Uniqueness of the Maximum Likelihood Estimates of the Parameters of A Mixed Autoregressive Moving Average Process

Åström, Karl Johan; Söderström, Torsten

1973

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

Link to publication

Citation for published version (APA):
Åström, K. J., & Söderström, T. (1973). *Uniqueness of the Maximum Likelihood Estimates of the Parameters of A Mixed Autoregressive Moving Average Process*. (Technical Reports TFRT-7026). Department of Automatic Control, Lund Institute of Technology (LTH).

Total number of authors:
2

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

UNIQUENESS OF THE MAXIMUM LIKELIHOOD ESTIMATES OF THE PARAMETERS OF A MIXED AUTOREGRESSIVE MOVING AVERAGE PROCESS

K J ÅSTRÖM
T SÖDERSTRÖM

Report 7306 March 1973
Lund Institute of Technology
Division of Automatic Control

UNIQUENESS OF THE MAXIMUM LIKELIHOOD ESTIMATES OF THE PARAMETERS OF A MIXED AUTOREGRESSIVE MOVING AVERAGE PROCESS.

K.J. Åström and T. Söderström

ABSTRACT.

Estimation of the parameters in a mixed autoregressive moving average process leads to a nonlinear optimization problem. The negative logarithm of the likelihood function, suitably normalized, converges to a deterministic function, called the loss function, as the sample length increases. The local and global extrema of this loss function are investigated. Conditions for the existence of a unique local minimum are given.

TABLE OF CONTENTS.

1. INTRODUCTION.

Let $\{y(t), t = 1, 2, \ldots\}$ be a stationary gaussian stochastic process with rational spectral density. It follows from the representation theorem, see e.g. Åström (1970), that the process can be representated as a mixed autoregressive moving average process, i.e.

$$A(q)y(t) = C(q)e(t) \tag{1.1}$$

where $e(t)$ is a sequence of independent normal (0,1) random variables. The operators $A(q)$ and $C(q)$ are given by

$$\begin{cases} A(q) = q^n + a_1 q^{n-1} + \ldots + a_n \\ C(q) = q^n + c_1 q^{n-1} + \ldots + c_n \end{cases} \tag{1.2}$$

where q is the forward shift operator.

It follows from the representation theorem that the polynomial $A(z)$ can be chosen so that it has all zeros inside the unit circle. The polynomial $C(z)$ may have zeros inside and on the unit circle. The number n can be chosen so that $A(q)$ and $C(q)$ have no common factors.

The estimation of the parameters $a_1, \ldots a_n, c_1, \ldots c_n$ with the maximum likelihood method leads to the problem of minimizing the function

$$V^N(\hat{a}_1, \ldots \hat{a}_n, \hat{c}_1, \ldots \hat{c}_n) = \frac{1}{2N} \sum_{t=1}^{N} \varepsilon^2(t) \tag{1.3}$$

See Åström-Bohlin (1966). The residual $\varepsilon(t)$ is a function of the observations $y(1), y(2), \ldots y(t)$. It is defined by

$$\varepsilon(t) = \frac{\hat{A}(q)}{\hat{C}(q)} y(t) = \frac{\hat{A}(q)C(q)}{\hat{C}(q)A(q)} e(t) \tag{1.4}$$

where

$$\begin{cases} \hat{A}(q) = q^n + \hat{a}_1 q^{n-1} + \ldots + \hat{a}_n \\ \hat{C}(q) = q^n + \hat{c}_1 q^{n-1} + \ldots + \hat{c}_n \end{cases} \tag{1.5}$$

Since $\hat{C}$ and A are assumed to have zeros strictly inside the unit circle and since we only are considering asymptotic properties the initial conditions of (1.4) are not important. They can e.g. be selected as zero.

The maximum likelihood estimates of the model parameters are obtained by finding the absolute minimum of V^N for each N. It can be shown that the estimate will converge to the true parameter values if the polynomials A(z) and $\hat{C}(z)$ have zeros strictly inside the unit circle. Since the function V^N is nonlinear in $\hat{c}_1, \ldots, \hat{c}_n$ the minimization must be done numerically. It may happen that the function V^N has several local minima. The existence of local minima may lead to wrong estimates and cause difficulties in the computations.

Since V^N is a random variable it is in general very difficult to analyse the existence of possible local minima. It can, however, be shown that V^N under mild conditions, see Hannan (1960), converges with probability one to the function V, defined by

$$V(\hat{a}_1 \ldots \hat{a}_n, \hat{c}_1 \ldots \hat{c}_n) = \lim_{N\to\infty} V^N(\hat{a}_1 \ldots \hat{a}_n, \hat{c}_1 \ldots \hat{c}_n) =$$

$$= \frac{1}{2} E\varepsilon^2(t) = \frac{1}{4\pi i} \oint \frac{\hat{A}(z)C(z)\hat{A}(z^{-1})C(z^{-1})}{A(z)\hat{C}(z)A(z^{-1})\hat{C}(z^{-1})} \frac{dz}{z} \qquad (1.6)$$

where the integral path is the unit circle.

The purpose of this report is to find all the local extrema of (1.6).

2. STATEMENT OF THE PROBLEM.

It was previously assumed that

- $n = \deg A(z) = \deg C(z) = \deg \hat{A}(z) = \deg \hat{C}(z)$
  $A(z)$ and $C(z)$ have no common factors (2.1)

This condition can be generalized somewhat. For technical reasons it will be suitable to assume that

- $n = \deg A(z) = \deg \hat{A}(z)$
  $m = \deg C(z) = \deg \hat{C}(z)$ (2.2)

and allow common factors in A and C. It is not necessary that $n = m$, although this may be a natural choice.

The apparently more general assumption

$$\deg A(z) \leq \deg \hat{A}(z)$$
$$\deg C(z) \leq \deg \hat{C}(z)$$

is easily obtained from (2.2) putting the last a_i and c_i parameters zero when necessary.

The polynomials involved are now rewritten as

$$\begin{cases} A(z) = z^n + a_1 z^{n-1} + \ldots + a_n = \prod_1^n (z-\alpha_i) \\ \hat{A}(z) = z^n + \hat{a}_1 z^{n-1} + \ldots + \hat{a}_n = \prod_1^n (z-\hat{\alpha}_i) \\ C(z) = z^m + c_1 z^{m-1} + \ldots + c_m = \prod_1^m (z-\gamma_i) \\ \hat{C}(z) = z^m + \hat{c}_1 z^{m-1} + \ldots + \hat{c}_m = \prod_1^m (z-\hat{\gamma}_i) \end{cases} \qquad (2.3)$$

To establish convergence of V^N it is furthermore assumed that

$$|\alpha_i| < 1 \qquad 1 \leq i \leq n \tag{2.4}$$

$$|\hat{\gamma}_j| < 1 \qquad 1 \leq j \leq m \tag{2.5}$$

To ensure that $\hat{a}_i = a_i$, $i = 1, \ldots, n$, $\hat{c}_j = c_j$, $i = 1, \ldots, m$, can be a local minimum the following conditions are assumed

$$|\hat{\alpha}_i| < 1 \qquad 1 \leq i \leq n \tag{2.6}$$

$$|\gamma_j| < 1 \qquad 1 \leq j \leq m \tag{2.7}$$

The conditions (2.4) and (2.5) are required to guarantee that the residuals will have finite variance. The condition (2.7) restricts all zeros of C to lie inside the unit circle.

The problem is to find all local extrema of the loss function V (1.6) subject to the constraints (2.4) - (2.7).

3. PRELIMINARIES.

The local extrema of the loss function will now be determined. The calculations are technical but straight-forward. The results are summarized as Lemma 4.1 in Section 4.

Introduce the reciprocials of the polynomials A, Â, C, and Ĉ defined as

$$\begin{cases} A^*(z) = 1 + a_1 z + \ldots + a_n z^n = z^n A(z^{-1}) \\ \hat{A}^*(z) = 1 + \hat{a}_1 z + \ldots + \hat{a}_n z^n = z^n \hat{A}(z^{-1}) \\ C^*(z) = 1 + c_1 z + \ldots + c_m z^m = z^m C(z^{-1}) \\ \hat{C}^*(z) = 1 + \hat{c}_1 z + \ldots + \hat{c}_m z^m = z^m \hat{C}(z^{-1}) \end{cases} \tag{3.1}$$

The stationary points of V are the solutions of

$$\begin{cases} \dfrac{\partial V}{\partial \hat{a}_i} = 0 & 1 \leq i \leq n \\ \dfrac{\partial V}{\partial \hat{c}_i} = 0 & 1 \leq i \leq m \end{cases} \tag{3.2}$$

After some computations we find that these conditions can be written as

$$\frac{1}{2\pi i} \oint z^i \frac{\hat{A}(z)C(z)C^*(z)}{A(z)A^*(z)\hat{C}(z)\hat{C}^*(z)} \frac{dz}{z} = 0 \qquad 1 \leq i \leq n$$

$$\frac{1}{2\pi i} \oint z^i \frac{\hat{A}(z)\hat{A}^*(z)C(z)C^*(z)}{A(z)A^*(z)\hat{C}(z)\hat{C}^*(z)^2} \frac{dz}{z} = 0 \qquad 1 \leq i \leq m \tag{3.3}$$

To avoid the formal difficulty that may arise if A and C have a common factor the polynomials A' and C' are now introduced.

$$
\begin{cases}
A(z) = A'(z)D(z) \\
C(z) = C'(z)D(z) \\
A'(z) = \prod_{1}^{n-k} (z-\alpha_i) \\
C'(z) = \prod_{1}^{m-k} (z-\gamma_i) \\
D(z) = \prod_{1}^{k} (z-\delta_i) \\
A'(z) \text{ and } C'(z) \text{ relatively prime} \\
\quad (\alpha_i \neq \gamma_j \quad 1 \leq i \leq n-k,\ 1 \leq j \leq m-k)
\end{cases}
\tag{3.4}
$$

The case $k = 0$ is permitted.

In the same way assume

$$
\begin{cases}
\hat{A}(z) = \hat{A}'(z)\hat{D}(z) \\
\hat{C}(z) = \hat{C}'(z)\hat{D}(z) \\
\hat{A}'(z) = \prod_{1}^{n-\hat{k}} (z-\hat{\alpha}_i) \\
\hat{C}'(z) = \prod_{1}^{m-\hat{k}} (z-\hat{\gamma}_i) \\
\hat{D}(z) = \prod_{1}^{\hat{k}} (z-\hat{\delta}_i) \\
\hat{A}'(z) \text{ and } \hat{C}'(z) \text{ relatively prime}
\end{cases}
\tag{3.5}
$$

Note that the value of $\hat{k}$ depends on the actual point $(\hat{a}_1, \ldots, \hat{a}_n, \hat{c}_1, \ldots, \hat{c}_m)$ in the parameter space.

The polynomials $A'^*(z)$, $\hat{A}'^*(z)$, $C'^*(z)$ and $\hat{C}'^*(z)$ are defined analogous with (3.1).

Furthermore introduce the function

$$f(z) = \frac{\hat{A}'(z)C'(z)C'^*(z)}{A'(z)A'^*(z)\hat{C}'(z)\hat{C}'^*(z)\hat{C}^*(z)} \tag{3.6}$$

Using (3.4) - (3.6) the equations (3.3) are rewritten

$$\begin{cases} \dfrac{1}{2\pi i} \oint z^i \hat{C}'^*(z) f(z) \dfrac{dz}{z} = 0 & 1 \leq i \leq n \\ \dfrac{1}{2\pi i} \oint z^i \hat{A}'^*(z) f(z) \dfrac{dz}{z} = 0 & 1 \leq i \leq m \end{cases} \tag{3.7}$$

The definition of $\hat{A}'^*(z)$ and $\hat{C}'^*(z)$ gives

$$\begin{cases} \displaystyle\sum_{j=0}^{m-\hat{k}} \hat{c}'_j \frac{1}{2\pi i} \oint z^{i+j} f(z) \frac{dz}{z} = 0 & 1 \leq i \leq n \\ \displaystyle\sum_{j=0}^{n-\hat{k}} \hat{a}'_j \frac{1}{2\pi i} \oint z^{i+j} f(z) \frac{dz}{z} = 0 & 1 \leq i \leq m \end{cases} \tag{3.8}$$

Define for $p \geq 1$

$$F_p = \frac{1}{2\pi i} \oint z^p f(z) \frac{dz}{z} \tag{3.9}$$

Then (3.8) becomes

$$\begin{bmatrix} 1 & \hat{c}'_1 & \cdots & \hat{c}'_{m-\hat{k}} & & & 0 \\ & \ddots & & & \ddots & & \\ 0 & & 1 & \hat{c}'_1 & & & \hat{c}'_{m-\hat{k}} \\ \hline 1 & \hat{a}'_1 & \cdots & \cdots & \hat{a}'_{n-\hat{k}} & & \\ & \ddots & & & & \ddots & 0 \\ 0 & & 1 & \hat{a}'_1 & \cdots & \cdots & \hat{a}'_{n-\hat{k}} \end{bmatrix} \begin{bmatrix} F_1 \\ \vdots \\ \vdots \\ \vdots \\ F_{n+m-\hat{k}} \end{bmatrix} = 0 \qquad (3.10)$$

The matrix in (3.10) is $(n+m) \times (n+m-\hat{k})$. Since $\hat{A}'(z)$ and $\hat{C}'(z)$ by assumption are relatively prime it follows from elementary algebra that the rank of the matrix is $n+m-\hat{k}$. See e.g. Dickson (1922).

Thus

$$\frac{1}{2\pi i} \oint z^i f(z) \frac{dz}{z} = 0 \qquad 1 \leq i \leq n+m-\hat{k} \qquad (3.11)$$

The poles of $f(z)$ inside the unit circle now are relabelled through

$$A'(z)\hat{C}'(z) = \prod_1^{n-k} (z-\alpha_i) \prod_1^{m-\hat{k}} (z-\hat{\gamma}_j) = \prod_1^{\ell} (z-u_i)^{t_i} \qquad (3.12)$$

where $u_i \neq u_j$ if $i \neq j$, $t_i \geq 1$ all i and

$$\sum_1^{\ell} t_i = n + m - k - \hat{k} \qquad (3.13)$$

This implies that $f(z)$ can be written

$$f(z) = \frac{g(z)}{\prod_{i=1}^{\ell} (z-u_i)^{t_i}} \qquad (3.14)$$

where

$$g(z) = \hat{A}'(z)C'(z) \cdot \frac{C'^*(z)}{A'^*(z)\hat{C}'^*(z)\hat{C}^*(z)} \tag{3.15}$$

is analytic inside the unit circle.

Using (3.14), the equation (3.11) can be replaced by

$$0 = \frac{1}{2\pi i}\oint \frac{z^{i-1}g(z)}{\prod_{1}^{\ell}(z-u_j)^{t_j}}\,dz = \sum_{k=1}^{\ell} \operatorname*{Res}_{z=u_k} \frac{z^{i-1}g(z)}{\prod_{j=1}^{\ell}(z-u_j)^{t_j}} =$$

$$= \sum_{k=1}^{\ell} \frac{1}{(t_k-1)!} D^{(t_k-1)}\left[\frac{z^{i-1}g(z)}{\prod_{\substack{j=1\\ j\neq k}}^{\ell}(z-u_j)^{t_j}}\right]_{z=u_k} =$$

$$= \sum_{k=1}^{\ell} \frac{1}{(t_k-1)!} \sum_{\nu=0}^{t_k-1} \binom{t_k-1}{\nu} D^{(\nu)}[z^{i-1}]_{z=u_k} D^{(t_i-1-\nu)}\left[\frac{g(z)}{\prod_{j\neq k}(z-u_j)^{t_j}}\right]_{z=u_k}$$

where D denotes differentiation with respect to z.

Hence

$$\sum_{k=1}^{\ell}\sum_{\nu=0}^{t_k-1} D^{(\nu)}[z^{i-1}]_{z=u_k} \cdot d_{k\nu} = 0 \qquad 1 \leq i \leq n+m-\hat{k} \tag{3.16}$$

where

$$d_{k\nu} = \frac{1}{\nu!(t_k-1-\nu)!} D^{(t_k-1-\nu)}\left[\frac{g(z)}{\prod_{j\neq k}(z-u_j)^{t_k}}\right]_{z=u_k} \tag{3.17}$$

Using matrix notation (3.16) can be expressed as

$$S \cdot G = 0 \tag{3.18}$$

where G is a $(m+n-k-\hat{k})$ vector,

$$G = \begin{bmatrix} d_{10} \\ \vdots \\ d_{1,t_1-1} \\ d_{20} \\ \vdots \\ d_{\ell,t_\ell-1} \end{bmatrix}$$

and S a $(m+n-\hat{k}) \times (m+n-k-\hat{k})$ matrix

$$S = \begin{bmatrix} 1 & 0 & \cdots & 1 & \cdots & D^{(t_\ell-1)}[z^0]_{z=u_\ell} \\ u_1 & 1 & & u_2 & & D^{(t_\ell-1)}[z^1]_{z=u_\ell} \\ u_1^2 & 2u_1 & & u_2^2 & & D^{(t_\ell-1)}[z^2]_{z=u_\ell} \\ \vdots & \vdots & & \vdots & & \\ u_1^{m+n-\hat{k}-1} & (m+n-\hat{k}-1)u_1^{m+n-\hat{k}-2} & \cdots & u_2^{m+n-\hat{k}-1} & \cdots & D^{(t_\ell-1)}[z^{m+n-\hat{k}-1}]_{z=u_\ell} \end{bmatrix}$$

The matrix S is a generalization of the van der Monde matrix. It follows from Kaufman (1969) that its upper, square part is non-singular. Thus (3.18) implies

$$G = 0 \tag{3.19}$$

A useful lemma will now be proved.

Lemma 3.1. Let $h(z)$ and $g(z)$ be analytic functions in a neighbourhood of z_0. Assume that $g(z_0) \neq 0$. Then

$$D^{(p)}[h(z)g(z)]_{z=z_0} = 0 \qquad 0 \leq p \leq n \tag{3.20}$$

is equivalent to

$$D^{(p)}[h(z)]_{z=z_0} = 0 \qquad 0 \leq p \leq n \tag{3.21}$$

Proof. The equation (3.20) implies

$$\sum_{i=0}^{p} \binom{p}{i} D^{(i)}[h(z)]_{z=z_0} D^{(p-i)}[g(z)]_{z=z_0} = 0 \qquad 0 \leq p \leq n$$

or equivalent in matrix form

$$\begin{bmatrix} D^{(0)}g(z) & & & & 0 \\ D^{(1)}g(z) & D^{(0)}g(z) & & & \\ D^{(2)}g(z) & 2D^{(1)}g(z) & D^{(0)}g(z) & & \\ \vdots & & & & \\ D^{(n)}g(z) & \cdot & \cdot & \cdot & D^{(0)}g(z) \end{bmatrix}_{z=z_0} \cdot \begin{bmatrix} D^{(0)}h(z) \\ D^{(1)}h(z) \\ \vdots \\ D^{(n)}h(z) \end{bmatrix}_{z=z_0} = 0 \tag{3.22}$$

According to the assumption $g(z_0) \neq 0$ the matrix is non-singular and the equivalence between (3.20) and (3.21) follows. □

Equations (3.19) and (3.17) give

$$D^{(t_k-1-\nu)}\left[\frac{g(z)}{\prod_{j\neq k}(z-u_j)^{t_k}}\right]_{z=u_k} = 0 \qquad (3.23)$$

$$0 \leq \nu \leq t_k-1, \; 1 \leq k \leq \ell$$

and it follows from Lemma 3.1 that

$$\begin{cases} D^{(i)}[g(z)]_{z=u_k} = 0 \\ 0 \leq i \leq t_{k-1}, \; 1 \leq k \leq \ell \end{cases} \qquad (3.24)$$

Using the Lemma 3.1 again [with $h(z) = \hat{A}'(z)C'(z)$ cf. (3.15)] the following equations are obtained.

$$\begin{cases} D^{(i)}[\hat{A}'(z)C'(z)]_{z=u_k} = 0 \\ 0 \leq i \leq t_{k-1}, \; 1 \leq k \leq \ell \end{cases} \qquad (3.25)$$

Hence

$$\hat{A}'(z)C'(z) \equiv \prod_{k=1}^{\ell}(z-u_k)^{t_k-1} \equiv A'(z)\hat{C}'(z) \qquad (3.26)$$

Thus it has been shown that the stationary points, i.e. the solutions of (3.7) must fulfil (3.26). Conversely, the calculations show that (3.26) implies (3.7). The latter assertion can be proven directly since (3.26) implies that $f(z)$ has no poles inside the unit circle.

4. MAIN RESULT.

The following lemma is a summary of the calculations in the previous section.

Lemma 4.1. Consider the loss function (1.6) subject to the conditions (2.4), (2.7) and the constraints (2.5), (2.6). Let $A'(z)$, $C'(z)$, $\hat{A}'(z)$, $\hat{C}'(z)$ be defined by (3.4), (3.5). Then the stationary points of V are the solutions of

$$\hat{A}'(z)C'(z) \equiv A'(z)\hat{C}'(z) \tag{4.1}$$

The next lemma deals with global minimum points.

Lemma 4.2. Consider the loss function V (1.6) subject to the conditions (2.4), (2.7) and the constraints (2.5), (2.6). Then the global minimum points of V are the solutions of

$$\hat{A}'(z)C'(z) \equiv A'(z)\hat{C}'(z) \tag{4.1}$$

Proof. Introduce

$$H^*(z) = \frac{\hat{A}^*(z)C^*(z)}{A^*(z)\hat{C}^*(z)} = 1 + \sum_{i=1}^{\infty} h_i z^i$$

where the infinite series converges in and on the unit circle. Put $h_0 = 1$, then

$$V = \frac{1}{4\pi i} \oint \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} h_j z^j h_k z^{-k} \frac{dz}{z} = \frac{1}{2}\left(1 + \sum_{j=1}^{\infty} h_j^2\right)$$

Thus

$$V \geq 1/2$$

with equality if and only if $H(z) \equiv 1$ or

$$\hat{A}(z)C(z) \equiv A(z)\hat{C}(z)$$

Invoking (3.4) and (3.5) we find that this equation is equivalent to (4.1).

□

It remains to analyze the solution of (4.1). The equation can be written as

$$\frac{C'(z)}{A'(z)} \equiv \frac{\hat{C}'(z)}{\hat{A}'(z)} \tag{4.2}$$

The number $\hat{k}$ has not been determined yet. To establish equality in (4.2) for all z it is necessary that both sides have the same poles and zeros. Since there are no common factor it is thus necessary and sufficient that $\hat{k} = k$, $\hat{A}'(z) \equiv A(z)$ and $\hat{C}'(z) = C(z)$.

Two cases can be separated:

1. $k = 0$. Then $\hat{k} = 0$ and the loss function has a unique local minimum

2. $k > 0$. Then $\hat{k} > 0$ and there are infinite many local minimum points. In fact, these minimum points form a manifold in the parameter space. On this manifold the loss function obtains its infimum. This case means that thatt the model contains too many parameters.

Another way of characterization is the following. The unknown parameters are $\hat{k}$, $\hat{a}'_1$, ..., $\hat{a}'_{n-\hat{k}}$, $\hat{c}'_1$, ..., $\hat{c}'_{m-\hat{k}}$, $\hat{d}_1$, ..., $\hat{d}_{\hat{k}}$. Of these must for all minimum points

$$\begin{cases} \hat{k} = k \\ \hat{a}'_i = a'_i & 1 \leq i \leq n-\hat{k} \\ \hat{c}'_i = c'_i & 1 \leq i \leq m-\hat{k} \end{cases}$$

while $\hat{d}_1 \ldots \hat{d}_{\hat{k}}$ (if $\hat{k} > 0$) are arbitrary.

The result of the calculation and the discussion is summed up in the following theorem.

<u>Theorem</u>. Consider the loss function (1.6) subject to the conditions (2.4), (2.7) and the constraints (2.5) and (2.6). Assume that $\deg A(z) = n$, $\deg \hat{A}(z) = n+\hat{n}$, $\deg C(z) = m$, $\deg \hat{C}(z) = m+\hat{m}$ where $\min(\hat{n},\hat{m}) \geq 0$ and that $A(z)$ and $C(z)$ are relatively prime.

i) If $\min(\hat{n},\hat{m}) = 0$ there is a unique local minimum, namely

$$\hat{a}_i = \begin{cases} a_i & 1 \leq i \leq n \\ 0 & \text{if } i > n \text{ and } \hat{n} > 0 \end{cases}$$

$$\hat{c}_i = \begin{cases} c_i & 1 \leq i \leq m \\ 0 & \text{if } i > m \text{ and } \hat{m} > 0 \end{cases}$$

ii) If $\min(\hat{n},\hat{m}) > 0$ there are infinitely many local minimum points given by the manifold

$$\left.\begin{aligned} \hat{A}(z) &\equiv L(z)A(z) \\ \hat{C}(z) &\equiv L(z)C(z)z^{\hat{m}-\hat{n}} \end{aligned}\right\} \quad \text{if } \hat{m} \geq \hat{n}$$

or

$$\hat{A}(z) \equiv L(z)A(z)z^{\hat{n}-\hat{m}}$$

$$\hat{C}(z) \equiv L(z)C(z) \qquad \text{if } \hat{m} \leq \hat{n}$$

$L(z)$ is an arbitrary unitary polynomial of degree $\min(\hat{n},\hat{m})$. Each point in the manifold also is a global minimum point.

iii) There are neither local maxima nor saddle points.

5. ACKNOWLEDGEMENTS.

The authors want to thank Mr. I. Gustavsson for many stimulating discussions.

It is also a pleasure to thank Mrs. G. Christensen who typed the manuscript. The partial support of the Swedish Board for Technical Development under contract 72-202/U137 is gratefully announced.

6. REFERENCES.

Åström, K.J. (1970).
Introduction to Stochastic Control Theory. Academic Press.

Åström, K.J. - Bohlin, T. (1966).
Numerical Identification of Linear Dynamical 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.

Dickson, L.E. (1922).
First Course in the Theory of Equations. Wiley.

Hannan, E.J. (1960).
Time Series Analysis. Meuthen and Co., London.

Kaufman, I. (1969).
The Inversion of the Vandermonde Matrix and the Transformation to the Jordan Canonical Form. IEEE Trans. Aut. Control, AC-14, p. 774-777.

CORRECTIONS

The abbreviation pa.b means page a line b.

p3.3 Read "the integration path"

p10.7 Replace "$D^{(t_i - 1 - \nu)}$" with "$D^{(t_k - 1 - \nu)}$"

p10.12 and p13.2 Replace "$\prod_{j \neq k} (z - u_j)^{t_k}$" with "$\prod_{j \neq k} (z - u_j)^{t_j}$"

p13.12 Read "$\prod_{k=1}^{\ell} (z - u_k)^{t_k}$"

p15.13 Read "$\hat{C}'(z) \equiv \hat{C}(z)$"

p15.20 Delete "that"