On the Convergence Properties of the Generalized Least Squares Identification Method

Söderström, Torsten

1972

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

Link to publication

Citation for published version (APA):
Söderström, T. (1972). *On the Convergence Properties of the Generalized Least Squares Identification Method*. (Research Reports TFRT-3048). Department of Automatic Control, Lund Institute of Technology (LTH).

Total number of authors:
1

3048

REPORT 7228
NOVEMBER 1972

On the Convergence Properties of the Generalized Least Squares Identification Method

TORSTEN SÖDERSTRÖM

LTH

Division of Automatic Control · Lund Institute of Technology

ON THE CONVERGENCE PROPERTIES OF THE GENERALIZED LEAST SQUARES IDENTIFICATION METHOD

T. Söderström

ABSTRACT.

Modelling of a discrete time system is often made by parametric identification. A linear difference equation is adapted to the dynamics of the system. The parameters of the equation can easily be estimated by the least squares method. This method has several advantages, but if the residuals are correlated, the estimates are biased. The method of generalized least squares proposed by Clarke is constructed to overcome this difficulty. This method is an iterative procedure. The dynamics of the system and the correlation of the residuals are estimated alternately.

The purpose of this report is to present an analysis of the convergence properties of the generalized least squares method. Two different variants are examined. They correspond to different ways of estimating the correlation of the residuals. It is shown that one of those variants is equivalent to a maximization of the likelihood function of the problem, when suitable assumptions are made. In this case the possible result of the method is closely related to the number of local minimum points of a corresponding loss function. Under the assumption of suitable regularity conditions of the input signal and the system dynamics the following is theoretically shown in the report.

For every given system the minimization gives the true values of the parameters if the signal to noise ratio is high enough. It is further shown that the minimization may give wrong values of the parameters if the signal to noise ratio is low enough. In this case the loss function has no unique local minimum point.

The second variant is the one proposed by Clarke. By counterexamples it is shown that also this variant may give wrong estimates for high noise levels.

The existence of wrong parameter estimates is illustrated by numerical examples. Plant measurements as well as simulated systems are used.

TABLE OF CONTENTS

I. INTRODUCTION

1.1 The structu

Consider a dynar
corresponding o
purpose of an i
the given data.
different ident

In order to dev
governed by som
be called the s
equation obtain

Assume that the
finite order. I
random processe

$A(q^{-1})y(t) = B($

where y(t) is t
v(t) a stationa
operator and

$A(q^{-1}) = 1 + a_1$

$B(q^{-1}) = b_1q^{-1}$

It is assumed t

For simplicity

i) e(t) is alw
equally dist

ii) σ^2 denotes

I. INTRODUCTION

1.1 The structure of the system.

Consider a dynamic process. A sequence of inputs $\{u(t)\}$ and corresponding outputs $\{y(t)\}$ are given from an experiment. The purpose of an identification is to fit a mathematical model to the given data. This can be done in many ways. A good survey of different identification methods is given in [4].

In order to develop some theory it is assumed that the process is governed by some equation. The process given by this equation will be called <u>the system</u> in this report, while <u>the model</u> refers to the equation obtained in some way from the given data.

Assume that the system is linear, discrete, time invariant and of finite order. If the disturbances can be represented by stationary random processes, the system can in general be represented by

$$A(q^{-1})y(t) = B(q^{-1})u(t) + v(t) \tag{1.1}$$

where $y(t)$ is the output at time t, $u(t)$ the input at time t and $v(t)$ a stationary stochastic process. q^{-1} is the backward shift operator and

$$A(q^{-1}) = 1 + a_1 q^{-1} + \ldots + a_n q^{-n} \tag{1.2}$$

$$B(q^{-1}) = b_1 q^{-1} + \ldots + b_n q^{-n} \tag{1.3}$$

It is assumed that the system is asymptotically stable.

For simplicity introduce the following conventions

i) $e(t)$ is always denoting white noise (a sequence of independent, equally distributed random variables with zero mean

ii) σ^2 denotes the variance of $Ee^2(t)$

iii) S denotes the ratio $\frac{Eu^2(t)}{\sigma^2}$, which is proportional to the signal to noise ratio.

In the following it will be assumed that the noise v(t) can be expressed as

$$v(t) = H(q^{-1})e(t) \tag{1.4}$$

where $H(q^{-1})$ is a stable filter and e(t) white noise.

Introduce the matrix notations

$$Y = \begin{bmatrix} y(n+1) \\ \cdot \\ \cdot \\ \cdot \\ y(N+n) \end{bmatrix} \qquad V = \begin{bmatrix} v(n+1) \\ \cdot \\ \cdot \\ \cdot \\ v(N+n) \end{bmatrix}$$

$$\phi = \begin{bmatrix} -y(n).. & -y(1) & u(n)... & u(1) \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ -y(N+n-1)... & -y(N) & u(N+n-1)... & u(N) \end{bmatrix}$$

$$\theta = \begin{bmatrix} a_1 \\ \cdot \\ \cdot \\ a_n \\ b_1 \\ \cdot \\ \cdot \\ b_n \end{bmatrix}$$

(1.1) can be written as

$$Y = \phi\theta + V \tag{1.5}$$

where N is arbitrary.

1.2 The least s

The least squar

$V_{LS}(\hat{\theta}) = ||Y -$

with the well-k

$\hat{\theta}_{LS} = \theta + (\phi^T\phi)$

assuming that t

Åström has show
v(t) is white r

Correlated nois
squares (GLS) n
come this situa

1.3 The Markov

Introduce the s

$R = \begin{bmatrix} r_v(0)... \\ \cdot \\ \cdot \\ \cdot \end{bmatrix}$

$r_v(\tau)$ denotes

If R is known

$V_M(\hat{\theta}) = ||Y-\phi\hat{\theta}$

with the resul

1.2 The least squares method

The least squares (LS) estimate $\hat{\theta}_{LS}$ of θ is obtained by minimizing

$$V_{LS}(\hat{\theta}) = ||Y - \phi\hat{\theta}||^2 = (Y - \phi\hat{\theta})^T(Y - \phi\hat{\theta})$$

with the well-known solution

$$\hat{\theta}_{LS} = \theta + (\phi^T\phi)^{-1}\phi^T V \tag{1.6}$$

assuming that the inverse exists.

Åström has shown [1] that this method gives consistent estimates if v(t) is white noise.

Correlated noise causes biased estimates. The generalized least squares (GLS) method introduced by Clarke [8] is intended to overcome this situation.

1.3 The Markov estimate

Introduce the symmetric matrix R, which is assumed to be non-singular

$$R = \begin{bmatrix} r_v(0) \dots & r_v(N+1) \\ & \ddots \\ & r_v(0) \end{bmatrix}$$

$r_v(\tau)$ denotes the covariance function of the noise v(t).

If R is known the Markov estimate $\hat{\theta}_M$ of θ is obtained by minimizing

$$V_M(\hat{\theta}) = ||Y-\phi\hat{\theta}||^2_{R^{-1}} = (Y-\phi\theta)^T R^{-1}(Y-\phi\theta)$$

with the result

$$\hat{\theta}_M = \theta + (\phi^T R^{-1} \phi)^{-1} \phi^T R^{-1} v \tag{1.7}$$

which is a consistent estimate.

This follows from the consistency of the LS estimate as shown below.

It is believed that the following description of the Markov estimation besides proving consistence will give some more insight in the method and motivation for the generalized least squares method (introduced in the next section) as well.

From the relation (1.4)

$$v = He \tag{1.8}$$

where

$$H = \begin{bmatrix} 1 & & & \\ h_1 & & & \\ \vdots & & 0 & \\ h_{N-1} & \cdots & h_1 & 1 \end{bmatrix}$$

and

$$e = \begin{bmatrix} e(n+1) \\ \vdots \\ e(N+n) \end{bmatrix}$$

Define the filter $F(q^{-1})$ by

$$F(q^{-1}) = H(q^{-1})^{-1} \tag{1.9}$$

and form a corresponding matrix

$$F = \begin{bmatrix} 1 & & & \\ f_1 & & & \\ \vdots & & 0 & \\ f_{N-1} & & f_1 & 1 \end{bmatrix} \tag{1.10}$$

(1.9) can then be written

$$F = H^{-1} \tag{1.11}$$

From (1.8) it follows that

$$R = Evv^T = \sigma^2 HH^T$$

and invoking (1.11)

$$R^{-1} = \frac{1}{\sigma^2} F^T F \tag{1.12}$$

Introduce the filtered signals

$$\begin{aligned} y^F(t) &= F(q^{-1}y(t) \\ u^F(t) &= F(q^{-1})u(t) \end{aligned} \tag{1.13}$$

or in matrix language

$$\phi^F = F\phi, \; Y^F = FY \tag{1.14}$$

Then $V_M(\hat{\theta}) = \frac{1}{\sigma^2}(Y^F-\phi^F\hat{\theta})^T(Y^F-\phi^F\hat{\theta})$ (1.15)

From (1.5), (1.8), (1.11) and (1.14)

$$Y^F = \phi^F\theta + e \tag{1.16}$$

From (1.15) and (1.16) it is seen that the consistency of $\hat{\theta}_M$ follows from the consistency of the LS estimate.

In figure 1 the configuration adapted to LS is shown

$v(t) = e(t)$ (white noise)

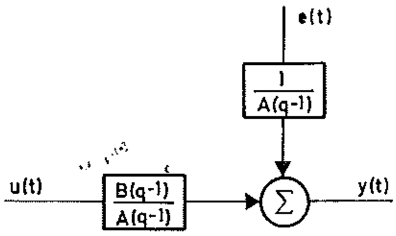


Figure 1

Figure 2 shows the general situation corresponding to (1.1)

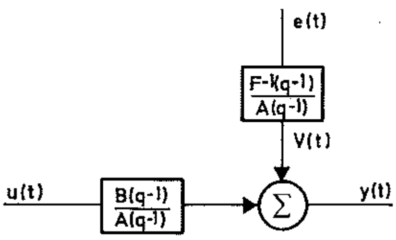


Figure 2

This system

the filter F

and $y^F(t)$ ha

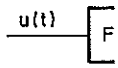


Figure 3

If R and the

obtained and

system. This

ing the cons

This system can, however, also be represented by figure 3, where the filter $F(q^{-1})$ has been moved and the filtered signals $u^F(t)$ and $y^F(t)$ have been introduced.

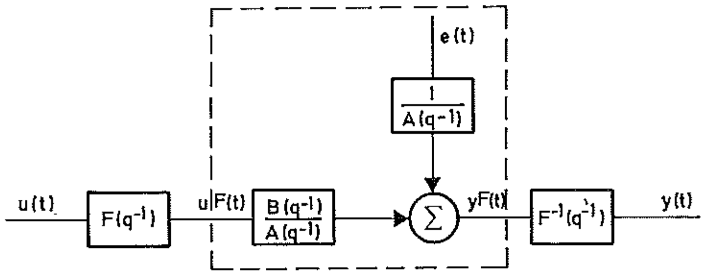


Figure 3

If R and then the filter $F(q^{-1})$ are known, $u^F(t)$ and $y^F(t)$ are easily obtained and it is sufficient to deal with the framed part of the system. This part, however, is quite similar to figure 1, thus indicating the consistency of the Markov estimate.

1.4. The generalized Least Squares method. Two versions.

The assumption of R known is highly unrealistic. In the general least squares (GLS) method θ and R are both estimated in an iterative way.

1. Guess a covariance matrix R_k.

2. Compute $\hat{\theta}_k$ from (1.7) with $R = R_k$.

3. Evaluate the residuals $\varepsilon_k = Y-\phi\hat{\theta}_k$ and use them to estimate a new covariance matrix R_{k+1}.

4. Put k=k+1 and 'repeat' from 2 until the estimate converges.

In this report two versions of the generalized least squares method are treated. In both versions the estimates of R are obtained by fitting an autoregression to the residuals.

Version 1:

This version can be described by the following scheme.

1. Guess a filter $\hat{C}_k(q^{-1}) = 1 + \hat{c}_{k1}q^{-1}+\ldots+\hat{c}_{kn}q^{-n}$

2. Compute $y_k^F(t)$ and $u_k^F(t)$ from

$$y_k^F(t) = \hat{C}_k(q^{-1})y(t) \qquad (1.17)$$

$$u_k^F(t) = \hat{C}_k(q^{-1})u(t)$$

and determine $\hat{\theta}_k$ by applying LS to the model

$$\hat{A}_k(q^{-1})y_k^F(t) = \hat{B}_k(q^{-1})u_k^F(t) + e(t)$$

3. Evaluate the residuals

$$\varepsilon_k(t) = \hat{A}_k(q^{-1})y(t) - \hat{B}_k(q^{-1})u(t) \tag{1.18}$$

Determine $\hat{C}_{k+1}(q^{-1})$ by fitting an autoregression to the residuals.

4. Put k=k+1 and repeat from 2 until convergence.

Clearly, this version corresponds to the model

$$\hat{A}(q^{-1})y(t) = \hat{B}(q^{-1})u(t) + \frac{1}{\hat{C}(q^{-1})} e(t) \tag{1.19}$$

with e(t) white noise.

<u>Version 2</u>:

This version coincides with Clarkes original proposal [8]. The iteration scheme is the following.

0. Put $y_o^F(t) = y(t)$, $u_o^F(t) = u(t)$, k=1

1. Guess a filter $\hat{C}_k(q^{-1}) = 1 + \hat{c}_{k1}q^{-1}+\ldots+ \hat{c}_{kn}q^{-n}$

2. Compute $y_k^F(t)$ and $u_k^F(t)$ from

$$\begin{aligned} y_k^F(t) &= \hat{C}_k(q^{-1})y_{k-1}^F(t) \\ u_k^F(t) &= \hat{C}_k(q^{-1})u_{k-1}^F(t) \end{aligned} \tag{1.17'}$$

and determine $\hat{\theta}_k$ by applying LS to the model

$$\hat{A}_k(q^{-1})y_k^F(t) = \hat{B}_k(q^{-1})u_k^F(t) + e(t)$$

3. Evaluate the residuals

$$\varepsilon_k(t) = \hat{A}_k(q^{-1})y_k^F(t) - \hat{B}_k(q^{-1})u_k^F(t) \tag{1.18'}$$

and determine a new filter $\hat{C}_{k+1}(q^{-1})$ by fitting an autoregression to the residuals.

4. Put k=k+1 and repeat from 2 until convergence.

With this version a successful iteration procedure ends when

$$\hat{C}_k(q^{-1}) \approx 1$$

The corresponding model is

$$\hat{A}(q^{-1})y(t) = \hat{B}(q^{-1})u(t) + \frac{1}{\prod\limits_{k=1}^{\infty} \hat{C}_k(q^{-1})} e(t) \tag{1.20}$$

For both the versions of GLS it is of course not necessary that the orders of the operators $\hat{A}$, $\hat{B}$ and $\hat{C}$ are the same. In this report the orders will in general be assumed to be the same, but the generalization is trivial.

The second version may be better if the noise v(t) is not generated as an autoregression. It will be shown, however, that both versions may fail (give biased estimates) at high noise levels.

The GLS method has some similarity with the repeated LS method as pointed out in [4].

In the repeated LS method (LS with successively higher order of the model) it is hoped that the A and B polynomials will have some factors in common. These factors are due to the correlation of the present noise.

In the GLS method there are always factors in common. To realized that, (1.19) is rewritten in the form

$$[\hat{A}(q^{-1})\hat{C}(q^{-1})]y(t) = [\hat{B}(q^{-1})\hat{C}(q^{-1})]u(t) + e(t)$$

The GLS method can thus be interpreted as a LS method with the constraint that the A and B polynomials have common factors.

In order to closer examine the properties of the two versions, the nature of the noise $v(t)$ or the covariance function $r_v(\tau)$ must be specified.

Some results in this report require only

$$v(t) = H(q^{-1})e(t)$$

where $H(q^{-1})$ is a stable filter and $e(t)$ is white noise.

Sometimes special interest will be paid to the following filter of finite order

$$H(q^{-1}) = \frac{1}{C(q^{-1})}$$

where

$$C(q^{-1}) = 1 + c_1 q^{-1} + \ldots + c_n q^{-n}$$

has all zeros outside the unit circle.

In these cases obviously

$$v(t) = \frac{1}{C(q^{-1})} e(t) \tag{1.21}$$

The reason for a study of (1.21) is its similarity in structure with the model (1.19).

It will be shown that under suitable regularity conditions on the input signal and the system dynamics the first version of the GLS method will always give consistent estimates, if the signal to noise ratio is high enough. However, if the noise level is high enough this version can give asymptotically biased estimates. It will also be shown that the second version can give biased estimates if the signal to noise ratio is low. All results hold asymptotically when the number of data tends to infinity.

II. MATHEMATI

2.1. Ergodic

It is the pur
of data tends

The least squ

$\hat{\theta}_{LS} = \theta + ($

The elements
is valuable t
in case of co

The questions
nature are co

The main resu

Theorem 2.1:

$y(t) = G(q^{-1}$

where

$G(q^{-1})$ and H(
orders.

e(t) is whit

$u(t) = u_1(t)$

$u_1(t)$ determ
is a periodi

$|u_1(t) - u_1'($

$u_2(t) = F(q^-$

II. MATHEMATICAL PRELIMINARIES

2.1. Ergodic properties of time series

It is the purpose to develop results which are valid as the number of data tends to infinity.

The least squares estimate $\hat{\theta}_{LS}$ (1.4) can be written

$$\hat{\theta}_{LS} = \theta + \left(\frac{\phi^T\phi}{N}\right)^{-1}\left(\frac{\phi^T v}{N}\right)$$

The elements of the matrices $\frac{\phi^T\phi}{N}$ and $\frac{\phi^T v}{N}$ are sample covariances. It is valuable to know when these sample covariances converge as $N\to\infty$, and in case of convergence the limits too.

The questions are answered by ergodic theory. Some results of this nature are collected in Appendix A.

The main result is the following.

Theorem 2.1: Consider the system

$$y(t) = G(q^{-1})u(t) + H(q^{-1})\, e(t)$$

where

$G(q^{-1})$ and $H(q^{-1})$ are asymptotically stable filters of finite orders.

$e(t)$ is white noise with finite fourth moment and independent of $u(t)$

$$u(t) = u_1(t) + u_2(t)$$

$u_1(t)$ deterministic and almost periodic, that is to every $\varepsilon > 0$ there is a periodic function $u_1'(t)$ such that

$$|u_1(t) - u_1'(t)| < \varepsilon \quad \text{all } t$$

$u_2(t) = F(q^{-1})\, v(t)$ with $F(q^{-1})$ an asymptotically stable filter of finite order and $v(t)$ white noise with finite fourth moment.

Let further $D_1(q^{-1})$ and $D_2(q^{-1})$ be asymptotically stable filters of finite orders.

Then

$$\lim \frac{1}{n} \sum_{t=1}^{n} (D_1(q^{-1})y(t) + D_2(q^{-1})u(t)) \begin{bmatrix} y(t) \\ u(t) \end{bmatrix}$$

$$= E(D_1(q^{-1})y(t) + D_2(q^{-1})u(t)) \begin{bmatrix} y(t) \\ u(t) \end{bmatrix} \tag{2.1}$$

with probability one and in mean square.

If $x(t)$ is deterministic, $E\,x(t)$ denotes $\lim\limits_{n\to\infty} \frac{1}{n} \sum\limits_{t=1}^{n} x(t)$.

2.2. Persist

Definition :

i) $\lim\limits_{N\to\infty} \frac{1}{N}$ ...

exist and

ii) the n b

$R_u = $ [$r_u(0$ ...

is positive

Some simple
characteriz
in [15]. In
(proved in

Lemma 2.1:
the spectra
tion is non
points.

If u(t) is
sist of a n
considered

Corr: Let y
order n and
then y(t) i

A simple ap

2.2. Persistently Exciting Signals.

Definition 2.1: u(t) is said to be persistently exciting of order n if

i) $\lim\limits_{N\to\infty} \frac{1}{N} \sum\limits_{t=1}^{N} u(t) = \bar{u}$ and $\lim\limits_{N\to\infty} \frac{1}{N} \sum\limits_{t=1}^{N} [u(t)-\bar{u}][u(t+\tau)-\bar{u}] = r_u(\tau)$

exist and

ii) the n by n symmetric matrix

$$R_u = \begin{bmatrix} r_u(0) & r_u(1) \ldots\ldots & r_u(n-1) \\ & & \\ & & r_u(1) \\ & & r_u(0) \end{bmatrix}$$

is positive definite.

Some simple properties of persistently exciting signals and a characterization of this concept in the frequency domain is given in [15]. In this report the following properties will be used (proved in [15]).

Lemma 2.1: u(t) is persistently exciting of order n if and only if the spectral density corresponding to the sample covariance function is non zero (in distributive sense) in at least n different points.

If u(t) is periodic, the spectral density will be discrete and consist of a number of δ-functions. The distribution $\delta(x)$ is here considered as non zero in $x = 0$.

Corr: Let $y(t) = H(q^{-1})u(t)$. If u(t) is persistently exciting of order n and $H(q^{-1})$ is stable and has no zeros on the unit circle, then y(t) is persistently exciting of order n.

A simple application of the definition is made in

<u>Lemma 2.2</u>: Let $y(t) = H(q^{-1})u(t) \qquad H(q^{-1}) = \sum_{i=0}^{n-1} h_i q^{-i}$

i) If $y(t) \equiv 0$ with probability one and $u(t)$ is persistently exciting, then $h_i = 0 \;\; i = 0,\dots,n-1$

ii) If $u(t)$ is not persistently exciting of order n, then there exists $H(q^{-1}) \not\equiv 0$ such that $y(t) \equiv 0$ with probability one.

<u>Proof</u>:

$$Ey^2(t) = \begin{bmatrix} h_o \cdots h_{n-1} \end{bmatrix} \begin{bmatrix} r_u(0)\dots\dots & & r_u(n-1) \\ & & \\ & & r_u(0) \end{bmatrix} \begin{bmatrix} h_o \\ \\ h_{n-1} \end{bmatrix}$$

$y(t) = 0$ with probability one if and only if $Ey^2(t) = 0$.

i) $Ey(t)^2 = 0$ and R_u non singular implies $h_i = 0 \;\; i = 0,\dots n-1$

ii) R_u is singular. Take the vector

$$\begin{bmatrix} h_o \\ \vdots \\ h_{n-1} \end{bmatrix}$$

in the null space of R_u. Then $E\,y(t)^2 = 0$.

Q.E.D.

2.3. The system covariance matrix

Consider the undisturbed linear system

$$y(t) = K(q^{-1})u(t)$$

Definition 2.2: The system covariance matrix of order 2k is understood as the 2k by 2k symmetric matrix

$$R = \begin{bmatrix} R_y & R_{yu} \\ R_{uy} & R_u \end{bmatrix}$$

$$= \begin{bmatrix} r_y(0)\ldots & r_y(k-1) & r_{yu}(0)\ldots & r_{yu}(k-1) \\ & r_y(0) & r_{yu}(1-k)\ldots & r_{yu}(0) \\ & & r_u(0) & r_u(k-1) \\ & & & \\ & & & r_u(0) \end{bmatrix}$$

$$= \lim_{N\to\infty} \frac{1}{N} \sum_{t=n+1}^{n+N} \begin{bmatrix} y(t-1) \\ \\ y(t-k) \\ u(t-1) \\ u(t-k) \end{bmatrix} [y(t-1)\ldots\ y(t-k)u(t-1)\ldots\ u(t-k)]$$

Lemma 2.1. Let R be the system covariance matrix of order k of

$$y(t) = K(q^{-1})u(t)$$

Then

$$x^T Rx = r_\varepsilon(0)$$

with

$$\varepsilon(t) = F(q^{-1})y(t) + G(q^{-1})u(t)$$

$$F(q^{-1}) = \sum_{1}^{k} f_i q^{-i}, \quad G(q^{-1}) = \sum_{1}^{k} g_i q^{-i}$$

$$x = [f_1 \ldots f_k g_1 \quad g_k]^T$$

Proof: Straight forward calculations give

$$x^T R x = \lim_{N\to\infty} \frac{1}{N} x^T \sum_{t=n+1}^{n+N} \begin{bmatrix} y(t-1) \\ \vdots \\ u(t-k) \end{bmatrix} [y(t-1) \ldots u(t-k)]x$$

$$= \lim_{N\to\infty} \frac{1}{N} \sum_{t=n+1}^{n+N} \left([f_1 \ldots . f_k g_1 \ldots g_k] \begin{bmatrix} y(t-1) \\ y(t-k) \\ u(t-1) \\ u(t-k) \end{bmatrix} \right)^2$$

$$= \lim_{N\to\infty} \frac{1}{N} \sum_{t=n+1}^{n+N} \varepsilon^2(t) = r_\varepsilon(0)$$

Q.E.D.

Theorem 2.2.: Let the controllable, asymptotically stable system

$$A(q^{-1})y(t) = B(q^{-1})u(t)$$

be of order n.

Consider the system covariance matrix R of order 2k.

i) Assume that $k \leq n$. If $u(t)$ is persistently exciting of order $n+k$, then R is positive definite.

ii) Assume that $k > n$. If $u(t)$ is persistently exciting of order $n+k$, then R is singular (positive semidefinite). Further the null space of R is spanned by vectors of the form

$$x = \begin{bmatrix} f_1 \\ \vdots \\ f_k \\ g_1 \\ \vdots \\ g_k \end{bmatrix} \tag{2.2}$$

where f_i and g_i fulfil the relations

$$F(q^{-1}) = \sum_{i=1}^{k} f_i q^{-i} = A(q^{-1}) L(q^{-1}) \tag{2.3a}$$

$$G(q^{-1}) = \sum_{i=1}^{k} g_i q^{-i} = -B(q^{-1}) L(q^{-1}) \tag{2.3b}$$

$$L(q^{-1}) = \sum_{i=1}^{k-n} l_i q^{-i} \quad \text{is arbitrary} \tag{2.4}$$

iii) Assume that $k \geq n$. If $u(t)$ is not persistently exciting of order $n+k$, then R is singular.

Remark: In the not described case, when $k < n$ and $u(t)$ is not persistently exciting of order $n+k$, nothing general can be stated.

Proof: Consider the equation

$$Rx = 0 \tag{2.5}$$

or equivalently

$$x^T R x = 0 \tag{2.6}$$

With notations from and use of lemma 2.3 this is written

$$r_\varepsilon(0) = 0 \tag{2.7}$$

Since then $r_\varepsilon(\tau) = 0$ all τ, it follows that (2.7) is equivalent to

$$r_{\varepsilon_1}(0) = 0 \quad \varepsilon_1(t) = A(q^{-1})\varepsilon(t)$$

Now $\varepsilon_1(t) = A(q^{-1})[F(q^{-1})y(t) + G(q^{-1})u(t)] =$

$$= [F(q^{-1})B(q^{-1}) + G(q^{-1})A(q^{-1})]u(t) \equiv H(q^{-1})u(t)$$

The original equation (2.1) is thus transformed into $h^T R_u h = 0$ or

$$R_u h = 0 \tag{2.8}$$

with

$$R_u = \begin{bmatrix} r_u(0)\ldots & & r_u(n+k-1) \\ & & \\ & & \\ & & r_u(0) \end{bmatrix}$$

$$h = \begin{bmatrix} h_1 \\ \\ h_{n+k} \end{bmatrix}$$

Separate two

Case a): Ass
(2.8) implie

If $F(q^{-1}) \not\equiv$

$\frac{B(q^{-1})}{A(q^{-1})} = -\frac{G}{F}$

where the le
order k-1.

If $k \leq n$ thi
is the only

If, on the c
form $G(q^{-1})$

$L(q^{-1}) = \sum_{i=1}^{k-r}$

The equatior
equations

$Tx = 0$

with x as be
$a_n, b_1, \ldots,$

$T = \begin{bmatrix} 0 \\ b_1 \\ \cdot \\ b_n \\ \cdot \\ 0 \end{bmatrix}$

From the di

(2.5)

(2.6)

is written

(2.7)

2.7) is equivalent

(t)

ed into $h^T R_u h = 0$ or

(2.8)

Separate two cases.

Case a): Assume that u(t) is persistently exciting of order n+k. (2.8) implies $h = 0$ or $H(q^{-1}) \equiv 0$.

If $F(q^{-1}) \not\equiv 0$ it is then concluded that

$$\frac{B(q^{-1})}{A(q^{-1})} = -\frac{G(q^{-1})}{F(q^{-1})} \tag{2.9}$$

where the left hand side is of order n and the right hand side of order k-1.

If $k \leq n$ this is a contradiction and $F(q^{-1}) \equiv G(q^{-1}) \equiv 0$, or $x = 0$ is the only solution of (2.5) which proves part i).

If, on the other hand, $k > n$, all solutions of (2.9) are of the form $G(q^{-1}) = -B(q^{-1})\,L(q^{-1})$, $F(q^{-1}) = A(q^{-1})\,L(q^{-1})$ where $L(q^{-1}) = \sum_{i=1}^{k-n} l_i q^{-i}$ is arbitrary. This proves part ii).

The equation $H(q^{-1}) \equiv 0$ can be transformed to a system of linear equations

$$Tx = 0$$

with x as before and T a (n+k) by 2k matrix, depending on $a_1, \ldots, a_n, b_1, \ldots, b_n$. More explicitly T is the matrix

$$T = \left[\begin{array}{cccc|cccc} 0 & & & 0 & 1 & & & 0 \\ b_1 & \ddots & & & a_1 & \ddots & & \\ & & \ddots & 0 & & & \ddots & 1 \\ b_n & & & b_1 & a_n & & & a_1 \\ & \ddots & & & & \ddots & & \\ 0 & & \ddots & & 0 & & \ddots & \\ & & & b_n & & & & a_n \end{array}\right]$$

From the discussion it is clear that the null space of T $N(T) = \{0\}$

if and only if $k \leq n$.

Case b): Assume that u(t) is not persistently exciting of order n+k. Then (2.5) is equivalent to $h \in N(R_u)$. Let r be an arbitrary vector in the null space $N(R_u)$. By transforming the equation as before

$$T\,x = r \tag{2.10}$$

If $k > n$ take $r = 0$ and x as (2.1) - (2.4).
If $k = n$, T is a square, invertible matrix and to every $r \neq 0$ there is a non trivial solution of (2.10). This proves part iii).

Q.E.D.

Interpretation: Consider $V = r_\varepsilon(0)$,

$$\varepsilon(t) = F(q^{-1})y(t)+G(q^{-1})u(t),\; F(q^{-1}) = \sum_1^k f_i q^{-i},\; G(q^{-1}) = \sum_1^k g_i q^{-i}$$

The system covariance matrix of order 2k is singular if and only if the minimum of V with respect to $\{f_i\}$ and $\{g_i\}$ is zero.
Loosely speaking the result of the theorem is:
If $k > n$, the filters $F(q^{-1})$ and $G(q^{-1})$ are of higher order than the system and it is possible to get V=0.

If $k \leq n$ it is not possible to get $V = 0$ if all modes of the system are excited.

III. MAIN RES

3.1. Introduc

In this chapt
it is shown (
ting the maxi
of convergenc
of the likeli
guarantee a u
3.1). As the
computer the
interest. In
mum points de
model (theore

The second ve
this chapter
version of GL

III. MAIN RESULTS

3.1. Introduction

In this chapter the first version of GLS is closer examined. First it is shown (theorem 3.1) that the method can be interpreted as adapting the maximum likelihood technique to this problem. The question of convergence is then reduced to an examination of local maximum points of the likelihood function. It is rather easy to give conditions which guarantee a unique global maximum of the likelihood function (lemma 3.1). As the computations of the GLS method must be carried out on a computer the possible existence of several local maximas is of greater interest. In three theorems it is shown that the number of local maximum points depends on the signal to noise ratio and the order of the model (theorems 3.2, 3.3 and 3.4).

The second version can be interpreted similarily. In the end of this chapter it is shown how to construct examples, where this version of GLS converges to biased estimates.

3.2. Maximum Likelihood Interpretation

In this section it is shown how the GLS method can be interpreted as the maximum likelihood method. Expressions for a corresponding loss function are given in matrix notations and using operators. Finally the limit of this function, as the number of samples tends to infinity, is studied.

Theorem 3.1: Assume that the disturbances are given by

$$v(t) = \frac{1}{C(q^{-1})} e(t) \tag{3.1}$$

$e(t)$ white Gaussian noise. The first version of the GLS method is equivalent to maximizing the likelihood function of this problem by a relaxation method.

Proof:

The probability function of y is given by

$$f(y) = \frac{1}{(2\pi)^{N/2}(\det R)^{1/2}} \exp\left[-\frac{1}{2}(Y-\phi\theta)^T R^{-1}(Y-\phi\theta)\right]$$

(3.1) is written by matrix notations

$$e = Fv$$

$$F = \begin{bmatrix} 1 & & & \\ c_1 & & 0 & \\ c_n & & & \\ & 0 & c_n \ldots\ldots & 1 \end{bmatrix}$$

From (1.12) it follows that

$$R = Evv^T = \left(\frac{1}{\sigma^2} F^T F\right)^{-1}$$

The likelihood

$-\log L = \frac{1}{2}(Y-\phi\theta$

Let

$W(\hat{\theta},F) = \frac{1}{2N}(Y-\phi$

so

$-\log L = \frac{N}{\hat{\sigma}^2} W(\hat{\theta}$

since $\det \hat{F} = 1$

$\frac{\partial L}{\partial \hat{\sigma}^2} = 0$ implies

so L is maximiz

$\hat{\sigma}^2 = 2W(\hat{\theta},\hat{F})$

Maximizing L is algorithm can b alternating bet

1. Minimize

2. Minimize

which is a rela

Remark 1: Denote $\hat{\varphi}_N$ when the rec $\hat{\varphi}_N$ has nice asy

The likelihood function is given by

$$-\log L = \frac{1}{2}(Y-\phi\hat{\theta})^T \frac{1}{\hat{\sigma}^2} F^T F(Y-\phi\hat{\theta}) + \frac{1}{2}\log\det(\hat{F}^{-1}\hat{\sigma}^2(\hat{F}^T)^{-1}) + \frac{N}{2}\log 2\pi \quad (3.2)$$

Let

$$W(\hat{\theta},F) = \frac{1}{2N}(Y-\phi\hat{\theta})^T \hat{F}^T\hat{F}(Y-\phi\hat{\theta}) \quad (3.3)$$

so

$$-\log L = \frac{N}{\hat{\sigma}^2} W(\hat{\theta},\hat{F}) + \frac{1}{2}\log(\hat{\sigma}^{2^N}) + \frac{N}{2}\log 2\pi$$

since $\det \hat{F} = 1$

$$\frac{\partial L}{\partial \hat{\sigma}^2} = 0 \text{ implies } -\frac{NW(\hat{\theta},\hat{F})}{(\hat{\sigma}^2)^2} + \frac{N}{2\hat{\sigma}^2} = 0$$

so L is maximized with respect to $\hat{\sigma}^2$ by

$$\hat{\sigma}^2 = 2W(\hat{\theta},\hat{F}) \quad (3.4)$$

Maximizing L is then equivalent to minimizing $W(\hat{\theta},\hat{F})$. The actual algorithm can be interpreted as a minimization of this function by alternating between

1. Minimize $W(\hat{\theta}_k,\hat{F}_k)$ with respect to $\hat{\theta}_k$

2. Minimize $W(\hat{\theta}_k,\hat{F}_{k+1})$ with respect to $\hat{F}_{k+1}$

which is a relaxation method.

Q.E.D.

Remark 1: Denote the estimate of $a_1,\ldots,a_n,b_1,\ldots,b_n,c_1,\ldots,c_n$ by $\hat{\varphi}_N$ when the record length is N. It follows from [3] and [6] that $\hat{\varphi}_N$ has nice asymptotic properties:

1. $\hat{\varphi}_N$ converges with probability one to the true parameter vector φ as N increases.

2. $\hat{\varphi}_N$ is asymptotic efficient (i.e. has minimal variance).

3. $\hat{\varphi}_N$ is asymptotic normal with the mean value φ and the covariance matrix

$$\frac{2W}{N} W_{\varphi\varphi}^{''-1}$$

Remark 2 $W(\hat{\theta}_k,\hat{F}_k)$ is a decreasing, bounded sequence, which implies convergence. Possible bounded limits must be stationary points of $W(\hat{\theta},\hat{F})$. They cannot be local maximum points. It is shown in Appendix B that saddle points have not to be considered either, since they are not "stable" points. By this concept it is meant that a start of the iteration sufficiently closed to a saddle point will not in general imply convergence to the point. Since the minimization of $W(\hat{\theta},\hat{F})$ has to be carried out on a computer, rounding errors must be introduced in the calculations, and the probability of convergence to a saddle point can for practical cases be regarded as zero. Local minimum points are thus the only "practically possible", bounded limits of $(\hat{\theta}_k,\hat{F}_k)$ as $k\to\infty$.

Remark 3 Note that the convergence of the minimization algorithm is very slow. It is shown in Appendix B that close to a minimum point $\hat{\theta}_k$ will converge linearly.

Remark 4 The second version of GLS can be interpreted in a similar way. Let $W(\hat{\theta},\hat{F})$ be defined from (3.3) and put

$\hat{F} = \prod_{i=1}^{\infty} \hat{F}_i$. The iteration procedure is a minimization with different constraints on $\hat{F}$. I.e. in step k, $\hat{F}_1,\ldots,\hat{F}_{k-1}$ are fixed. $W(\hat{\theta}_k,\hat{F})$ is minimized with respect to $\hat{F}_k$. $\hat{F}_{k+1} = \hat{F}_{k+2} = \ldots I$. This step corresponds to the estimation of the filter $\hat{C}_k(q^{-1})$. From this interpretation it is clear that $W(\hat{\theta},\hat{F})$ is decreased in each step. This fact is shown by straight forward calculations in [18].

From the discussion in 1.3 it is clear that the loss function $W(\hat{\theta},\hat{F})$ can be expressed as

$$W(\hat{a}_1,\dots,\hat{a}_n\ \hat{b}_1,\dots,\hat{b}_n\hat{c}_1,\dots,\hat{c}_n) = \frac{1}{2N}\sum_{t=1}^{N}\varepsilon^F(t)^2 \tag{3.5}$$

$$\varepsilon^F(t) = \hat{C}(q^{-1})\varepsilon(t) = \varepsilon(t) + \hat{c}_1\varepsilon(t-1)+\dots+\hat{c}_n\varepsilon(t-n) \tag{3.6}$$

$$\varepsilon(t) = \hat{A}(q^{-1})y(t) - \hat{B}(q^{-1})u(t) \tag{3.7}$$

so

$$\varepsilon^F(t) = \hat{C}(q^{-1})\,\frac{\hat{A}(q^{-1})B(q^{-1})-A(q^{-1})\,B\,(q^{-1})}{A(q^{-1})}u(t) +$$

$$+\,\frac{\hat{A}(q^{-1})\hat{C}(q^{-1})}{A(q^{-1})}\,v(t) \tag{3.8}$$

Clearly W is a polynomial in $\hat{a}_1\dots\dots\hat{c}_n$ where the coefficients are different sample covariances. An analysis of W and especially the local minimum points of this function must therefore be done in a probabilistic setting. In order to do the analysis reasonable asymptotic theory will be used.

In the following some assumptions are made

i) $u(t) = u_1(t) + G(q^{-1})e_1(t)$

$u_1(t)$ is deterministic, and almost periodic. $G(q^{-1})$ is a stable filter of finite order.
$e_1(t)$ is white noise.

ii) $v(t) = H(q^{-1})e_2(t)$

$H(q^{-1})$ is a stable filter of finite order
$e_2(t)$ is white noise

iii) $e_1(t)$ and $e_2(t)$ ($u(t)$ and $v(t)$) are independent.

Under these assumptions it follows from Theorem 2.1 that W has a limit $V(\hat{a}_1 \ldots \hat{a}_n\ \hat{b}_1 \ldots \hat{b}_n\ \hat{c}_1 \ldots \hat{c}_n)$ with probability one, and that

$$V(\hat{a}_1 \ldots \hat{c}_n) = V_1(\hat{a}_1 \ldots \hat{c}_n) + V_2(\hat{a}_1 \ldots \hat{c}_n) \tag{3.9}$$

$$V_i(\hat{a}_1 \ldots \hat{c}_n) = \frac{1}{2} E\varepsilon_i^F(t)^2 \tag{3.10}$$

$$\varepsilon_1^F(t) = \hat{C}(q^{-1}) \frac{\hat{A}(q^{-1})B(q^{-1}) - A(q^{-1})\hat{B}(q^{-1})}{A(q^{-1})} u(t) \tag{3.11}$$

$$\varepsilon_2^F(t) = \frac{\hat{A}(q^{-1})\hat{C}(q^{-1})\hat{H}(q^{-1})}{A(q^{-1})} e(t) \tag{3.12}$$

The notation $Eu_1^2(t)$ denotes

$$\lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} u_1^2(t).$$

It is the purpose of Sections 3.3 - 3.7 closer to examine the loss function V. The main interest will be an investigation when the loss function has a unique local minimum.

In order to simplify the analysis a bit only "interesting" values of the parameter estimates will be considered.

In many cases the following compact set in the parameter space will be reasonable:

i) $\hat{A}(z)$ has all zeros inside the circle $|z| \leq r<1$.

ii) $C(z)$ " " " " " " $|z| \leq r<1$.

iii) $\hat{b}_i$ bounded.
r close to 1.

This restric

i) means

ii) is mot
finite

iii) must b

3.3. Global

This section
ning the glo
cial case

$v(t) = \frac{1}{C(q^{-1}}$

Lemma 3.1: C
(3.13). Denc
of the syste

i) Global

$\hat{A}(q^-$

$\varepsilon_1^F(t$

with

ii) $\hat{a}_i$ = a
$\hat{b}_i$ = b
$\hat{c}_i$ = c

is alw

This restriction is well justified by physical reasons.

i) means that a stable model is required,

ii) is motivated by the representation theorem [2] and finite variance of the output,

iii) must be fulfilled if the model has finite gain.

3.3. Global properties of the loss function.

This section contains some simple considerations concerning the global minimum of the loss function in the special case

$$v(t) = \frac{1}{C(q^{-1})}\, e(t) \tag{3.13}$$

Lemma 3.1: Consider the loss function (3.9) with $v(t)$ as (3.13). Denote the order of the model by m and the order of the system by n. Assume $m \geq n$.

i) Global minimum points are the solution of

$$\begin{cases} \hat{A}(q^{-1})\hat{C}(q^{-1}) = A(q^{-1})C(q^{-1}) \\ \varepsilon_1^F(t) = \hat{C}(q^{-1})\, \dfrac{\hat{A}(q^{-1})B(q^{-1}) - A(q^{-1})\hat{B}(q^{-1})}{A(q^{-1})}\, u(t) = 0 \\ \text{with probability one.} \end{cases} \tag{3.14}$$

ii)

$$\begin{array}{llll} \hat{a}_i = a_i & i = 1, \dots n & \hat{a}_i = 0 & i = n+1, \dots m \\ \hat{b}_i = b_i & i = 1, \dots n & \hat{b}_i = 0 & i = n+1, \dots m \\ \hat{c}_i = c_i & i = 1, \dots n & \hat{c}_i = 0 & i = n+1, \dots m \end{array} \tag{3.15}$$

is always a global minimum point.

iii) If $u(t)$ is persistently exciting of order $n+m$, $m=n$, and the system is controllable, then (3.15) is the unique global minimum point.

iv) If $u(t)$ is not persistently exciting of order $n+m$, $m=n$, there may exist other global minimum points than (3.15).

v) If $u(t)$ is persistently exciting of order $n+m$, $m>n$, there are in general several global minimum points. These points are equivalent in the sense that they all satisfy

$$\frac{\hat{B}(q^{-1})}{\hat{A}(q^{-1})} = \frac{B(q^{-1})}{A(q^{-1})} \tag{3.16}$$

$$\frac{1}{\hat{A}(q^{-1})\hat{C}(q^{-1})} = \frac{1}{A(q^{-1})C(q^{-1})} \tag{3.17}$$

Proof: Clearly $\inf V_1 = 0$. Further $\inf V_2 = \frac{1}{2} Ee^2(t)$. To realize that define

$$G(q^{-1}) = \frac{\hat{A}(q^{-1})\hat{C}(q^{-1})}{A(q^{-1})C(q^{-1})} = 1 + \sum_{i=1}^{\infty} g_i q^{-i}$$

Then

$$V_2 = \frac{1}{2} Ee^2(t)\left[1 + \sum_{i=1}^{\infty} g_i^2\right]$$

and $\inf V_2 = \frac{1}{2} Ee^2(t)$ for $g_i = 0$, $i = 1, \dots$

i) The equations

$$V_1 = \inf V_1$$
$$V_2 = \inf V_2$$

have the solutions (3.14).

ii) The assertion follows directly from i).

iii) From Lemma 2.2 it is concluded that

$$\hat{A}(q^{-1})B(q^{-1}) \equiv A(q^{-1})\hat{B}(q^{-1})$$

and by arguments as in the proof of Theorem 2.1 the assertion follows.

iv) An example for a first order system with $a \neq c$

$$u(t) = \lambda^t, \quad \lambda = \pm 1$$

$$\hat{a} = c, \qquad \hat{b} = b\,\frac{1 + c\lambda}{1 + a\lambda}, \quad \hat{c} = a$$

v) Lemma 2.2 implies

$$\hat{A}(q^{-1})B(q^{-1}) \equiv A(q^{-1})\hat{B}(q^{-1})$$

so (3.16) is proved. (3.17) follows from (3.14). In general the factor in common between $\hat{A}(q^{-1})$ and $\hat{B}(q^{-1})$ can be chosen in several ways to satisfy (3.17).

Q.E.D.

Remark: The assumption that (3.13) holds is essential for the result.

3.4. Estimates at high signal to noise ratios. Models of correct order.

In this section a theorem of uniqueness is given and discussed. The essential part of the proof is found in Appendix C as a series of lemmas.

Theorem 3.2: Let the system of order n

$$A(q^{-1})y(t) = B(q^{-1})u(t) + v(t), \quad v(t) = H(q^{-1})e(t) \tag{3.18}$$

be controllable and the input u(t) persistently exciting of order 2n. Assume that the order of the model is n. Consider parameter estimates in Ω, an arbitrary compact set.

Then there is a constant S_o such that if $S_o \leq S < \infty$ then the loss function (3.9) has exactly one stationary point in Ω. This point is a local minimum and satisfies

$$\begin{cases} \hat{a}_i = a_i + O(1/S) & i = 1 \dots n \\ \hat{b}_i = b_i + O(1/S) & i = 1 \dots n \\ \hat{c}_i = \bar{c}_i + O(1/S) & i = 1 \dots n \end{cases} \tag{3.19}$$

where $\bar{C}(q^{-1}) = 1 + \bar{c}_1 q^{-1} + \dots + \bar{c}_n q^{-n}$ and $(\bar{c}_1, \dots \bar{c}_n)$ is the minimum point of

$$E\left[\hat{C}(q^{-1})v(t)\right]^2 \tag{3.20}$$

Proof: Introduce the vectors x and y by

$$x = \begin{bmatrix} \hat{a}_1 - a_1 \\ \vdots \\ \hat{a}_n - a_n \\ \hat{b}_1 - b_1 \\ \vdots \\ \hat{b}_n - b_n \end{bmatrix} \qquad y = \begin{bmatrix} \hat{c}_1 \\ \vdots \\ \hat{c}_n \end{bmatrix}$$

Then the loss function (3.9) can be written

$$V(x,y) = \frac{1}{2}\, x^T P(y) x + \varepsilon h(x,y)$$

with P(y) as the covariance matrix of the system

$$A(q^{-1}) y^F(t) = - B(q^{-1}) u^F(t)$$

$$u^F(t) = \hat{C}(q^{-1}) u(t)$$

This fact follows from Lemma 2.3. From corr of Lemma 2.1 and Theorem 2.2 follows that P(y) is non singular for all y. Further the loss function is assumed to be scaled so that ε denotes the quantity 1/S.

The function $h(0,y) = 2\, E[\hat{C}(q^{-1}) v(t)]^2$ is quadratic in y. It has a unique minimum point y_o, which fulfils $h''_{yy}(0,y_o)$ positive definite. Invoking Theorem C.1 the proof is finished.

Q.E.D.

What sense have the different assumptions?

i) The restriction on the input signal is very natural. This condition is necessary for the result (Lemma 3.1).

ii) The study of only parameter estimates in Ω is motivated before.

iii) The restriction on the signal to noise ratio is crucial as is shown in Theorem 3.3.

iv) The assumption of controllability is essential. If the system is non controllable, there is a factor in common between $A(q^{-1})$ and $B(q^{-1})$. Equation (3.18) can be divided by this factor, obtaining a controllable system of lower order than the original and with another correlation of the noise. If the system is not controllable, it is thus equivalent to regard the order of the model as higher than the order of the (controllable part of the) system. This situation is treated in Section 3.7, where it is shown that non controllable systems in general will give no unique local minimum.

3.5. Estimates at low signal to noise ratios.

This section deals with the case of low signal to noise ratios. It turns out that a possible property of the noise plays an essential role for non uniqueness.

Definition 3.2. The noise $v(t) = H(q^{-1})e(t)$ fulfils the "noise condition" (NC) if there exist at least two different pairs of polynomials $\hat{A}_1(q^{-1})$, $\hat{C}_1(q^{-1})$ and $\hat{A}_2(q^{-1})$, $\hat{C}_2(q^{-1})$, such that

$$V_2(\hat{a}_1 \ldots \hat{a}_n,\ \hat{c}_1 \ldots \hat{c}_n) = E\left[\frac{\hat{A}(q^{-1})\hat{C}(q^{-1})H(q^{-1})}{A(q^{-1})}\, e(t)\right]^2 \qquad (3.9)$$

has a local minimum point with a positive definite matrix of second order derivatives in $(\hat{a}_{11} \ldots \hat{a}_{1n},\ \hat{c}_{11} \ldots \hat{c}_{1n})$ and $(\hat{a}_{21} \ldots \hat{a}_{2n},\ \hat{c}_{21} \ldots \hat{c}_{2n})$.

y is essential. If
there is a factor
$^{-1}$). Equation (3.18)
btaining a controll-
the original and
noise. If the sys-
hus equivalent to
higher than the
of the) system. This
3.7, where it is
ems in general will

atios.

signal to noise
operty of the noise
ess.

e(t) fulfils the
t least two diffe-
$^{-1}$) and $\hat{A}_2(q^{-1})$,

$$\left. q^{-1}) \; e(t)\right]^2 \qquad (3.9)$$

ve definite matrix
$_{1n}$, $\hat{c}_{11} \ldots \hat{c}_{1n}$) and

Remark:

$$w(t) = \frac{H(q^{-1})}{A(q^{-1})} e(t)$$

is the measurement noise if all noise of the process is interpreted as measurement noise.

Corr 1: $v(t)$ fulfils (NC) if there exists a minimum point with V_2'' positive definite, $\hat{A}(q^{-1}) \nmid \hat{C}(q^{-1})$.

Proof: Take $\hat{A}_2 = \hat{C}$, $\hat{C}_2 = \hat{A}$. By symmetry this corresponds to another point satisfying the predescribed conditions.

Corr 2: If

$$v(t) = \frac{1}{C(q^{-1})} e(t)$$

it is sufficient that there is a factorization of $A(q^{-1})$ and $C(q^{-1})$ such that $A(q^{-1})C(q^{-1}) = A_1(q^{-1})C_1(q^{-1})$ where $A_1(q^{-1})$ and $C_1(q^{-1})$ have no factors in common.

Proof: $\hat{A}_1(q^{-1}) = A_1(q^{-1})$, $\hat{C}_1(q^{-1}) = C_1(q^{-1})$ and $\hat{A}_2(q^{-1}) = C_1(q^{-1})$, $\hat{C}_2(q^{-1}) = A_1(q^{-1})$ define two different (local and global) minimum points. The matrix of second order derivatives is given by

$$\frac{\partial^2 V_2}{\partial \hat{a}_i \partial \hat{a}_j} = 2E\left[\left(q^{-i}\hat{C}(q^{-1})v(t)\right)\left(q^{-j}\hat{C}(q^{-1})v(t)\right)\right]$$

$$\frac{\partial^2 V_2}{\partial \hat{a}_i \partial \hat{c}_j} = 2E\left[\left(q^{-i}\hat{C}(q^{-1})v(t)\right)\left(q^{-j}\hat{A}(q^{-1})v(t)\right)\right] +$$

$$+ \; 2E\left[\left(q^{-i-j}v(t)\right)\left(\hat{A}(q^{-1})\hat{C}(q^{-1})v(t)\right)\right]$$

$$\frac{\partial^2 V_2}{\partial \hat{a}_i \partial \hat{c}_j} = 2E\left[\left(q^{-i}\hat{A}(q^{-1})v(t)\right)\left(q^{-j}\hat{A}(q^{-1})v(t)\right)\right]$$

With $\hat{A}(q^{-1})\hat{C}(q^{-1}) = A(q^{-1})C(q^{-1})$ the second term of $\partial^2 V_2/\partial \hat{a}_i \partial \hat{c}_j$ vanishes and $\frac{1}{2}\, V_2''$ becomes the system covariance matrix of

$$\hat{A}(q^{-1})y(t) = \hat{C}(q^{-1})u(t), \quad u(t) = \hat{A}(q^{-1})v(t)$$

From Theorem 2.1 it follows that V_2'' is positive definite.

Q.E.D.

It would be valuable to know, when (NC) is fulfilled in general. However, (NC) is depending on the orders of $\hat{A}(q^{-1})$ and $\hat{C}(q^{-1})$ and the correlation of the noise. Some results for the simple case of first order models are given in Section 3.6.

The concept of (NC) is now used in a theorem of non uniqueness.

Theorem 3.3. Assume that the noise $v(t)$ fulfils (NC). Then there is a number $S_1 > 0$ such that $0 < S \leq S_1$ implies that the loss function V (3.9) has more than one local minimum.

Remark: The result of the theorem holds only for sufficiently small values of the signal to noise ratio. Simulations show, however, see Chapter 4, that the result may be true also for reasonable values of S.

Proof: It will be shown that V has (at least) two local minimum points satisfying

$$\begin{cases} \hat{A}(q^{-1}) = A_i(q^{-1}) + O(S) \\ \hat{C}(q^{-1}) = C_i(q^{-1}) + O(S) \end{cases} \qquad i = 1,\ 2 \tag{3.21}$$

It follows from the proof of Theorem 3.2 that the equations

$$\frac{\partial V}{\partial \hat{b}_i} = 0 \qquad i = 1,\ \ldots,\ n$$

are a system of linear equations in the unknown parameters. The system has always a unique solution, depending on $\hat{a}_i$ and $\hat{c}_i$ but not on S.

Put this solution into the remaining equations.

$$\begin{cases} \dfrac{\partial V}{\partial \hat{a}_i} = 0 \qquad i = 1,\ \ldots,\ n \\[2ex] \dfrac{\partial V}{\partial \hat{c}_i} = 0 \qquad i = 1,\ \ldots,\ n \end{cases} \tag{3.22}$$

(3.22) is now written in the form

$$0 = V_2'(x) + S\tilde{V}_1'(x) \tag{3.23}$$

where it has been assumed that $\sigma^2 = Ee^2(t) = 1$. x denotes the vector $[\hat{a}_1 \ldots \hat{a}_n,\ \hat{c}_1 \ldots \hat{c}_n]^T$.

(NC) implies the existence of two points x_1 and x_2 satisfying

$$\begin{cases} V_2'(x_i) = 0 \\ V_2''(x_i) \text{ positive definite} \end{cases} \qquad i = 1,\ 2 \tag{3.24}$$

From Lemma C.3 it follows that the solutions (3.21) exist.

When the variables are ordered as

$$[\hat{a}_1 \ldots \hat{a}_n \; \hat{c}_1 \ldots \hat{c}_n \; \hat{b}_1 \ldots \hat{b}_n]$$

the matrix of second order derivatives will be

$$V'' = \begin{bmatrix} V_2'(x_i + O(S) & O(S) \\ O(S) & SP \end{bmatrix}$$

where P is a positive definite matrix. From Lemma B.5 it follows that V" is positive definite and that the obtained solutions of V' = 0 are local minimum points.

Q.E.D.

Bohlin [5] has given results, which can be used to test if an estimate is the true maximum likelihood estimate. The test quantity involves sample covariances of $\varepsilon(t)$ and u(t). If, however, the noise level is high this method cannot be used successfully in the case described here. The minimum points of the loss function will give residuals $\varepsilon_1(t)$ and $\varepsilon_2(t)$ satisfying $\varepsilon_1(t) - \varepsilon_2(t) = O(S)$ so also all possible test quantities will differ just a little if S is small.

3.6. Analy
order

The noise
first orde
loss funct

$V_2(\hat{a}_1,\hat{c})$ =

where

$r_\tau = r_w(\tau)$

$w(t) = \frac{H(q}{A(q}$

An analysis

Lemma 3.2.
if and only

$D^* = r_1^2(r_2$

Proof: See

The follow
noise cond
the measure

Example 1:

$w(t) = \frac{}{(1+}$

olutions (3.21)

es will be

‹. From Lemma B.5
:te and that the
ıl minimum points.

Q.E.D.

:an be used to test
.kelihood estimate.
'ariances of $\varepsilon(t)$
l is high this
the case described
function will give
$\varepsilon_1(t) - \varepsilon_2(t) = O(S)$
will differ just a

3.6. Analysis of the "noise condition" (NC) for first order models.

The noise condition (NC) will be closer analysed for first order models in this section. In this case the loss function (3.9) reduces to

$$V_2(\hat{a}_1,\hat{c}) = [1 + (\hat{a}+\hat{c})^2 + \hat{a}^2\hat{c}^2]r_o + $$

$$+ [2(\hat{a}+\hat{c})(1+\hat{a}\hat{c})]r_1 + [2\hat{a}\hat{c}]r_2 \tag{3.25}$$

where

$$r_\tau = r_w(\tau)$$

$$w(t) = \frac{H(q^{-1})}{A(q^{-1})}\, e(t)$$

An analysis of this function is made in

Lemma 3.2. For models of order one (NC) is fulfilled if and only if

$$D^* = r_1^2(r_2-r_o)^2 - 4(r_o^2-r_1^2)(r_1^2-r_o r_2) > 0 \tag{3.26}$$

Proof: See Appendix D.

The following examples illustrate the fact that the noise condition depends on the covariance function of the measurement noise $w(t)$.

Example 1:

$$w(t) = \frac{1}{(1+aq^{-1})(1+cq^{-1})}\, e(t)$$

(NC) is fulfilled if and only if $a \neq c$ (Corr 2 of Def. 3.1).

Example 2:

$$r_2 = 0$$

Then $D^* > 0$ if and only if

$$|r_1| > \frac{\sqrt{3}}{2} r_o$$

For the special structure $w(t) = (1 + cq^{-1})e(t)$ this is never fulfilled.

Example 3:

$$r_1 = 0$$

Then $D^* = 4\, r_o^3 r_2$ and the sign of D^* is equal to the sign of r_2. For the special structure $w(t) = (1 + \gamma q^{-2}) \cdot e(t)$. $D^* > 0$ if and only if $\gamma > 0$, i.e. $(z^2 + \gamma)$ has zeros on the imaginary axis.

Example 4:

$$w(t) = \frac{1 + cq^{-1}}{1 + aq^{-1}}\, e(t)$$

Up to second order terms in a and c D^* is given by

$$D^* = (a - c)(3a + 7c) + \dots$$

This expression indicates that a rather involved relation between a and c determines if (NC) is fulfilled or not.

3.7. Estimates at high signal to noise ratios. Models of too high an order.

Since the true order of a system seldom is known in practice, it is valuable to know what will happen if the model has higher order than the system. In this section it is shown how the result for models of correct order (Theorem 3.2) can be generalized.

The result of Theorem 3.2 can be described as follows. Neglecting terms $O(1/S)$ the (unique) minimum satisfies

i) it minimizes $V_1(\hat{a}_1 \ldots \hat{b}_i \ldots \hat{c}_n)$,

ii) with the remaining degrees of freedom it minimizes $V_2(\hat{a}_1 \ldots \hat{c}_n)$.

If the order of the model is greater than the order of the system it will turn out that there may be more than one minimum point, but the characterization above still applies if local minimum points are concerned under part ii).

<u>Theorem 3.4</u>: Let the system

$$A(q^{-1})y(t) = B(q^{-1})u(t) + v(t), \quad v(t) = H(q^{-1})e(t) \tag{3.27}$$

be controllable and of order n.

Assume that the order of the model is $n+k$, $k > 0$ and that $u(t)$ is persistently exciting of order $2n+k$ Consider parameter estimates in Ω, an arbitrary compact set. Then there is a constant S_o such that if $S_o \leqslant S < \infty$.

i) All local minimum points of the loss function (3.9) fulfil

$$\hat{A}(q^{-1}) = A(q^{-1})L(q^{-1}) + o(1), \quad S \to \infty \tag{3.28}$$

$$\hat{B}(q^{-1}) = B(q^{-1})L(q^{-1}) + o(1), \quad S \to \infty \tag{3.29}$$

where $L(q^{-1}) = 1 + \ell_1 q^{-1} + \ldots + \ell_k q^{-k}$.

Further $L(q^{-1})$ and $\hat{C}(q^{-1})$ fulfil

$$L(q^{-1}) = \bar{L}(q^{-1}) + o(1), \quad S \to \infty \tag{3.30}$$

$$\hat{C}(q^{-1}) = \bar{C}(q^{-1}) + o(1), \quad S \to \infty \tag{3.31}$$

where $(\bar{\ell}_1, \ldots, \bar{\ell}_k, \bar{c}_1, \ldots, \bar{c}_{n+k})$ is a stationary point of

$$V_3(\ell_1, \ldots, \ell_k, c_1, \ldots, c_{n+k}) =$$

$$= E[L(q^{-1})\hat{C}(q^{-1})v(t)]^2 \tag{3.32}$$

The matrix of second order derivatives of V_3 in $(\bar{\ell}_1, \ldots, \bar{c}_{n+k})$ must be positive definite or positive semidefinite.

ii) If the matrix of second order derivatives of V_3 in $(\bar{\ell}_1, \ldots, \bar{c}_{n+k})$ is positive definite, then there <u>exists</u> a <u>unique</u> local minimum point of the form (3.28) - (3.31) and the terms o(1) can be replaced by O(1/S). Further the matrix V" is positive definite in this point.

<u>Proof</u>: See Appendix E.

<u>Remark 1</u>: The number of stationary points of V_3 and the

number of local minimum points of V are coupled to the condition (NC) introduced in Section 3.5.

Remark 2: All possible local minimum points have the property

$$\frac{\hat{B}(q^{-1})}{\hat{A}(q^{-1})} = \frac{B(q^{-1})}{A(q^{-1})} + o(1), \quad S \to \infty$$

Remark 3: If V_3'' is singular in $(\bar{\ell}_1, \ldots, \bar{c}_{n+k})$ nothing general can be stated. In the special case

$$v(t) = \frac{1}{C(q^{-1})} e(t)$$

all points $(\ell_1, \ldots, \ell_k, \hat{c}_1, \ldots, \hat{c}_{n+k})$ satisfying

$$L(q^{-1})\hat{C}(q^{-1}) \equiv C(q^{-1}) \tag{3.33}$$

are global minimum points of V_3 and for some of them V_3'' is singular. This follows from the proof of Corr 2 of Def. 3.2. However, from Lemma 3.1, part i), it follows that the points satisfying (3.33) correspond to global minimum points of the loss function.

3.8. Counter examples to convergence of the second version of GLS.

In this section an example illustrating the possible behaviour of the second version of GLS is described. The question of convergence of this version under suitable conditions cannot be answered easily, and it has not been studied by the author.

The following case will be taken into consideration. The system and the model are both of first order. The iteration is started with the LS estimate of a and b. Conditions for convergence in the next step are examined. If the estimated operator $\hat{C}(q^{-1}) \equiv 1$ then the following estimation of a and b will give the same result as before.

The interesting equations are thus:

$$\hat{c} = -\frac{r_{\varepsilon}(1)}{r_{\varepsilon}(0)} = 0 \tag{3.34}$$

$$\varepsilon(t) = (1 + \hat{a}q^{-1})y(t) - \hat{b}q^{-1}u(t) \tag{3.35}$$

$$\begin{bmatrix} r_y(0) & -r_{yu}(0) \\ -r_{yu}(0) & r_u(0) \end{bmatrix} \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = \begin{bmatrix} -r_y(1) \\ r_{yu}(1) \end{bmatrix} \tag{3.36}$$

<u>Example</u>: Consider the system

$$(1 + aq^{-1})y(t) = u(t) + v(t), \quad v(t) = \frac{1}{1 + cq^{-1}} e(t)$$

where u(t) is white noise. There is a number $S_o > 0$ such that if $0 < S \leqslant S_o$ then (3.34) has two solutions w.r.t.a, which satisfy

$$a = O(S), \qquad S \to 0$$

$$a = -c + O(S), \qquad S \to 0$$

In Appendix F the existence of these solutions are proved.

Note that in this example of first order systems, only systems with a special value of a and a low signal to noise ratio will converge to biased estimates. Nevertheless, the examples indicate that the method may yield "wrong" results. If the iterations procedure is studied in more steps, several more cases of convergence to uncorrect estimates may be detected.

If there is no restriction on the input signal, other examples can be constructed. For example, if the input signal is not persistently exciting of order 2, there are values of a, <u>independent</u> of S, such that the system in the example above will yield biased estimates.

IV. NUMERICAL ILLUSTRATION.

4.1. Introduction.

The theory of the GLS method in Chapter 3 requires an infinite number of data. For practical purposes it is interesting to know if the result holds with "good approximation" for a finite number of data.

The loss function (3.9) is a polynomial in the variables $\hat{a}_1, \ldots, \hat{a}_n, \hat{b}_1, \ldots, \hat{b}_n, \hat{c}_1, \ldots, \hat{c}_n$. The coefficients are different sample covariances, which converge with probability one to the corresponding covariances, as $N \to \infty$. A sufficiently small deviation of the coefficients from their limits can only move the minimum points a little bit, but the probability for a drastical change of the character of the loss function is very small.

This means that for a "sufficiently large" number of data the results of Chapter 3 will hold with probability close to one. However, it is not practically possible to analyze what sufficiently large exactly means.

In order to examine the situation of a finite number of data simulations were used. These simulations are illustrating the results of Chapter 3 as well.

The simulations were carried out on a UNIVAC 1108. A description of the used programs is given in Appendix G.

The results of the simulations are presented in the next sections.

ter 3 requires an
al purposes it is
lds with "good app-
ata.

ial in the vari-
..., $\hat{c}_n$. The coef-
ances, which con-
rresponding cova-
all deviation of
an only move the
probability for
of the loss func-

large" number of
old with probabi-
t practically
y large exactly

a finite number
simulations are
3 as well.

a UNIVAC 1108. A
given in Appendix

resented in the

All the simulated systems were generated by the equation

$$A(q^{-1})y(t) = B(q^{-1})u(t) + v(t)$$

$$v(t) = H(q^{-1})e(t)$$

The number of samples were 500 in all cases and the input signal was a PRBS with amplitude 1.0.

4.2. Illustration of Theorem 3.2.

These examples are intended to demonstrate that when the conditions of Theorem 3.2 are fulfilled there is a solution, which satisfies

$$\hat{a}_i \approx a_i$$

$$\hat{b}_i \approx b_i$$

$$\hat{c}_i \approx \bar{c}_i$$

where $\bar{C}(q^{-1})$ corresponds to the minimum point of

$$E[\hat{C}(q^{-1})v(t)]^2$$

The following systems were studied.

Table 4.1 - Generated systems.

System	a_1, (a_2)	b_1, (b_2)	$H(q^{-1})$	$Ee^2(t)$
S1	-0.8	1.0	$\frac{1}{1 + 0.7q^{-1}}$	1.0
S2	-0.8	1.0	$(1 + 0.7q^{-1})$	0.01
S3	-0.8	1.0	$(1-1.0q^{-1}+0.2q^{-2})$	0.01
S4	-1.5 0.7	1.0 0.5	$(1 + 0.7q^{-1})$	0.01

The results of the identifications are given in Table 4.2. The iterations were started with the LS estimation of the a_i and the b_i parameters. The results are very well in accordance with the expectations.

Table 4.2 - Identification results.

System	$\hat{a}_1$, ($\hat{a}_2$)	$\hat{b}_1$, ($\hat{b}_2$)	$\hat{c}_1$, ($\hat{c}_2$)	$\bar{c}_1$, ($\bar{c}_2$)
S1	-0.804	1.010	0.697	0.7
S2	-0.803	1.000	-0.449	-0.469
	-0.803	1.003	-0.607 0.352	-0.603 0.284
S3	-0.799	1.005	0.555	0.589
S4	-1.505 0.704	1.001 0.498	-0.444	-0.469

4.3. Illustration of Theorem 3.3.

For the following systems 500 samples were generated with a PRBS as input signal. The iterations were started with expected values of $\hat{c}_i$.

The system S7 requires a comment. The equation

$$V' = V'(\hat{\theta}, \sigma) = 0 \tag{4.1}$$

where $\hat{\theta}^T = [\hat{a}_1 \ldots \hat{b}_i \ldots \hat{c}_n]$ and $\sigma^2 = Ee^2(t)$ was solved (using analytic expressions for the covariances) with successively decreasing values of the parameter σ. A change $d\sigma$ of σ causes a change in $\hat{\theta}$ approximately

$$d\hat{\theta} = - V''(\hat{\theta}, \sigma)^{-1} \frac{\partial}{\partial \sigma} V'(\hat{\theta}, \sigma) d\sigma$$

Starting with this new value of $\hat{\theta}$ (4.1) was solved with respect to $\hat{\theta}$ by Newton-Raphson technique. This procedure of computing solutions for different, decreasing values of σ stops when V'' is not positive definite or when $\hat{\theta} \approx \theta_o$ is obtained as solution.

Table 4.3 - Generated systems.

System	a_1	b_1	$H(q^{-1})$	$Ee^2(t)$
S5	-0.8	1.0	$\frac{1}{1 - 0.2q^{-1}}$	100.0
S6	0.0	1.0	$(1 + 0.7q^{-2})$	100.0
S7 (=S1)	-0.8	1.0	$\frac{1}{1 + 0.7q^{-1}}$	1.0

The results presented in Table 4.4.coincide with the predicted values. The last system shows that it is not necessary that the noise has unrealistic high variance for Theorem 3.3 to hold. The expected value of $\hat{b}_1$ is computed from the equation

$$\frac{\partial}{\partial \hat{b}_1} V(\hat{a}_1,\hat{b}_1,\hat{c}_1) = 0$$

where the values of $\hat{a}_1$ and $\hat{c}_1$ are inserted.

Table 4.4 - Identification results.

System	$\hat{a}_1$	$\hat{b}_1$	$\hat{c}_1$	Expected values of $\hat{a}_1$	$\hat{b}_1$	$\hat{c}_1$
S5	-0.774	1.051	-0.233	-0.8	1.0	-0.2
	-0.243	0.740	-0.767	-0.2	1.0	-0.8
S6	-0.676	0.705	0.676	-0.69	0.68	0.69
	0.674	0.882	-0.678	0.69	0.68	-0.69
S7	-0.804	1.010	0.697	-0.8	1.0	0.7
	0.327	0.461	-0.771	0.35	0.44	-0.81

4.4. Illustration of Theorem 3.4.

The illustration of Theorem 3.4 has turned out for the author to be more difficult than the previous examples. The reason for this difficulty is probably that the properties (as existence of several minimum points) of the loss function are rather sensitive for the number of data and the realization. This fact is also the reason why the examples in this section require more iterations for convergence.

Analogously to the previous examples the iterations were started with the expected values of the $\hat{c}_i$ parameters.

Table 4.5 - Generated systems.

System	a_1	b_1	$H(q^{-1})$	$Ee^2(t)$
S8	-0.8	1.0	$(1 + 0.8q^{-2})$	0.01
S9	-0.4	1.0	$\frac{1}{(1 - 0.8q^{-1})(1 + 0.8q^{-1})}$	1.0

The result of the identifications (see Table 4.6) are well coinciding with the theory.

Table 4.6 - Identification results.

System	From simulation $\hat{a}_1$	$\hat{a}_2$	$\hat{b}_1$	$\hat{b}_2$	$\hat{c}_1$	$\hat{c}_2$
S8	-0.94	0.11	1.00	-0.14	0.11	-0.46
	-0.06	-0.60	1.00	0.74	-0.78	0.11
	-1.39	0.47	1.00	-0.59	0.55	-0.15
S9	-0.45	0.06	0.97	-0.04	0.03	-0.66
	0.40	-0.35	1.00	0.79	-0.82	0.05
	-1.11	0.25	1.01	-0.69	0.69	-0.11
	Expected values					
S8	-0.80	0.00	1.00	0.00	0.00	-0.45
	-0.13	-0.54	1.00	0.67	-0.67	0.00
	-1.47	0.54	1.00	-0.67	0.67	0.00
S9	-0.40	0.00	1.00	0.00	0.00	-0.64
	0.40	-0.32	1.00	0.80	-0.80	0.00
	-1.20	0.32	1.00	-0.80	0.80	0.00

4.5. Illustration of Section 3.8.

The following examples illustrate that the second version of GLS can converge to "wrong" values of the estimates. The third example, System S12, is constructed in a way similar to System S7. Of course, there is another equation to be solved. (More exactly $\hat{c}(a,c_1\sigma) = 0$ is solved with respect to a for decreasing values of the parameter σ and with fixed value of the parameter c.)

Table 4.7 - Generated systems.

System	a_1	b_1	$H(q^{-1})$	$Ee^2(t)$
S10	-0.5	1.0	$\frac{1}{1 + 0.5q^{-1}}$	100.0
S11	0.0	1.0	$\frac{1}{1 - 0.8q^{-1}}$	100.0
S12	-0.7	1.0	$\frac{1}{1 + 0.9q^{-1}}$	1.2

The results of the identifications, given in Table 4.8, confirm the theory. $\sigma_{\hat{c}_1}$ denotes the estimated standard deviation of $\hat{c}_1$. The PRBS which is used as input signal is with "good approximation" white noise.

Table 4.8 - Identification results.

System	$\hat{a}_1$	$\hat{b}_1$	$\hat{c}_1$	$\sigma_{\hat{c}_1}$	Exp. values of $\hat{a}_1$	$\hat{b}_1$	$\hat{c}_1$
S10	-0.01	0.98	0.007	0.045	0.0	1.0	0.0
S11	-0.79	1.08	-0.026	0.045	-0.8	1.0	0.0
S12	0.16	0.94	0.043	0.045	0.09	1.0	0.0

V. EXAMPLES OF LACK OF UNIQUENESS FOR INDUSTRIAL DATA.

5.1. Introduction.

In this chapter identification results using the GLS method of real data are presented. The main purpose of the identifications was to investigate the possible existence of more than one minimum point of the loss function. A straight forward application of a test of order [3] would in general result in more complex models. However, the orders of models in the presented cases are not unreasonable.

The results of the identifications are compared with models obtained with the "ordinary" maximum likelihood model

$$\hat{A}(q^{-1})y(t) = \hat{B}(q^{-1})u(t) + \hat{C}(q^{-1})e(t) \qquad (5.1)$$

It is to be noted that for a "wrong" minimum point the covariance matrix

$$\frac{2V}{N} V''^{-1}$$

of the parameter estimates has dubious meaning.

5.2. Identification of dynamics of a heat rod process.

The system is a copper rod, which acts as a one dimensional heat diffusion process. The system is located at Div. of Automatic Control, Lund Institute of Technology. Identification results using the ML model (5.1) as well as a short description of the process are given in [14]. (The data used here is Serie S1, output $x = 3\ell/4$.)

The test quantity for comparing models of orders 4 and 5, [3], is $F(862,3)$ and has the value 109. Since the ML identification [14] indicates a model of order 4 as reasonable, this order was considered in spite of the great value of the test quantity.

The loss function turns out to have (at least) two minimum points for fourth order models. The results are presented in Tables 5.1 - 5.2 and Figures 5.1 - 5.4.

The theoretical value of the static gain is 0.25, which indicates that model 1 is the most correct one.

In Figures 5.1, 5.2 the following signals are plotted:

1. the input $u(t)$,
2. the output $y(t)$,
3. the model output $y_m(t) = \dfrac{\hat{B}(q^{-1})}{\hat{A}(q^{-1})} u(t)$
4. the model error $e_m(t) = y(t) - y_m(t)$
5. the residuals $\varepsilon(t)$.

In Figures 5.3, 5.4 normalized covarious functions are plotted. The criterion by Bohlin [5] can be formulated as: the estimate is true if and only if

$r_\varepsilon(\tau) = 0$

$r_{\varepsilon u}(\tau) = E\varepsilon($

The second c

$r_{e_m} u(\tau) = 0$

Discussion o

Already from
expected tha
is very much

A comparison
shows little

In the model
due to noise

From Figure
are not whit
nal and the
for the seco

a heat rod process.

cts as a one dimen-
system is located at
itute of Technology.
model (5.1) as well
s are given in [14].
put $x = 3\ell/4$.)

els of orders 4 and
e 109. Since the
odel of order 4 as
d in spite of the

(at least) two mini-
The results are pre-
es 5.1 - 5.4.

gain is 0.25, which
orrect one.

gnals are plotted:

) u(t)
)

$y_m(t)$

ious functions are
can be formulated
if

$$r_\varepsilon(\tau) = 0 \qquad \tau > 0$$

$$r_{\varepsilon u}(\tau) = E\varepsilon(t)u(t+\tau) = 0 \quad \text{all } \tau$$

The second condition can also be written as

$$r_{e_m u}(\tau) = 0 \quad \text{all } \tau$$

Discussion of the results:

Already from the values of the static gain it can be expected that model 1 is superior to model 2. This fact is very much confirmed by the plotted signals.

A comparison with plots of the ML model (see [14]) shows little difference between that model and model 1.

In the model 2 the output is "interpreted" as mainly due to noise.

From Figure 5.3 and 5.4 it is seen that the residuals are not white in any of the two models. The input signal and the residuals are considerably more correlated for the second model.

	Model 1	Model 2	Corresponding model in [14]
$\hat{a}_1$	-2.4307±0.0424	-1.2374±0.0510	-2.9563±0.0017
$\hat{a}_2$	1.8776±0.1118	0.6056±0.0888	3.2694±0.0049
$\hat{a}_3$	-0.37271±0.0999	-0.5161±0.0723	-1.6134±0.0047
$\hat{a}_4$	-0.071016±0.0303	0.3525±0.0332	0.3025±0.0015
$\hat{b}_1 \cdot 10^3$	0.14908±0.0559	-0.6393±0.0664	0.0
$\hat{b}_2 \cdot 10^3$	-0.016125±0.1233	-0.3379±0.0788	0.1297±0.0124
$\hat{b}_3 \cdot 10^3$	-0.84271±0.1261	-0.7060±0.0719	-0.4166±0.0252
$\hat{b}_4 \cdot 10^3$	1.5127±0.0822	-0.4971±0.0775	0.7942±0.0138
$\hat{c}_1$	1.3882±0.0516	-0.6593±0.0528	Comparison is impossible
$\hat{c}_2$	1.2794±0.0468	-1.0558±0.0622	
$\hat{c}_3$	0.95739±0.0504	0.2150±0.0486	
$\hat{c}_4$	0.39324±0.0330	0.5017±0.0464	
V	$2.16 \cdot 10^{-7}$	$5.13 \cdot 10^{-7}$	$0.92 \cdot 10^{-7}$
$\hat{\sigma}$	$0.658 \cdot 10^{-3}$	$1.013 \cdot 10^{-3}$	$0.428 \cdot 10^{-3}$

Table 5.1 - Parameter estimates from GLS identification of the heat rod.

Poles	
Zeros	
Static gain	

Table 5.2 -

2	Corresponding model in [14]
0510	-2.9563±0.0017
0888	3.2694±0.0049
0723	-1.6134±0.0047
0332	0.3025±0.0015
0664	0.0
0788	0.1297±0.0124
0719	-0.4166±0.0252
0775	0.7942±0.0138
0528	Comparison is impossible
0622	
0486	
0464	
-7	$0.92 \cdot 10^{-7}$
0^{-3}	$0.428 \cdot 10^{-3}$

GLS identification of

	Model 1	Model 2	Corresponding model in [14]
Poles	-0.435	-0.189+i 0.672	0.602+i 0.184
	0.810+i 0.140	-0.189-i 0.672	0.602-i 0.184
	0.810-i 0.140	0.806+i 0.269	0.810
	0.952	0.806-i 0.269	0.943
Zeros	-1.166	0.064+i 1.089	-1.606+i 1.883
	1.442+i 0.827	0.064-i 1.089	-1.606-i 1.883
	1.442-i 0.827	-0.652	-
Static gain	0.2528	-0.0106	0.2440

Table 5.2 - Poles, zeros and static gain of the models of the heat rod process.

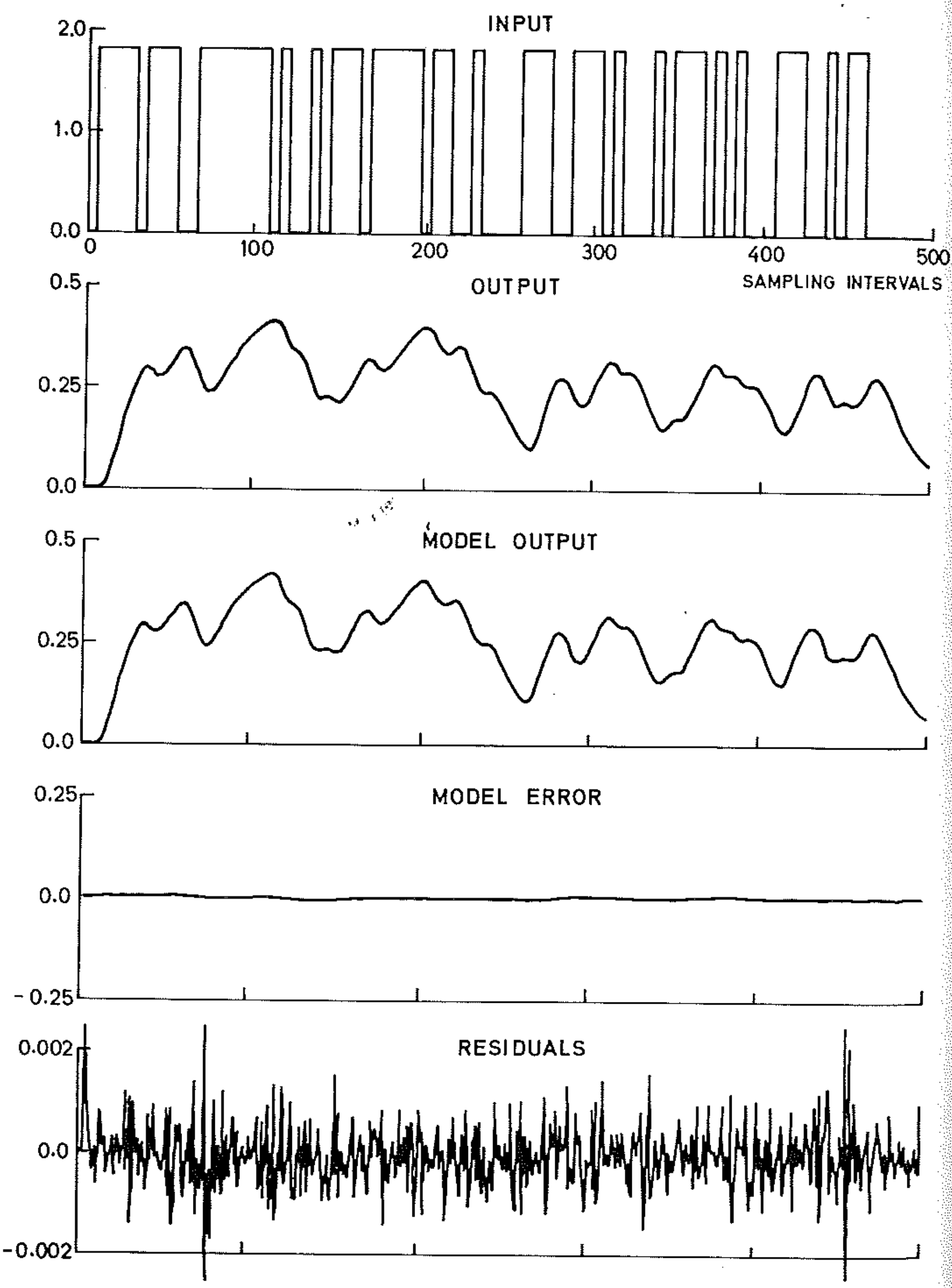


Fig. 5.1 - Model 1 of the heat-rod process. All variables are given in ^{o}C. (Constants are added to the input, the output and the model output.) The sampling period is 10 sec.

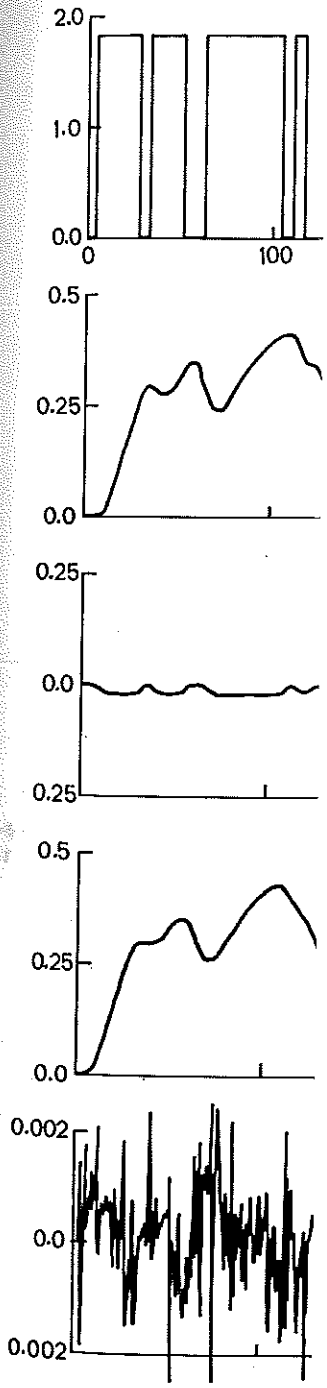
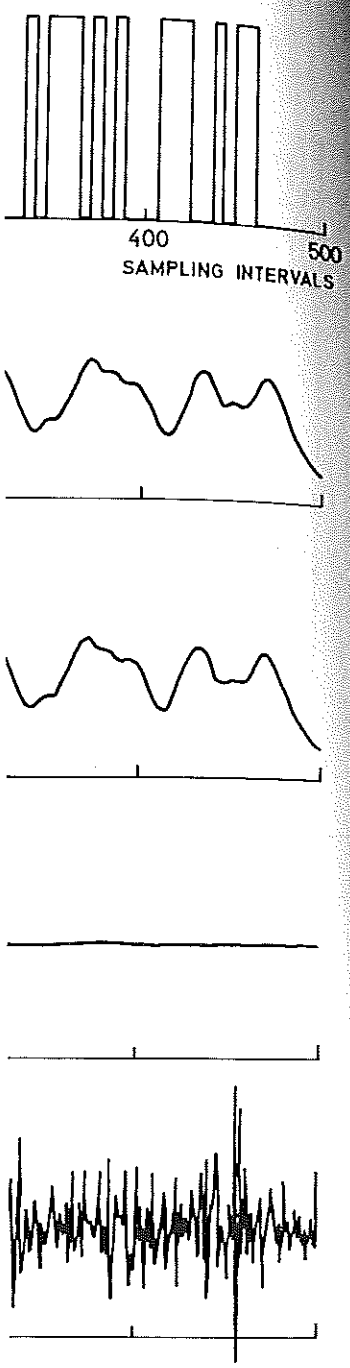


Fig. 5.2 - Mo
(C
ou



.l variables are given in ^{o}C.
the output and the model
sec.

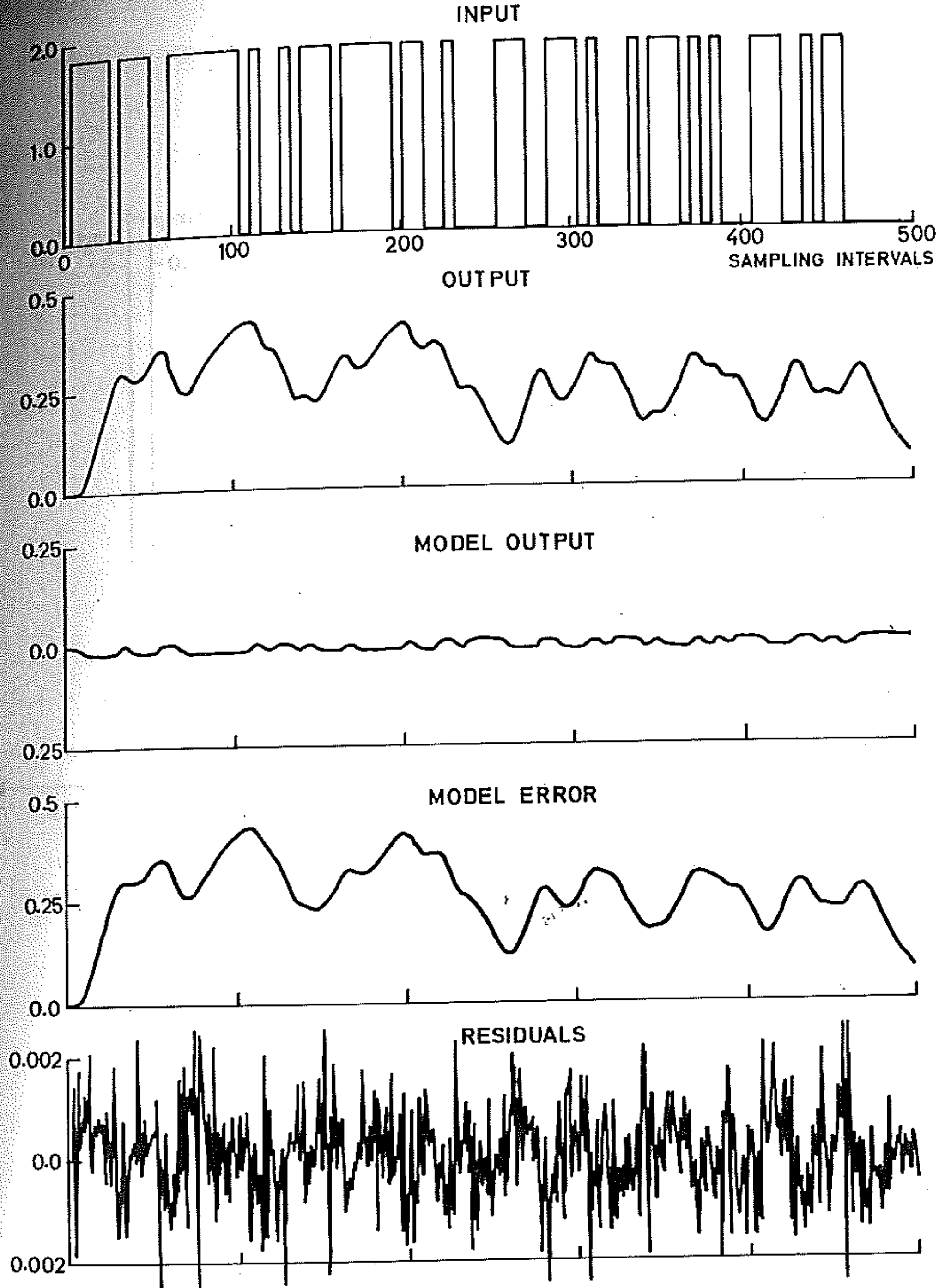


Fig. 5.2 - Model 2 of the heat-rod process. All variables are given in ^{o}C. (Constants are added to the input, the output and the model output.) The sampling period is 10 sec.

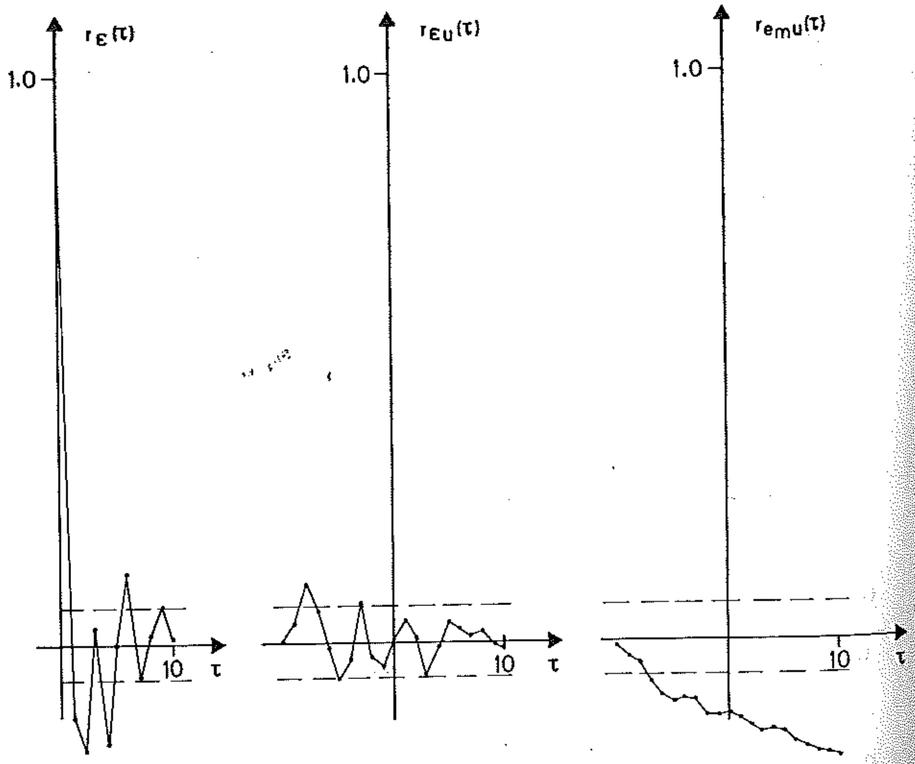


Fig. 5.3 - Normalized sample covariance functions for the heat-rod model 1.
The dashed lines give the 5% confidence interval.
The time is given in sampling periods.

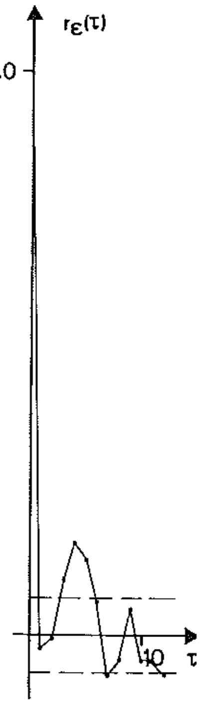
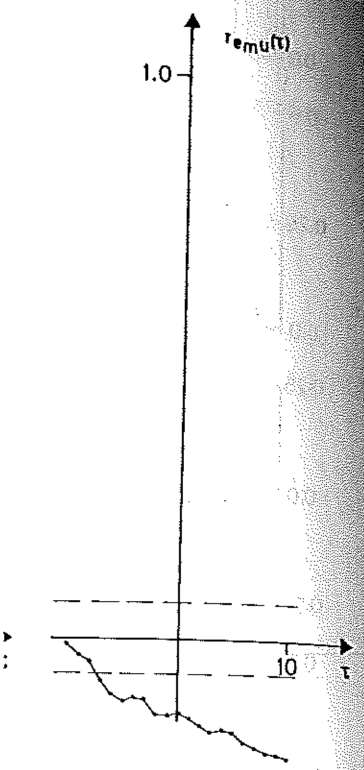


Fig. 5.4 - No
he
Th
Th



ce functions for the

5% confidence inter-

ing periods.

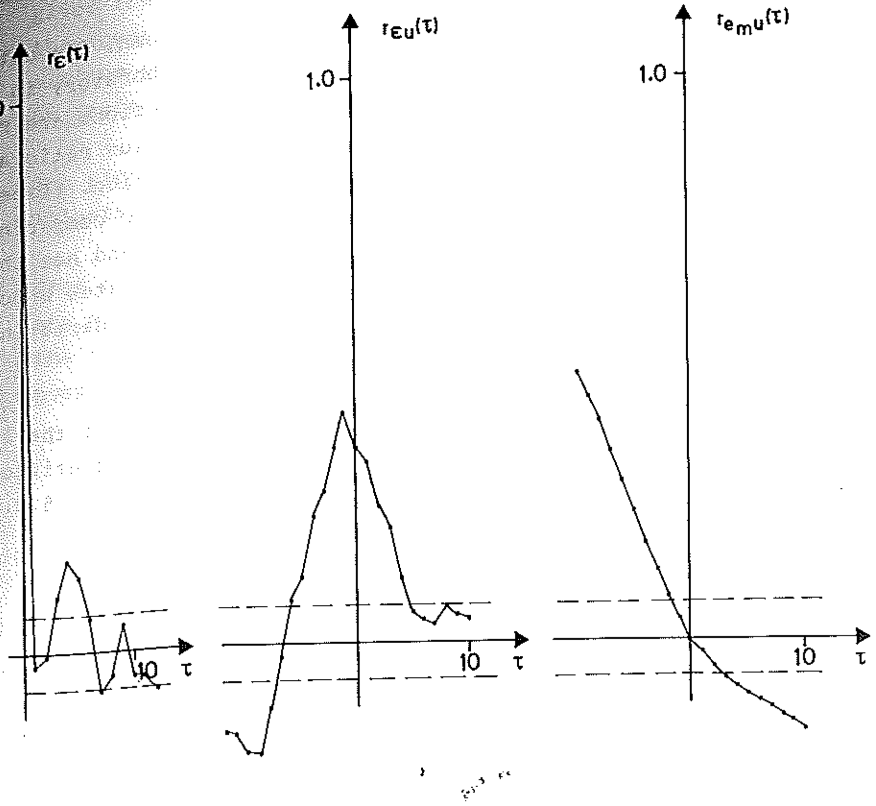


Fig. 5.4 - Normalized sample covariance functions for the heat-rod model 2.
The dashed lines give 5% confidence interval.
The time is given in sampling periods.

5.3. Identification of dynamics of a distillation column.

The system is a binary distillation column. The data have been received from National Physical Laboratory in London. Results of maximum likelihood identifications are reported in [12]. The input signal is the reflux ratio and the output signal is the top product composition. (Experiment 4B, [12], was used.) The test quantity for comparing models of orders 2 and 3, [3], is F(240,3) and has the value 36. Since the ML identification [12] indicates a model of order 2 as reasonable, this order was considered in spite of the great value of the test quantity.

The second order models two minimum points of the loss function were found. The results from the identification are given in Tables 5.3, 5.4 and Figures 5.5 - 5.8.

Discussion of the result.

From Table 5.3 it is seen that $\hat{C}(q^{-1})$ of model 2 is very like $\hat{A}(q^{-1})$ of model 1. With Theorem 3.3 in mind, this is not astonishing.

The model from [12] is very like the model 1, which means that the noise can be well modelled as

$$v(t) = \hat{C}_{ML}(q^{-1})e(t)$$

as well as

$$v(t) = \frac{1}{\hat{C}_{GLS}(q^{-1})} e(t)$$

with e(t) white noise.

The values
gives the b
the lower v
minimum poi
be preferre

The plots o
illustratio

From Figure
are most wh
the input f
possible) t

distillation column.

column. The data have
Laboratory in London.
fications are repor-
reflux ratio and the
position. (Experiment
ty for comparing mo-
0,3) and has the value
indicates a model of
s considered in spite
ity.

points of the loss
m the identification
ures 5.5 - 5.8.

) of model 2 is very
3.3 in mind, this is

model 1, which means
as

The values of the static gain indicate that model 1 gives the best description of the process. Also from the lower value of loss function at the corresponding minimum point, it can be expected that this model is to be preferred.

The plots of the results, Figures 5.5 - 5.6, are a nice illustration of the expected differences.

From Figures 5.7 and 5.8 it is noted that the residuals are most white for model 2 and most uncorrelated with the input for model 1. That means it is hard (or impossible) to choose the "best" model from these figures.

Parameter	Model 1	Model 2	Corresp. ML model in [12]
$\hat{a}_1$	-1.5275±0.0187	0.1865±0.0820	-1.5369±0.0180
$\hat{a}_2$	0.5473±0.0188	-0.0227±0.0529	0.5535±0.0180
$\hat{b}_1$	0.2447±0.0193	0.3730±0.0218	0.2251±0.0190
$\hat{b}_2$	-0.6164±0.0194	0.2174±0.0317	-0.5979±0.0214
$\hat{c}_1$	0.8213±0.062	-1.5076±0.0754	Comparison
$\hat{c}_2$	0.4088±0.059	0.5358±0.0747	impossible
V	69.04	164.37	71.68
$\hat{\sigma}$	11.75	18.13	11.97

Table 5.3 - Parameter estimates from GLS identification of the distillation column data

	Model 1	Model 2	Corresponding model in [12]
Poles	0.95	-0.093+i 0.118	0.96
	0.57	-0.093-i 0.118	0.58
Zero	2.52	-0.584	2.66
Static gain	-18.77	0.507	-22.46

Table 5.4 - Poles, zero and static gain of the GLS models of the distillation column.

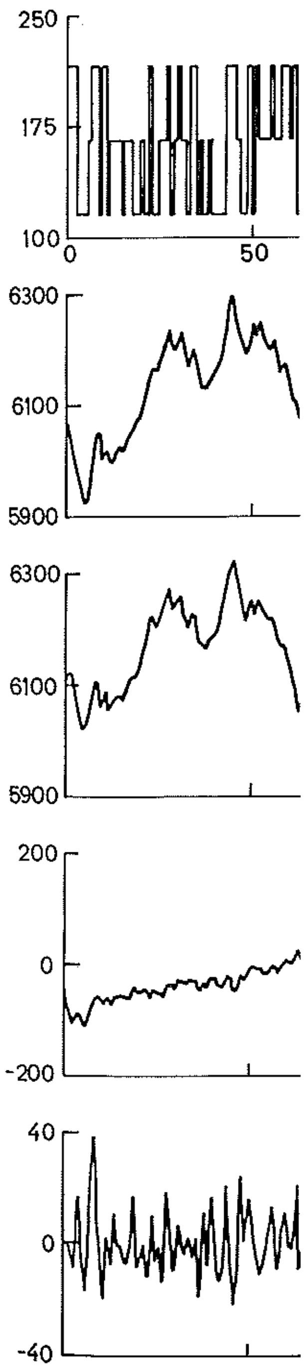


Fig. 5.5 -

2	Corresp. ML model in [12]
0820	-1.5369±0.0180
0529	0.5535±0.0180
0218	0.2251±0.0190
0317	-0.5979±0.0214
0754	Comparison
0747	impossible
	71.68
	11.97

GLS identification of
data

2	Corresponding model in [12]
.118	0.96
.118	0.58
	2.66
	-22.46

ain of the GLS models
mn.

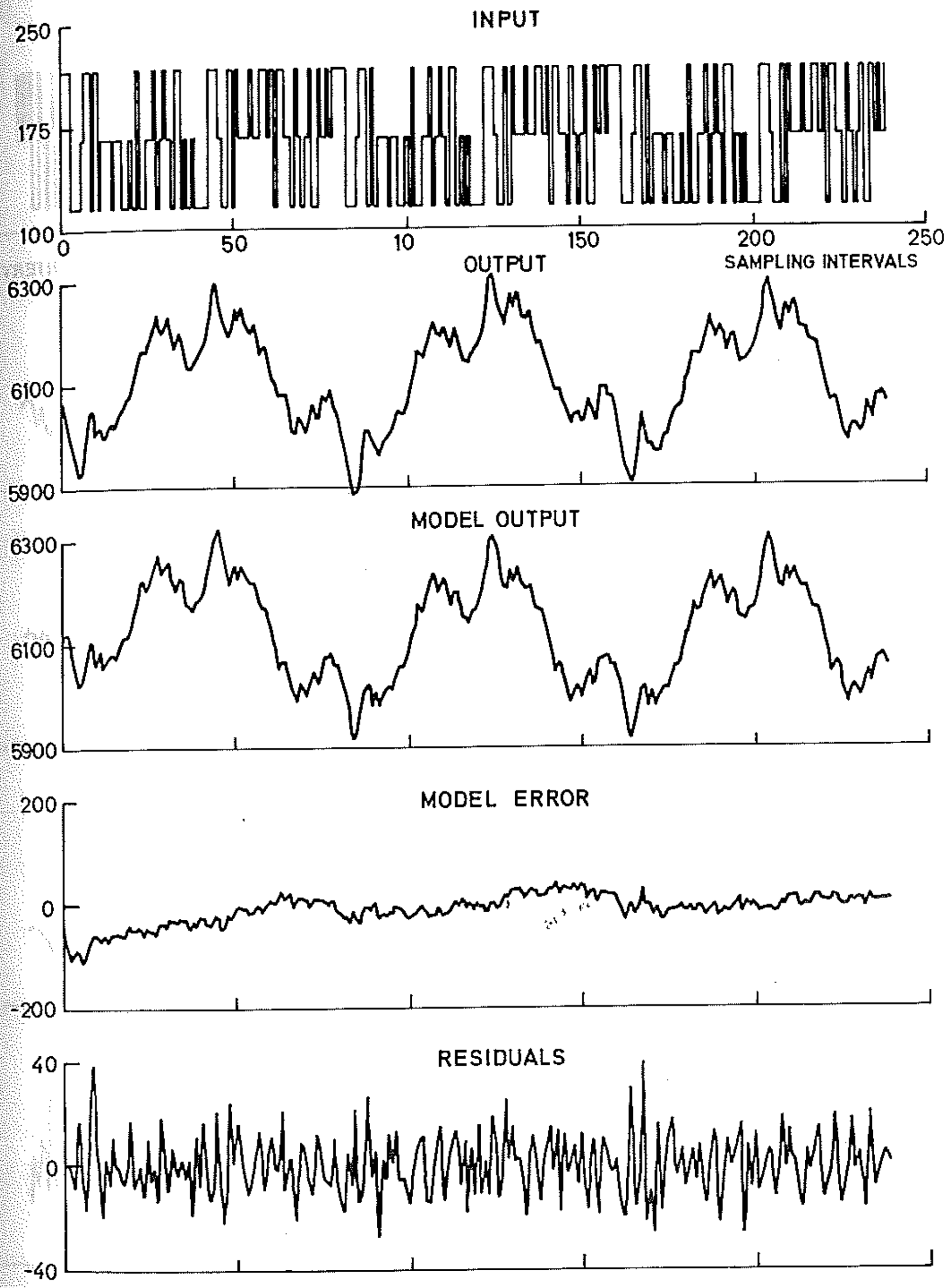


Fig. 5.5 - Model 1 of the distillation column. Digital units are used. The sampling period is 96 sec.

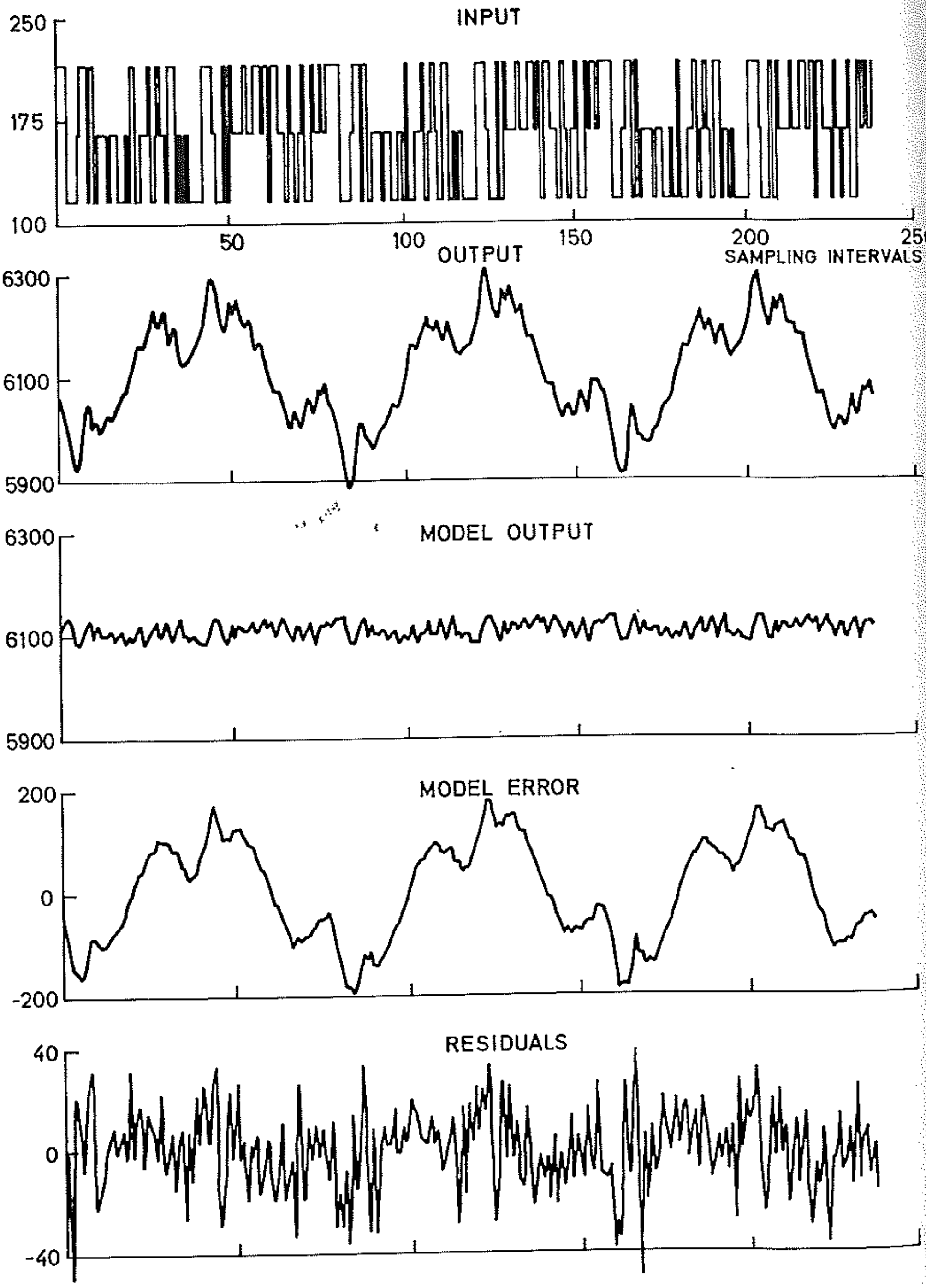


Fig. 5.6 - Model 2 of the distillation column. Digital units are used. The sampling period is 96 sec.

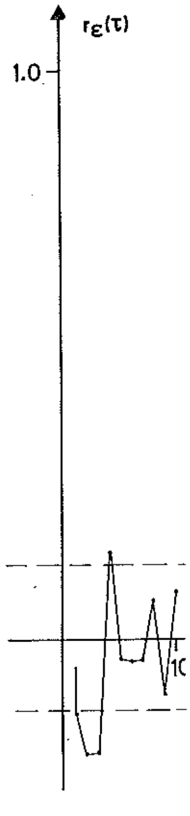
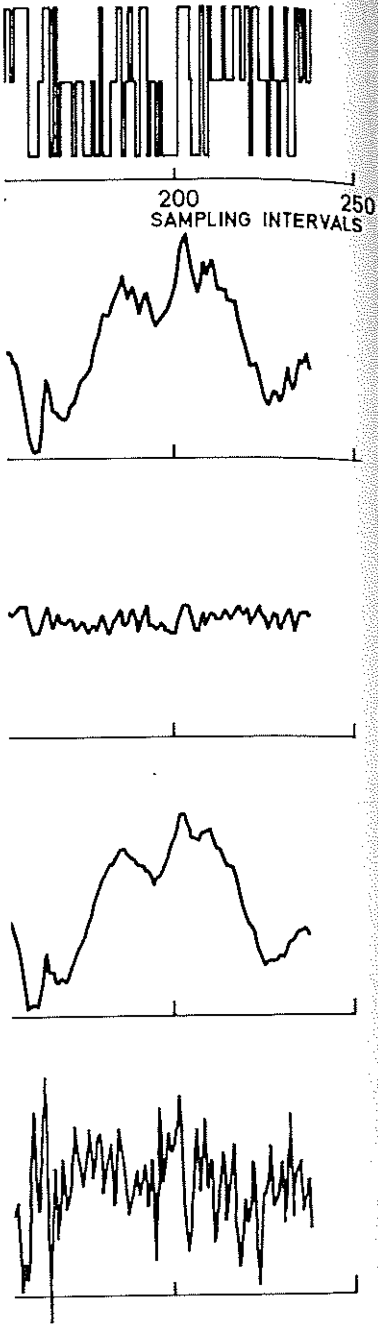


Fig. 5.7 -



n column. Digital
ng period is 96 sec.

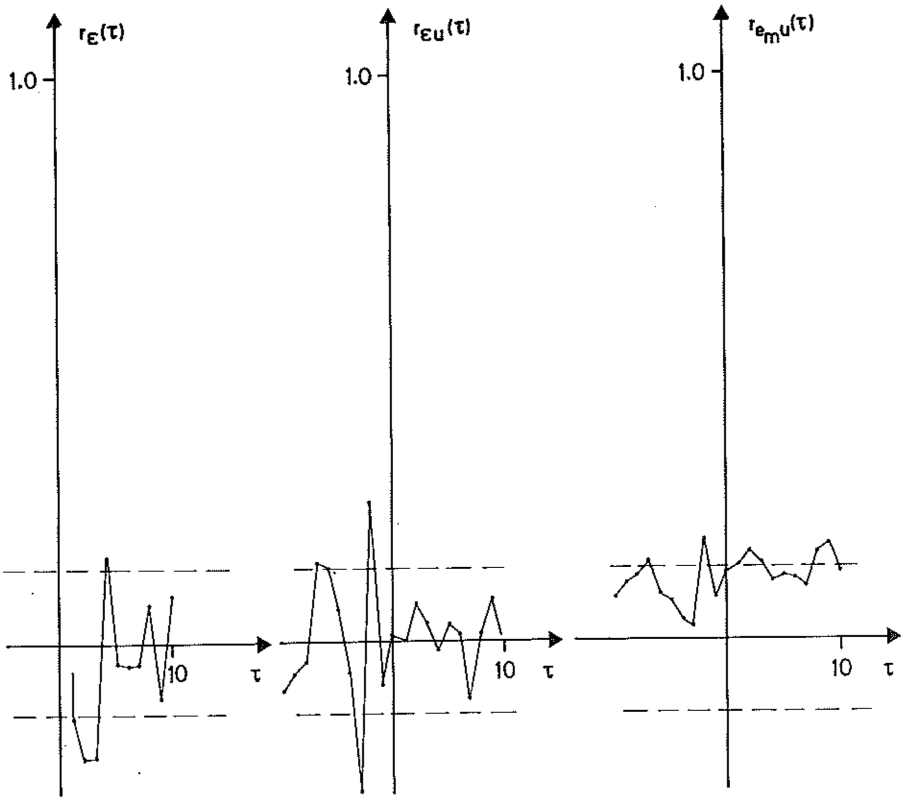


Fig. 5.7 - Normalized sample covariance functions for the distillation column, model 1.
The dashed lines give the 5% confidence interval.
The time is given in sampling periods.

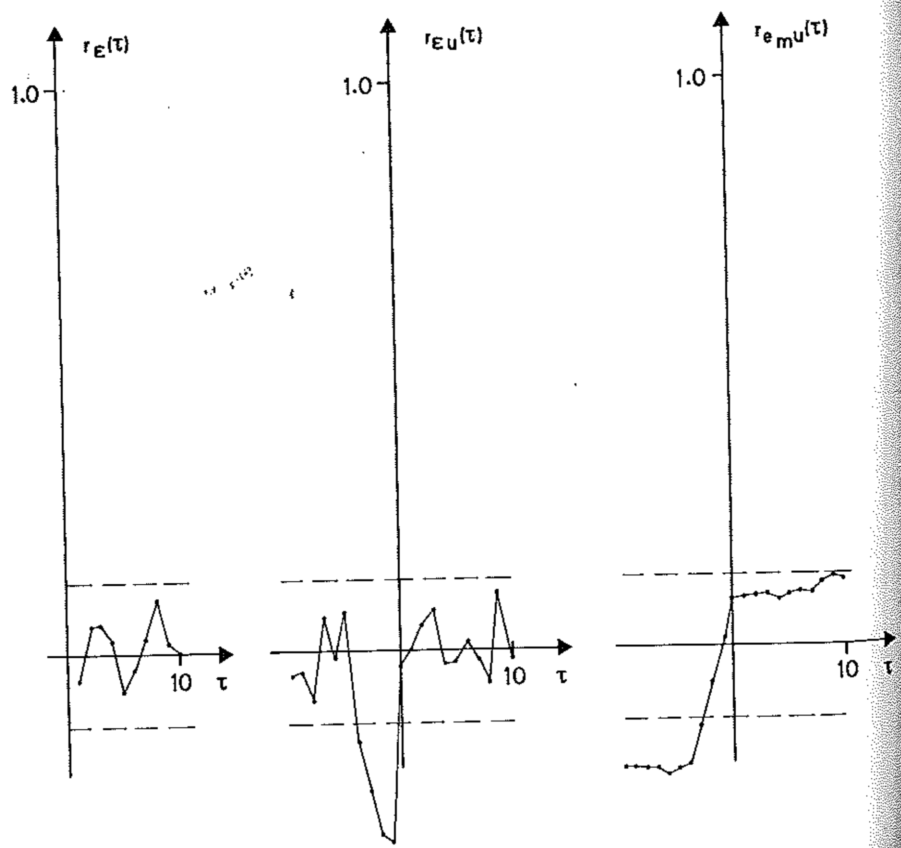
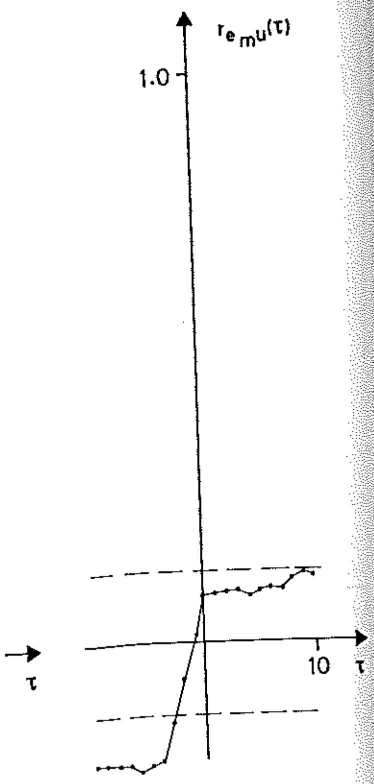


Fig. 5.8 - Normalized sample covariance functions for the distillation column, model 2.
The dashed lines give the 5% confidence interval.
The time is given in sampling periods.



ance functions for the
el 2.
e 5% confidence inter-

pling periods.

5.4. Identification of dynamics of a nuclear reactor.

The system is a nuclear reactor where the input is reactivity created by control rod movement and the output is the nuclear power, measured by fission chamber. Measurements have been received from OECD Halden Reactor Project in Norway.

The experiment is described in [16] and is called RUN 11 EP 714B. The first 1000 data were used.

The system contains a direct term. This is easily estimated by shifting the input signal. The used $\hat{B}(q^{-1})$ polynomials were of the form

$$\hat{B}(q^{-1}) = \hat{b}_o + \ldots + \hat{b}_n q^{-n}$$

Test quantities for comparing order are F(1000,3). The value when models of orders 1 and 2 are compared is 11.4, while the value is 1.1 when models of orders 2 and 3 are compared. Thus the order two seems to be good.

Two minimum points of the loss function were found for this order. The result of the identifications is given in Tables 5.5, 5.6 and Figures 5.9 - 5.13.

ML identification using the model (5.1) has been done [7], [16].

Discussion of the result.

It is seen from the figures that the differences between the models are small. Further $(\hat{a}_1, \hat{a}_2, \hat{c}_1, \hat{c}_2)$ of model 1 is close to $(\hat{c}_1, \hat{c}_2, \hat{a}_1, \hat{a}_2)$ of model 2. In fact, both models as well as the model in [7] may be simplified to a first order system

$$y(t) = 2.4(1 + 2.6q^{-1})u(t) + \frac{1}{1 - 0.9q^{-1}} e(t) \qquad (5.2)$$

if approximate factors in common and small zeros are omitted.

An identification of a first model gave the result:

$$y(t) = \frac{2.396 + 6.234q^{-1}}{1 - 0.00012q^{-1}} u(t) + \frac{1}{1 - 0.918q^{-1} + 0.0001q^{-2}} e(t)$$

and $\lambda = 2.6660 \cdot 10^{-2}$ which differs just a little from the simplified model.

Since the two models do not differ very much it is impossible to call any of them the "best" or most "correct" one.

If (5.2) is an adequate description of the dynamical behaviour of the process then it is expected with Theorem 3.4 in mind, that there will be (at least) two different but equivalent models of second order. The models obtained by identification are in fact close to these expected models. Of course, this is a very loose discussion according to the assumption that (5.2) describes the system adequately enough.

Parameter	M
$\hat{a}_1$	-0.
$\hat{a}_2$	0.
$\hat{b}_o$	2
$\hat{b}_1$	5
$\hat{b}_2$	-1
$\hat{c}_1$	-0.
$\hat{c}_2$	-0.
$V \cdot 10^4$	
$\hat{\sigma} \cdot 10^2$	

Table 5.5 - F
t

	Mo
Poles	0.088
	0.088
Zeros	-
Static gain	

Table 5.6 - P
of

have more than one minimum point. In this case the result of the GLS identification depends on the start values of the parameter estimates. The existence of several minimum points can be shown theoretically for low signal to noise ratios. In practice it can happen also for reasonable values of this ratio. It is not always easy without intimate knowledge of the actual process to decide which of the models that will be the "best" or most "correct".

VII. ACKNOWLEDGEMENTS.

The author wants to express his great gratitude to Prof. K.J. Åström for suggesting the subject and for his valuable guidance.
He also wants to thank his colleagues, tekn.lic. Ivar Gustavsson and civ.ing. Lennart Ljung, for many stimulating discussions.
It is a pleasure to thank Mrs. G. Christensen, who typed the manuscript, and Mrs. B. Tell, who drew the figures.
The author is grateful for the measurements, which were supplied by tekn.lic. Bo Leden and National Laboratory, London, and OECD Halden Reactor Project.

VIII. REFERENCES.

[1] K.J. Åström: Lectures on the Identification Problem - the Least Squares Method. Report 6806, 1968, Division of Automatic Control, Lund Institute of Technology.

[2] K.J. Åström: Introduction to Stochastic Control Theory. Academic Press, 1970.

[3] K.J. Åström and T. Bohlin: Numerical Identification of Linear Dynamic Systems from Normal Operating Records. Paper, IFAC Symposium on Theory of Self-Adaptive Systems, Teddington, England. In Theory of Self-Adaptive Control Systems (Ed. P.H. Hammond), Plenum Press, New York, 1966.

[4] K.J. Åström and P. Eykhoff: System Identification - A Survey. Automatica 7, 123 - 162, 1971.

[5] T. Bohlin: On the Problem of Ambiguites in Maximum Likelihood Identification. Automatica 7, 199 - - 210, 1971.

[6] P.E. Caines: The Parameter Estimation of State Variable Models of Multivariable Linear Systems. Control Systems Centre Report No. 146, The University of Manchester, Institute of Science and Technology, 1971.

[7] S. Carlsson: Maximum Likelihood identifiering av reaktordynamik från flervariabla experiment. Master Thesis, Division of Automatic Control, Lund Institute of Technology.

[8] D.W. Clarke: Generalized Least Squares Estimation of the Parameters of a Dynamic Model. 1st IFAC Symposium on Identification in Automatic Control Systems, Prague, 1967.

[9] I.H. Cramér and M.R. Leadbetter: Stationary and Related Stochastic Processes. John Wiley & Sons, New York, 1967.

[10] D.K. Faddeev and V.N. Faddeeva: Computational Methods of Linear Algebra. W.H. Freeman and Company, San Fransisco, 1963.

[11] B.V. Gnedenko: The Theory of Probability. Chelsea Publishing Company, New York, 1963.

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

[13] E. Kreindler and A. Jameson: Conditions for Nonnegativeness of Partioned Matrices. IEEE Trans. Aut. Control, Feb., 1972, 147 - 148.

[14] B. Leden: Identification of Dynamics of a One Dimensional Heat Diffusion Process. Report 7121, 1971, Division of Automatic Control, Lund Institute of Technology.

[15] L. Ljung: Characterization of the Concept of "Persistently Exciting" in the Frequency Domain. Report 7119, 1971, Division of Automatic Control, Lund Institute of Technology.

[16] G. Olsson: Identification of the Halden Boiling Water Reactor Dynamics. Report, Division of Automatic Control, Lund Institute of Technology. To appear.

[17] J.M. Ortega, W.C. Rheinboldt: Iterative Solution of Nonlinear Equations in Several Variables. Academic Press, New York, 1970.

[18] P.H. Phillipson: Convergence of Clarke's Generalized Least Squares Method in Process Parameter Estimation. 1970, Dep. of Eng., Univ. of Leicester, England.

[19] R. Rao and S.K. Mitra: Generalized Inverse of Matrices and its Applications. John Wiley & Sons, New York, 1971.

[20] N. Wiener: Generalized Harmonic Analysis and Tauberian Theorems. The MIT Press, MIT, Mass., 1964.

APPENDIX A
A SUMMARY OF ERGODICITY THEOREMS.

The purpose of this appendix is a study of expressions of the type

$$\frac{1}{n}\sum_{t=1}^{n} z_1(t)z_2(t)$$

and their limits as $n \to \infty$. $z_i(t)$ will be deterministic signals or stationary stochastic processes of the type

$$z(t) = H(q^{-1})e(t)$$

where $H(q^{-1})$ is a stable filter and $e(t)$ a sequence of independent, equally distributed random variables (white noise). For the study of such expressions some well-known ergodicity theorems will be used. In order to show how these are exploited, the theorems will be stated here in form of two lemmas.

Lemma A.1: Assume that $x(t)$ is a stationary process with discrete time and finite variance. If the covariance function $r_x(\tau) \to 0$ as $|\tau| \to \infty$ then

$$\frac{1}{n}\sum_{t=1}^{n} x(t) \to Ex(t)$$

with probability one and in mean square.

Proof: See [11].

Lemma A.2: Assume that $x(t)$ is a stochastic process with zero mean. If the covariance function fulfils

$$|r(t,s)| \leq K \frac{t^\alpha + s^\alpha}{1 + |t-s|^\beta}$$

with $K > 0$, $0 \leq 2\alpha < \beta < 1$ then

$$\frac{1}{n} \sum_{t=1}^{n} x(t) \to 0$$

with probability one and in mean square.

Proof: See [9].

Some kinds of conditions for deterministic signals are also needed. Inspired of the theory of almost periodic functions, see [20], almost periodic sequences will be used. In the time discrete case the results are much more simple than for time continuous functions.

Definition A.1: The sequence $\{u(t)\}_{t=1}^{\infty}$ is said to be almost periodic if to every $\varepsilon > 0$ there exists a periodic sequence $\{v(t)\}_{t=1}^{\infty}$ (that is $v(t) = v(t+T)$ some T, all t) with finite period T, such that

$$|v(t) - u(t)| < \varepsilon \text{ all } t$$

It is now possible to start the analysis.

Lemma A.3: Let the stationary stochastic processes $z_1(t)$ and $z_2(t)$ be given by

$$z_1(t) = G(q^{-1})e(t)$$

$$z_2(t) = H(q^{-1})e(t)$$

where $e(t)$ is white noise with zero mean, unit variance and finite fourth moment μ.

$$G(q^{-1}) = \sum_{i=0}^{\infty} g_i q^{-i}$$

and

$$H(q^{-1}) = \sum_{i=0}^{\infty} h_i q^{-i}$$

If

$$\sum_{i=0}^{\infty} g_i^2 < \infty, \quad \sum_{i=0}^{\infty} h_i^2 < \infty$$

then

$$\frac{1}{n} \sum_{t=1}^{n} z_1(t)z_2(t) \to Ez_1(t)z_2(t) = \sum_{i=0}^{\infty} h_i g_i, \quad n \to \infty$$

with probability one and in mean square.

Remark: The condition on $G(q^{-1})$ and $H(q^{-1})$ means just that $z_1(t)$ and $z_2(t)$ have finite variances.

Proof: Define $v(t) = z_1(t) \cdot z_2(t)$.

$v(t)$ is a stationary stochastic process with

$$Ev(t) = \sum_{i=0}^{\infty} h_i g_i$$

The convergence of this sum is an immediate consequence of the assumptions and Schwartz' lemma.

In order to use Lemma A.1 the covariance function must be computed.

$$r_v(\tau) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} g_i g_j h_k h_\ell Ee(t-i)e(t+\tau-j) \cdot$$

$$\cdot\ e(t-k)e(t+\tau-\ell) - \left(\sum_{i=0}^{\infty} h_i g_i\right)^2$$

But

$$Ee(t-i)e(t+\tau-j)e(t-k)e(t+\tau-\ell) =$$

$$= \delta_{j,\tau+i}\delta_{\ell,\tau+k} + \delta_{i,k}\delta_{j,\ell} + \delta_{\ell,\tau+i}\delta_{j,\tau+k} +$$

$$+ (\mu-3)\delta_{j,\tau+i}\delta_{k,i}\delta_{\ell,\tau+i}$$

which gives

$$r_v(\tau) = \sum_{i=0}^{\infty} \sum_{k=0}^{\infty} g_i g_{i+\tau} h_k h_{k+\tau} +$$

$$+ \sum_{i=0}^{\infty} \sum_{k=0}^{\infty} g_i g_{k+\tau} h_k h_{i+\tau} + (\mu-3) \sum_{i=0}^{\infty} g_i g_{i+\tau} h_i h_{i+\tau} =$$

$$= \left(\sum_{i=0}^{\infty} g_i g_{i+\tau}\right)\left(\sum_{k=0}^{\infty} h_k h_{k+\tau}\right) +$$

$$+ \left(\sum_{i=0}^{\infty} g_i h_{i+\tau}\right)\left(\sum_{k=0}^{\infty} g_{k+\tau} h_k\right) +$$

$$+ (\mu-3) \sum_{i=0}^{\infty} g_i g_{i+\tau} h_i h_{i+\tau}$$

From this expression the following inequalities are obtained using Schwartz' lemma.

$$|r_v(\tau)| \leq \sqrt{\sum_{i=0}^{\infty} g_i^2 \sum_{j=0}^{\infty} g_{i+\tau}^2 \sum_{k=0}^{\infty} h_k^2 \sum_{\ell=0}^{\infty} h_{\ell+\tau}^2} +$$

$$+ \sqrt{\sum_{i=0}^{\infty} g_i^2 \sum_{j=0}^{\infty} h_{j+\tau}^2 \sum_{k=0}^{\infty} g_{k+\tau}^2 \sum_{\ell=0}^{\infty} h_\ell^2} +$$

$$+ |\mu-3| \sqrt{\sum_{i=0}^{\infty} g_i^2 h_i^2 \sum_{j=0}^{\infty} g_{j+\tau}^2 h_{j+\tau}^2}$$

But

$$\sum_{i=0}^{\infty} g_{i+\tau}^2 = \sum_{i=0}^{\infty} g_i^2 - \sum_{i=0}^{\tau-1} g_i^2 \to 0$$

as $\tau \to \infty$, which implies $|r_v(\tau)| \to 0$, as $\tau \to \infty$.

Invoking Lemma A.1 the proof is finished.

Q.E.D.

Lemma A.4: Let the stationary stochastic processes $z_1(t)$ and $z_2(t)$ be given by

$$z_1(t) = G(q^{-1}) \cdot e_1(t)$$

$$z_2(t) = H(q^{-1}) \cdot e_2(t)$$

Here $e_1(t)$ and $e_2(t)$ are independent white noises with zero means and unit variances.

$$G(q^{-1}) = \sum_{i=0}^{\infty} g_i q^{-i}$$

and

$$H(q^{-1}) = \sum_{i=0}^{\infty} h_i q^{-i}$$

If

$$\sum_{i=0}^{\infty} g_i^2 < \infty, \quad \sum_{i=0}^{\infty} h_i^2 < \infty$$

then

$$\frac{1}{n} \sum_{t=1}^{n} z_1(t) z_2(t) \to 0, \quad n \to \infty$$

with probability one and in mean square.

<u>Proof</u>: Define $v(t) = z_1(t) \cdot z_2(t)$, a stochastic process with zero mean.

The covariance function of $v(t)$ is

$$r_v(\tau) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} g_i g_j h_k h_\ell E e_1(t-i) e_1(t+\tau-j) \cdot$$

$$\cdot\, e_2(t-k) e_2(t+\tau-\ell) =$$

$$= \sum_{i=0}^{\infty} g_i g_{i+\tau} \sum_{k=0}^{\infty} h_k h_{k+\tau}$$

From this expression

$$|r_v(\tau)| \leq \sqrt{\sum_{i=0}^{\infty} g_i^2 \sum_{j=0}^{\infty} g_{j+\tau}^2 \sum_{k=0}^{\infty} h_k^2 \sum_{\ell=0}^{\infty} h_{\ell+\tau}^2} \to 0, \quad \tau \to \infty$$

The assertion of the lemma now follows from Lemma A.1.

Q.E.D.

<u>Lemma A.5</u>: Let $z_1(t)$ be a deterministic, bounded sequence and $z_2(t)$ a stationary, stochastic process, given by

$$z_2(t) = G(q^{-1}) e(t)$$

where $e(t)$ is white noise with zero mean and unit variance,

$$G(q^{-1}) = \sum_{i=0}^{\infty} g_i q^{-i}, \quad \sum_{i=0}^{\infty} g_i^2 < \infty$$

If the covariance function of $z_2(t)$ fulfils

$$|r_{z_2}(\tau)| \leq C\tau^{-\gamma} \quad \gamma > 0 \quad \tau \geq 1$$

then

$$\frac{1}{n} \sum_{t=1}^{n} z_1(t) z_2(t) \to 0, \quad n \to \infty$$

with probability one and in mean square.

<u>Proof</u>: Define $v(t) = z_1(t) \cdot z_2(t)$, a stochastic process with zero mean. By the assumptions $z_1(t)$ is bounded, say $|z_1(t)| \leq D$. The covariance function of $v(t)$ fulfils

$$|r_v(t,s)| = |E z_1(t) z_2(t) z_1(s) z_1(s)| \leq D^2 |r_{z_2}(t-s)| \leq$$

$$\leq D^2 C |t-s|^{-\gamma} \quad \text{for } |t-s| \geq 1$$

Lemma A.2 can now be used with $\alpha = 0$, $\beta = \gamma$ and

$$K = D^2 \max\left(\frac{r_{z_2}(0)}{2}, C\right)$$

Q.E.D.

Lemma A.6: Let $z_1(t)$ and $z_2(t)$ be two almost periodic sequences. Then

$$\frac{1}{n}\sum_{t=1}^{n} z_1(t) \cdot z_2(t)$$

converges as $n \to \infty$.

Proof: Define $v(t) = z_1(t) \cdot z_2(t)$. Clearly $v(t)$ is also almost periodic. Let $u(t)$ be a periodic sequence such that

$$|v(t) - u(t)| < \varepsilon \quad \text{(all t)}$$

The convergence of

$$\frac{1}{n}\sum_{t=1}^{n} u(t)$$

is trivial. Put

$$s_n = \frac{1}{n}\sum_{t=1}^{n} v(t)$$

Using the Cauchy criterion for the sequence

$$\frac{1}{n}\sum_{t=1}^{n} u(t)$$

$$|s_n - s_m| = \left|\frac{1}{n}\sum_{t=1}^{n} (v(t) - u(t) + u(t)) - \right.$$

$$\left. - \frac{1}{m}\sum_{t=1}^{m} (v(t) - u(t) + u(t))\right| \leq$$

$$\leq 2\varepsilon + \left|\frac{1}{n}\sum_{t=1}^{n} u(t) - \frac{1}{m}\sum_{t=1}^{m} u(t)\right| < 3\varepsilon$$

$$\text{if } \min(m,n) > N(\varepsilon)$$

Using the same criterion for the sequence s_n the convergence is proved.

Q.E.D.

The following example shows that $x(t)$ bounded does not imply convergence of

$$\frac{1}{n}\sum_{t=1}^{n} x(t)$$

This means especially that $z_i(t)$ bounded is a too weak condition in Lemma A.6.

Example: Define $x(t)$ by

$$\begin{aligned} x(t) &= 1 \quad & t &= 1 \\ &= -1 & t &= 2,.4 \\ &= 1 & t &= 3\ldots12 \\ &= -1 & t &= 13\ldots36 \end{aligned}$$

and

$$x(t) = (-1)^{m-1}, \quad 4 \cdot 3^{m-1} + 1 \leq t \leq 4 \cdot 3^m$$

Put

$$s_n = \frac{1}{n}\sum_{t=1}^{n} x(t)$$

Then $s_n = 1/2$ if $n = 4 \cdot 3^m$ m odd

and $s_n = -1/2$ if $n = 4 \cdot 3^m$ m even

From this it follows that $\lim\inf s_n < \lim\sup s_n$ and thus $\lim s_n$ does not exist.

It is now possible to prove Theorem 2.1.

<u>Proof of Theorem 2.1</u>: An inspection of the kind of terms in (2.1) shows that the proof follows immediately from Lemmas A.3 - A.6.

If e(t) and/or v(t) has not zero mean, it is rewritten as $e(t) = [e(t) - Ee(t)] + Ee(t)$ and the lemmas are applied twice. In this case the following easily proved property is required as well.

If $v(t) = H(q^{-1}) \cdot e(t)$, e(t) white noise with zero mean and

$$\sum_{i=0}^{\infty} h_i^2 < \infty$$

then

$$\frac{1}{n}\sum_{t=1}^{n} v(t) \to 0, \quad n \to \infty$$

with probability one and in mean square.

Q.E.D.

<u>APPENDIX B</u>

ANALYSIS OF THE MINIMIZATION ALGORITHM.

The purpose of this appendix is to examine the properties of the minimization algorithm. To get reasonable work it is assumed that the loss function is a quadratic form, which is a good approximation close to a stationary point.

Define

$$W(x,y) = \frac{1}{2}[x^T \quad y^T]\begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22}\end{bmatrix}\begin{bmatrix} x \\ y\end{bmatrix} \tag{B.1}$$

where Q is a symmetric matrix.

The vector x corresponds to $[\hat{a}_1-a_1, \ldots, \hat{b}_n-b_n]^T$ and the vector y corresponds to $[\hat{c}_1-c_1, \ldots, \hat{c}_n-c_n]^T$.

The minimization procedure is given by

$$\begin{cases} Q_{11}x_{k+1} + Q_{12}y_k = 0 \\ Q_{21}x_{k+1} + Q_{22}y_{k+1} = 0 \end{cases} \tag{B.2}$$

It is assumed in the following that $Q_{11} > 0$, $Q_{22} > 0$ (are positive definite) which always can be assumed to be true for the actual loss function (3.3). An exception is the case of no noise and too high an order of the model, but this case can be excluded. This means that (B.2) has always a unique solution.

Introduce

$$\begin{cases} P_1 = Q_{11}^{-1}Q_{12}Q_{22}^{-1}Q_{21} \\ P_2 = Q_{22}^{-1}Q_{21}Q_{11}^{-1}Q_{12} \end{cases} \tag{B.3}$$

Then from (B.2)

$$\begin{bmatrix} x_{k+1} \\ y_{k+1} \end{bmatrix} = \begin{bmatrix} P_1 & 0 \\ 0 & P_2 \end{bmatrix} \begin{bmatrix} x_k \\ y_k \end{bmatrix} \tag{B.4}$$

It is of great interest to analyze the eigenvalues of the matrix

$$\begin{bmatrix} P_1 & 0 \\ 0 & P_2 \end{bmatrix}$$

Lemma B.1: Let A and B be two matrices, such that AB and BA are defined. If $\lambda \neq 0$ is an eigenvalue of AB then λ is also an eigenvalue of BA.

Proof: $ABe = \lambda e$ gives $BABe = \lambda Be$.
If $Be \neq 0$ then λ is an eigenvalue of BA with the eigenvector Be.
If $Be = 0$ then $\lambda = 0$, a contradiction.

Q.E.D.

Corr: P_1 and P_2 have the same non-zero eigenvalues.

The following well-known lemma will be used below and in Appendix C.

Lemma B.2: The symmetric matrix

$$Q = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix}$$

is positive definite if and only if $Q_{22} > 0$ and $Q_{11} - Q_{12}Q_{22}^{-1}Q_{21} > 0$.

Further, if $Q \geq 0$ (positive semidefinite) and $Q_{22} > 0$ then $Q_{11} - Q_{12}Q_{22}^{-1}Q_{21}$ is positive semidefinite.

Proof: See [13].

Introduce

$$\begin{cases} \tilde{Q}_1 = Q_{11} - Q_{12}Q_{22}^{-1}Q_{21} \\ \tilde{Q}_2 = Q_{22} - Q_{21}Q_{11}^{-1}Q_{12} \end{cases} \tag{B.5}$$

Then the criterion (with the assumptions above of Q_{11} and Q_{22}) can be written

$Q > 0$ if and only if $\tilde{Q}_1 > 0$ if and only if $\tilde{Q}_2 > 0$

(B.3) and (B.5) give easily

$$\begin{cases} P_1 = I - Q_{11}^{-1}\tilde{Q}_1 \\ P_2 = I - Q_{22}^{-1}\tilde{Q}_2 \end{cases} \tag{B.6}$$

Let P_1 have an eigenvalue λ with the associated eigenvector e

$$P_1 e = \lambda e$$

(B.6) gives

$$e - Q_{11}^{-1}\tilde{Q}_1 e = \lambda e, \quad \tilde{Q}_1 e = (1 - \lambda)Q_{11}e$$

and

$$1 - \lambda = \frac{e^T\tilde{Q}_1 e}{e^T Q_{11} e} \tag{B.7}$$

Lemma B.3: All eigenvalues of P_1 are positive.

Proof:

$$e^T\tilde{Q}_1 e = e^T Q_{11} e - e^T Q_{12} Q_{22}^{-1} Q_{21} e \leq e^T Q_{11} e$$

which gives

$$1 - \lambda \leq 1 \quad \text{or} \quad \lambda \geq 0$$

Q.E.D.

Lemma B.4: P_1 has a basis of eigenvectors.

Proof: Follows from [19] (Thm. 6.2.3) since P_1 is a product of a positive definite matrix and a positive (semi-) definite matrix.

Lemma B.5: Let λ denote an eigenvalue of P_1.

i) $Q > 0$ if and only if $\lambda < 1$ (all λ).

ii) $Q \geq 0$ if and only if $\lambda \leq 1$ (all λ) with equality for at least one λ.

iii) Q indefinite if and only if $\lambda > 1$ some λ.

APPENDIX C

ON CONDITIONS FOR LOCAL MINIMUM POINTS OF A SPECIAL FUNCTION.

In this appendix a special function is studied and its possible minimum points are examined. The reason for studying this function is that it can be interpreted as the loss function of the GLS method.

When the variance of the noise is small the equation $V' = 0$ will lead to equations of the type

$$f(x) + g(x) = 0, \quad g(x) = O(\varepsilon)$$

where ε is a small number. Some of the following lemmas deal with the properties of the solution of such equations.

The first lemma is the well-known principle of contraction mapping. It is stated here in order to later show how it can be used for the actual problems.

Lemma C.1: Let $B_\delta(x_o)$ denote the set $\{x; ||x-x_o|| \leq \delta\}$. Consider a map $S(x)$. If

i) $$||S(x_o) - x_o|| \leq (1-\alpha)\delta \qquad \alpha < 1 \tag{C.1}$$

ii) $$||S(x') - S(x'')|| \leq \alpha||x' - x''||, \quad x',x'' \in B_\delta(x_o) \tag{C.2}$$

then $S(x)$ has a unique fixpoint (a solution of $x = S(x)$) in $B_\delta(x_o)$.

Proof: See [17].

The next lemma deals with necessary properties of solutions. It does not guarantee existence or uniqueness of solutions.

<u>Lemma C.2</u>: Consider the equation

$$F(x,\varepsilon) = f(x) + g(x,\varepsilon) = 0 \tag{C.3}$$

where f and g are continuous functions.

Denote the null space of f by Nf $\{Nf = \{x;\ f(x) = 0\}\}$
Let Ω be an arbitrary, compact set, which may depend on ε. Assume that

i) Ω - Nf is non empty

ii) there are constants $\varepsilon_1 > 0$ and $K < \infty$ such that $0 \leq \varepsilon \leq \varepsilon_1$ implies

$$\sup_{x \in \Omega} ||g(x, \varepsilon)|| \leq K\varepsilon$$

Then there is a number $\varepsilon_o > 0$ such that if $0 \leq \varepsilon \leq \varepsilon_o$ and $\bar{x}$ is a solution of $F(x,\varepsilon) = 0$ then

$$\inf_{x_o \in Nf} ||\bar{x} - x_o|| \to 0, \quad \varepsilon \to 0 \tag{C.4}$$

<u>Proof</u>: Define a set $M(\varepsilon')$, a neighbourhood of Nf by

$$M(\varepsilon') = \{x;\ \inf_{x_o \in Nf} ||x - x_o|| \leq \varepsilon'\}$$

By the construction and the continuity of f

$$\inf_{x \in \Omega - M(\varepsilon')} ||f(x)|| = \alpha(\varepsilon') > 0 \quad \text{if } \varepsilon' > 0$$

$\{$where it is assumed that $\Omega - M(\varepsilon')$ is non empty$\}$.

Let $0 \leq \varepsilon \leq \varepsilon_1$. Then

$$\inf_{x \in \Omega - M(\varepsilon')} ||F(x,\varepsilon)|| \geq \inf_{x \in \Omega - M(\varepsilon')} ||f(x)|| -$$

$$- \sup_{x \in \Omega - M(\varepsilon')} ||g(x,\varepsilon)|| \geq \alpha(\varepsilon') - K\varepsilon$$

Define now $\varepsilon_o = \min\left\{\varepsilon_1, \frac{1}{2}\frac{\alpha(\varepsilon')}{K}\right\}$ which is strictly positive.

Let $0 \leq \varepsilon \leq \varepsilon_o$. Then

$$\inf_{x \in \Omega - M(\varepsilon')} ||F(x,\varepsilon)|| \geq \frac{1}{2}\alpha(\varepsilon') > 0$$

If $\bar{x}$ is a solution of (B.3) then $\bar{x} \in M(\varepsilon')$ and

$$\inf_{x_o \in Nf} ||\bar{x} - x_o|| \leq \varepsilon'$$

However, ε' can be chosen arbitrary small, so all solutions of (C.3) fulfil (C.4).

Q.E.D.

<u>Corr</u>: If $g(x,\varepsilon) = \varepsilon h(x,\varepsilon)$ where $h(x,\varepsilon)$ is a continuous function, the compact set Ω can be chosen arbitrarily.

The following lemma gives a sufficient condition for existence of a unique solution of the form (B.4).

<u>Lemma C.3</u>: Consider the equation

$$F(x,\varepsilon) = f(x) + g(x,\varepsilon) = 0 \tag{C.3}$$

where f and g are twice differentiable functions and $\dim f = \dim g = \dim x$.

Let x_o be a zero of f(x) such that

i) $f'_x(x_o)$ is non singular

ii) there is a set $B_\delta(x_o) = \{x;\ ||x - x_o|| \leq \delta\}$ with δ (independent of ε) > 0, and constants ε_1, C_1 and C_2 such that

a) x_o is the only zero of $f(x)$ in $B_\delta(x_o)$,

b) $0 \leq \varepsilon \leq \varepsilon_1$ implies

$$\sup_{x \in B_\delta(x_o)} ||g(x,\varepsilon)|| \leq C_1\varepsilon$$

$$\sup_{x \in B_\delta(x_o)} ||g'_x(x,\varepsilon)|| \leq C_2\varepsilon$$

Then there is a number $\varepsilon_o > 0$ such that $0 \leq \varepsilon \leq \varepsilon_o$ implies

i) $F(x,\varepsilon) = 0$ has a unique solution $\bar{x}$ in $B_\delta(x_o)$

ii) $\bar{x}$ fulfils

$$\bar{x} - x_o = O(\varepsilon), \quad \varepsilon \to 0 \qquad \text{(C.5)}$$

<u>Proof</u>: Study solutions of (C.3) in $B_{\delta_o}(x_o)$ where δ_o is an arbitrary constant satisfying $0 < \delta_o \leq \delta$.

Consider the function

$$S(x,\varepsilon) = x - f'_x(x_o)^{-1}F(x,\varepsilon)$$

If $S(x,\varepsilon)$ is a contraction mapping its fixpoint is the solution of $x - f'_x(x_o)^{-1}F(x,\varepsilon) = x$ of $F(x,\varepsilon) = 0$.
Put $C_o = ||f'_x(x_o)^{-1}||$.

Let $0 \leq \varepsilon \leq \varepsilon_1$. Then

$$||S(x_o,\varepsilon) - x_o|| \leq ||f'_x(x_o)^{-1}|| \cdot ||F(x_o,\varepsilon)|| \leq C_oC_1\varepsilon$$

Let x' and x'' be two arbitrary, different points in $B_{\delta_o}(x_o)$.
With use of the mean value theorem [17]

$$\frac{||S(x',\varepsilon) - S(x'',\varepsilon)||}{||x' - x''||} = \sup_{0 \leq t \leq 1} ||S'_x(tx' + (1-t)x'',\varepsilon)||$$

Assume that the supremum is obtained at $x = x'''$.

$$\frac{||S(x',\varepsilon) - S(x'',\varepsilon)||}{||x' - x''||} \leq ||S'_x(x''',\varepsilon)|| =$$

$$= ||I - f'_x(x_o)^{-1}[f'_x(x''') + g'_x(x''',\varepsilon)]||$$

$$\leq C_o||f'_x(x''') - f'_x(x_o)|| + C_oC_1\varepsilon \leq C_oC_3\delta_o + C_oC_1\varepsilon$$

for some constants C_3 (depending on δ but not on δ_o).

Now (C.1) and (C.2) are fulfilled if

$$C_oC_1\varepsilon \leq (1-\alpha)\delta_o$$

$$C_oC_3\delta_o + C_oC_1\varepsilon \leq \alpha$$

Choose a value of α. Let δ_o satisfy

$$\delta_o = K\varepsilon$$

where

$$K \geq \frac{C_oC_1}{1-\alpha}$$

Define then

$$\varepsilon_o' = \min\left[\varepsilon_1,\ \frac{\delta}{K},\ \frac{\alpha}{C_o(C_1 + C_3K)}\right] \tag{C.6}$$

Then (C.1), (C.2) and $\delta_o \leq \delta$ are fulfilled if $0 \leq \varepsilon \leq \varepsilon_o'$.

Now consider the set $\Omega = B_\delta(x_o) - B_{\delta_o}(x_o)$.

It has to be shown that $F(x,\varepsilon) = 0$ has no solutions in Ω if ε is small enough.

If δ_o is small enough

$$\inf_\Omega ||f(x)|| = \inf_{||x-x_o||=\delta_o} ||f(x)|| =$$

$$= \inf_{||x-x_o||=\delta_o} ||f(x_o) + f_x'(x_o)(x-x_o) +$$

$$+ O(||x-x_o||^2)|| = \alpha\delta_o + O(\delta_o^2)$$

α denotes the smallest singular value of $f_x'(x_o)$.

Thus there are constants ε_1' and C_4 such that $0 \leq \varepsilon \leq \varepsilon_1'$ implies

$$\inf_{x\in\Omega} ||F(x,\varepsilon)|| \geq \inf_{x\in\Omega} ||f(x)|| - \sup_{x\in\Omega} ||g(x,\varepsilon)|| \geq$$

$$\geq \alpha\delta_o - C_4\delta_o^2 - C_1\varepsilon$$

This expression should be positive. Insert $\delta_o = K\varepsilon$.

$$\varepsilon[(\alpha K - C_1) - C_4K^2\varepsilon] > 0$$

Now choose finally

$$K = \max\left[\frac{C_oC_1}{1-\alpha},\ \frac{3C_1}{\alpha}\right]$$

and

$$\varepsilon_o = \min\left[\varepsilon_o',\ \varepsilon_1',\ \frac{C_1}{C_4K^2}\right]$$

With these values of K and ε_o and with $\delta_o = K\varepsilon$ it can be seen by going through the proof once more that $F(x,\varepsilon) = 0$ has a unique solution in $B_{\delta_o}(x_o)$ and no solution in $B_\delta(x_o) - B_{\delta_o}(x_o)$.

Q.E.D.

Remark: If $f_x'(x_o)$ is singular, nothing general can be stated. Consider the scalar examples $F_1(x) = x^2 - \varepsilon$ and $F_2(x) = x^2 + \varepsilon$. $F_1(x)$ has zeros close to $x_o = 0$, but these do not satisfy (C.5). $F_2(x)$ has no real zeros at all.

Near a local extremum the matrix of second order derivatives plays a fundamental role for determining the character of the extremum. The following lemmas which deal with quadratic forms will be useful in the analysis of this matrix.

Lemma C.4: Consider the symmetric matrix

$$Q = \begin{bmatrix} A + \varepsilon A_1 & \varepsilon B \\ \varepsilon B^T & \varepsilon C \end{bmatrix} \tag{C.7}$$

and the vector

$$r = \begin{bmatrix} \varepsilon b \\ 0 \end{bmatrix} \tag{C.8}$$

Assume that A and C are positive definite. Then if $0 < \varepsilon \leq \varepsilon_o$ where $1/\varepsilon_o >$ the largest eigenvalue of $A^{-1}[A_1 - BC^{-1}B^T]$

i) Q is positive definite

ii) $Q^{-1}r = O(\varepsilon), \ \varepsilon \to 0$

<u>Proof</u>:

i) By Lemma C.2 $Q > 0$ is equivalent to

$$A + \varepsilon A_1 - \varepsilon B(\varepsilon C)^{-1}\varepsilon B^T > 0 \quad \text{or}$$

$$A + \varepsilon D > 0 \tag{C.9}$$

where $D = A_1 - BC^{-1}B^T$.

(C.9) is apparently true for small values of ε (since the eigenvalues of $A + \varepsilon D$ are continuous functions of ε). ε must only be smaller than the smallest number δ such that

$$\det[A + \delta D] = 0 \tag{C.10}$$

(C.10) is rewritten as

$$\det[A\delta(\tfrac{1}{\delta} I + A^{-1}D)] = \det(A\delta)\det(\tfrac{1}{\delta} I + A^{-1}D) = 0$$

From this equation it is seen that $1/\delta$ = the largest eigenvalue of $A^{-1}D$.

ii) Using formulas for the inverse of a partionated matrix [10]

$$Q^{-1}r = \begin{bmatrix} (A + \varepsilon D)^{-1}\varepsilon b \\ -C^{-1}B(A + \varepsilon D)^{-1}\varepsilon b \end{bmatrix}$$

If $\varepsilon < \delta$ then $[A + \varepsilon D]^{-1} = A^{-1} + O(\varepsilon)$ and $Q^{-1}r = O(\varepsilon)$ follows easily.

Q.E.D.

<u>Lemma C.5</u>: Consider the function

$$V(x,\varepsilon) = \frac{1}{2} x^T Q(\varepsilon)x + x^T r(\varepsilon) \tag{C.11}$$

with

$$Q(\varepsilon) = \begin{bmatrix} A + \varepsilon A_1 & \varepsilon B \\ \varepsilon B^T & \varepsilon C \end{bmatrix} \qquad r(\varepsilon) = \begin{bmatrix} \varepsilon b \\ 0 \end{bmatrix}$$

with A_1 in a symmetric matrix, A and C are symmetric and positive definite matrices. There is a constant $\varepsilon_o >$ such that if $0 < \varepsilon \leq \varepsilon_o$ then:

To every $K_2 > 0$ there is a constant K_1 (depending on K_2 and ε_o but not on ε) such that

$$\inf_{||x||=K_1\varepsilon} V(x,\varepsilon) \geq K_2\varepsilon^2 \tag{C.12}$$

<u>Proof</u>: Consider the set

$$\Omega(V_o,\varepsilon) = \{x;\ V(x,\varepsilon) \leq V_o\}$$

Define

$$x_o(\varepsilon) = -Q(\varepsilon)^{-1}r(\varepsilon)$$

Then $\Omega(V_o,\varepsilon)$ is given by

$$\frac{1}{2}\,(x - x_o(\varepsilon))^TQ(\varepsilon)(x - x_o(\varepsilon)) \leq V_o + \frac{1}{2}\,x_o(\varepsilon)^TQ(\varepsilon)x_o(\varepsilon) \quad (C.13)$$

$\Omega(V_o,\varepsilon)$ is non empty if

$$0 < \varepsilon < \delta$$

$$V_o \geq -\frac{1}{2}\,x_o(\varepsilon)^TQ(\varepsilon)x_o(\varepsilon)$$

where δ is the largest eigenvalue of $A^{-1}[A_1 - BC^{-1}B^T]$.

Let x_i denote the i:th component of x.
Define a new set

$$\Omega_1(V_o,\varepsilon) = \{x;\ |x_i - x_o(\varepsilon)_i| \leq \sup_{x\in\Omega(V_o,\varepsilon)} |x_i - x_o(\varepsilon)_i| \text{all } i\}$$

Clearly $\Omega(V_o,\varepsilon) \subset \Omega_1(V_o,\varepsilon)$.

What is $\sup\limits_{x\in\Omega(V_o,\varepsilon)} |x_i - x_o(\varepsilon)_i|$?

Let e_i denote a unit vector, which i:th component is 1.

Then the maximum of $e_i^T(x - x_o(\varepsilon))$ under the constraint

$$(x - x_o(\varepsilon))^TQ(\varepsilon)(x - x_o(\varepsilon)) = 2V_o + x_o(\varepsilon)^TQ(\varepsilon)x_o(\varepsilon)$$

is sought.

Using a Lagrange multiplier

$$e_i + \lambda 2Q(\varepsilon)(x - x_o(\varepsilon)) = 0$$

$$(x - x_o(\varepsilon))^TQ(\varepsilon)(x - x_o(\varepsilon))^T = 2V_o + x_o(\varepsilon)^TQ(\varepsilon)x_o(\varepsilon)$$

from which

$$\sup_{x\in\Omega(V_o,\varepsilon)} |x_i - x_o(\varepsilon)_i| = \sqrt{\frac{2V_o + x_o^T(\varepsilon)Q(\varepsilon)x_o(\varepsilon)}{[Q(\varepsilon)^{-1}]_{ii}}} \qquad (C.14)$$

is obtained by straight forward calculations.

The sphere

$$S_1(V_o,\varepsilon) = \left\{x;\ ||x - x_o(\varepsilon)|| \leq \sqrt{\sum_i \left[\frac{2V_o + x_o^T(\varepsilon)Q(\varepsilon)x_o(\varepsilon)}{[Q(\varepsilon)^{-1}]_{ii}}\right]^2}\right\}$$

contains the set $\Omega(V_o,\varepsilon)$ and so does the sphere

$$S_2(V_o,\varepsilon) = \left\{x;\ ||x|| \leq ||x_o(\varepsilon)|| + \sqrt{\sum_i \left[\frac{2V_o + x_o^T(\varepsilon)Q(\varepsilon)x_o(\varepsilon)}{[Q(\varepsilon)^{-1}]_{ii}}\right]^2}\right\}$$

A graphical illustration of the sets $\Omega(V_o,\varepsilon)$, $\Omega_1(V_o,\varepsilon)$, $S_1(V_o,\varepsilon)$ and $S_2(V_o,\varepsilon)$ for a two dimensional example is given in Fig. C.1.

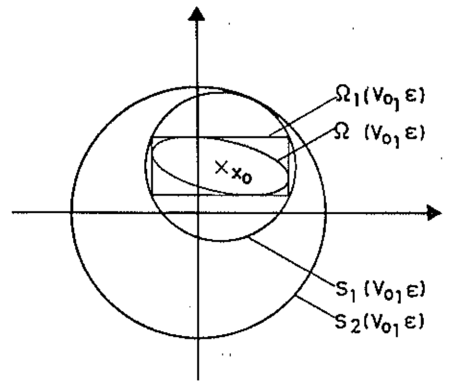


Fig. C.1.

The function $V(x,\varepsilon)$ has the following property. Let M_1 and M_2 be two convex and compact sets, containing $x_o(\varepsilon)$ and with boundaries ∂M_1 and ∂M_2. If $M_1 \subset M_2$ then

$$\inf_{x \in \partial M_1} V(x,\varepsilon) \leq \inf_{x \in \partial M_2} V(x,\varepsilon).$$

This is true since $V(x,\varepsilon)$ is a convex function. Define $\bar{x}_2 \in \partial M_2$ by

$$V(\bar{x}_2,\varepsilon) = \inf_{x \in \partial M_2} V(x,\varepsilon)$$

There is at least one point $\bar{x}_1 \in \partial M_1$ such that

$$\bar{x}_1 = tx_o(\varepsilon) + (1-t)\bar{x}_2 \qquad 0 \leq t \leq 1$$

so

$$\inf_{x \in \partial M_1} V(x,\varepsilon) \leq V(\bar{x}_1,\varepsilon) \leq tV\big(x_0(\varepsilon),\varepsilon\big) + (1-t)V(\bar{x}_2,\varepsilon) \leq$$

$$\leq V(\bar{x}_2,\varepsilon) = \inf_{x \in \partial M_2} V(x_2,\varepsilon)$$

Put now $M_1 = \Omega(V_o,\varepsilon)$ and $M_2 = S_2(V_o,\varepsilon)$.

Applying this property

$$\inf_{||x||=R(V_o,\varepsilon)} V(x,\varepsilon) \geq V_o \tag{C.15}$$

where

$$R(V_o,\varepsilon) = ||x_o(\varepsilon)|| + \Big[\big(2V_o + x_o^T(\varepsilon)Q(\varepsilon)x_o(\varepsilon)\big)q(\varepsilon)\Big]^{1/2} \tag{C.16}$$

$$q(\varepsilon) = \sum_i \frac{1}{[Q(\varepsilon)^{-1}]_{ii}} \tag{C.17}$$

There are constants ε_o, C_1, C_2 and C_3 (with $\varepsilon_o < \delta$ and C_1, C_2, C_3 independent of ε) such that $0 < \varepsilon \leq \varepsilon_o$ implies

$$||x_o(\varepsilon)|| \leq C_1\varepsilon$$

$$||x_o^T(\varepsilon)Q(\varepsilon)x_o(\varepsilon)|| \leq C_2\varepsilon^2$$

$$q(\varepsilon) \leq C_3$$

The last inequality follows from the expression for the inverse of a partionated matrix [10].

Define

$$R_1(\varepsilon,V_o) = C_1\varepsilon + [C_3(2V_o + C_2\varepsilon^2)]^{1/2} \tag{C.18}$$

Let now $0 < \varepsilon \leq \varepsilon_o$. Then $R(\varepsilon,V_o) \leq R_1(\varepsilon,V_o)$ and from the property of $V(x,\varepsilon)$ described above

$$\inf_{||x||=R_1(\varepsilon,V_o)} V(x) \geq V_o$$

Now take $K_2 > 0$ arbitrary and put $V_o = K_2\varepsilon^2$.
Then $R_1(\varepsilon,V_o) = K_1\varepsilon$ with

$$K_1 = C_1 + [C_3(2K_2 + C_2)]^{1/2}$$

and the lemma is proved.

Q.E.D.

In the following theorem the results of the foregoing lemmas are applied to a function of special structure. It will later turn out that the loss function of the GLS method has this structure.

<u>Theorem C.1</u>: Consider the function

$$V(x,y,\varepsilon) = \frac{1}{2}\, x^T P(y)x + \varepsilon h(x,y) \tag{C.19}$$

where $P(y)$ is a positive definite matrix for all y, twice differentiable with respect to y and $h(x,y)$ a twice differentiable function. ε is considered as a fix parameter.

Then there are necessary and sufficient conditions for local minimum points in an arbitrary compact set Ω.

There is a constant $\varepsilon_o > 0$ such that if $0 < \varepsilon \leq \varepsilon_o$ the following is true.

i) Every stationary point of $V(x,y,\varepsilon)$ in Ω fulfils

$$(x,y) = (0,y_o) + [O(\varepsilon),o(1)], \quad \varepsilon \to 0 \tag{C.20}$$

where y_o is a solution of

$$h'_y(0,y) = 0 \tag{C.21}$$

If (x,y) is a local minimum point it is necessary that $h''_{yy}(0,y_o)$ is positive definite or positive semidefinite.

ii) If y_o is a solution of (C.21) and $h''_{yy}(0,y_o)$ is positive definite then there <u>exists</u> a <u>unique</u> local minimum of the form (C.20), and the point will in fact satisfy

$$(x,y) = (0,y_o) + [O(\varepsilon),O(\varepsilon)], \quad \varepsilon \to 0 \tag{C.22}$$

The matrix of second order derivatives is positive definite in the minimum point.

<u>Proof</u>: The equation $V' = 0$ turns out to be

$$\begin{bmatrix} P(y)x \\ \frac{\partial}{\partial y}\left(\frac{1}{2}\, x^T P(y)x\right) \end{bmatrix} + \varepsilon \begin{bmatrix} h'_x(x,y) \\ h'_y(x,y) \end{bmatrix} = 0 \tag{C.23}$$

and the matrix of second order derivatives

$$V'' = \begin{bmatrix} V''_{xx} & V''_{xy} \\ V''_{yx} & V''_{yy} \end{bmatrix} =$$

$$= \begin{bmatrix} P(y) & \frac{\partial}{\partial y}[P(x)y] \\ \frac{\partial}{\partial y}[P(x)y]^T & \frac{\partial^2}{\partial y^2}\left[\frac{1}{2}\, x^T P(y)x\right] \end{bmatrix} +$$

$$+ \varepsilon \begin{bmatrix} h''_{xx}(x,y) & h''_{xy}(x,y) \\ h''_{yx}(x,y) & h''_{yy}(y,y) \end{bmatrix} \qquad (C.24)$$

The first part of (C.23) yields the necessary condition

$$||x|| = \varepsilon||P(y)^{-1}h'_x(x,y)|| \leq K\varepsilon \qquad (C.25)$$

where

$$K = \sup_{(x,y)\in\Omega} ||P(y)^{-1}h'_x(x,y)||$$

Apply Lemma C.2 to the second part of (C.23) putting

$$f(y) = h'_y(0,y)$$

$$g(y,\varepsilon) = \frac{1}{\varepsilon}\frac{\partial}{\partial y}\left(\frac{1}{2}\, x^T P(y)x\right) + h'_y(x,y) - h'_y(0,y)$$

Assume that (C.25) holds. Then there is a number $\varepsilon'_o > 0$ such that if $0 < \varepsilon \leq \varepsilon'_o$ the following condition is necessary

$$y - y_o = o(1), \quad \varepsilon \to 0 \qquad (C.26)$$

where y_o is some solution of

$$h'_y(0,y) = 0$$

If (x,y) is a minimum point, it is necessary that V'' is positive definite or positive semidefinite. From this it follows that the same must be true to V''_{yy} and further that there is a number ε''_o such that $0 < \varepsilon \leq \varepsilon''_o$ implies the same condition for $h''_{yy}(0,y_o)$.

The first part of the theorem is proved.

If $h''_{yy}(0,y_o)$ is positive definite, it follows from Lemma C.3 that there is a number $\varepsilon'''_o > 0$ such that $0 < \varepsilon \leq \varepsilon'''_o$ implies that (C.26) can be replaced by

$$y = y_o + O(\varepsilon) \qquad (C.27)$$

When ε is small

$$V(x,y,\varepsilon) - V(0,y_o,\varepsilon) = [x^T \quad (y-y_o)^T] \begin{bmatrix} \varepsilon h'_x(0,y_o) \\ 0 \end{bmatrix} +$$

$$+ \frac{1}{2}[x^T \quad (y-y_o)^T] \begin{bmatrix} P(y_o) + \varepsilon h_{xx}(0,y_o) & \varepsilon h_{xy}(0,y_o) \\ \varepsilon h_{yx}(0,y_o) & \varepsilon h_{yy}(0,y_o) \end{bmatrix} \cdot$$

$$\cdot \begin{bmatrix} x \\ y-y_o \end{bmatrix} + r(x,y,\varepsilon)$$

where $r(x,y,\varepsilon) = O(||(x,y) - (0,y_o)||^3)$.

A straight forward application of Lemma C.5 gives: there are constants ε_o^{1V}, K_1 and K_2 such that $0 < \varepsilon \leq \varepsilon_o^{1V}$ implies

$$\inf_{||(x,y)-(0,y_o)||=K_1\varepsilon} V(x,y,\varepsilon) - V(0,y_o,\varepsilon) - r(x,y,\varepsilon) \geq K_2\varepsilon^2$$

But there are constants ε_o^{v} and K_3 such that $0 < \varepsilon \leq \varepsilon_o^{v}$ implies

$$\sup_{||(x,y)-(0,y_o)||=K_1\varepsilon} r(x,y,\varepsilon) \leq K_3\varepsilon^3$$

Thus

$$\inf_{||(x,y)-(0,y_o)||=K_1\varepsilon} V(x,y,\varepsilon) \geq V(0,y_o,\varepsilon) + K_2\varepsilon^2 - K_3\varepsilon^3$$

is greater than $V(0,y_o,\varepsilon)$ if $K_2 - K_3\varepsilon > 0$.

Put

$$\varepsilon_o^{vi} = \frac{K_2}{2K_3}$$

Then $0 < \varepsilon \leq \min(\varepsilon_o^{iv},\varepsilon_o^{v},\varepsilon_o^{vi})$ implies the existence of a local minimum point in the set

$$S(\varepsilon) = \{(x,y);\ ||(x,y) - (0,y_o)|| \leq K_1\varepsilon\}$$

When $(x,y) \in S(\varepsilon)$

$$V'' = \begin{bmatrix} P(y_o) + O(\varepsilon) & O(\varepsilon) \\ O(\varepsilon) & \varepsilon h''_{yy}(0,y_o) + O(\varepsilon^2) \end{bmatrix}$$

By Lemma C.4 it follows that there is a constant ε_o^{vii} such that $0 < \varepsilon \leq \varepsilon_o^{vii}$ and $(x,y) \in S(\varepsilon)$ imply that V'' is positive definite. From this it follows that $V(x,y,\varepsilon)$ has a unique minimum point in $S(\varepsilon)$.

Finally, choose $\varepsilon_o = \min(\varepsilon_o', \varepsilon_o'', \varepsilon_o''', \varepsilon_o^{iv}, \varepsilon_o^{v}, \varepsilon_o^{vi}, \varepsilon_o^{vii})$. Going through the proof once more, it is seen that all parts hold.

Q.E.D.

<u>Remark 1</u>: The greatest possible value of ε_o may depend on Ω. It is in general not possible to take Ω as the whole space. A simplified example: $V_\varepsilon(x) = x^2 + \varepsilon(x^3-x)$ has two stationary points: $x_1(\varepsilon) = -\frac{2}{\varepsilon} + O(\varepsilon)$ and $x_2(\varepsilon) = \frac{\varepsilon}{2} + O(\varepsilon^3)$ while $V_o(x)$ has one stationary point, $x = 0$.

<u>Remark 2</u>: If $h''_{yy}(0,y_o)$ is positive semidefinite (singular) nothing general can be stated. An illustrative example is

$$V(x,y) = \frac{1}{2}x^2 + \varepsilon\left[\frac{1}{2}x^2 + xy + Ky^n\right]$$

where the integer $n \geq 3$.

The equation (C.21) has the only solution $y = 0$ and $h''_{yy}(0,0) = 0$. For this function

$$V' = \begin{bmatrix} x + \varepsilon x + \varepsilon y \\ \varepsilon x + \varepsilon K n y^{n-1} \end{bmatrix}$$

$$V'' = \begin{bmatrix} 1 + \varepsilon & \varepsilon \\ \varepsilon & \varepsilon K n(n-1)y^{n-2} \end{bmatrix}$$

The stationary points are the solutions of

$$x = -\frac{\varepsilon}{1+\varepsilon} y$$

$$y\left(y^{n-2} - \frac{\varepsilon}{(1+\varepsilon)Kn}\right) = 0$$

$(x,y) = (0,0)$ is always a stationary point and a saddle point. If

$$y^{n-2} = \frac{\varepsilon}{(1+\varepsilon)Kn}$$

has a solution then V" is positive definite in that point. This implies

i) If n is odd, there is one minimum point and $x = O(\varepsilon)$, $y = O\left(\varepsilon^{1/(n-2)}\right)$.

ii) If n is even and $K > 0$, there are two minimum points and $x = O(\varepsilon)$, $y = O\left(\varepsilon^{1/(n-2)}\right)$.

iii) If n is even and $K < 0$, there are no minimum points.

APPENDIX D.

ANALYSIS OF THE NOISE CONDITION (NC) FOR FIRST ORDER MODELS.

In order to prove Lemma 3.2 it is necessary to study the derivatives of V_2 and the solutions of $V_2' = 0$.

First order derivatives.

Direct computations give

$$\begin{cases} \frac{1}{2} V_{\hat{a}}' = (\hat{a} + \hat{c} + \hat{a}\hat{c}^2)r_o + (1 + 2\hat{a}\hat{c} + \hat{c}^2)r_1 + \hat{c}r_2 \\ \frac{1}{2} V_{\hat{c}}' = (\hat{a} + \hat{c} + \hat{a}^2\hat{c})r_o + (1 + 2\hat{a}\hat{c} + \hat{a}^2)r_1 + \hat{a}r_2 \end{cases} \tag{D.1}$$

The equations $V_2' = 0$ are rewritten

$$\begin{cases} (\hat{a} + \hat{c} + \hat{a}\hat{c}^2)r_o + (1 + 2\hat{a}\hat{c} + \hat{c}^2)r_1 + \hat{c}r_2 = 0 \\ [(\hat{a} - \hat{c})][\hat{a}\hat{c}r_o + (\hat{a} + \hat{c})r_1 + r_2] = 0 \end{cases} \tag{D.2}$$

Case i): A possible solution fulfils

$$\begin{cases} \hat{a} = \hat{c} \\ (2\hat{a} + \hat{a}^3)r_o + (1 + 3\hat{a}^2)r_1 + \hat{a}r_2 = 0 \end{cases} \tag{D.3}$$

Let $f(x) = (2x + x^3)r_o + (1 + 3x^2)r_1 + xr_2$.
With use of the relation

$$r_2 > -r_o + 2\frac{r_1^2}{r_o}$$

which holds since $w(t)$ is persistently exciting of order 3,

$$f(1) = 3r_o + 4r_1 + r_2 > \frac{2}{r_o}(r_o + r_1)^2 > 0$$

$$f(-1) = -3r_o + 4r_1 - r_2 < -\frac{2}{r_o}(r_o - r_1)^2 < 0$$

$$f'(x) = (2 + 3x^2)r_o + 6xr_1 + r_2 >$$

$$> \frac{1}{r_o}[(r_o^2 - r_1^2) + 3(xr_o + r_1)^2] > 0$$

From these inequalities it is concluded that (D.3) has a unique solution, which satisfies $|\hat{a}| < 1$.

Case ii): The other possibility can be written

$$\begin{cases} (\hat{a} + \hat{c})r_o + (1 + \hat{a}\hat{c})r_1 = 0 \\ \hat{a}\hat{c}r_o + (\hat{a} + \hat{c})r_1 + r_2 = 0 \end{cases} \tag{D.4}$$

Introducing the new variables $\hat{d}_1 = \hat{a} + \hat{c}$, $\hat{d}_2 = \hat{a}\hat{c}$ it is found that $\hat{a}$ and $\hat{c}$ are the roots of

$$z^2 - \hat{d}_1 z + \hat{d}_2 = 0 \tag{D.5}$$

$$\begin{bmatrix} r_o & r_1 \\ r_1 & r_o \end{bmatrix} \begin{bmatrix} \hat{d}_1 \\ \hat{d}_2 \end{bmatrix} + \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = 0 \tag{D.6}$$

Real valued solutions of (D.5) exist when the discriminant $\hat{d}_1^2 - 4\hat{d}_2 \geq 0$ or invoking (D.6)

$$D^* = r_1^2(r_2 - r_o)^2 - 4(r_o^2 - r_1^2)(r_1^2 - r_o r_2) \geq 0 \tag{D.7}$$

Proof of Lemma 3.2: From the analysis above it is clear that

i) if $D^* < 0$ then $V_2' = 0$ has one solution

ii) if $D^* = 0$ then $V_2' = 0$ has three coincident solutions

iii) if $D^* > 0$ then $V_2' = 0$ has three different solutions.

Only the case $D^* > 0$ has to be considered closer.

The change of variables means that the function

$$E[\hat{D}(q^{-1})w(t)]^2, \qquad \hat{D}(q^{-1}) = 1 + \hat{d}_1 q^{-1} + \hat{d}_2 q^{-2}$$

is minimized. This function has a unique minimum with a positive definite matrix of second order derivatives.

When $D^* > 0$ the solutions of (D.5) satisfy $\hat{a} \neq \hat{c}$ and the Jacobian of the transformations of variables is non singular. This fact implies that V_2'' is positive definite for solutions of (D.5) if $D^* > 0$.

Q.E.D.

APPENDIX E.

PROOF OF THEOREM 3.4.

In this appendix it will be shown that by changing variables, Theorem 3.4 follows from Theorem C.1.

Proof of Theorem 3.4: Introduce the vectors (as in the proof of Theorem 3.2)

$$x = \begin{bmatrix} \hat{a}_1 - a_1 \\ \vdots \\ \hat{a}_n - a_n \\ \vdots \\ \hat{a}_{n+k} \\ \hat{b}_1 - b_1 \\ \vdots \\ \hat{b}_n - b_n \\ \vdots \\ \hat{b}_{n+k} \end{bmatrix} \qquad y = \begin{bmatrix} \hat{c}_1 \\ \vdots \\ \hat{c}_{n+k} \end{bmatrix} \tag{E.1}$$

The loss function can be written

$$V(x,y) = \frac{1}{2} x^T P(y) x + \varepsilon h(x,y) \tag{E.2}$$

with $P(y)$ as the covariance matrix of the system

$$A(q^{-1}) y^F(t) = - B(q^{-1}) u^F(t), \qquad u^F(t) = \hat{C}(q^{-1}) u(t)$$

$P(y)$ is, however, always singular, but the null space of $P(y)$ is independent of y. This is obvious, since from Theorem 2.2 the null space is spanned by vectors of the form

$$\begin{bmatrix} f_1 \\ \vdots \\ f_{n+k} \\ g_1 \\ \vdots \\ g_{n+k} \end{bmatrix} \tag{E.3}$$

with

$$F(q^{-1}) = \sum_{i=1}^{n+k} f_i q^{-i} = A(q^{-1}) L'(q^{-1}) \tag{E.4}$$

$$G(q^{-1}) = \sum_{i=1}^{n+k} g_i g^{-i} = B(q^{-1}) L'(q^{-1}) \tag{E.5}$$

$$L'(q^{-1}) = \sum_{i=1}^{k} \ell_i' q^{-i} \quad \text{arbitrary} \tag{E.6}$$

Introduce now the new variables

$$x' = \begin{bmatrix} x_1' \\ x_2' \end{bmatrix}$$

where x_1' is of dimension k and x_2' of dimension $2n+k$. The vector x' is defined by

$$x = Qx' = [Q_1 \mid Q_2] \begin{bmatrix} x_1' \\ x_2' \end{bmatrix} \tag{E.7}$$

where

$$Q_1 = \left[\begin{array}{ccc} 1 & & 0 \\ a_1 & \ddots & \\ \vdots & & \\ a_n & & 1 \\ & \ddots & \vdots \\ 0 & & a_n \\ \hdashline 0 & & 0 \\ b_1 & & \\ \vdots & \ddots & \\ b_n & & b_1 \\ & \ddots & \vdots \\ 0 & & b_n \end{array}\right] \tag{E.8}$$

Q_1 is a $(2n+2k) \times k$ matrix and Q an arbitrary $(2n+2k) \times$ $\times$ $(2n+k)$ matrix with the properties $Q_1^T Q_2 = 0$ and Q non singular. Q_2 can for instance be constructed by Gram Schmidt orthogonalization.

From the discussion it follows that

$Q_1 x_1'$ is a typical element in the null space $N(P(y))$
$Q_2 x_2'$ is a typical element in the space $N(P(y))^{\perp}$

From these facts it is concluded that

$$P(y)Q_1 = 0$$

and that the matrix

$$R(y) = Q_2^T P(y) Q_2 \tag{E.9}$$

of order $(2n+k) \times (2n+k)$ is non singular for all y.

The loss function is now written as

$$V(x_2', z) = \frac{1}{2} x_2'^T R(z) x_2' + \varepsilon k(x_2', z) \tag{E.10}$$

where z denotes the vector

$$\begin{bmatrix} x_1' \\ y \end{bmatrix}$$

Write the vector x_1' as

$$x_1' = \begin{bmatrix} \ell_1 \\ \vdots \\ \ell_k \end{bmatrix} \tag{E.11}$$

Then $x = Q_1 x_1'$ is equivalently expressed as

$$\hat{A}(q^{-1}) = A(q^{-1})\hat{L}(q^{-1}), \quad \hat{B}(q^{-1}) = B(q^{-1})\hat{L}(q^{-1}) \tag{E.12}$$

with

$$L(q^{-1}) = 1 + \hat{\ell}_1 q^{-1} + \ldots + \hat{\ell}_k q^{-k} \tag{E.13}$$

The function $k(0,z)$ is written by operators as

$$k(0,z) = E[\hat{L}(q^{-1})\hat{C}(q^{-1})v(t)]^2$$

Invoking Theorem C.1 the proof is finished.

Q.E.D.

APPENDIX F.

CONSTRUCTION OF COUNTER EXAMPLES TO THE SECOND VERSION OF GLS.

The equations (3.34) - (3.64) for the example of Section 3.8 are examined in this appendix.

(3.36) has the solution

$$\hat{a} = -\frac{r_y(1)}{r_y(0)}$$

$$\hat{b} = 1$$

from which

$$\varepsilon(t) = \frac{1 + \hat{a}q^{-1}}{1 + aq^{-1}} \cdot q^{-1}u(t) + \frac{1 + \hat{a}q^{-1}}{1 + aq^{-1}} v(t)$$

Define the functions F, f and g by

$$F(a,c) = r_{\varepsilon}(1) = f(a,c) + Sg(a,c)$$

$$f(a,c) = E\left[\frac{1 + \hat{a}'q^{-1}}{1 + aq^{-1}} v(t) \cdot \frac{1 + \hat{a}'q^{-1}}{1 + aq^{-1}} v(t+1)\right]$$

with

$$\hat{a}' = -\frac{r_y'(1)}{r_y'(0)}, \qquad y'(t) = \frac{1}{1 + aq^{-1}} v(t)$$

g(a,c) is a differentiable function.

Consider now especially

$$v(t) = \frac{1}{1 + cq^{-1}} e(t)$$

Then

$$\hat{a}' = \frac{a + c}{1 + ac} \qquad \text{and} \qquad f(a,c) = \frac{-ac(a+c)}{(1-ac)(1+ac)^2}$$

Further

$$f(0,c) = 0 \qquad f_a'(0,c) = \frac{-c^2}{(1-c)(1+c)^2}$$

$$f(-c,c) = 0 \qquad f_a'(-c,c) = \frac{c^2}{(1-c^2)^2(1+c^2)}$$

if $c \neq 0$ the existence of solutions of the forms

$a = O(S)$

$a = -c + O(S)$

now follow from Lemma C.3.

APPENDIX G.

DESCRIPTION OF PROGRAMS.

The main structure of the program package for the GLS identification is given in the table below. In the following pages a more detailed description of every subroutine is given.

Program or subroutine	Purpose	Called subroutines
TGLS	Main program	SIMUL GLS
SIMUL	Simulates the system	PRBSTA PRB NODI
GLS	Performs the GLS identification	LS FILT RESID VGLS
PRBSTA, PRB	Generates a PRBS	-
NODI	Generates white noise	-
LS	Performs a LS identification	LSQ
LSQ	Computes a least squares solution	-
FILT	Filters data	-
RESID	Computes the residuals	-
VGLS	Computes the loss function and related variables	FILT DSYMIN EIGS
DSYMIN	Invertes a symmetric matrix	-
EIGS	Computes eigenvalues and eigenvectors of a symmetric matrix	-

In subroutine VGLS there is a possibility to improve the solution by making some (approximative) Newton Raphson iterations.

```
C     PROGRAM TGLS
C
C     MAIN PROGRAM FOR GENERALIZED LEAST SQUARES IDENTIFICATION
C     OF SIMULATED DATA
C     AUTHOR TORSTEN SODERSTROM 1971-10-01
C
C     THE FOLLOWING DATA ARE READ FROM CARDS
C  1- M,ISYST,IMOD,INF - 4I10
C     M - NUMBER OF SAMPLES (MAX 1000)
C     IF M=0 THE PROGRAM STOPS
C     ISYST=10000*NA+100*NB+NC  ORDER OF TRUE OPERATORS
C     IMOD=10000*MNA+100*MNB+MNC  ORDER OF ESTIMATED OPERATORS
C     INF=10000*ITER+1000*IFILT+100*INIT+10*IPRINT+ISIM
C     ITER - MAX NUMBER OF ITERATIONS
C     IFILT  =0-FILTER ORIGINAL DATA   =1-FILTER FILTERED DATA
C     INIT  =0 START WITH LS ESTIMATE OF A AND B
C           =1 START WITH VALUES OF A AND B FROM CARD
C           =2 START WITH VALUES OF C       FROM CARD
C     IPRINT  =0-LITTLE OUTPRINT  =1 GREAT OUTPRINT
C     ISIM =0 U(T) IS A PRBS
C          =1 U(T) IS A WHITE NOISE INDEPENDENT OF E(T)
C  2- (T(I),I=1,(NA+NB+NC)),AL -  8F10.5
C     T - PARAMETER VECTOR  (TRUE VALUES)
C     AL - STANDARD DEVIATION OF THE NOISE
C  3- /IF INIT=1/ (T(I),I=1,(MNA+MNB)) - 8F10.5 START VALUES OF A AND B
C  3- /IF INIT=2/ (T(I)+MNA+MNB),I=1,MNC) -8F10.5  START VALUES OF C
C     SUBROUTINE REQUIRED
C             SIMUL
C             PRBSTA
C             PRB
C             NODI
C             GLS
C             LS
C             LSQ
C             VGLS
C             DSYMIN
C             FILT
C             RESID
C             EIGS
C
      DIMENSION U(1000),Y(1000),DAT(3000),AB(1000,11)
      DIMENSION TSYST(30),TMOD(30)
C
```

```
      SUBROUTINE SIMUL(U,Y,T,AMPL,AL,M,NA,NB,NC,IU1,IU2,IE)
C
C     COMPUTES A SIMULATION OF THE SYSTEM
C     A(Q)*Y(T)=B(Q)*U(T)+C(Q)*E(T)     E(T) GAUSSIAN WHITE NOISE
C     A(Q)=1 + A(1)*Q**(-1) +...+ A(NA)*Q**(-NA)
C     B(Q)=    B(1)*Q**(-1) +...+ B(NB)*Q**(-NB)
C     C(Q)=1 + C(1)*Q**(-1) +...+ C(NC)*Q**(-NC)
C     STARTVALUES OF U(T) AND E(T) ARE ZERO.
C     AUTHOR TORSTEN SODERSTROM, 1971-10-01
C
C     U - VECTOR OF ORDER M CONTAINING THE INPUT
C     Y - VECTOR OF ORDER M CONTAINING THE OUTPUT
C     T - VECTOR OF ORDER (NA+NB+NC) CONTAINING THE PARAMETERS
C     T=(A(1)...A(NA),B(1)...B(NB),C(1)...C(NC))
C     AMPL - AMPLITUDE OR STANDARD DEVIATION OF THE INPUT SIGNAL
C     AL - STANDARD DEVIATION OF THE NOISE
C     M - ORDER OF U,Y  (MIN 1,NO MAX)
C     NA - ORDER OF A
C     NB - ORDER OF B
C     NC - ORDER OF C
C     (NA+NB+NC) (MIN 0,MAX 30)
C     IU1 - =1 THE INPUT SIGNAL IS A PRBS.
C            2 THE INPUT SIGNAL IS A STEP AT TIME T=1
C            3 THE INPUT SIGNAL IS AN IMPULSE AT TIME T=1
C            4 THE INPUT SIGNAL IS WHITE NOISE INDEPENDENT OF E(T)
C            5 THE INPUT SIGNAL IS CONTAINED IN U
C     IU2 - NUMBER OF BITS IN THE SHIFTREGISTER FOR THE PRBSGENERATOR
C     (MIN3,MAX17)
C     IE - STARTVALUES TO NODI  IE MUST BE AN ODD INTEGER
C
C     ATTENTION. FOR BEST RESULT THE VALUE OF IE MUST BE CHOSEN WITH CARE
C
C     SUBROUTINE REQUIRED
C             NODI
C             PRBSTA
C             PRB
C
      DIMENSION U(1),Y(1),T(1)
      DIMENSION FI(30),LA(17),LX(17)
C
```

```
      SUBROUTINE GLS(DAT,T,AB,M,NA,NB,NC,ITER,ITER1,IFILT,INIT,IPRINT,
     *EPST,IA,IB)

      COMPUTES THE GENERALIZED LEAST SQUARES ESTIMATE
      A(Q)*C(Q) Y(T) = B(Q)*C(Q) U(T) + E(T)
      A(Q)=1 + A(1)*Q**(-1) +...+ A(NA)*Q**(-NA)
      B(Q)=    B(1)*Q**(-1) +...+ B(NB)*Q**(-NB)
      C(Q)=1 + C(1)*Q**(-1) +...+ C(NC)*Q**(-NC)

      AUTHOR  TORSTEN SODERSTROM  1971-10-01
      DAT - VECTOR OF ORDER 3*M, CONTAINING THE DATA IN THE FOLLOWING FORM
      TIME(1),U(1),Y(1),TIME(2),... Y(M)
      T - VECTOR OF ORDER (NA+NB+NC) AT RETURN CONTAINING THE PARAMETER
      ESTIMATES
      T = (A(1),...A(NA),B(1),...B(NB.,C(1),...C(NC))
      AB - MATRIX OF ORDER M*(NA+NB+NC) USED INTERNLY
      M - ORDER OF U AND Y (NUMBER OF SAMPLES) (MIN 31,MAX 1000)
      NA,NB,NC - ORDER OF A,B,C RESP.
      (NA+NB+NC)   (MIN 0,MAX 30)
      ITER - MAX NUMBER OF ITERATIONS (MIN 0,NO MAX)
      ITER1 - MAX NUMBER OF VGLS-CALLS (MIN 1,NO MAX)
      IFILT - IFILT=0 THE FILTER C(Q) IS APPLIED TO ORIGINAL DATA
            - IFILT=1 THE FILTER C(Q) IS APPLIED TO FILTERED DATA
      INIT - INIT=0 THE ITERATION IS STARTED WITH THE LS-ESTIMATES OF A AND B
             INIT=1 THE ITERATION IS STARTED WITH GIVEN VALUES OF A AND B
             INIT=2 THE ITERATION IS STARTED WITH GIVEN VALUES OF C
      IPRINT - IPRINT =0 MINIMAL RESULTS ARE PRINTED
               IPRINT =1 MEDIUM  RESULTS ARE PRINTED
               IPRINT =2 MUCH    RESULTS ARE PRINTED
      EPST - TEST QUANTITY FOR STOP OF ITERATIONS
      IA,IB DIMENSION PARAMETERS OF AB

      THE VECTOR DAT IS NOT DESTROYED

      SUBROUTINE REQUIRED
             LS
             LSQ
             RESID
             FILT
             VGLS
             DSYMIN
             EIGS

      DIMENSION DAT(1),T(1),AB(IA,IB)
      DIMENSION U(1000),UF(1000),Y(1000),YF(1000),RES(1000),DATA(3000)
      DIMENSION T1(30),T2(30),TT(30),NNB(1)
      COMMON/LSCOM/ V,SS,P(50,50),C(50),Q(50)
```

```
      SUBROUTINE PRBSTA(LA,NA)
C
C     SUBROUTINE TO START UP THE PRB-SUBROUTINE
C
C     REFERENCES, W. W. PETERSON, ERROR-CORRECTING CODES
C     B. ROSENGREN AND I. NORDH, KONSTRUKTION AV PRBS-GENERATOR
C     M. RUDEMO, ON PSEUDO-RANDOM NOISE GENERATED BY SHIFT REGISTERS
C     AUTHOR, STURE LINDAHL 1970-02-10
C     REVISE, STURE LINDAHL 1970-11-24
C
C     LA VECTOR, CONTAINING THE FEEDBACK-POLYNOMIAL
C     NA NUMBER OF BITS IN THE SHIFTREGISTER
C
C     NA MUST BE IN THE RANGE 3.LE.NA.LE.17
C
C     SUBROUTINE REQUIRED
C            NONE
      DIMENSION LA(1)
```

```
      SUBROUTINE PRB(LA,LX,Y,NA,AMP)
C
C     SUBROUTINE TO GENERATE A NEW STATE IN A PRBS-GENERATOR
C     REFERENCES, W. W. PETERSON, ERROR CORRECTING CODES
C     B. ROSENGREN AND I. NORDH, KONSTRUKTION AV PRBS-GENERATOR
C     M. RUDEMO, ON PSEUDO-RANDOM NOISE GENERATED BY SHIFT REGISTERS
C     AUTHOR, STURE LINDAHL 1970-02-10
C     REVISED, STURE LINDAHL 1970-11-23
C
C     LA VECTOR, CONTAINING THE FEEDBACK-POLYNOMIAL
C     LX VECTOR, CONTAINING THE ACTUAL STATE
C     Y  OUTPUT FROM PRBS-GENERATOR
C     NA NUMBER OF BITS IN THE SHIFTREGISTER
C     AMP SPECIFIED AMPLITUDE OF OUTPUT-SIGNAL
C
C     LA CAN BE ASSIGNED VALUES IN A STARTROUTINE PRBSTA
C
C     SUBROUTINE REQUIRED
C            NONE
C
      DIMENSION LA(1),LX(1)
```

```
      SUBROUTINE NODI(NODD,GAUSS)
C
C      GENERATES RANDOM NUMBERS N(0,1).SUITED FOR REPEATED USE.
C     REFERENCE B JANSON, RANDOM NUMBER GENERATORS.
C     AUTHOR K EKLUND 9/9 1970
C
C
C     GAUSS-RETURNED CONTAINING A RANDOM NUMBER N(0,1)
C     NODD -BY FIRST CALL OF NODI, NODD MUST EQUAL AN ODD INTEGER
C           NODD IS RETURNED CONTAINING A NEW ODD INTEGER WHICH
C           IS USED BY REPEATED CALLS
C
C     SUBROUTINES REQUIRED
C         NONE
C
```

```
      SUBROUTINE LS(DAT,T,AB,M,NU,NA,NB,IA,IB,IPRINT)
C
C     COMPUTES LEAST SQUARES MODEL
C     Y(T)+A(1)*Y(T-1)+..+A(NA)*Y(T-NA)=
C     B1(1)*U1(T-1)+...B1(NB(1))*U1(T-NB(1))+...
C     BNU(1)*UNU(T-1)+...BNU(NB(NU))*UNU(T-NB(NU)))+E(T)
C     AUTHOR, TORSTEN SODERSTROM, 1970-03-03
C     REVISED, TORSTEN SODERSTROM, 1971-10-01
C
C     DAT-VECTOR OF ORDER M*(NA+NB(1)+...+NB(NU)+1)
C     CONTAINING THE DATA IN THE FOLLOWING FORM
C     TIME(1),U1(1),U2(1),...UNU(1),Y(1)...
C     TIME(2),U1(2),U2(2),...UNU(2),Y(2),...
C     TIME(M),U1(M),U2(M),...UNU(M),Y(M)
C     T-VECTOR OF ORDER (NA+NB(1)+...NB(NU))
C     T=(A(1),..A(NA),B1(1),..B1(NB(1),B2(1),..BNU(NB(NU)))
C     AB-MATRIX OF ORDER M*(NA+NB(1)+...+NB(NU)+1) USED INTERNLY
C     M-NUMBER OF SAMPLES (NO MAX)
C     NA-NUMBER OF A-PARAMETERS.
C     NU-NUMBER OF INPUTS
C     NB-VECTOR OF ORDER NU
C     NB(I) IS THE NUMBER OF BI-PARAMETERS
C     THE FOLLOWING RESTRICTIONS ON M,NA,NU,NB MUST HOLD
C     (NA+NB(1)+...NB(NU)) (MIN 0,MAX 50)
C     NA+NB(1)+...NB(NU)+MAX(NA,NB(1),...NB(NU)) .LT. M
C     IA,IB - DIMENSION PARAMETERS OF AB
C     IPRINT-PRINT PARAMETER.
C     IPRINT=0-NOTHING IS PRINTED.
C     IPRINT=1 THE PARAMETERS ESTIMATES AND STANDARD DEVIATIONS
C              THE LOSS FUNCTION AND THE SINGULAR VALUES ARE PRINTED
C     IPRINT=2 AS IPRINT=1 + THE COVARIANCE MATRIX OF THE PARAMETER
C              ESTIMATES IS PRINTED
C
C     THE FOLLOWING VARIABLES LIE IN A COMMON BLOCK CALLED /LSCOM/
C     V-THE LOSS FUNCTION
C     S-ESTIMATED STANDARD DEVIATION OF THE NOISE
C     P-MATRIX OF DIMENSION 50*50 - THE COVARIANCE MATRIX OF
C     THE PARAMETER ESTIMATES
C     C-VECTOR OF DIMENSION 50 - THE STANDARD DEVIATION OF
C     THE PARAMETER ESTIMATES
C     Q-VECTOR OF DIMENSION 50 CONTAINING THE SINGULAR VALUES
C
C     THE VECTOR DAT IS NOT DESTROYED
C
C     SUBROUTINE REQUIRED
C             LSQ
C
      DIMENSION AB(IA,IB)
      DIMENSION DAT(1),T(1),NB(1)
      COMMON /LSCOM / V,S,P(50,50),C(50),Q(50)
      DIMENSION XX(50,1)
C
```

```
      SUBROUTINE LSQ(AB,XX,Q,EPS,MM,NN,JJP,IM,IN,IP,INP)
C
C     COMPUTES THE LEAST SQUARES SOLUTION  OF THE SYSTEM A*X=B USING
C     SINGULAR VALUE DECOMPOSITION.
C     REFERENCE, GOLUB-REINSCH,SINGULAR VALUE DECOMPOSITION AND
C     LEAST SQUARES SOLUTIONS.
C     AUTHOR,TORSTEN SODERSTROM,11/06-70.
C
C     AB-MATRIX OF ORDER MM*(NN+JJP). THE FIRST NN COLUMNS CONTAIN
C     THE MATRIX A.  THE LAST JJP COLUMMNS CONTAIN THE MATRIX B.
C     XX-MATRIX OF ORDER NN*JJP,RETURNED CONTAINING THE LEAST
C     SQUARES SOLUTION.
C     Q-VECTOR OF ORDER NN, RETURNED CONTAINING THE SINGULAR VALUES OF A.
C     EPS-IF ANY ELEMENT OF Q IS .LT. EPS*MAX Q(I), IT IS
C     CONSIDERED AS ZERO.
C     MM-NUMBER OF ROWS OF A (NO MAX).
C     NN-NUMBER OF COLUMNS OF A (MAX 50). NN .LE. MM.
C     JJP-NUMBER OF COLUMNS OF B (NO MAX).
C     IM,IN,IP,INP-DIMENSION PARAMETERS.
C
C     ATTENTION. THE MATRIX AB IS DESTROYED.
C
C     SUBROUTINE REQUIRED
C             NONE
C
      DIMENSION AB(IM,INP),XX(IN,IP),Q(IN)
      DIMENSION E(50)
C
```

```
      SUBROUTINE FILT(U,UF,X,M,N)
C
C     COMPUTES THE FILTERED SIGNAL
C     UF(T)=U(T)+X(1)*U(T-1)+...+X(N)*U(T-N)
C     STARTVALUES OF U(T) ARE ASSUMED TO BE ZERO
C
C     AUTHOR, TORSTEN SODERSTROM 1971-10-15
C
C     U - VECTOR OF ORDER M, CONTAINING THE SIGNAL TO BE FILTERED
C     UF- VECTOR OF ORDER M, CONTAINING THE FILTERED SIGNAL
C     X - VECTOR OF ORDER N, CONTAINING THE FILTER
C     M - ORDER OF U (MIN 1,NO MAX)
C     N - ORDER OF X (MIN 0,MAX 20)
C     N.LE.M
C
C     SUBROUTINE REQUIRED
C            NONE
C
      DIMENSION U(1),UF(1),X(1)
      DIMENSION FI(20)
C
```

```
      SUBROUTINE RESID(U,Y,RES,X,M,NA,NB)
C
C     COMPUTES THE RESIDUALS
C     RES(T)=Y(T)+A(1)*Y(T-1)+...A(NA)*Y(T-NA)-
C     -B(1)*U(T-1)-...-B(NB)*U(Y-NB)
C     RES(T)=0  T=1,... MAX(NA,NB)
C
C     AUTHOR  TORSTEN SODERSTROM 1971-10-15
C
C     U - VECTOR OF ORDER M, CONTAINING THE INPUT SIGNAL
C     Y - VECTOR OF ORDER M, CONTAINING THE OUTPUT SIGNAL
C     RES - VECTOR OF ORDER M ,CONTAINING THE RESIDUALS
C     X - VECTOR OF ORDER (NA+NB)
C     X=(A(1),...A(NA),B(1),...B(NB))
C     M- NUMBER OF SAMPLES  (MIN 1,NO MAX)
C     NA,NB - ORDER OF A RESP B
C     (NA+NB) (MIN 0,MAX 20)
C     MAX(NA,NB) .LT. M
C
C     SUBROUTINE REQUIRED
C            NONE
C
      DIMENSION U(1),Y(1),RES(1),X(1)
      DIMENSION FI(21)
C
```

```
      SUBROUTINE VGLS(U,UF,Y,YF,RES,T,M,NA,NB,NC,IFILT,IPRINT,ITMAX)
C
C     COMPUTES THE LOSS FUNCTION ETC FOR THE GLS PROBLEM
C
C     AUTHOR  TORSTEN SODERSTROM 1971-10-01
C
C     U - VECTOR OF ORDER M CONTAINING THE INPUT
C     UF- VECTOR OF ORDER M CONTAINING THE FILTERED INPUT
C     Y - VECTOR OF ORDER M CONTAINING THE OUTPUT
C     YF- VECTOR OF ORDER M CONTAINING THE FILTERED OUTPUT
C     RES-VECTOR OF ORDER M CONTAINING THE RESIDUALS  RES(T)=A(Q)*Y(T)-B(Q)*U(T
C     T - VECTOR OF ORDER (NA+NB+NC) CONTAINING THE ACTUAL PARAMETER VALUES
C     M-ORDER OF U AND Y (NUMBER OF SAMPLES) (MIN 31,MAX 1000)
C     NA,NB,NC - NUMBER OF A,B,C PARAMETERS RESP
C     (NA+NB+NC) (MIN 0,MAX 30)
C     IFILT - IFILT=0 THE FILTER C(Q) IS APPLIED TO ORIGINAL DATA
C          - IFILT=1 THE FILTER C(Q) IS APPLIED TO FILTERED DATA
C     IPRINT -PRINT PARAMETER
C     THE FOLLOWING VARIABLES ARE PRINTED
C     IPRINT=0 THE LOSS FUNCTION AND THE GRADIENT
C              STANDARD DEVIATIONS OF THE PARAMETERS AND THE NOISE
C              EXTRAPOLATED PARAMETER ESTIMATES BASED ON NEWTON-RAPHSON
C     ELSE     AS IPRINT=0 +
C              THE MATRIX OF SECOND ORDER DERIVATIVES
C              ITS EIGENVALUES AND EIGENVECTORS
C              THE ESTIMATED COVARIANCE MATRIX OF THE PARAMETER ESTIMATES
C     ITMAX - MAX NUMBER OF NEWTON RAPHSON STEPS.
C
C     SUBROUTINE REQUIRED
C             FILT
C             RESID
C             DSYMIN
C             EIGS
C
      DIMENSION U(1),UF(1),Y(1),YF(1),RES(1),T(1)
      DIMENSION RESF(1000),VT(30),VTT(30,30),P(30,30),DT(30),
     FT2(30),R(30,30),EV(30),C(20)
      DOUBLE PRECISION P
C
```

```
      SUBROUTINE DSYMIN(N,IA,IFAIL,A)
C
C     DOUBLE PRECISION VERSION OF SUBROUTINE SYMIN.
C     SUBROUTINE FOR INVERSION OF SYMMETRIC MATRICES.
C     REFERENCE,RUTISHAUSER,CACM,ALG.NR.150.
C     AUTHOR,K.MORTENSSON 04/04-68.
C
C     A-MATRIX TO BE INVERTED.UPON RETURN A CONTAINS A-1 IF THE
C     INVERSION HAS SUCCEEDED.
C     N-ORDER OF A.
C     IFAIL-RETURNED 0 IF THE SUBROUTINE HAS EXECUTED CORRECTLY,
C     1 IF NOT.
C     IA-DIMENSION PARAMETER.
C     CAUTION.NEAR-SINGULAR MATRICES MAY GIVE MISLEADING RESULTS.
C     MAXIMUM ORDER OF A=40.
C
C     SUBROUTINE REQUIRED
C             NONE
C
      DOUBLE PRECISION A,BIG,TEST,Q,P
C
      DIMENSION A(IA,IA),P(40),Q(40),IR(40)
C
```

```
      SUBROUTINE EIGS(A,R,EV,N,IA,MV)
C
C     COMPUTES EIGENVALUES AND EIGENVECTORS OF A REAL SYMMETRIC MATRIX
C     USING THRESHOLD JACOBI METHOD.
C     REFERENCE, RALSTON AND WILF, MATHEMATICAL METHODS FOR DIGITAL
C     COMPUTERS, CHAPTER 7.
C     AUTHOR, C.KALLSTROM 1970-07-16.
C
C     A -ORIGINAL MATRIX (SYMMETRIC). DESTROYED IN COMPUTATION.
C        RESULTANT EIGENVALUES ARE DEVELOPED IN DIAGONAL OF MATRIX IN
C        DESCENDING ORDER.
C     R -RESULTANT MATRIX OF EIGENVECTORS (STORED COLUMNWISE, IN SAME
C        SEQUENCE AS EIGENVALUES).
C     EV VECTOR CONTAINING THE EIG. VALUES IN DESCENDING ORDER.
C     N -ORDER OF MATRICES A AND R.
C     IA DIMENSION PARAMETER.
C     MV-INPUT CODE
C        0  COMPUTE EIGENVALUES AND EIGENVECTORS.
C        1  COMPUTE EIGENVALUES ONLY (R MUST STILL APPEAR IN CALLING
C           SEQUENCE).
C     THE OFF-DIAGONAL ELEMENTS IN A ARE SET EQUAL TO 0 BEFORE RETURN.
C     THERE ARE NO MAXIMUM ORDER OF THE MATRICES A AND R.
C
C     SUBROUTINE REQUIRED
C           NONE
C
      DIMENSION A(IA,IA),R(IA,IA),EV(1)
C
```