Selected Transl. in Math.
Statist. and Probability
Volume 14, 1978

ON THE STATISTICAL EXAMINATION OF CONTINUOUS STATE MARKOV PROCESSES. IV*

MÁTYÁS ARATÓ

Introduction

In this paper we shall deal with results which are connected with estimates of parameters of n-dimensional discrete stationary Gaussian Markov processes and their density functions. Textbooks (e.g. Grenander and Rosenblatt [8]) give these estimates and their densities under assumptions concerning the eigenvalues of a matrix A (see below, formula (1.2)). This means that we have to have some knowledge concerning the process itself, and particularly concerning the matrix A, beforehand; this, in practice, is usually not the case. In this paper I shall outline a treatment which, on the one hand, enables us to reduce certain special cases of the general case to the one-dimensional real or complex case (with simple eigenvalues), and, on the other hand, enables us to treat the solution of the general case more easily. To write the matrix A (or Q in the case of differential equations) in the Jordan form requires some preliminary nonstatistical investigations. To the extent that this is not possible, the solution of particular problems becomes very complicated, and a very long sequence of observations is necessary to get a reliable estimate of the parameters (let us recall here that the calculation of eigenvalues of a matrix involves serious numerical difficulties). I shall show which of the parameters can be well estimated (i.e. from fewer observations with greater accuracy). I am also including results I obtained earlier concerning sufficient statistics of the processes we are dealing with. The results concerning the theory of n-dimensional stationary processes can be found in the recent book by Rozanov [11], which will often be cited, sometimes without special reference. The basic results of this paper formed part of my dissertation [4], but so far they have not been published.

This paper does not refer to any practical task; I mention first of all the significance of application for communication theory and for economics (see e.g. the examples given in the work by Quenouille [10]).

Discrete n-Dimensional Stationary Normal Case

§1. Systems of stochastic differential or difference equations with constant coefficients

Any n-dimensional regular stochastic process $\xi(t) = \{\xi_k(t)\}_{k=1}^n$ of the Gaussian Markov type—similar to the one- or two-dimensional case—satisfies the following stochastic differential equation:

$$d\xi(t) = Q \cdot \xi(t)dt + d\zeta(t), \tag{1.1}$$

where $Q = \{q_{ij}\}_{i,j=1}^n$ denotes a square matrix whose eigenvalues λ_i satisfy the conditions $\text{Re}\, \lambda_i < 0$ and $\zeta(t) = \{\zeta_k(t)\}_{k=1}^n$ denotes an n-dimensional Wiener process; $\mathbf{E}\Delta\zeta_k(t) = 0$ and $\mathbf{E}\Delta\zeta_i(t)\zeta_j(t) = \Delta t \cdot s_{ij}$, where the matrix $\{s_{ij}\}_{i,j=1}^n$ is positive definite. The matrix Q is uniquely determined (see Doob [6]).

The process $\xi(k\epsilon)$ ($\epsilon > 0$; $k = 0, \pm 1, \dots$) represents a discrete stationary Gaussian Markov process, and therefore it satisfies (for the sake of simplicity we put $\epsilon = 1$)

$$\xi(k+1) = A\xi(k) + \zeta(k+1), \tag{1.2}$$

where $\zeta(k)$ is a sequence of independent Gaussian random vectors. The eigenvalues ρ_i of the matrix A and the eigenvalues λ_i of Q are connected by $\rho_i = e^{\lambda_i}$.

When a linear transformation of R^n is effected by means of a nonsingular matrix S, we obtain matrices $Q' = SQS^{-1}$ and $A' = SAS^{-1}$ corresponding to Q and A respectively, and the corresponding random variables $\xi' = S\xi$ will satisfy

$$d\xi'(t) = Q'\xi'(t)dt + d\zeta'(t), \tag{1.1'}$$

$$\xi'(k+1) = A'\xi'(k) + \zeta'(k+1), \tag{1.2'}$$

respectively, where $\zeta' = S\zeta$. With an appropriate choice of S we can bring Q' (A') to its simplest form. In general, an arbitrary matrix cannot be transformed into a diagonal matrix, but any matrix can be transformed into the so-called Jordan form.

A matrix of order m is called a Jordan elementary matrix if it reads as follows:

$$\begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & 1 \\ 0 & \cdots & \cdots & \cdots & \lambda \end{pmatrix}.$$

A matrix B is called a matrix of Jordan type if it consists of one or more Jordan

elementary matrices, all placed along the main diagonal, and the remaining elements are equal to zero.

We shall call a sequence of vectors $h_1, \dots, h_m$, $h_i \in R^n$, $i = 1, \dots, m$, a chain belonging to the eigenvalue λ of the matrix B if $h_1 \neq 0$, $Bh_1 = \lambda h_1$, $Bh_2 = \lambda h_2 + h_1, \dots, Bh_m = \lambda h_m + h_{m-1}$. The following theorem, showing that any matrix can be transformed into Jordan form, holds true (see, for example, Pontrjagin [9], §34).

THEOREM. *There exists a basis of R^n consisting of all the vectors belonging to one or more chains defined by the matrix B. If B is a real matrix, the chains forming the basis can be chosen so that those corresponding to real eigenvalues are real and those corresponding to complex conjugate eigenvalues are complex conjugate.*

The mapping defined by B can be rewritten in the coordinate system described by the above theorem; the new matrix describing this mapping is now of the Jordan type. After rewriting equations (1.1′) and (1.2′) in the Jordan form, these equations read (after deleting the primes)

$$\begin{aligned} d\xi_1(t) &= \lambda_1\xi_1(t)dt + \xi_2(t)dt + d\zeta_1(t), \\ d\xi_2(t) &= \lambda_1\xi_2(t)dt + \xi_3(t)dt + d\zeta_2(t), \\ &\vdots \\ d\xi_{k_1}(t) &= \lambda_1\xi_{k_1}(t)dt + d\zeta_{k_1}(t), \\ d\xi_{k_1+1}(t) &= \lambda_2\xi_{k_1+1}(t)dt + \xi_{k_1+2}(t)dt + d\zeta_{k_1+1}(t), \\ &\vdots \\ d\xi_n(t) &= \lambda_r\xi_n(t)dt + d\zeta_n(t); \end{aligned} \tag{1.3}$$

$$\begin{aligned} \xi_1(k+1) &= \rho_1\xi_1(k) + \xi_2(k) + \zeta_1(k+1), \\ &\vdots \\ \xi_{k_1}(k+1) &= \rho_1\xi_{k_1}(k) + \zeta_{k_1}(k+1), \\ \xi_{k_1+1}(k+1) &= \rho_2\xi_{k_1+1}(k) + \xi_{k_1+2}(k) + \zeta_{k_1+1}(k+1), \\ &\vdots \\ \xi_n(k+1) &= \rho_r\xi_n(k) + \zeta_n(k+1), \end{aligned} \tag{1.4}$$

respectively. In such cases we may use the results established in the one- or two-dimensional cases without any difficulties, because our system of equations has been decomposed into equations describing real or complex processes. Further, supposing that the above transformation into Jordan form is always possible, we shall deal with one Jordan elementary matrix only.

§2. Sufficient statistics

Let us calculate the density function of a finite realization of the stationary Gaussian process satisfying the following stochastic difference equation:

$$\xi(t) + a_1\xi(t-1) + \cdots + a_p\xi(t-p) = \zeta(t). \tag{2.1}$$

Let us suppose that $\mathsf{E}\xi(t) = 0$ and that the process $\zeta(t)$ is a sequence of independent random variables (which are of course normally distributed). The covariance matrix of the random variables $\xi(1), \ldots, \xi(N)$ is symmetric with respect to both diagonals because of the stationarity of the process; the same holds true for its inverse matrix, denoted by R_N^{-1}. The density function of the variables $\xi(1), \ldots, \xi(N)$ will be

$$p_{\xi(1),\ldots,\xi(N)}(x_1, \ldots, x_N) = (2\pi)^{-N/2}|R_N|^{-1/2}\exp\{-\tfrac{1}{2}(X_N R_N^{-1} X_N^*)\}.$$

The transformation

$$\begin{aligned} x_1 &= x_1, \\ &\vdots \\ x_p &= x_p, \\ x_{p+1} + a_1 x_p + \ldots + a_p x_1 &= z_{p+1}, \\ &\vdots \\ x_N + a_1 x_{N-1} + \ldots + a_p x_{N-p} &= z_N \end{aligned}$$

(the determinant of which equals 1) gives, when the independence of the two groups of variables ($\xi(1), \ldots, \xi(p)$ and $\zeta(p+1), \ldots, \zeta(N)$) is taken into consideration and the assumption of independence is used,

$$p_{\xi(1),\ldots,\xi(p),\zeta(p+1),\ldots,\zeta(N)}(x_1, \ldots, x_p, z_{p+1}, \ldots, z_n)$$
$$= (2\pi)^{-\frac{N}{2}}|R_p|^{-\frac{1}{2}}\sigma_\zeta^{-(N-p)}\exp\left\{-\frac{1}{2}\left[(X_p R_p^{-1} X_p^*) + \frac{1}{\sigma_\zeta^2}\sum_{p+1}^{N} z_i^2\right]\right\}$$

from which it follows that

$$p_{\xi(1),\ldots,\xi(N)}(x_1, \ldots, x_N) = (2\pi)^{-N/2}|R_p|^{-1/2}\sigma_\zeta^{-(N-p)}$$
$$\times \exp\left\{-\frac{1}{2}\left[(X_p R_p^{-1} X_p^*) + \frac{1}{\sigma_\zeta^2}\sum_{p+1}^{N}(x_i + a_1 x_{i-1} + \cdots + a_p x_{i-p})^2\right]\right\},$$

From (2.1) we can calculate all the elements of the matrix R_N^{-1} if we use its above-mentioned symmetry. Let $a_0 = 1$; then

$$R_N^{-1} = \begin{pmatrix} a_0^2 & a_0a_1 & a_0a_2 & \ldots & a_0a_p & 0 & 0 & \ldots & 0 \\ a_0a_1 & a_0^2+a_1^2 & a_0a_1+a_1a_2 & \ldots & a_0a_{p-1}+a_1a_p & a_0a_p & 0 & \ldots & 0 \\ a_0a_2 & a_0a_1+a_1a_2 & a_0^2+a_1^2+a_2^2 & \ldots & & & & & \\ \vdots & & & & & & & & \\ a_0a_p & a_0a_{p-1}+a_1a_p & & \ldots & \sum_1^p a_i^2 & & \sum_1^{p-1} a_ia_{i+1} & \ldots & \\ 0 & & & & & \ddots & & & \\ \vdots & & & & & & \ddots & & \\ 0 & & & & & & & & a_0^2 \end{pmatrix} \tag{2.2}$$

which means that the r_{ij} element of R_N^{-1} has the form

$$r_{ij} = \begin{cases} 0, & \text{for } |i-j| > p+1 \\ \sum_{k=0}^{p-|j-i|} a_k a_{k+|j-i|}, & \text{for } |i-j| \leq p+1,\ p < i, j < N-p+1 \\ \sum_{k=0}^{i} a_k a_{k+j-i}, & \text{for } i < p,\ i \leq j \\ \sum_{k=0}^{i} a_k a_{k+(i-j)}, & \text{for } j \leq i \leq p. \end{cases} \tag{2.3}$$

Moreover,

$$|R_N^{-1}| = |R_p^{-1}| \cdot \sigma_\zeta^{2(N-p)}. \tag{2.4}$$

According to a well-known theorem of Dynkin [7],

$$\log p(X, a) - \log p(X; a^0)$$

forms a minimal sufficient statistic for the family of distributions $p(X, a)$ when a^0 is fixed. Taking this theorem and the above form of R_N^{-1} into consideration, we obtain

THEOREM 2.1. *If a stationary Gaussian process $\xi(t)$ satisfies (2.1), then the system*

$$\left(\sum_{p+1}^{N-p} x_i^2,\ \sum_{p+1}^{N-p+1} x_i x_{i-1}, \ldots, \sum_{p+1}^{N} x_i x_{i-p},\ x_1^2 + x_N^2,\ x_1x_2 + x_Nx_{N-1}, \ldots x_1x_p \right.$$
$$\left. + x_N x_{N-p+1},\ x_1^2 + x_{N-1}^2, \ldots, x_2x_p + x_{N-1}x_{N-p+1}, \ldots, x_p^2 + x_{N-p+1}^2\right)$$

of the sample $x_1, \ldots, x_N$ ($N > p$) forms a minimal sufficient statistic.

We may quote the well-known fact (see, for example, Rozanov [11]) that the spectral density function of a process satisfying (2.1) is a rational function of $e^{i\lambda}$ whose numerator is a constant. We may ask whether in general it is possible to find a sufficient statistic of a process with a rational spectral density function. The following example shows that such a statistic (containing less then N elements) in general does not exist. Indeed, let us consider the process

$$\xi(t) = a_0\zeta(t) + a_1\zeta(t-1),$$

where $\zeta(t)$ is a sequence of independent normal random variables. Let $\mathsf{E}\xi(t) = 0$ and

$$\sigma_\xi^2 = \mathsf{E}\xi^2(t) = (a_0^2 + a_1^2)\,\mathsf{E}\zeta^2(t), \quad \varrho = \frac{\mathsf{E}\xi(t)\xi(t-1)}{\sigma_\xi^2} = \frac{a_0 a_1}{a_0^2 + a_1^2}.$$

Evidently $\mathsf{E}\xi(t)\xi(t-\tau) = 0$ for $|\tau| > 1$. The density of the random variables $\xi(1), \ldots, \xi(N)$ has the following form:

$$p(y_1, \ldots, y_N) = \sigma_\xi^{-N}(2\pi)^{-\frac{N}{2}}|B_N|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma_\xi^2}\sum_{i,j=1}^{N} b_{ij}^* y_i y_j\right\},$$

where $B_N^{-1} = \{b_{ij}^*\}_{i,j=1}^N$ is the inverse of the correlation matrix B_N. It can easily be found that

$$b_{ij}^* = (-1)^{j-i}\varrho^{|j-i|}|B_{i-1}||B_{N-j}|\frac{1}{|B_N|},$$

for $i < j$ and $B_i = (u_1^{i+1} - u_2^{i+1})/(u_1 - u_2)$ (the inverse matrix is, of course, again symmetric with respect to both diagonals). Note that $|B_N|$ satisfies the difference equation

$$|B_N| = |B_{N-1}| - \varrho^2|B_{N-2}|$$

and u_1 and u_2 are the roots of $u^2 - u + \rho^2 = 0$, i.e. $u_1 = (1 + \sqrt{1-4\rho^2})/2$ and $u_2 = (1 - \sqrt{1-4\rho^2})/2$. Since for example the functions b_{iN}^* $(i = 1, \ldots, N)$ as functions of ρ are independent, such a system of functions with $x_1, \ldots, x_N$ as independent variables, which form a sufficient statistic, cannot exist (see Dynkin [7], §2).

In the n-dimensional case let us consider the homogeneous Gaussian Markov process satisfying

$$\xi(t) = A\xi(t-1) + \zeta(t),$$

where $A = \{\alpha_{ij}\}_{i,j=1}^n$, $\xi(t) = \{\xi_j(t)\}_{j=1}^n$, $\zeta(t) = \{\zeta_j(t)\}_{j=1}^n$ and

$$\mathsf{E}\zeta_j(t) = 0$$

$$\mathsf{E}\zeta_i(t)\zeta_j(t+\tau) = \begin{cases} s_{ij}, & \text{for } \tau = 0, \\ 0, & \text{for } \tau \neq 0, \end{cases}$$

where the sequence $\zeta(t)$ is independent.

The correlation function $B(\tau)$ of the process $\xi(t)$ has the form

(2.5) $$B(\tau) = A^\tau \cdot B(0),$$

where

(2.6) $$B(0) = A \cdot B(0) \cdot A^* + B_\zeta(0).$$

If A consists of only one Jordan elementary matrix and $B_\zeta(0)$ is a diagonal matrix, then $B(0)$ can be determined by using the following method: Under the above assumptions $\xi(t)$ satisfies

$$\xi_n(t) = \rho\xi_n(t-1) + \zeta_n(t),$$
$$\xi_{n-1}(t) = \rho\xi_{n-1}(t-1) + \xi_n(t-1) + \zeta_{n-1}(t),$$
$$\vdots$$
$$\xi_1(t) = \rho\xi_1(t-1) + \xi_2(t-1) + \zeta_1(t).$$

By multiplying the first equation by $\xi_n(t-1)$ and by $\xi_n(t)$, and calculating the mean in both cases, we obtain

$$\mathsf{E}\xi_n(t)\zeta_n(t) = s_n$$
$$\mathsf{E}\xi_n(t)\xi_n(t-1) = \varrho\beta_{nn}^2, \quad \text{where} \quad \beta_{nn} = \mathsf{E}\,\xi_n^2,$$
$$\beta_{nn} = \varrho^2 \cdot \beta_{nn} + s_n, \quad \beta_{nn} = \frac{s_n}{1-\varrho^2}.$$

By again multiplying the first equation by $\xi_{n-1}(t)$ and $\xi_{n-1}(t-1)$ respectively, and forming the mean, we obtain

$$\mathsf{E}\xi_n(t)\xi_{n-1}(t) = \varrho\,\mathsf{E}\xi_n(t-1)\xi_{n-1}(t),$$
$$\mathsf{E}\xi_n(t)\xi_{n-1}(t-1) = \varrho\;\;\mathsf{E}\xi_n(t-1)\xi_{n-1}(t-1),$$

respectively. By using these results, multiplying the second equation by $\zeta_{n-1}(t)$, $\xi_n(t)$, $\xi_{n-1}(t)$ and $\xi_n(t-1)$ and forming the mean value again, we obtain

$$\mathsf{E}\xi_{n-1}(t)\zeta_{n-1}(t) = s_{n-1},$$
$$\mathsf{E}\xi_{n-1}(t)\xi_n(t-1) = \frac{s_n}{(1-\varrho^2)^2},$$
$$\mathsf{E}\xi_{n-1}(t-1)\xi_n(t) = \frac{\varrho^2 s_n}{(1-\varrho^2)^2},$$
$$\mathsf{E}\xi_{n-1}(t)\xi_{n-1}(t-1) = \varrho\beta_{n-1,n-1} + \frac{\varrho s_n}{(1-\varrho^2)^2},$$
$$\beta_{n-1,n-1}(1-\varrho^2) = s_n\frac{1+\varrho^2}{(1-\varrho^2)^2} + s_{n-1}.$$

By this method all the elements of $B(0)$ can be calculated.

We have been assuming that the components of the process $\zeta(t)$ are independent, and at the same time that the determinant of the transformation of the variables $\xi(1), \ldots, \xi(N)$ into $\xi(1), \zeta(2), \ldots, \zeta(N)$ according to (1.1) is equal to one; therefore

$$p_{\xi(1),\dots,\xi(N)}(x_{11},\dots,x_{1N};\,x_{21},\dots,x_{2N};\,\dots;\,x_{n1},\dots,x_{nN})$$

$$= (2\pi)^{-\frac{Nn}{2}} |R_{Nn}|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2} X R_{Nn}^{-1} X^*\right\}$$

(2.7)

$$= p_{\xi(1)}(x_{11},\dots,x_{n1})(2\pi)^{-\frac{(N-1)n}{2}} \left(\prod_{i=1}^{n} s_i\right)^{-\frac{N-1}{2}}$$

$$\cdot \exp\left\{-\sum_{j=1}^{N-1} \sum_{i+1}^{n} \frac{1}{2s_i}(x_{ij+1}-\alpha_{i1}x_{1j}-\dots-\alpha_{in}x_{nj})^2\right\}$$

where

(2.8) $$p_{\xi(1)}(x_{11},\dots,x_{n1}) = (2\pi)^{-\frac{n}{2}} |B(0)|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2} X_1 B(0)^{-1} X_1^*\right\}.$$

The above relations give us

(2.9) $$|R_{Nn}| = |B(0)| \cdot \left(\prod_{i=1}^{N} s_i\right)^{N-1},$$

where the matrix R_{Nn}^{-1} is

$$\begin{pmatrix}
a_{10} & b_{11} & 0 & \cdots\cdots & 0 & a'_{12} & b_{12} & 0 & \cdots\cdots & 0 & \cdots & a'_{1n} & b_{1n} & 0 & \cdots\cdots & 0 \\
b_{11} & a_{11} & b_{11} & \cdots\cdots & 0 & b'_{12} & a_{12} & b_{12} & \cdots\cdots & 0 & \cdots & b'_{1n} & a_{1n} & b_{1n} & \cdots\cdots & 0 \\
\vdots & & & & & \vdots & & & & & & & & & & \vdots \\
0 & & & a_{11} & b_{11} & 0 & & & & a_{12} & b_{12} & & & & a_{1n} & b_{1n} \\
0 & & & b_{11} & a''_{12} & 0 & & & & a'_{12} & a''_{12} & & & & b'_{1n} & a''_{1n} \\
a'_{12} & b'_{12} & 0 & \cdots & 0 & 0 & a_{20} & b_{22} & 0 & \cdots & 0 & \cdots & & & & \\
b_{12} & a_{12} & b'_{12} & \cdots & 0 & 0 & b_{22} & a_{22} & b_{22} & \cdots & 0 & & & & & \\
\vdots & & & & & \vdots & & & & & \vdots & & & & & \\
0 & & & & & a''_{12} & & & & a_{22} & b_{22} & & & & & \\
 & & & & & & & & & b_{22} & a''_{22} & \cdots & & & & \\
\vdots & & & & & & & & & & & & & & & \\
a'_{1n} & b'_{1n} & 0 & \cdots & 0 & 0 & & & & & & a_{n0} & b_{nn} & 0 & \cdots & 0 \\
b_{1n} & a_{1n} & b'_{1n} & \cdots & 0 & 0 & & & & & & b_{nn} & a_{nn} & b_{nn} & \cdots & 0 \\
\vdots & & & & & & & & & & & \vdots & & & & \\
0 & & & \cdots & a_{1n} & b'_{1n} & & & & & & 0 & & \cdots & a_{nn} & b_{nn} \\
0 & & & \cdots & b_{1n} & a''_{1n} & & & & & & 0 & & \cdots & b_{nn} & a''_{n0}
\end{pmatrix}$$

(2.10)

with the elements

(2.11)

$$a_{i0} = \beta_{ii} + \sum_{j=0}^{n} \frac{(\alpha_{ij})^2}{s_j}; \quad a''_{i0} = \frac{1}{s_i};$$

$$a'_{ik} = \beta_{ik} + \sum_{j=1}^{n} \frac{\alpha_{ij}\alpha_{ik}}{s_j}; \quad a''_{ik} = 0, \quad \text{for } k \neq i,\ k \neq 0;$$

$$a_{ii} = \frac{1+(\alpha_{ii})^2}{s_i} + \sum_{j\neq i} \frac{(\alpha_{ji})^2}{s_j}; \quad a_{ik} = \sum_{j=1}^{n} \frac{\alpha_{ji}+\alpha_{jk}}{s_j}, \quad k \neq i;$$

$$b_{ik} = -\frac{\alpha_{ki}}{s_k}; \quad b'_{ik} = -\frac{\alpha_{ik}}{s_i}; \quad B(0) = \{\beta_{ik}\}_{i,k=1}^{n}.$$

Hence the system

$$\Bigg(\sum_{j=2}^{N} x_{ij}^2,\ \sum_{j=1}^{N} x_{i_1 j}x_{i_2 j},\ \sum_{j=1}^{N-1} x_{i_1 j}x_{i_2 j+1}, \quad i, i_1, i_2 = 1,\dots,n,$$

$$x_{11}^2,\dots,x_{n1}^2, x_{1N}^2,\dots,x_{nN}^2, x_{11}x_{21}, x_{11}x_{31},\dots,x_{11}x_{n1},$$

$$x_{21}x_{31},\dots,x_{21}x_{n1},\dots,x_{n-1\,1}x_{n1}\Bigg)$$

will be a sufficient statistic.

Let us investigate the special case $A = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}$ as an example. Here evidently

$$A^t = e^{-\lambda t}\begin{pmatrix} \cos\omega\tau & \sin\omega\tau \\ -\sin\omega\tau & \cos\omega\tau \end{pmatrix}$$

and

$$B(0) = \frac{1}{1-\alpha^2-\beta^2}\begin{pmatrix} s & 0 \\ 0 & s \end{pmatrix},$$

where

$$\beta_{11} = \frac{\gamma[s_1(1-\alpha^2+\beta^2)+2\beta\alpha s_{12}]+\delta[s_2(1-\alpha^2+\beta^2)-2\alpha\beta s_{12}]}{\gamma^2-\delta^2},$$

$$\beta_{12} = \frac{\beta_{22}\alpha\beta-\beta_{11}\alpha\beta+s_{12}}{1-\alpha^2+\beta^2},$$

$$\beta_{22} = \frac{\delta[s_1(1-\alpha^2+\beta^2)+2\alpha\beta s_{12}]+\gamma[s_2(1-\alpha^2+\beta^2)-2\alpha\beta s_{12})]}{\gamma^2-\delta^2},$$

$$e^{-2\lambda} = \alpha^2+\beta^2, \quad \tan\omega = \frac{\beta}{\alpha}, \quad \gamma = (1-\alpha^2)^2+\beta^2(1+\alpha^2), \quad \delta = \beta^2(1+\alpha^2+\beta^2),$$

In particular, for $s_{12} = 0$ and $s_{11} = s_{22} = s$

$$B(0) = \frac{1}{1-\alpha^2-\beta^2}\begin{pmatrix} s & 0 \\ 0 & s \end{pmatrix}.$$

The density function could be written down at least formally, even if the components of the random vector $\zeta(k)$ are not independent, but the corresponding relations are too complicated to be given here.

Later we shall have to determine the density functions of processes satisfying a differential equation of the Jordan type (with one block only).

Let $\xi(t) = \{\xi_k(t)\}_{k=1}^n$ be such a process. Then the matrix R_{Nn}^{-1} has the form

$$\begin{pmatrix} A_{11} & A_{12} & 0 & 0 \ldots & 0 & 0 \\ A'_{12} & A_{22} & A_{23} & 0 \ldots & 0 & 0 \\ \vdots & & & & & \\ 0 & & & & A_{n-1\,n-1} & A_{n-1\,n} \\ 0 & & & & A'_{n-1\,n} & A_{nn} \end{pmatrix}, \tag{2.12}$$

where the matrices A_{ik} are

$$A_{ii+1} = \begin{pmatrix} a_{ii+1} & 0 & 0 \ldots & & 0 \\ -\frac{1}{s_i} & \frac{\varrho}{s_i} & 0 \ldots & & 0 \\ \cdot & & \cdot & & \\ \cdot & & & \cdot\ \frac{\varrho}{s_i} & 0 \\ \cdot & & & & \\ 0 & & & -\frac{1}{s_i} & 0 \end{pmatrix}, \tag{2.13}$$

$$A_{ii} = \begin{pmatrix} a_{ii} & -\frac{\varrho}{s_i} & 0 & \ldots & & 0 \\ -\frac{\varrho}{s_i} & \frac{1+\varrho^2}{s_i} & -\frac{\varrho}{s_i} & \ldots & & 0 \\ 0 & -\frac{\varrho}{s_i} & \frac{1+\varrho^2}{s_i} & \ldots & & 0 \\ \vdots & & & & & \vdots \\ \vdots & & & & \frac{1+\varrho^2}{s_i} & -\frac{\varrho}{s_i} \\ 0 & & & & -\frac{\varrho}{s_i} & \frac{1}{s_i} \end{pmatrix}. \tag{2.14}$$

In particular, for $n = 2$ we can readily calculate

$$a_{11} = \frac{(1-\rho^2)^3}{s_1(1-\rho^2)+s_2} + \frac{\rho^2}{s_1},$$

$$a_{12} = -\frac{\rho(1-\rho^2)^2}{s_1(1-\rho^2)^2+s_2} + \frac{\rho}{s_1},$$

$$a_{22} = \frac{(1-\rho^2)[s_1(1-\rho^2)^2+s_2(1+\rho^2)]}{s_2[s_1(1-\rho^2)^2+s_2]} + \frac{1}{s_1} + \frac{\rho^2}{s_2}.$$

In general, the investigation of the n-dimensional process $\xi(t)$ satisfying the difference equation

$$\xi(t) + C_1\xi(t-1) + \cdots + C_p\xi(t-p) = \xi(t)$$

seems to be quite natural as a generalization of the one-dimensional autocorrelation scheme. We shall not deal with such processes in general, although the relations for the correlation matrix can be formulated similarly to the case of $p = 1$.

§3. Regression problems

Actually we most often observe not the n-dimensional process $\xi(t)$, but the process

$$\eta(t) = C \cdot h(t) + \xi(t)$$

where the functions $h(t) = \{h_k(t)\}_{k=1}^m$ are known while the elements c_{ij} of the matrix $C = \{c_{ij}\}_{i=1,\ldots,n}^{j=1,\ldots,m}$ are unknown.

When observing the process $\eta(t)$, the estimation of regression coefficients c_{ij} forms one of the central problems of the theory of time series. In practice the case occurring most often is that of polynomials or trigonometric polynomials $h_k(t)$. Using the above results, we shall deal now with maximum likelihood estimators, and we shall investigate some of their properties.

Assuming that the equation describing a discrete process already has the form (1.4), let us investigate the density of a process described by one Jordan elementary matrix with real eigenvalue. Let us assume that the components of the variable $\zeta(t)$ are independent, although such an assumption must always be verified in practice. We shall not deal separately with the complex eigenvalue case. Under the assumptions we have made, the conditional density of the random variables $\xi(2), \ldots, \xi(N)$ under the condition $\xi(1) = x(1)$ has the form

$$C_N \exp -\left\{\sum_{j=1}^{N-1}\left[\frac{1}{2s_1}(x_{1j+1}-\varrho x_{1j}-x_{2j})^2+\frac{1}{2s_2}(x_{2j+1}-\varrho x_{2j}-x_{3j})^2+\ldots\right.\right.$$
$$\left.\left.+\frac{1}{2s_{n-2}}(x_{n-1\,j+1}-\varrho x_{n-1\,j}-x_{nj})^2+\frac{1}{2s_n^2}(x_{nj+1}-\varrho x_{nj})^2\right]\right\}.$$

The density function of $\xi(1)$ can be calculated by using (2.6) and by determining $B(0)$ ($\mathsf{E}\xi_j^2 = s_j$). In the preceding section we have seen that the conditional density function is

$$C_N \exp -\frac{1}{2}\left\{\frac{\varrho^2}{s_1}x_{11}^2+\left(\frac{1}{s_1}+\frac{\varrho^2}{s_2}\right)x_{21}^2+\left(\frac{1}{s_2}+\frac{\varrho^2}{s_3}\right)x_{31}^2+\ldots+\left(\frac{1}{s_{n-1}}+\frac{\varrho^2}{s_n}\right)x_{n1}^2+\right.$$
$$+\frac{1}{s_1}x_{1N}^2+\frac{1}{s_2}x_{2N}^2+\ldots \qquad\qquad +\frac{1}{s_n}x_{nN}^2+$$
$$+\frac{2\varrho}{s_1}x_{11}x_{21}+\frac{2\varrho}{s_2}x_{21}x_{31}+\ldots \qquad\qquad +\frac{2\varrho}{s_{n-1}}x_{n-11}x_{n1}+$$
$$+\frac{1+\varrho^2}{s_1}\sum_2^{N-1}x_{1i}^2+\ldots \qquad\qquad +\frac{1+\varrho^2}{s_n}\sum_2^{N-1}x_{ni}^2+$$
$$-\frac{2\varrho}{s_1}\sum_1^{N-1}x_{1i+1}x_{1i}+\ldots \qquad\qquad -\frac{2\varrho}{s_n}\sum_1^{N-1}x_{ni+1}x_{ni}+$$
$$\left.-\frac{2}{s_1}\sum_1^{N-1}x_{1i+1}x_{2i}+\ldots \qquad\qquad -\frac{2}{s_{n-1}}\sum_1^{N-1}x_{n-1i+1}x_{ni}\right\}.$$

For the sake of simplicity we shall only deal with the conditional maximum likelihood equations (under the condition $\xi(1) = x(1)$). In particular, let $h(t) \equiv 1 = (1, \ldots, 1)$, which means that we want to estimate the expected values $m_1, \ldots, m_n$ of the process $\xi(t)$. The conditional maximum likelihood equations are the following:

$$-\frac{2\rho^2}{s_1}(x_{11}-m_1)-\frac{2}{s_1}(x_{1N}-m_1)-\frac{2\varrho}{s_1}(x_{21}-m_2)-\frac{2(1+\varrho^2)}{s_1}\sum_2^{N-1}(x_{1i}-m_1)$$
$$+\frac{2\varrho}{s_1}\sum_{i=1}^{N-1}[(x_{1i}-m_1)+(x_{1i+1}-m_1)]+\frac{2}{s_1}\sum_{i=1}^{N-1}(x_{2i}-m_2)=0,$$
$$-2\left(\frac{1}{s_1}+\frac{\varrho^2}{s_2}\right)(x_{21}-m_2)-\frac{2}{s_2}(x_{2N}-m_2)-\frac{2\varrho}{s_1}(x_{11}-m_1)$$
$$-\frac{2\varrho}{s_2}(x_{31}-m_3)-\frac{2(1+\varrho^2)}{s_2}\sum_2^{N-1}(x_{2i}-m_2)+\frac{2\varrho}{s_2}\sum_1^{N-1}[(x_{2i}-m_2)+(x_{2i+1}-m_2)]$$
$$+\frac{2}{s_1}\sum_1^{N-1}(x_{1i+1}-m_1)+\frac{2}{s_2}\sum_1^{N-1}(x_{3i}-m_3)=0,$$

$$\vdots$$
$$-2\left(\frac{1}{s_{k-1}}+\frac{\varrho^2}{s_k}\right)(x_{k1}-m_k)-\frac{2}{s_k}(x_{kN}-m_k)-\frac{2\varrho}{s_{k-1}}(x_{k-11}-m_{k-1})-\frac{2\varrho}{s_k}(x_{k+1,1}-m_{k+1})$$
$$-\frac{2(1+\varrho^2)}{s_k}\sum_2^{N-1}(x_{ki}-m_k)+\frac{2\varrho}{s_k}\sum_1^{N-1}[(x_{ki}-m_k)+(x_{k-1i+1}-m_k)]$$
$$+\frac{2}{s_{k-1}}\sum_1^{N-1}(x_{k-1i+1}-m_{k-1})+\frac{2}{s_k}\sum_1^{N-1}(x_{k+1i}-m_{k+1})=0,$$
$$\vdots$$
$$-2\left(\frac{1}{s_{n-1}}+\frac{\rho^2}{s_n}\right)(x_{n1}-m_n)-\frac{2}{s_n}(x_{nN}-m_n)-\frac{2\rho}{s_{n-1}}(x_{n-1,1}-m_{n-1})$$
$$-\frac{2(1+\rho^2)}{s_n}\sum_2^{N-1}(x_{ni}-m_n)+\frac{2\rho}{s_n}\sum_1^{N-1}[(x_{ni}-m_n)+(x_{ni+1}-m_n)]$$
$$-\frac{2}{s_{n-1}}\sum_1^{N-1}(x_{n-1i+1}-m_{n-1})=0.$$

After simple but elaborate calculations we obtain the following system of equations:

$$\frac{\varrho^2-\varrho}{s_1}x_{11}+\frac{1-\varrho}{s_1}x_{1N}+\frac{\varrho-1}{s_1}x_{21}+\frac{(1-\varrho)^2}{s_1}\sum_2^{N-1}x_{1i}-\frac{1}{s_1}\sum_2^{N-1}x_{2i}$$
$$=\frac{(1-\varrho)^2}{s_1}m_1+\frac{\varrho-1}{s_1}m_2+\frac{(1-\varrho)^2}{s_1}(N-2)m_1-\frac{(N-2)}{s_1}m_2,$$
$$\left(\frac{1}{s_1}+\frac{\varrho^2-\varrho}{s_2}\right)x_{21}+\frac{1-\varrho}{s_2}x_{2N}+\frac{\varrho}{s_1}x_{11}-\frac{1}{s_1}x_{1N}+\frac{\varrho-1}{s_2}x_{31}+\frac{(1-\varrho)^2}{s_2}\sum_2^{N-1}x_{2i}$$
$$-\frac{1}{s_1}\sum_2^{N-1}x_{1i}-\frac{1}{s_2}\sum_2^{N-1}x_{3i}=\left(\frac{1}{s_1}+\frac{(1-\varrho)^2}{s_2}\right)m_2+\frac{\varrho-1}{s_1}m_1+\frac{\varrho-1}{s_2}m_3$$
$$+\frac{(1-\varrho)^2}{s_2}(N-2)m_2-\frac{(N-2)}{s_1}m_1-\frac{(N-2)}{s_2}m_3,$$
$$\vdots$$
$$\left(\frac{1}{s_{k-1}}+\frac{\varrho^2-\varrho}{s_k}\right)x_{k1}+\frac{1-\varrho}{s_k}x_{kN}+\frac{\varrho}{s_{k-1}}x_{k-11}-\frac{1}{s_{k-1}}x_{k-1N}+\frac{\varrho-1}{s_k}x_{k+1,1}$$
$$+\frac{(1-\varrho)^2}{s_k}\sum_2^{N-1}x_{ki}-\frac{1}{s_{k-1}}\sum_2^{N-1}x_{k-1i}-\frac{1}{s_k}\sum_2^{N-1}x_{k+1i}=\left(\frac{1}{s_{k-1}}+\frac{(1-\varrho)^2}{s_k}\right)m_k$$
$$+\frac{\varrho-1}{s_{k-1}}m_{k-1}+\frac{\varrho-1}{s_k}m_{k+1}+\frac{(1-\varrho)^2}{s_k}(N-2)m_k-\frac{N-2}{s_{k-1}}m_{k-1}-\frac{N-2}{s_k}m_{k+1},$$
$$\vdots$$
$$\left(\frac{1}{s_{n-1}}+\frac{\varrho^2-\varrho}{s_n}\right)x_{n1}+\frac{1-\varrho}{s_n}x_{nN}+\frac{\varrho}{s_{n-1}}x_{n-11}-\frac{1}{s_{n-1}}x_{n-1N}+\frac{(1-\varrho)^2}{s_n}\sum_2^{N-1}x_{ni}$$
$$-\frac{1}{s_{n-1}}\sum_2^{N-1}x_{n-1i}=\left(\frac{1}{s_{k-1}}+\frac{(1-\varrho)^2}{s_k}(N-1)\right)m_n+\frac{\varrho-(N-1)}{s_{n-1}}m_{n-1},$$

which, in fact, has the form

$$\begin{aligned} C_1X &= a_{11}m_1 + a_{12}m_2, \\ C_2X &= a_{21}m_1 + a_{22}m_2 + a_{23}m_3, \\ &\vdots \\ C_kX &= a_{kk-1}m_{k-1} + a_{kk}m_k + a_{kk+1}m_{k+1}, \\ &\vdots \\ C_nX &= a_{nn-1}m_{n-1} + a_{nn}m_n, \end{aligned}$$

where the meaning of the constants is clear from the above system.

When all the variances are equal, for instance $s_i = s$, the matrix A of the system $b = A \cdot m$ has the form

$$A = \begin{pmatrix} a_0 & b & 0 & 0 & \dots & 0 \\ b & a & b & 0 & \dots & 0 \\ 0 & b & a & b & \dots & 0 \\ \vdots & & \ddots & & \ddots & \\ \vdots & & & & a & b \\ 0 & & & & b & a \end{pmatrix}.$$

The inverse of A can be calculated by again using probabilistic reasoning. In particular, when $n = 2$ the system outlined above takes the form

$$\frac{\rho^2 - \rho}{s_1} x_{11} + \frac{1-\rho}{s_1} x_{1N} + \frac{\rho - 1}{s_1} x_{21} + \frac{(1-\rho)^2}{s_1} \sum_2^{N-1} x_{1i} - \frac{1}{s_1} \sum_2^{N-1} x_{2i}$$

$$= \frac{(1-\rho)^2(N-1)}{s_1} m_1 - \left(\frac{N-2}{s_1} + \frac{1-\rho}{s_1}\right) m_2,$$

$$\left(\frac{1}{s_1} + \frac{\rho^2 - \rho}{s_2}\right) x_{21} + \frac{1-\rho}{s_2} x_{2N} + \frac{\rho}{s_1} x_{11} - \frac{1}{s_1} x_{1N} - \frac{(1-\rho)^2}{s_2} \sum_2^{N-1} x_{2i}$$

$$- \frac{1}{s_1} \sum_2^{N-1} x_{1i} = \frac{\rho - N - 1}{s_1} m_1 + m_2 \left(\frac{1}{s_1} + \frac{(N-1)(1-\rho)^2}{s_2}\right).$$

It is worth mentioning that, contrary to the one-dimensional case, the expected value can be well estimated by the arithmetical mean of the observations even in the cases when ρ has a value close to one.

In this special case ($n = 2$) the unconditional maximum likelihood equations are the following (omitting detailed calculations):

$$\left\{\frac{\varrho^2 - \varrho}{s_1} + \frac{(1-\varrho^2)^3}{s_1(1-\varrho^2)^2 + s_2}\right\} x_{11} + \frac{1-\varrho}{s_1} x_{1N} + \left\{\frac{\varrho - 1}{s_1} - \frac{\varrho(1-\varrho^2)^2}{s_1(1-\varrho^2)^2 + s_2}\right\} x_{21}$$

$$+ \frac{(1-\varrho)^2}{s_1} \sum_2^{N-1} x_{1i} - \frac{1}{s_1} \sum_2^{N-1} x_{2i} = \left[\frac{(1-\varrho)^2(N-1)}{s_1} + \frac{(1-\varrho^2)^3}{s_1(1-\varrho^2)^2 + s_2}\right] m_1$$

$$- \left[\frac{N-2}{s_1} + \frac{1-\varrho}{s_1} + \frac{\varrho(1-\varrho^2)^2}{s_1(1-\varrho^2)^2 + s_2}\right] m_2,$$

$$\left\{\frac{1}{s_1} + \frac{\rho^2 - \rho}{s_2} + \frac{(1-\rho^2)[s_1(1-\rho^2)^2 + s_2(1+\rho^2)]}{s_2[s_1(1-\rho^2)^2 + s_2]}\right\} x_{21} + \frac{(1-\rho)}{s_2} x_{2N}$$

$$+ \left(\frac{\rho}{s_1} - \frac{\rho(1-\rho^2)^2}{s_1(1-\rho^2)^2 + s_2}\right) x_{11} - \frac{1}{s_1} x_{1N} + \frac{(1-\rho)^2}{s_2} \sum_2^{N-1} x_{2i} - \frac{1}{s_1} \sum_2^{N-1} x_{1i}$$

$$= \left[\frac{\rho}{s_1} - \frac{\rho(1-\rho^2)^2}{s_1(1-\rho^2)^2 + s_2} - \frac{N-1}{s_1}\right] m_1$$

$$+ \left[\frac{1}{s_1} + \frac{(1-\rho^2)[s_1(1-\rho^2)^2 + s_2(1+\rho^2)]}{s_2[s_1(1-\rho^2)^2 + s_2]} + \frac{(N-1)(1-\rho)^2}{s_2}\right] m_2.$$

If $\rho \sim 1$, then from the approximate equations the following estimations can be derived:

$$m_2 = \frac{1}{N-2} \sum_2^{N-1} x_{2i},$$

$$m_1 = \frac{x_{1N} - x_{11}}{N-2} + \frac{1}{N-2} \sum_2^{N-1} x_{1i} + \frac{1}{(N-2)^2} \sum_2^{N-1} x_{2i} - \frac{x_{21}}{N-2}.$$

From this result it can be seen that, contrary to the one-dimensional case, the estimate by the arithmetic mean is always good.

§4. Estimation of the elements of the matrix A

Let us suppose that $\mathsf{E}\xi(t) = 0$ and the Gaussian process $\xi(t)$ satisfies equation (1.2). In principle, the maximum likelihood estimates of the elements of the unknown matrix A can easily be made from observations of the process $\xi(t)$ at $t = 1, \dots, N$ by using the corresponding density function. For the sake of simplicity let us again consider the conditional density function, which can be handled more easily, as well as the conditional maximum likelihood estimates (we have in mind the n-dimensional analog of estimates of the so-called serial correlation; cf. [2], §1). Let $A = \{\alpha_{ij}\}_{i,j=1}^n$ and let the component variables $\zeta(i)$ be independent, with $\mathsf{E}(\zeta_j(i))^2 = s_j$ and $\mathsf{E}(\zeta_{j_1}(i)\zeta_{j_2}(k)) = 0$ for $j_1 \neq j_2$; then the conditional density of the variables $\xi(2), \dots, \xi(N)$ (under the condition $\xi(1) = x(1)$) is

$$p_{\xi(2),\dots,\xi(N)}(x_{21}, \dots, x_{2N}; \dots; x_{n1}, \dots, x_{nN} \mid \xi(1) = x(1))$$

$$= (2\pi)^{-\frac{n(N-1)}{2}} \left(\prod_1^n s_i\right)^{-\frac{N-1}{2}} \exp\left\{-\sum_{i=1}^n \frac{1}{s_i} \sum_{j=1}^{N-1} (x_{ij+1} - \alpha_{i1}x_{1j} - \alpha_{i2}x_{2j} - \dots - \alpha_{in}x_{nj})^2\right\}.$$

The maximum likelihood equations are

$$
\begin{aligned}
&\frac{\partial \log p}{\partial \alpha_{i1}} = \sum_{j=1}^{N-1}(x_{ij+1}-\alpha_{i1}x_{1j}-\alpha_{i2}x_{2j}-\ldots-\alpha_{in}x_{nj})x_{1j} = 0,\\
&\quad\vdots \\
(4.1)\qquad &\frac{\partial \log p}{\partial \alpha_{in}} = \sum_{j=1}^{N-1}(x_{ij+1}-\alpha_{i1}x_{1j}-\alpha_{i2}x_{2j}-\ldots-\alpha_{in}x_{nj})x_{nj} = 0,\\
&\qquad i = 1, \ldots, n.
\end{aligned}
$$

By considering new random variables

$$
\begin{aligned}
\eta_{i1}(N) &= \frac{1}{\sqrt{N-1}} \sum_{j=1}^{N-1} (\xi_i(j+1) - \alpha_{i1}\xi_1(j) - \alpha_{i2}\xi_2(j) - \cdots - \alpha_{in}\xi_n(j))\xi_1(j) \\
&= \frac{1}{\sqrt{N-1}} \sum_{1}^{N-1} \zeta_i(j+1)\xi_1(j), \\
(4.2)\qquad \eta_{i2}(N) &= \frac{1}{\sqrt{N-1}} \sum_{j=1}^{N-1} \zeta_i(j+1)\xi_2(j), \\
&\ \ \vdots \\
\eta_{in}(N) &= \frac{1}{\sqrt{N-1}} \sum_{j=1}^{N-1} \zeta_i(j+1)\xi_n(j)
\end{aligned}
$$

and denoting the solutions of the system (4.1) for the observations $\xi(1), \ldots, \xi(N)$ by $\hat{\alpha}_{ij}$, these variables can be written as follows:

$$
\begin{aligned}
\eta_{ik}(N) &= \frac{1}{\sqrt{N-1}}\sum_{j=1}^{N-1}[(\hat{\alpha}_{i1}-\alpha_{i1})\xi_1(j)+(\hat{\alpha}_{i2}-\alpha_{i2})\xi_2(j)+\ldots+(\hat{\alpha}_{in}-\alpha_{in})\xi_n(j)]\xi_k(j) \\
(4.3)\qquad & i,k=1,\ldots,n
\end{aligned}
$$

using the relation

$$\sum_{j=1}^{N-1}\xi_i(j+1)\xi_k(j) = \sum_{j=1}^{N-1}(\hat{\alpha}_{i1}\xi_1(j)+\ldots+\hat{\alpha}_{in}\xi_n(j))\xi_k(j)$$

which follows from (4.1).

The equation (4.3) can be written in the form

$$
\begin{aligned}
\eta_{ik}(N) &= \sqrt{N-1}\,(\hat{\alpha}_{i1}-\alpha_{i1})\sum_{j=1}^{N-1}\frac{\xi_k(j)\xi_1(j)}{N-1}+\ldots \\
(4.3')\qquad &\ldots+\sqrt{N-1}\,(\hat{\alpha}_{in}-\alpha_{in})\sum_{j=1}^{N-1}\frac{\xi_k(j)\xi_n(j)}{N-1} \qquad i,k=1,\ldots,n.
\end{aligned}
$$

Equations (4.2) enable us to derive

$$(4.4)\qquad \mathsf{E}\eta_{il_1}(N)\eta_{il_2}(N) = s_i\,\mathsf{E}\xi_{l_1}(j)\xi_{l_2}(j)$$

using the properties of the variables $\zeta_i(j)$.

On the other hand, when the eigenvalues of A are less than 1 the process is ergodic, and therefore

$$(4.5)\qquad \frac{1}{N-1}\sum_{j=1}^{N-1}\xi_i(j)\xi_k(j) \to \mathsf{E}\xi_i(j)\xi_k(j), \qquad i,k=1,\ldots,n,$$

with probability 1.

The variables $\eta_{ik}(N)$ have asymptotically normal distribution (when $N \to \infty$) with covariance matrix

$$s_i\{\mathsf{E}\xi_{l_1}(j)\xi_{l_2}(j)\}_{l_1,l_2=1}^{n} = \{\mathsf{E}\eta_{il_1}(N)\eta_{il_2}(N)\}$$

for α_{ik} ($i, k = 1, \ldots, n$) fixed (see, for example, Rozanov [10]). By taking (4.3′) into consideration and using (4.5) we can see that the variables $\eta_{ik}(N)$ ($k = 1, \ldots, n$) depend linearly on the variables $\sqrt{N-1}(\hat{\alpha}_{ik} - \alpha_{ik})$ ($k = 1, \ldots, n$) and, therefore, that they are also asymptotically normally distributed with covariance matrix

$$s_i \cdot \{\mathsf{E}\xi_{l_1}(j)\xi_{l_2}(j)\}^{-1}.$$

The estimators $\hat{\alpha}_{ij}$ are evidently unbiased as well as consistent.

Let us note that the system

$$\sum_{j=1}^{N-\tau-1}[\xi_i(j+\tau) - \alpha_{i1}\xi_1(j+\tau-1) - \cdots - \alpha_{in}\xi_n(j+\tau-1)]\,\xi_k(j) = 0,$$
$$i, k = 1, \ldots, n,$$

for $\tau = 1, \ldots, r$ can be used instead of (4.1) to determine the estimates of α_{ik}. We shall denote by $\hat{\alpha}_{ik}$ the corresponding estimators. In this way we obtain different estimates for different values of τ, and thus the following question arises: For which systems shall we obtain better results when calculating expected values of these estimates? As long as the system is of stationary Gaussian Markov type, the estimates do not improve, but in systems which are not very different from the ones mentioned (e.g. in the case of non-Gaussian systems) this approach could be useful.

So far we have dealt with the known procedure of obtaining estimates of unknown parameters α_{ik}. We shall now consider the problem of reliability of these estimates and the difficulties which may arise in investigating them. More precisely, the problem is whether the eigenvalues of the matrix A can be determined using the estimates described above. This means, for example, whether we can bring the system to the Jordan form for large values of N (and how large a value of N for what accuracy).

For this purpose let us investigate, as a first step, the estimate of a single parameter ρ of the n-dimensional process $\xi(t)$ satisfying a system of the form (1.4). From (2.12) it follows that the maximum likelihood estimate based on the conditional density function $\hat{\rho}$ under the condition $\xi(1) = x(1)$ will be

$$\frac{\sum_1^{N-1} x_{1i}x_{1i+1} - \sum_1^{N-1} x_{1i}x_{2i} + \sum_1^{N-1} x_{2i}x_{2i+1} - \cdots - \sum_1^{N-1} x_{n-1i}x_{ni} + \sum_1^{N-1} x_{ni}x_{ni+1}}{\sum_{k=1}^{n} \sum_{i=1}^{N-1} x_{ki}^2}.$$

From (1.4) it can easily be seen that

$$\hat{\varrho} - \varrho = \frac{\sum_1^{N-1} \zeta_{1i}\zeta_{1i+1} + \cdots + \sum_1^{N-1} \xi_{n1}\zeta_{ni+1}}{\sum_{k=1}^{n} \sum_{i=1}^{N-1} \xi_{ki}^2}$$

and the expected value of the numerator on the right-hand side equals zero (because $\mathsf{E}\xi_{ki}\zeta_{ki+1} = 0$). Therefore the estimate is asymptotically unbiased. For large values of N, because of ergodicity, the denominator asymptotically equals $(N-1)\Sigma_1^n \sigma_k^2$; on the other hand, the expected value of the squared numerator is $\Sigma s_k \sigma_k^2$. Hence

$$\mathsf{E}\left(\sqrt{N-1}(\hat{\varrho}-\varrho)\right)^2 \sim \frac{\Sigma s_k \sigma_k^2}{(\Sigma \sigma_k^2)^2} \quad \text{for} \quad N \to \infty.$$

In particular, when $n = 2$ and $s_1 = s_2$

$$\mathsf{E}(\sqrt{N-1}(\hat{\rho} - \rho))^2 \sim \frac{(1-\rho^2)^3}{2(1-\rho^2)^2 + (1+\rho^2)}$$

(see [2], §2, where we had $(1-\rho^2)$ for the variance of $\sqrt{N-1}(\hat{\rho}-\rho)$ in the one-dimensional case).

As in the one-dimensional case, we have to investigate the asymptotic distribution of the above estimate for values of ρ which are close to 1. We have seen in the one-dimensional case that for $\sigma_\xi^2 = 1$ the estimate of ρ was uniformly asymptotically normally distributed for $-1 < \rho < 1$, while in the case of $s = 1$ this is by no means so. Especially for this reason we had to determine the distribution of the estimates of the parameters in the case $s = 1$ for continuous-time processes, and in the case $\sigma_\xi^2 = 1$ we had to determine the asymptotic distribution of the estimates of parameters of a discrete process. This distinction has to be kept in the n-dimensional case as well, but performing the corresponding calculations is indeed very elaborate. On the basis of heuristic considerations we obtain the following: the estimate of ρ is uniformly asymptotically normal for $-1 < \rho < 1$

in the case when $\sigma_1^2 = 1$ and $\sigma_2^2 = c\ (\neq 1)$, while in the case $s_1 = s_2 = 1$ this is not true. The corresponding distribution of the estimates in the two-dimensional continuous-time case will be given in another paper.

Turning back to the general case, we may assert:

From Theorem 3.3 of [1] and from the corollary of Theorem 5.4 in [2] it follows that when all the eigenvalues of the matrix A in (1.2) are real and simple and we have observed the process $\eta(t) = \xi(t) + m$, then *it is not possible to construct finite confidence intervals for the expected value m; for eigenvalues $\rho_1, \ldots, \rho_n$ nonzero lower confidence limits using continuous functionals cannot be constructed.*

This means, at the same time, that the values $\rho_1, \ldots, \rho_n$ cannot be distinguished.

To the extent that the matrix A has multiple and complex eigenvalues, the arithmetic means are good estimates of the corresponding expected values, and there is no need of giving infinite confidence intervals.

We do not give the corresponding theorems.

BIBLIOGRAPHY

1. M. Arató, *On the statistical examination of continuous state Markov processes.* I, Magyar Tud. Akad. Mat. Fiz. Oszt. Közl. **14** (1964), 13–34; English transl., Selected Transl. Math. Statist. and Probability, vol. 14, Amer. Math. Soc., Providence, R. I., 1978, pp. 203–225. MR 37 #3645.

2. ———, *On the statistical examination of continuous state Markov processes.* II, Magyar Tud. Akad. Mat. Fiz. Oszt. Közl. **14** (1964), 137–159; English transl., Selected Transl. Math. Statist. and Probability, vol. 14, Amer. Math. Soc., Providence, R. I., 1978, pp. 227–251. MR 37 #5998.

3. ———, *On the statistical examination of continuous state Markov processes.* III, Magyar Tud. Akad. Mat. Fiz. Oszt. Közl. **14** (1964), 317–330; English transl., Selected Transl. Math. Statist. and Probability, vol. 14, Amer. Math. Soc., Providence, R. I., 1978, pp. 253–267. MR 37 #5999.

4. ———, *On statistical problems concerning stationary Markov processes,* Dissertation, Moscow State Univ., Moscow, 1962. (Russian)

5. ———, *Sufficient statistics for stationary Gaussian processes,* Teor. Verojatnost. i Primenen. **6** (1961), 216–218 = Theor. Probability Appl. **6** (1961), 199–200. MR **27** #829.

6. J. L. Doob, *The elementary Gaussian processes,* Ann. Math. Statist. **15** (1944), 229–282. MR **6**, 89.

7. E. B. Dynkin, *Necessary and sufficient statistics for a family of probability distributions,* Uspehi Mat. Nauk **6** (1951), no. 1 (41), 68–90; English transl., Selected Transl. Math. Statist. and Probability, vol. 1, Amer. Math. Soc., Providence, R. I., 1961, pp. 17–40. MR **12**, 839; **22** #7195.

8. Ulf Grenander and Murray Rosenblatt, *Statistical analysis of stationary time series,* Wiley, New York, 1957. MR **18**, 959.

9. L. S. Pontrjagin, *Ordinary differential equations,* Fizmatgiz, Moscow, 1961; English transl., Addison-Wesley, Reading, Mass., 1962. MR **23** #A362; **25** #4156.

10. M. H. Quenouille, *The analysis of multiple time series,* Hafner, New York, 1957. MR **19**, 1205.

11. Ju. A. Rozanov, *Stationary random processes,* Fizmatgiz, Moscow, 1963; English Holden-Day, San Francisco, Calif., 1967. MR 28 #2580; 35 #4985.

Translated by J. GREGOR