z_{[N\times 1]} = 0_{[N\times 1]} \tag{28}$$

where

$$Z^*_{[N\times N]} = \begin{bmatrix} Z^*_{11} & Z^*_{12} & \cdots & Z^*_{1N} \\ & Z_{22} & \cdots & Z^*_{2N} \\ & & & \vdots \\ 0 & & & Z^*_{NN} \end{bmatrix}, \qquad Z^*_{[N\times 1]} = \begin{bmatrix} Z^*_{1,N+1} \\ Z^*_{2,N+1} \\ \vdots \\ Z_{N,N+1} \end{bmatrix} \tag{29,30}$$

At the same time we obtain the minimum of the sum of squares

$$Q_{min} = z^{*2}_{N+1,N+1} \tag{31}$$

Since matrix Z^* (29) is a triangular one, the solution of the system of equations (29) is very simple :

$$\begin{aligned} \hat{r}_N &= \frac{Z^*_{N,N+1}}{Z^*_{N,N}} \\ \hat{r}_{N-1} &= \frac{1}{Z^*_{N-1,N-1}}\left(Z^*_{N-1,N+1} + \hat{r}_N\, Z^*_{N-1,N}\right) \\ \hat{r}_{N-k} &= \frac{1}{Z^*_{N-k,N-k}}\left(Z^*_{N-k,N+1} + \sum_{i=0}^{k-1} \hat{r}_{N-i}\, Z^*_{N-k,N-i}\right) \end{aligned} \tag{32}$$

Let us consider now the situation where the data from L observations have been reduced into a triangular matrix with elements z^*_{ij} ($i = 1, 2,...N+1$; $j = i, i+1,...N+1$), and new data are obtained forming a futher L+1 row in the matrix on the left side of equation (26). All elements in this new row can be annuled by the gradual application of the described transformation, and this simultaneously means the correction of the upper triangular matrix. This step is more accurately described by the following procedure in ALGOL-60:

```
procedure REDUCE (matrix) new data:(row) order:(N);
 value N, row; array matrix, row; integer N;
 begin real c, s, de; integer i, k;
   for i:= 1 step 1 until N+1 do
    if row[i] ≠ 0 then
    begin de:= sqrt(row[i] ↑2 + matrix[i,i] ↑2);
          c:= matrix[i,i]/de; s:= row[i]/de;
     for k:= i step 1 until N+1 do
      begin de:= c x row[k] - s x matrix[i,k];
        matrix[i,k]:= s x row[k] + c x matrix[i,k];
        row[k]:= de
      end
    end
 end REDUCE;
```

The successive application of this procedure permits the processing of growing data without the necessity to memorize them.All necessary information on observed past history accumulates in the triangular matrix.

Let us now consider how this unified algorithm could also be used for processing the first N+1 rows of data in matrix Z (14). It is obvious that the sum of squares (10) will not change, if the matrix in (13) is extended by a zero matrix of dimensions (N+1)x(N+1)

$$\begin{bmatrix} 0 \\ Z \end{bmatrix}_{[(L+N+1)\times(N+1)]} \cdot \underset{[(N+1)\times 1]}{\tilde{r}} = \begin{bmatrix} 0 \\ e \end{bmatrix}_{[(L+N+1)\times 1]}$$

This zero matrix (more exactly, its upper triangular portion) can thus be considered as the initial state of matrix Z^*, when it does not contain yet any information on the process. This approach permits the use of the unified algorithm for all the processed data (i= 1,2,...L).

Triangular matrix (29) is non-singular, if reduction was applied to N linearly indipendent data rows at least. Beginning from this instant equations (32) can be used for computing in any arbitrary step estimates $\hat{r}$ (j = 1,2,...N) which are optimal in the sense of least squares for the whole past history.